THE DEVELOPING BILINGUAL LEXICON

Annabelle David

Speech and Language Sciences Section
School of Education, Communication and Language Sciences
University of Newcastle upon Tyne

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Newcastle upon Tyne.
July, 2004

DECLARATION OF ORIGINALITY

The material presented in this thesis is the original work of the candidate except as otherwise acknowledged. It has not been submitted previously, in part or whole, for any award, at any university, at any other time.

Annabelle David

ACKNOWLEDGMENTS

Financial support was provided by a Research Studentship from the ESRC (R42200124115), as well as studentships from the University of Newcastle's Faculty of Humanities and Social Sciences and Alumni Fund;

I would like to sincerely thank the many people who have been involved, directly or indirectly, in the various stages of the research reported in this thesis and have contributed to making the past few years a great learning experience:

Barbara Pearson for her long distance support and help with some parts of the data analysis; Sophie Kern for allowing me to use the French adaptation of the CDI and part of her norming data;

Barbara Dodd, Thomas Klee, Zhu Hua and Ghada Khattab for their willingness to listen, comment, criticise and/or encourage; Sharon, Will, Jenni, Nicole, Juliane, John and all the others in the Speech section for their help and support; Graeme for checking my English;

Li Wei for his all round support, advice and help with every stage of the doctoral process. In addition to being a great source of inspiration, he has taught me a lot and given me numerous opportunities to try my hands at other aspects of academic life;

Thomas, Floriane, Antonia, Emma, Duncan, Pénélope, Elisa, Jack, Oliver R, Liam, Rebecca, Oliver B and, last but by no means least, Anna and their parents and family for being so patient and cooperative over such a long period of time. Needless to say that without the children and their family, this thesis would not have been possible;

My parents and brother for their constant interest, encouragement, love and unfailing confidence in all I do; Paulo, for his "consultancy services", chocolate supply, patience, love and unfading support. Sure, I could have written this thesis without him (actually, I *did* write this thesis without him! Not a word in here about micro-simulation!), but life without him would not be as exciting.

ABSTRACT

It is often said that bilinguals are not two monolinguals in one person. But what does this really mean, especially in the context of bilingual acquisition? Despite the upsurge of case studies of bilingual children since the 1990's, the main central issue within the literature has largely remained focused on the one-vs.-two-system debate. Earlier studies focused on the question of whether bilingual children had a single/fused system or two separate/differentiated ones. There are a growing number of more recent studies focusing, instead, on the relationship between the two languages in the developing language system of the child.

The study on which this thesis is based is the first longitudinal group study of lexical development of French-English bilingual children. The study aims to investigate the nature of the developing bilingual lexicon and its impact on the development of syntax. The key questions addressed in this new body of research include: are bilingual children developing in the same way and at the same rate as their monolingual peers; are there cross-linguistic influences on bilingual acquisition; are there features, patterns or processes specific to bilingual acquisition?

We report findings from a longitudinal group study of 13 children between 1;4 and 3;0 who are acquiring French and English simultaneously within the one person – one language framework.

The originality of this study lies in several main points. First of all, a larger number of children have been studied systematically than in traditional longitudinal studies, which are usually based on either cross-sectional sample or on single cases. Secondly, the children in this study have been systematically selected according to a set of sociolinguistic variables. This allows meaningful comparisons of the results as well as possible future replications of the study with even larger samples or with other language pairs. Furthermore, the methods used in the study are innovative in that both quantitative and qualitative methods have been used longitudinally as opposed to only longitudinal qualitative data or only quantitative cross-sectional data.

The profiling of the bilingual lexicon reports that bilingual children's lexical categories in each language develop in a parallel manner whether or not the children are dominant in a language. The results also show that their development is very similar to previously reported data for monolingual children. Despite current theories,

the evidence suggests that bilingual children produce translation equivalents before the 50-word stage. However, I attempt to bring forward the idea that cross-linguistic equivalents are different from synonyms within a language and so bilinguals cannot be compared to monolinguals in that respect. This thesis also sets the age of first word combinations for bilingual children to around 1;8 while claiming that this is only achieved after each language has reached the 50-word milestone. Finally, great variability is noted throughout the thesis in terms of lexical development amongst the children. Some of the differences are explained by socio-linguistic factors such as parental strategies and language exposure. Therefore, the importance of accounting for such factors when studying bilingual language development is underlined.

Our understanding of bilingual acquisition centrally contributes to our understanding of language acquisition in general. Similar features of bilingual and monolingual acquisition have been highlighted throughout this thesis. Thus, the bilingual lexicon has shown to develop at a similar rate and in a similar manner as the monolingual one despite being strongly influenced by individual socio-linguistic factors.

PUBLICATIONS ARISING FROM THIS THESIS

Forthcoming publications:

1. David, A. & Li Wei (2004). The bilingual lexicon. In Cohen, J., McAlister, K., Rolstad, K., & MacSwan, J. (Eds.). *ISB4: Proceedings of the 4th International Symposium on Bilingualism.* Somerville, MA: Cascadilla Press.

2. David, A. & Li W. (2004). To what extent is code-switching dependent on a bilingual child's lexicon? *Sociolinguistica,* 18.

Parts of this thesis have been presented at conferences:

1. The bilingual lexicon, *4th International Symposium on Bilingualism*, Arizona State University, USA (April 2003).

2. The Principle of Contrast: a strategy or a constraint? Evidence from bilingual children, *Child Language Seminar*, University of Newcastle, UK (July 2003).

3. Who can we trust? - On using parental reports with bilingual children, *Vocabulary knowledge and use: measurements and applications workshop*, University of the West of England, Bristol, UK (January 2004).

4. Bilingual acquisition: theoretical implications of a longitudinal group study. *Bilingualism in the 21st century*, Welsh Institute for Social and Cultural Affairs at Bangor University (guest lecture), Wales (February 2004).

5. The impact of the lexicon on code-switching, *2nd Lisbon Meeting on Language Acquisition with Special Reference to Romance Languages*, Universidade de Lisboa, Portugal (June 2004).

6. Can bilingual children produce two-word utterances with 50 words? *7th PG Conference in Theoretical and Applied Linguistics*, University of Durham, UK (June 2004).

7. The role of parental strategies in a bilingual child's language development, *Child Language Seminar*, University of the West of England, Bristol, UK (July 2004).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

The following abbreviations were used in the text of this thesis:

ANN: Anna

ANT: Antonia

BCDI: The MacArthur Communicative Development Inventory: toddlers (British English adaptation).

CDI(s): The MacArthur Communicative Development Inventory

D: Measure of vocabulary diversity as calculated by the *vocd* function of the CHILDES database.

DUN: Duncan

ELI: Elisa

EMM: Emma

FCDI: Inventaire Français du développement communicatif chez le nourrisson: mots et phrases (French adaptation of CDI).

FLO: Floriane

GAM: Grammatical Analysis Module

JAC: Jack

LIA: Liam

ML: Matrix Language

MLU: Mean Length of Utterance

OLB: Oliver B

OLR: Oliver R

OPOL: One Person-One Language parental strategy

PEN: Pénélope

REB: Rebecca

TE(s): Translation equivalent(s)

THO: Thomas

TTR: Type-Token Ratio

SECTION I:

PRESENTATION OF THE STUDY

CHAPTER 1:

INTRODUCTION AND LITERATURE REVIEW

1.1. INTRODUCTION

This study is a longitudinal group study of 13 children between the ages of 1;4 and 3;0 who are acquiring French and English simultaneously. Data has been collected using both quantitative (parental checklists) and qualitative methods (diaries, naturalistic productions). The focus of this study is on the nature of the developing bilingual lexicon and its impact on the emergence of syntax, especially language mixing. The study draws on insights from two different but interconnected disciplines: language acquisition especially lexical development and childhood bilingualism.

This study is innovative in many different ways. The first characteristic that makes this study different from others is the topic and its approach to it. The topic is unusual in the way that it presents data on lexical acquisition of bilingual children. Bilingual acquisition literature has recently largely focused on morphosyntax and language choice (e.g. Meisel, 1994a; Genesee, Boivin & Nicoladis, 1996). Furthermore, the topic will be approached from a different perspective. Earlier studies focused on the question of whether bilingual children had a single/fused system or two separate/differentiated ones. In particular, bilingual lexical development has recently mostly focused on the existence (or not) of cross-linguistic equivalents in the bilingual lexicon (e.g. Pearson, Fernandez & Oller, 1995) and the implications for the two-system hypothesis. There are a growing number of studies focusing, instead, on the relationship and influences between the two languages in the developing language system of the child. Therefore, this study will examine the lexical development of bilingual children, focusing not only on the presence or absence of translation equivalents (chapter 4) but also observing and analysing the development of lexical categories in bilingual children (chapter 3), a subject limited to monolingual acquisition up to now, the relationship between the lexicon and the emergence of grammar, especially language mixing (chapter 6) and the factors accounting for individual differences amongst bilingual children (chapters 7 and 8).

The originality of this study also lies in its methods and design. Larger numbers of children in the field of bilingual language acquisition have been systematically studied through longitudinal case studies (e.g. De Houwer, 1990; Lanza, 1997) or

cross-sectional language samples (e.g. Garcia, 1983; Goodz, 1994). I report findings from a longitudinal group study of 13 children between 1;4 and 3;0 who are acquiring French and English simultaneously. To my knowledge, no study has ever before been carried out on this scale looking at detailed lexical development of bilingual children over time. Further, the children in this study have been systematically selected according to a set of sociolinguistic variables. A group study allows the researcher to control variables such as gender, siblings, parents' language that a single case study would be unable to do. This allows meaningful comparisons of the results as well as providing a basis for further replication of the study with even larger samples with other language pairs or comparisons with different types of bilinguals. Subsequently, this piece of research will be one of the first to follow a group of children who are living in different countries but who otherwise have a very similar linguistic background and history. Moreover, a recent trend in bilingual language acquisition research has been to combine methods of data collection (e.g. Deuchar & Quay, 2000). Following this trend, the methods used in the study are innovative in that both quantitative and qualitative methods have been used longitudinally as opposed to only qualitative data being collected longitudinally or only quantitative cross-sectional data. The study uses traditional methods of data collection such as parental diaries and naturalistic recordings but also more recent methods, i.e. parental reports (see chapter 2 for more details on the methods used in this study).

In this chapter, I discuss the general issues related to bilingual lexical development. Section 1.2 presents some of the definitions of bilingualism that are available in the literature, while section 1.3 provides a brief overview of the most influential studies in bilingual first language acquisition research. Section 1.4 addresses the main question that has occupied bilingual research for the last three decades, whether bilingual children start with one system or two for their languages in the early stages of development. The debate has been reassessed on numerous occasions. Researchers nowadays agree that it is rather problematic because, among other factors, there is a lack of a single definition of what is called 'system' and what really counts as evidence. Therefore, in this thesis an attempt is being made to examine bilingual language development from a new different perspective. Section 1.5 focuses therefore on the rising number of studies reporting on the relationship and influences between the two languages in the developing language system of the child and the issues raised

by this new debate. The next section (1.6) briefly reviews the literature on monolingual and bilingual lexical acquisition, raising specific issues and topics not yet covered in studies of bilingual acquisition. Finally, as most research on language development reports great variability amongst children, the last section of this chapter (section 1.7) presents the possible factors accounting for these individual differences. The goal of this chapter is to review the relevant literature while the next chapter will present the methodology used in this study.

1.2. DEFINITIONS OF CHILDHOOD BILINGUALISM

Let us start this overview with a definition of bilingualism and what it means, in the present study, at least, to be bilingual. There is no single definition of the word bilingualism in the literature. Instead each author gives their own account of what being bilingual means to them in their context. Li, in a review of current literature, explained that a bilingual person is "someone with the possession of two languages" (2000:7). This, rather broad definition, highlights the difficulty researchers have been facing when trying to define bilingualism and who is (or is not) bilingual. The most important factors in determining the types of bilinguals we are dealing with are: the context and age of acquisition of the two languages. Firstly, the age of acquisition is probably the single most important factor. Two scenarios have emerged from the literature. The first is when children acquire their two languages from birth or very early in their language development. The phenomenon is called simultaneous acquisition. The expression "Bilingual First Language Acquisition" (e.g. De Houwer, 1990) is also frequently used in the literature for this type of bilingualism. If children are exposed to the second language later on in life, they are classified as successive or sequential bilinguals. The second important factor, linked to the age of the person, is the context of acquisition. In most cases, children acquiring a second language after the age of three, do it through more formal teaching, while a child acquiring two languages before three tends to have a less formal acquisition process. The two categories are also sometimes referred to as 'early' bilinguals and 'late' bilinguals. An immediate issue with this categorisation is the age limit. What is meant by 'early'? There is no consensus on age limits indicating early and late. Other definitions attempt to define bilinguals in terms of their competence in each language. A

'balanced' bilingual, for example, would be somebody who has mastered two languages to the same (equal) extent. Similarly, a semi-lingual is somebody who is said to have insufficient knowledge or abilities in either of the two languages. In any case, the individual circumstances surrounding the acquisition of the two languages are different for every child and are therefore extremely important factors to be accounted for.

The category of bilinguals that is the focus of this study is that of children being exposed to two languages before the age of three. Romaine (1995) described six main types of early childhood bilingualism. The main distinctions for these different situations were based on the languages known and spoken to the child by the parents as well as the language(s) of the community around them. The six types were described as follows:

- One person - One language: the parents have different native languages with each having some degree of competence in the other's language. One of these languages is also that of the community. The parents each speak their own language to the child.
- Non-dominant home language: the parents have different native languages. One of these languages is the dominant language of the community. Both parents speak the non-dominant language to the child. Only when outside home, is the child exposed to the dominant language.
- Non-dominant home language without community support: the parents share the same native language. But the language of the community is not that of the parents. The parents speak their own language to the child. Only when outside home is the child exposed to the dominant language.
- Double non-dominant home language without community support: the parents have different native languages. The language of the community is also different from either of the parents' language. The parents each speak their own language to the child from birth. The child may in this case be trilingual.
- Non-native parents: the parents share the same native language. The community language is the same as that of the parents. But one of the parents always addresses the child in a language which is not his or her native one.

- Mixed languages: the parents are bilingual. Some groups of the community may also be bilingual. The parents code-switch and mix languages while speaking to the child.

These types are not exclusive; other combinations of strategies have been reported. Type 1: One Person-One Language is the most commonly reported acquisitional context in the literature. The children, in this study, can all be considered as early or simultaneous bilinguals as they were exposed to two languages from birth in an informal manner (for more details see chapter 2) according to that (One person- one language) family pattern. Therefore, being bilingual will refer, here, to the simultaneous acquisition of two languages from birth. That does not imply that anybody acquiring two languages in any other way is not to be considered as bilingual.

1.3. OVERVIEW OF BILINGUAL LANGUAGE ACQUISITION RESEARCH

In almost every single book or article about bilingual language acquisition, a brief mention is given of the two earliest studies looking at bilingual children. Probably the earliest detailed study of bilingual language acquisition dates from 1913 when Ronjat described and analysed his son's language development in French and German. His was one of the first longitudinal studies based on a parental diary. He, amongst other things, described how he was advised to use the One Person – One Language (or "une personne – une langue") strategy by the linguist Grammont (Ronjat, 1913:3). Early studies also include Leopold (1939-1949) who studied his daughter acquiring German and English simultaneously. Both Ronjat and Leopold believed very strongly that bilingualism did not hinder their child's language development and that it could only be a positive phenomenon, which was quite extraordinary at the time. Though these studies certainly provided a good starting point for further debates and investigations, neither of them is replicable as they both involved single case studies carried out by linguist parents.

In the last twenty years the amount of research has grown dramatically. The last decade, especially, has seen a rise in the number of studies within the field of bilingual language acquisition. These studies are of crucial importance to the larger field of language acquisition research both from a theoretical and descriptive point of view. Bilinguals are the ideal subjects for cross-linguistic research investigating the

relative impact of language specific versus more universal aspects of language acquisition (De Houwer, 1990). This rapid increase is also due to the realisation that growing up with two languages is not so uncommon and also that if language acquisition theories are to provide a comprehensive overview of the process, they need to include bilingual language acquisition (Genesee, 2003). Studies of bilingual language acquisition can indeed contribute significantly to the general theory of language acquisition. In a recent overview of bilingual studies, De Houwer (1999) finds around 50 different studies carried out on more than 110 bilingual children between the ages of 5 months and six years that can, unlike Ronjat's and Leopold's, "stand up to close methodological scrutiny" (De Houwer, 1998: 250). They cover 13 different languages. Languages studied are mainly Indo-European languages (English, French, German and Spanish). Most of the children studied come from a middle-class environment.

The methodologies for data collection have also evolved. While most studies are still longitudinal case studies using naturalistic samples of language (e.g. De Houwer, 1990; Lanza, 1997; Deuchar & Quay, 2000), variety appears to be rising. A notable exception to the case-study approach is the work by Pearson and her colleagues in Miami (e.g. Pearson, Fernandez, Lewedeg & Oller, 1997) who looked at the lexical development of bilingual Spanish-English children in Miami based on parental reports. Another exception to the usual case study is the work by Meisel and colleagues through the DUFDE project looking mostly at morphosyntax (e.g. Meisel, 1990). Experimental work with bilinguals has also been carried out, although to a lesser extent. Davidson, Jergovic, Imami & Theodos (1997), for example, compared the use of the mutual exclusivity constraint in monolingual and bilingual children using experimental methods. Paradis (2001) compared phonological acquisition of monolinguals and bilinguals using experimental methods and a larger number of subjects (i.e. 35).

The different aspects of language have not all been studied in equal measure. Far more interest has been paid to production rather than perception. Recent work on language perception in bilinguals includes that of Bosch & Sebastian-Gallés, (1997) who examined very young bilinguals sound perception abilities in Catalan and Spanish.

Studies focusing on language production have covered many different aspects. Morphosyntax is probably one of the topics that have been written most about in bilingual language acquisition research. De Houwer (1990), Meisel (1994a), Paradis & Genesee (1996), Deuchar & Quay (2000), to name but a few, have all studied the specific features of bilingual syntax. One of the most important and widely covered topics in morphosyntactic bilingual research is that of code-switching or language mixing. Mixed utterances have recently received a lot of attention (e.g. Lanza, 1997; Vihman, 1998). Pragmatics is another topic widely covered. Language choice, for example, relates to the sociolinguistic choices children have to make to be understood and to conform to the rules set out. Studies designed to analyse the how, when and why of these choices include Saunders (1988), De Houwer (1990), Lanza (1997), and Deuchar & Quay (2000). Phonetics/phonology has received rather less attention. Phonology/phonetics include studies by Deuchar & Quay (2000) looking at voice onset time of a Spanish-English bilingual, Johnson & Lancaster (1998) and Paradis (2001), mentioned above. Lexical development has received even less attention. Pearson, Fernandez & Oller (1993) and Deuchar & Quay (2000) have respectively examined the distribution of the lexicon over the two languages and the existence of translation equivalents. Finally, language input has only just recently been acknowledged as an important issue in bilingual acquisition though bilinguals have to balance two different language inputs. Lanza (1992) analysed parental discourse strategies, while Pearson *et al.* (1997) examined the influence of language exposure on the bilingual lexicon. Some of the issues raised here will be dealt with in more detail in the relevant sections below.

1.4. THE ONE-VERSUS-TWO-SYSTEMS DEBATE

The progress of recent research on bilingual acquisition has been motivated by some important questions. The most prominent debate researchers have been focusing on since the 1970s is the one-versus-two-systems debate. Researchers have been divided between two main options: (i) the one or unitary system hypothesis which claims that bilingual children start by having only one system with elements from both languages and gradually differentiate them (e.g. Leopold, 1939-1949; Volterra & Taeschner, 1978; Redlinger & Park, 1980); and (ii) the two or separate systems hypothesis where

children do have a separate system for each language from the earliest stages of language development (e.g. De Houwer; 1990; Lanza, 1997; Deuchar & Quay, 2000). This is a very complex debate as it relies on unresolved issues relating to bilingual language acquisition such as the cerebral organisation of language(s) in the brain, and the influence of the sociolinguistic environment on the development of the child's language(s). In addition, other methodological issues arise when considering the debate: (i) the definition of what has been called a system is less than clear; (ii) what counts as evidence for either hypothesis is not clearly defined; and finally (iii) how and when the researcher is able to claim language awareness on the child's part. There will follow a description of both sides of the argument as well as a critical view of the issues raised by each of the models.

1.4.1. The one-/unitary-system hypothesis

The single system hypothesis relies on one major principle: monolingualism is the default language faculty and the organisation of the brain of a child who acquires more than one language from birth or shortly after must undergo a certain number of important changes due to the acquisition of more than one language (Genesee, 2003). These changes take time and appear gradually (i.e. lexical differentiation might be complete before the process of syntactic differentiation starts). So, bilingual children are initially monolinguals with undifferentiated syntactic, lexical and phonological systems and subsystems. Most of the theories that made a claim for this one-system hypothesis did so on the evidence that most children mix languages early in their language development and that the amount of mixing decreases over time.

Ronjat (1913) and Leopold (1939-1949) both believed that initially the bilingual child does not make the distinction between the two languages involved, unlike an adult, who makes a judgement based on the interlocutor or the context. It was only through repeated language contact that the children acquired their pragmatic skills and were able to distinguish between one language and the other one.

Perhaps the most convincing and certainly the most detailed study formulating the one-system hypothesis was Volterra & Taeschner's study (1978) of two bilingual Italian-German girls. They proposed that the differentiation between the two systems did appear gradually. They divided the process into three main stages as follows:

- Stage 1: The child's lexicon is hybrid in the sense that the child does not seem to differentiate between the lexicons of the two languages. No translation equivalents or synonyms are found in the child's lexicon at that stage which last from about 1;6 to 2;1. The child also has only one phonological system for both languages.
- Stage 2: The slow separation of the lexicons of the two languages takes place. Cross-linguistic synonyms are introduced but the child still uses the same set of syntactic rules for both languages. For example, a French/English bilingual child might say: *a car blue* instead of *a blue car*, as in French the order in this case would be noun followed by adjective. Children may reach that stage from 2;5 up until 3;3.
- Stage 3: The two languages are separated in terms of vocabulary and syntax. Children enter this last stage from about 2;9. It is from this point onwards that they can be considered as true bilinguals and that their pragmatic skills also develop so that they are able to associate one person with one language.

Most of Volterra & Taeschner's evidence for this hypothesis, and their very detailed model, comes from code-switching and the presence (or not) in the lexicon of cross-linguistic synonyms. Volterra & Taeschner were not the only ones proposing a one-system hypothesis.

Redlinger & Park (1980) also provided the field with evidence for the single-system hypothesis. They claimed that the child's amount of code-switching decreased over time and that this was due to the gradual differentiation of the two linguistic systems.

This one system hypothesis, regardless of how much it might have been recently criticised, was one of the most explicit and precise theories of bilingual language development to date. Volterra & Taeschner's claim sparked further useful work on the issue.

1.4.2. Issues with the unitary system hypothesis

Several issues arise from the one-system hypothesis. First of all, the ages at which children reach each stage are not clearly defined and do seem to overlap with each other (as between stages 2 and 3). Secondly, studies claiming a one-system development have been criticised for their methodology: not being replicable or not

presenting the data in a context which would allow other researchers to verify their validity and reliability. Deuchar & Quay (2000), for example, stated that the number of translation equivalents produced by the girls in Volterra & Taeschner's study might have been inhibited by the fact that German speakers were present for most recordings and so the girls were less willing to produce Italian words even if they knew them. Further, data with their Italian-speaking father was not presented.
Genesee (1989) also put forward the idea that context and input might play a role in the absence of synonyms. Genesee (1989) explained that this is a language acquisition process underlying bilingualism which also occurs in monolingual acquisition. For example, just as the monolingual English child will over-extend the meaning *dog* to refer to any four-legged animal, so will the bilingual French-English child 'over-extend' the word *chien* to refer to a dog when speaking English. The only difference is that bilingual children over-extend "inter-linguistically" as well as "intra-linguistically" while monolingual children can only over-extend "intra-linguistically".

Probably the most important criticism the one-system hypothesis has received, so far, deals with language mixing. Recent studies (e.g. Meisel, 1990) have shown that language mixing is actually very complex and rule-governed even in children. Genesee (1989) provided very complete and comprehensive alternative reasons behind some of the issues raised by Volterra & Taeschner (1978) in relation to code-mixing. Genesee (1989) believed that young bilingual children can differentiate two languages at the very beginning of their language development, and can use their two languages in functionally different ways. According to him, the reason for bilingual language mixing lies in the fact that the language a child uses at any one point in time is incomplete and does not include the vocabulary and grammar needed to express certain meanings. Therefore, in order to achieve efficient communication they have to borrow from, or mix with another language in which they are more competent. The second reason is that although the vocabulary or syntax required to express the intended meaning is available in the language currently in use, the vocabulary and syntax in the other language is simpler or easier to produce. Finally, the third reason is that mixed input produced by others results in the child's mixed output. He points out that language mixing is also a learned behaviour. For example, one may expect children exposed to frequent and general mixing to mix frequently. However if each of the child's parents speaks only one language to the child, the number of mixed

utterances used by the child will be much less. Genesee (1989) also explained that mixing may decline with age not because the separation of language is taking place but because children are acquiring a more complete linguistic system and therefore, do not need to "borrow" from the other language. In addition, Deuchar & Quay (1998) also argued that the one-system hypothesis "is not tenable because of the paucity of lexical resources when the child begins to produce two-word utterances" (1998: 231). They claimed that the child they studied mixed lexical items because of a lack of "contextually appropriate" ones in her lexicon (Deuchar & Quay, 1998: 236). Therefore, the existence of mixed sentences as such cannot be taken as evidence for the one-system hypothesis.

Further, Genesee (2003) argued that in order for children to be said to have only one system at the earliest stages of their language development, one would need to show that they are able to use any language in any context or "indiscriminately in all contexts of use" (Genesee, 2003: 167). Since most of the data used in the one-system hypothesis claim did not include both language contexts, it appears to be difficult to demonstrate. Genesee (2003) argued that if one could show that children do use language appropriately in different language contexts, then that would provide evidence for the two-system hypothesis. In addition to that, we have no indication as to how frequent the mixing reported by those studies was. Studying mixing in isolation is inadequate (Genesee, 2003). Once again, context is important and data should be presented with contextual information and in two different language contexts if possible.

An additional shortcoming of some of the earlier work was the use of language awareness to support claims about children having a single system. Volterra & Taeschner (1978) based their claims on the fact that the children were, at first, not able to use the right language with the right person. It was argued that awareness of the two languages played a major role in the issue of differentiation and that appropriate language choices reflected the child's awareness of their languages. According to MacLaughlin (1984), however, language awareness needs to be defined more clearly in order to be able to account for its role in language differentiation. Zhu & Li (in press) claimed that the two issues, language awareness and language differentiation, were actually quite distinct from each other. They argued that the distinction between the two concepts was similar to the competence/performance distinction. Indeed, it is not because a child is unable to perform that s/he does not

have the competence, or the ability, to do it. Similarly, if a child is unable to choose between two languages to be used with a given interlocutor that does not automatically mean that s/he has only one, undifferentiated linguistic system.
The misinterpretation of language awareness, raised in relation to the one-system hypothesis, has lead researchers to build upon their own studies in order to show that the evidence for the gradual differentiation did not stand up on a number of points.

1.4.3. The two-/separate-systems hypothesis

Unlike the unitary system hypothesis, the two systems hypothesis proposes that children start out as bilinguals, that there is no monolingual default setting as previously suggested by the one-system hypothesis. The languages of the bilingual child are represented in differentiated ways from at least the beginning of early language production, and possibly earlier (Genesee, 2003). It is now generally accepted that the unitary system hypothesis is not consistent with most recent evidence, and thus that bilingual children do not go through an initial unitary stage.
Meisel (1989) criticised Volterra & Taeschner's description of the syntactic mixing stage for being too vaguely defined. He recorded two French-German bilingual children between the ages of 2 and 4, investigating word order and subject-verb agreement since they were distinct in the two languages and could thus provide evidence for (non-)differentiated syntax. Meisel (1989) reported that both subjects preferred SV(O) word order, in both languages. However, language-specific patterns emerged. VOS or AdvSV(O) structures were present in French while German structures included AdvVS or OVS. The fronting of the adverb or object in German triggered the subject-verb inversion. Regarding subject-verb agreement, the children did start using verb inflections to mark tense, gender and number. According to Meisel (1989), this shows that children are able to differentiate their two languages as soon as they begin to use "syntactic means of expression".
De Houwer (1990) reported on the development of a young Dutch-English bilingual, Kate, who grew up in Belgium. She provided evidence against the one-system hypothesis and its claims regarding language choice. She observed that Kate used mainly Dutch when interacting with monolingual Dutch speakers but would sometimes switch to English if interacting with Dutch-English bilinguals. According

to De Houwer, she was aware of her linguistic abilities and those of her interlocutors. Overall, she would use language similarly to age-matched monolinguals. Regarding her language mixing, although, it was not the focus of the study, De Houwer reported that she mostly inserted single lexical items in utterances in the other language. She stated that these mixed utterances were structurally grammatical. Consequently, the child must have known enough about the grammar of the two languages to be able to avoid ungrammatical sentences.

Genesee, Nicoladis & Paradis (1995) observed five French-English bilingual children living in Montreal. They observed them in three different language contexts: with the English-speaking mother, with their French-speaking father and finally with both parents. The children were between 22 and 26 months old and all at the one- or two-word stage. They reported that the children were able to use their two languages in a context sensitive manner. All of the children used more French with their fathers (than English) and more English with their mothers. Even in situations where both parents were present, the children were able to differentiate the two languages and use them appropriately. Questioning their own methods, Genesee, Boivin and Nicoladis (1996) embarked on an additional study with similar children. They wondered whether the fact that the children in the previous study used the appropriate languages was due to them associating a lexical item with a parent rather than with a language. So, they introduced interactions with a "monolingual" stranger, forcing the children to identify language characteristics and then use the appropriate language. Out of the four children, three adjusted to the stranger's language by using more of his language than the other, and also using more of that language than with the parent who spoke the same language. That confirmed the idea that children often do realise that their parent understands both languages. According to Genesee *et al.* (1996), this also means that the children differentiate the two languages, and provides further evidence for the two-system hypothesis.

There are other studies that have shown that the unitary system hypothesis does not seem to hold in light of their own results. Deuchar & Quay (2000) have, for example, dismissed the claim that bilingual children do not produce any cross-linguistic equivalents (see below). While this debate continues to attract new studies, few researchers have proposed alternatives to the one or two systems.

1.4.4. Other possibilities

Recently, some researchers have started asking whether it is appropriate to continue to investigate the one-vs.-two system debate. Deuchar & Quay (2000) argued, throughout their book, that the polarisation of the issue into one or two systems may be oversimplifying the whole debate. They hinted that the other alternative to the one or two-system hypothesis might be that the child has no initial system at all "in either language" (Deuchar & Quay, 2000: 45). They based that claim on the fact that they did not find any evidence for a voicing system. The child appeared to show a progression from "the absence of contrast based on VOT towards the establishment of a contrast based on short versus long lag" (Deuchar & Quay, 2000: 44), in utterance initial stops at least. Deuchar & Quay preferred to focus on "how and when language differentiation occurs" (2000: 113). The possibility that different aspects of language are gradually differentiated at different points in time arose from their data. Lexical differentiation was established early for their subject, M (around 1;7-1;8). That was followed by morphosyntactic differentiation (around 1;11) and the emergence of two different voicing contrasts (from 1;11 to 2;3).

Other researchers have focused on one aspect of language and attempted to determine the age of differentiation. According to Nicoladis (1998), pragmatic and lexical differentiations are distinct phenomena. Pragmatic differentiation refers to appropriate language choice and lexical differentiation refers to the use of translation equivalents. Based on her data from a single Portuguese-English bilingual child, she claimed that pragmatic differentiation occurred earlier than lexical differentiation. Paradis (1996) re-examined Leopold's data and claimed that Hildegard's prosodic development followed different paths and thus appeared to be differentiated from between 1;6 and 2;0. Lanza (1997) claimed that Siri and Thomas, the two Norwegian-English bilinguals she followed, were able to use language in a context-sensitive manner from as young as two years old.

Another possibility is proposed by Paradis (2004). He claimed that there were four possibilities in which the bilingual brain could be organised. His research is mostly based on bilinguals with aphasia. However, most of these options can certainly be applied to children. The first two are called the extended system and the dual system which relate to the two options described above. The third option is that of the

tripodite system. With this system, the bilingual speaker has in fact three systems: one system where everything that is common to the two languages is represented only once and two other systems with elements that differ in both languages. Finally, the subsystem hypothesis which, according to Paradis (2004), is the one that holds the more truth. This concept is based on the notion of neurofunctional modularity, i.e. "that language is represented as a neurofunctional system divided into a number of neurofunctional modules" (Paradis, 2004: 119). These modules (i.e. phonology, morphosyntax and lexical semantics) are divided into further sub-modules that represent the languages the person is able to speak (see figure 1.1). This theory allows for the attrition of some parts of the language abilities in aphasic bilinguals as well as the attrition of one language over another one.

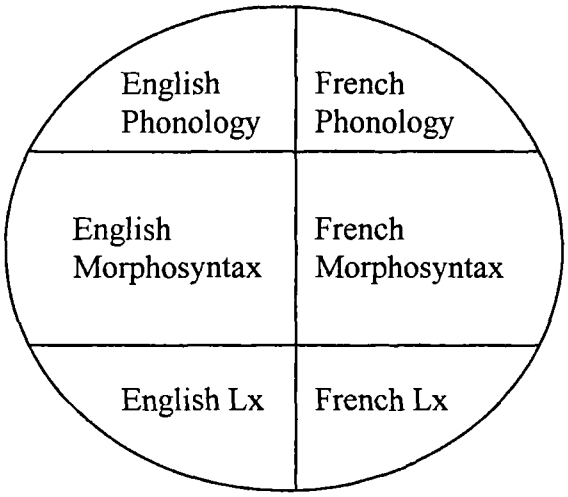


Figure 1.1 The neurofunctional subsystems (adapted from Paradis, 2004)

One could conclude that the one-vs.-two-system debate is therefore too simplistic in the way that aspects of language might not develop all at the same time and might not be differentiated at the same stage. Instead there might be different systems with different aspects of language that evolve at different points in time. Some of these aspects might start as one system for both languages and later become two while others might be differentiated from a very early age, if not from birth.

1.4.5. Inherent issues with the current debate

Finally, a number of issues linked with the one-vs.-two-system debate have arisen and have, up until today, remained unsolved. Despite the large number of recent studies pointing towards the separate-system hypothesis from a very early age, the lack of precise conceptualisation on the nature of 'system' remains a largely unresolved question. What does 'system' refer to? Researchers have used different adjectives in front of the word 'system' in order to specify what they were referring to: lexical system, voicing system, linguistic system, etc. However much of the debate is pointless unless the notion of system is defined and made clear. Deuchar & Quay (2000) suggested that generalising about one versus two linguistic systems could not be done without taking the evidence available at all levels of the language into account.

The failure to define what is meant by 'system', in turn makes it impossible to determine what type of data would constitute support for single versus differentiated systems. Deuchar & Quay (2000) claimed that lexical differentiation could not be claimed only on the basis of the presence of translation equivalents. According to them, lexical differentiation involves the presence of translation equivalents, which occurred very early on in their subject as well as the gradual appropriate language choice.

Finally, Paradis (2001) argued that the dichotomy between the one-or-two-system debate is too simplistic. If bilingual adults never reach a point of full separation and are not thought of as two monolinguals in one (see Grosjean, 1998), then why should we expect bilingual children to have separate systems in early childhood? Going further, some researchers have even asked whether it is actually appropriate to debate on the issue for children under the age of two (De Houwer, 1995). So, it may be more appropriate to approach the study of bilingual first language acquisition from another point of view.

Most recent research sees language development in bilinguals as belonging to independent but interactive systems. Bialystok (2001) noted that studies on simple dichotomies do not usually have empirical support. She argued that there is no evidence to show why bilingual children should behave differently than bilingual adults in terms of mental representation of their two languages. Any description of

languages' mental representation in adults must include the level of proficiency and the circumstances of second-language learning, amongst other factors. Therefore, might not children experience similar complexity in the organisation of languages in the brain? They might need to adjust and restructure the linguistic systems as their proficiency and competence in both languages evolve.

The one-versus-two system debate continues to engender new empirical studies; however, it is now widely believed that the two-system hypothesis holds more truth. Despite the predominance of the two systems hypothesis, the detail of how language differentiation occurs and what constitutes evidence for it, are still unclear.

1.5. MONOLINGUAL VERSUS BILINGUAL LANGUAGE ACQUISITION

While the bilingual literature primarily focuses on the one-versus-two-system debate, another question has emerged to attract new empirical research: the nature of bilingual language development. Researchers have started looking at the debate from another angle. Do bilingual children develop in the same way as monolingual children? Are the two linguistic systems (as it is now assumed that bilinguals do have two from an early age) completely independent of one another? Are bilinguals two monolinguals in one? All these questions have recently been raised within the field of bilingual language acquisition.

Within this debate of monolingual acquisition being similar (or not) to bilingual acquisition, a dichotomy based on two key concepts, autonomy and interdependence, has emerged. Autonomous bilingual development (Paradis, 2001) would mean that bilingual children acquire language in the same way as monolinguals, so bilingual children would indeed be two monolinguals in one. The opposite would be interdependence, where two languages influence each other during acquisition, meaning bilingual children's language development is different to that of monolinguals. Paradis & Genesee (1996) have argued for autonomous language development in bilinguals, at least as far as syntactic development is concerned. Genesee (2003) added that language development in bilinguals does not need to be either completely autonomous or completely interdependent. According to Genesee, certain aspects of language can develop autonomously and others interdependently. In addition, languages that share certain characteristics might provide more

interdependence than those with fewer similarities when acquired simultaneously by a bilingual child.

1.5.1. Overall similar development: autonomous development

The claim by Volterra & Taeschner (1978) that children start life with a fused lexical and syntactic system, led to the conclusion by some that bilingual language acquisition was different to monolingual acquisition. It was widely believed that bilinguals are late reaching developmental milestones because they have to deal with more than one language and have to master all of the extra abilities mentioned above (i.e. mixing, appropriate language choice, etc). The assumption that bilinguals are weaker language learners than monolingual speakers due to the demands of learning two languages simultaneously is referred to as the bilingualism deficit hypothesis (Oller, Eilers, Urbano, & Cobo-Lewis, 1997). It implies that the young child has to split cognitive resources between the two languages and consequently has only limited resources for each language. However, people like Meisel (1990) claimed that bilinguals tended to focus more on formal aspects of language and were therefore able to acquire certain grammatical constructions faster and with fewer errors than many monolinguals.

Although, theoretically speaking, separate language development is possible without there being any kind of similarity with monolingual acquisition, recent research has shown that by and large bilingual language acquisition is similar to monolingual acquisition. As far as the general course of language acquisition is concerned, bilingual children are no different from monolingual children. All children (monolingual or bilingual) go through the same stages of language development: pre-vocalic stage, babbling stage, one-word stage, multi-word utterances, etc. These stages are usually reported around the same time in both populations. There has been no study reporting major delays concerning the ages at which bilingual children start combining words, for example (De Houwer, 1990: 50). In addition, recent studies comparing bilinguals and monolinguals have shown that, for the most part, bilingual children do acquire language–specific features appropriately. Most of the research and literature in the area of bilingual language acquisition's similarities to monolingual acquisition comes in the study of morphosyntax.

De Houwer (1990), in her longitudinal case study of a bilingual girl acquiring Dutch and English simultaneously, noted that Kate closely resembled any age-matched monolingual. From the sentence types, clause types and clause constituents she used, De Houwer concluded that Kate's syntactic development was language-dependent. In addition, Kate was using other structures of both English and Dutch similarly to monolinguals e.g. tag questions, interrogative structures, complex and compound sentences.

Further evidence comes from French-English bilinguals in Paradis & Genesee's (1996) study. The two and three-year-olds they followed displayed very similar patterns of acquisition to those characterising age-matched monolinguals who acquire either French or English. Genesee (2003) argued that it was now well documented that monolingual French-speaking children acquired finite verb forms earlier than their English-speaking counterparts. So, one could assume that French-English bilinguals would show earlier acquisition of finite forms in English than monolingual English-speaking children due to the early emergence of finiteness in French. However, Paradis & Genesee (1996) did not find any evidence pointing in that direction. On the contrary, they found that the bilinguals used finite verbs earlier in French than in English reflecting monolingual development in those languages. Although, (unfortunately) the authors did not give any detailed comparison of monolingual and bilingual acquisition of finite forms in English, there did not appear to be any evidence showing that bilinguals were using finites earlier in English than monolinguals. Bilinguals also produced constructions in accordance with the target language such as in the use of pronouns with finite forms and the order of negation markers with lexical verbs. These patterns were acquired within the same age range as monolinguals.

Other studies focusing on different aspects of language have also shown similar results (i.e. that bilingual children are, in general, very similar to monolinguals in their linguistic development). A study by Pearson and her colleagues investigating the lexical development of bilingual toddlers showed no "statistical basis for concluding that the bilingual children were slower to develop early vocabulary" (Pearson, Fernandez & Oller, 1993: 93) than monolingual children. They claimed that bilingual vocabulary assessments should take into account language dominance as well as the performance in the two languages together in order to provide valid measures. Petitto, Katerelos, Levy, Gauna, Tétreault & Ferraro (2001) observed that their bilingual

children (English-French and French-Quebec sign language) attained the classic lexical milestones along the same maturational time course as monolingual infants. Similar results on a similar population were reported by Holowka, Brosseau-Lapré & Petitto (2002).

Studies on phonetics and phonology have found that bilingual children have a similar speech onset time to monolinguals and they demonstrate monolingual-like competence. Oller, Eilers, Urbano & Cobo-Lewis (1997) showed that infants exposed to one or two languages showed similar ages for onset of canonical babbling (production of well-formed syllables). The production of babbling is an event known to be fundamentally related to speech development. Further, quantitative measures of vocal performance (proportion of usage of well-formed syllables and vowel-like sounds) showed additional similarities between monolingual and bilingual infants. The similarities were true regardless of socio-economic status. It is important to emphasise at this point that these results were all based on children acquiring the two languages simultaneously. Most of the children had been exposed to the two languages from birth.

1.5.2. Evidence for different language development from monolinguals

Although most of the studies comparing monolingual language development to bilingual first language acquisition have looked at syntax, there have been a few studies on the interface between phonetics and phonology and they seem to have revealed a different picture. There are indications that bilingual children do seem to develop differently from monolinguals in at least three different aspects: (i) the overall rate of occurrence of developmental speech errors; (ii) the type of speech errors; and (iii) the quality of sounds. Zhu & Dodd (2004) reported on a number of studies of different language pairs (i.e. Punjabi-English, Arabic-English) suggesting that bilingual children do tend to make more speech errors when compared with age-matched monolinguals. They also made different types of errors that would be considered atypical and as signs of a disorder had they occurred in the speech of monolingual children. Khattab (2002) found that Arabic-English bilinguals did acquire distinct voice onset time patterns for each language, but the patterns did not

always resemble monolingual ones. However, she claimed that these differences also appeared in monolingual acquisition and thus could not be interpreted as a sign of cross-linguistic influences. One of the major reasons behind these differences between monolinguals and bilinguals in these cases is that the studies these conclusions are based on all reported on successive bilinguals. The vast majority of the children had been exposed to one language at home and the other language from their environment later on in life (i.e. nursery or school). To my knowledge, no study of simultaneous bilingual children has shown such results.

1.5.3. Specific features of bilingual development

Apart from the evidence of a few studies, it seems bilingual language acquisition is similar to monolingual, at least in terms of morphosyntax. However, caution needs to be used when drawing conclusions from studies showing similarities between monolingual and bilingual language acquisition. Firstly, those similarities do not necessarily or automatically mean that the two languages of the bilingual child develop at the same speed or even in the same way as each other. Secondly, it does not mean that the two languages are completely independent and do not interact or influence each other. Evidence of cross-linguistic influence has been widely reported (see section 1.5.3.2). Finally, bilingual children have to deal with extra aspects of language due to the fact that they are acquiring not one but two languages. In addition to the ability to formulate correct grammatical strings, bilingual children have the added capacity to co-ordinate two languages on-line in accordance with the structural constraints of both languages during language mixing (Genesee, 2003). They also need to be able to produce cross-linguistic synonyms appropriately. On top of two separate 'monolingual' competencies, bilingual children have extra capabilities and constraints to deal with two languages, i.e. a bilingual competence.

1.5.3.1.Differences between the two languages

The fact that bilingual children do, overall, develop at the same rate as their monolingual peers does not mean that their two languages develop at the same speed or in the same way. Evidence comes from different aspects of language development.

Schlyter (1993) argued that in the two languages of a bilingual, there is always one language that is dominant and one that is weaker. Therefore, she claimed that the two languages do not follow the same linguistic development. According to Schlyter (1993), the stronger language develops very much like a monolingual's first language. Her claims were based on morphosyntactic data and especially the phenomena of finiteness, word order and placement of negation. On the other hand, the weaker language "exhibits great variation" (Schlyter, 1993: 305): from non-existence to lower occurrence of these grammatical features. That weaker language develops similarly to the second language of second language learners. Thus, there are both qualitative and quantitative differences between the two languages. A further study looking at the differences between languages is Paradis & Genesee (1996) who have found that the French-English bilinguals did not use finite verbs in the same way in both languages. Finite verbs appeared earlier in French than in English; they used subject pronouns in French exclusively with finite verbs, but used them with both finite and non-finite verbs in English, thus following the rules of both languages. Finally, the children did place verbal negatives after the lexical verb in French but before the lexical verb in English, (e.g. *n'aime pas* and *do not like*) once again respecting the appropriate language structures. Lexically speaking, Holowka *et al.* (2002) reported that for all the children they followed, the lexical growth in one language was more rapid than in the other one. Many studies have shown that bilingual children are usually dominant in one language and that perfect balance of their two languages is rare (e.g. Lanza, 1997; Bosch & Sebastián-Gallés, 2001).

1.5.3.2.Cross-linguistic influences

As well as differences in the rate of development between the two languages, bilingual children's language development differs from that of monolinguals on a second issue, namely cross-linguistic influence or linguistic transfer. Cross-linguistic influence or transfer differs from language mixing. It is simply one language influencing the other one in one way or another (e.g. structure). Linguistic transfer has been defined by Paradis & Genesee (1996: 3) as the "incorporation of a grammatical property into one language from the other".

Cross-linguistic influences in simultaneous bilingual acquisition have been noted by several studies. However, their exact nature is far from being clearly settled. Some

researchers have argued that the degree of variation between different cross-linguistic influences studied depends on the language combination (Döpke, 2000). That would mean that closely related languages would allow more cross-linguistic interaction or transfer than typologically different languages would. According to Döpke (2000), that has already been shown in the literature. She claimed that children acquiring German and English simultaneously have shown more cross-linguistic transfer than children acquiring French and English or French and German. The reason appeared to be that German and English both have a surface SVO structure and thus provide the child with a false impression of the similarities of the structures at least during the first stages of language acquisition. Döpke's own study (2000) revealed that the English-German bilinguals tended to over-generalise the –VO word order of both English and German which instantiates both –VO and –OV word orders depending on the clausal structure of the utterance.

Similarly, Hulk & Müller (2000) claimed that for cross-linguistic influence to take place there has to be "a certain overlap of the two systems at the surface level" (Hulk & Müller, 2000: 229). Having shown that cross-linguistic influence occurred in the domain of object drop, they concluded that cross-linguistic transfer was due to internal language factors and not external factors such as language dominance or input as cross-linguistic influences were noted in areas where the two languages overlapped each other. Despite their conclusion, the study did not account for language dominance or input. Müller (2004) claimed that the course of acquisition of morphosyntax in bilingual children is largely independent of cross-linguistic influence and language dominance.

Another study showing cross-linguistic influence in a Cantonese-English bilingual was carried out by Yip & Matthews (2000). Although the two languages are typologically different, they observed a certain amount of transfer. They claimed that the child they studied showed both quantitative and qualitative differences from monolinguals that could be attributed to transfer. These differences were observed at a time when the child was clearly dominant in Cantonese (measured by MLU) and the child favoured transfer of Cantonese structures into English. Consequently, they attributed most of the transfer to language dominance.

Finally, Paradis (2001) looked at the issue of autonomous language development from a different perspective. She looked at the phonological development and the cross-

linguistic influences of French-English bilinguals. She found that language dominance could be a factor for cross-linguistic influences.

The differences in interpretation of the factors influencing these cross-linguistic influences might be due to the exposure to the two languages. In cases where proficiency in each language is not balanced, cross-linguistic transfer might be more important. Thus, the capacity to actually acquire two languages simultaneously might be connected with the conclusions such studies draw (Genesee, 2003). Cross-linguistic transfer might not happen in every case, at least not in the same manner.

It is important also to note that these cross-linguistic influences are temporary (Genesee, 2003) as we know that bilingual adults do produce monolingual-like structures and are thus able to acquire the appropriate forms at some stage.

1.5.3.3.Language mixing

The one domain in which monolinguals and bilinguals clearly differ is code-mixing. Mixing or switching has been studied extensively from different perspectives (e.g. social, psychological, or linguistic) in adults. However, these studies will not be reviewed in any detail here as we are dealing with language development.

First of all, it is necessary to define the expression code-mixing (or language mixing). Mixing has been recently preferred to code-switching when dealing with language acquisition. Both terms mean the alternation between two languages both within and across utterances or the juxtaposition of two languages within the same speech exchange. However, the term code-switching implies a crucial awareness as well as a fully developed grammatical knowledge which the child does not have at the earliest stages of language development. Meisel (1994b) argued that language mixing before the age of 2;6 does differ in function and form from what we call code-switching. At that stage the researcher is usually not quite sure yet whether the child is switching (according to the interlocutor, topic or context and is aware of it) or whether the mixes are elements of an unsuccessful construction. However by using the term mixing researchers do not necessarily mean that the child is violating the constraints of adult code-switching. I will mainly be using the expression mixing throughout this study using one of Meisel's definitions: "instances where features of two languages are juxtaposed, within a clause or across clause boundaries" (Meisel, 1994b: 414).

Furthermore, code-switching will be used to refer to adults' behaviour and especially that of the parents of the bilingual children studied.

Language mixing is a typical phenomenon arising in bilingual language acquisition. It is typical in the sense that monolinguals do not mix languages, as by definition, they only have to deal with one language. It is not something typical of children, however, since many bilingual adults also code-switch. This behaviour is often seen as a deviant form of language by many researchers, professionals and parents. Yet, recent research has shown that code-switching in adults is certainly not random and follows grammatical constraints (e.g. Poplack, 1980; Myers-Scotton, 1993; Myers-Scotton & Lake, 2000). Myers-Scotton's (1993) language frame model is widely used in the field of bilingualism. The model was mostly based on her work with Swahili-English bilinguals in Eastern Africa. It relies on two simple distinctions. First of all, in every act of code-switching, one language is dominant (Matrix language or ML) and the other one is embedded (EL). The matrix language is the one which determines the word order in the sentence. The second distinction is between content and system morphemes. Content morphemes are words that have a lexical meaning, or content: i.e. nouns, verbs and adjectives. System morphemes are all the others. They do not have a lexical meaning. They are elements like tense markers and agreement inflections, modals, etc... The distinction is similar to, but not to be confused with, the open- versus closed-class distinction used later (see chapter 3). Amongst others, one of the main principles of the model is that the ML contributes to more morphemes at a discourse level; the surface morpheme order is that of the ML; and all relevant system morphemes must come from the matrix language. According to that model, content morphemes, lexical words, can be inserted into the matrix language, but not system morphemes. One could also find utterances like "Yes, I want some CHOCOLAT", but you could not find a construction like "There is PAS juice" as *pas* [not] is a French system morpheme marking negation that has to come from the matrix language (assuming that in both cases English is the matrix language). The model has received some criticism on a number of points. First, the idea of a matrix language is fine when using the model to analyse adults' speech but it can prove very problematic with two-word utterances from children. If a child says: "pas juice", how can one decide which language is the matrix one? Very often the context can provide clues and researchers have used it as the main factor influencing their decisions.

Secondly, the distinction between content and system morphemes is ambiguous and very tricky to make, especially with regard to children's utterances.

Despite the important contribution of such models to the field of bilingual language acquisition, there is no consensus on the exact nature of these grammatical constraints. Most researchers, however, agree that these constraints and rules do avoid grammatically deviant or illicit structures. Just like adults, bilingual children can switch or mix between languages very easily at any level of the grammatical structure, both within and across sentence boundaries. Research has examined constraints on intra- and inter-sentential code-mixing by bilingual children in a number of different studies.

Vihman (1985) studied her own son, Raivo, who acquired Estonian and English simultaneously from birth. She argued that her son's mixing was qualitatively different from that of adult bilinguals as he was mixing more function words than content ones (i.e. nouns, adjectives, verbs). She argued that adults rarely mix or switch function words. Nouns are most commonly inserted in mixed utterances produced by adults (Romaine, 1995). Lanza's (1997) results were very similar to Vihman's while following two English-Norwegian bilinguals. However, she claimed that the mixing of function words was evidence of language dominance and so concluded that adult switching was similar to children's mixing. Research highlighting the difference between adult and children's mixing is therefore a topic that needs further investigation.

Other studies looking at the qualitative aspects of language mixing include Meisel (1994b) and Köppe & Meisel (1995) which both found a decline in the proportion of function words in French-German bilinguals between the ages of 2 and 2;3. Deuchar & Quay (2000) reported a similar trend in a Spanish-English bilingual. Children do mix languages and the literature has tried to show what the constraints governing these structures might be. Both Vihman (1985) and Lanza (1997) have shown that bilingual children do follow the surface features of grammar (e.g. word order) and abstract notions of grammatical knowledge (e.g. tense and gender agreements). The first aspect is apparent from very early on (i.e. from the two-word stage) while the second aspect is apparent from the time children show overt knowledge of these notions (i.e. usually around 2;6, according to Meisel, 1994b).

Regarding quantitative aspects of language mixing in young children, Vihman (1985) and Köppe & Meisel (1995) have both reported a sharp decrease in mixed language utterances at the onset of inflectional grammar (2;0 to 2;6).
For more details on the literature of bilingual language mixing, I will review the most influential studies directly relating to this study in chapter 6. Bilingual first language acquisition is clearly different from monolingual acquisition as it involves the capacity to "co-ordinate two grammars during on-line production" (Genesee, 2003), a specific bilingual competence.

Overall, while most of the studies reviewed above do provide more evidence for the separate development hypothesis, they also highlight the qualitative and quantitative differences between bilingual and monolingual language acquisition. Bilingual children do seem to be able to demonstrate the same general patterns of language development as monolinguals while expressing their bilinguality in unique ways. It is widely believed that bilinguals are late reaching developmental milestones because they have to deal with more than one language and have to master all of these extra abilities mentioned above. However, people like Meisel (1990) have claimed that bilinguals tend to focus more on formal aspects of language and are therefore able to acquire certain grammatical constructions faster and with fewer errors than many monolinguals.

One aspect of language acquisition in bilinguals which has received little attention over the past few years is that of lexical development. The next section will deal with a brief review of the literature of monolingual and bilingual lexical acquisition studies.

1.6. LEXICAL ACQUISITION

Lexical acquisition has received a lot of attention in monolingual language development but relatively little in bilingual language acquisition in recent years. I will first provide a brief overview of the literature on monolingual development and some of the issues raised there, then will focus on bilingual lexical acquisition and finally, I will highlight one theory accounting for the transition from words to grammar, namely the critical mass hypothesis from Locke (1997).

1.6.1. Monolingual lexical acquisition

When looking at lexical development, there are many different angles from which researchers have approached the question. Through different methods (e.g. diaries, cross-sectional experiments, parental reports, recordings of naturalistic productions), they have highlighted the processes children use in acquiring new lexical items, the stages they go through and the challenges they face. Overall, cross-linguistically, the same developmental lexical patterns have been described by the literature. Children start producing their first words around the age of 12 months. They reach the 50-word stage around 1;7 (e.g. Goldfield & Reznick, 1990). Most charts describing milestones of early lexical development have suggested that between the ages of one and two years, a child learns typically from 30 to 300 new words (see Dromi, 1996 for a discussion of these estimates).

Regarding the issue of which words are acquired first and what proportion each category represents in the child's early lexicon, there was until recently a consensus that nouns were acquired first (Benedict, 1979). However, recent research has provided evidence against that noun-first prediction showing that verbs can be learnt as early as nouns (e.g. de León, 2001; Bassano, 2000). Cross-linguistic studies show that overall lexical development is similar in French and English (Hickmann, 1997). Four stages in the order of acquisition of lexical categories and the proportion of each have recently been cross-linguistically observed (see chapter 3 for more details).

One important topic in monolingual lexical acquisition has been the meaning of early words. The type of relationships that exist between the growing conceptual abilities of children and the processes that underlie early lexical acquisition are not yet fully understood. Most research assumes that children acquire word meanings through repeated hearing of the same word in different contexts (Dromi, 1996). One aspect of early lexical/semantic development that will interest us, as it has direct implications for bilingual children, is the synonymy issue. Language as a system rarely allows two complete or full synonyms, as a general rule, for economy purposes. Studies looking at early lexical development have found that children do tend to avoid synonyms (e.g. Clark, 1987). This issue will be raised further in chapter 4.

1.6.2. Bilingual lexical acquisition

Bilingual lexical acquisition has received relatively little interest in the last few years compared with other aspects of bilingual language development. Until now lexical development has been almost exclusively studied from a bilingual point of view (De Houwer, 1999). This means that the emphasis has been placed on the distribution and proportion of each of the child's lexicon. Researchers have used that distribution to determine language dominance. If a child has more lexical items in language A, then s/he must be dominant in A (Nicoladis & Genesee, 1997; Nicoladis & Secco, 1998). Another typical bilingual issue concerns cross-linguistic equivalents. Only rare studies have looked at the distribution and development of the bilingual lexicon as a whole and/or in comparison with monolinguals.

As mentioned earlier, most bilingual language development studies have used longitudinal data often based on diaries. Ronjat (1913), Leopold (1939-1949), and more recently Deuchar & Quay (2000) have all kept diaries of their children that included a detailed account of each new lexical item appearing in the vocabulary of the child up to a certain point in time.

Ronjat (1913) described his son's lexical development in French and German focusing on phonetic and phonological acquisition. Ronjat explained that his son's German vocabulary appeared and developed more quickly than his French one. One of the points that Ronjat raised that will be relevant to us, in this study, was the fact that Louis (his son) acquired equivalents across languages. Although, Ronjat did not devote a lot of explanation to the issue, he mentioned several examples of the child's knowledge of two distinct vocabularies from at least 20 months on. From that age, the child would provide the two equivalent words for the same object or concept. For example, he would say "oeil Auge" [*eye* (fr.) *eye* (ger.)] pointing at his own eye (Ronjat, 1913: 81). In this example, the two words clearly refer to the same concept. However, his diary does not contain all of his son's new words and equivalents were not the focus of the study, so one is unable to tell whether earlier examples might have existed.

Leopold (1939-1949) also compiled a detailed diary of his daughter's lexical development amongst other aspects. His daughter was being raised bilingually in English and German while living in the USA. Leopold (1939-1949) studied the

question of equivalents across languages. He described three different types of equivalents found in her lexicon. (i) Successive bilingual synonyms; (ii) competing bilingual synonyms; and (iii) permanent bilingual synonyms. In the first category, when the second word was acquired, it usually replaced the first one. The first one ceased to exist. The two words did not exist at the same time. In the second category, the two equivalents or synonyms co-existed side by side for a while until one of them displaced the second one. Finally, the third category was composed of synonyms that were used concurrently. These two studies can be seen as rather limited as they relied on diary notes only to account for the child's lexicon. Consequently, data is limited to that observed by the parent recording the diary. Neither Ronjat nor Leopold would have been able to observe how their child behaved and spoke when they were not present.

Volterra & Taeschner (1978) and their research focusing on the one-vs.-two systems debate has already been mentioned above. Their account of two girls' language development involved the analysis of their lexicon. According to them, the girls produced few equivalents in the earliest stages of language development. In these cases where the tables given did show some equivalents, such as *da* (ger.) and *là* (it.) [there] or *danke* (ger.) and *grazie* (it.) both meaning *thank you*, Volterra & Taeschner explained that the children did "not appear to consider such words as exactly corresponding to each other" (1978: 314).

Up until recently diaries were the main and/or only source of information on which lexical analyses of bilingual children's development were based. An added method used by Deuchar and Quay (2000) provides us with extra information and probably more reliable data than the earlier diary studies. Deuchar & Quay (2000) looked at the first author's daughter M, acquiring Spanish and English simultaneously. In addition to a very detailed diary, they also recorded the child in different language settings interacting with others and sometimes using the same toys. This allowed them to check the context in which lexical items were used. They reported that by 2;0, 53% of the lexical items in M's lexicon were clearly in English. The focus of the lexical analysis was, once again, on translation equivalents. They found that from 0;11, the child had equivalents in her lexicon: *tatai* [goodbye] and *bye*. According to the authors, the data clearly showed evidence against the Principle of Contrast as described by Clark (1987; see chapter 4 for a discussion of the Principle of Contrast) that claimed that children avoid synonyms.

One of the first studies to examine bilingual lexical acquisition from a monolingual point of view was by Pearson and her colleagues. Working with children acquiring Spanish and English in Miami, they used different methods and a different perspective on the lexical development of bilingual children. Their studies were large scale ones so the diary approach would not have been an appropriate method of data collection. Instead, they chose to collect data using parental checklists. In 1993, they claimed that bilingual children were not slower than monolinguals to develop early vocabulary (Pearson *et al.*, 1993). They found that bilinguals had on average the same number of words in their lexicon as monolinguals. They also found evidence against the Principle of Contrast and the presence of translation equivalents in the early bilingual lexicon (Pearson, Fernandez & Oller, 1995). This study will be discussed in more detail in chapter 4. They also examined the role of input and language exposure in bilinguals and found a significant correlation between a child's language exposure and the size of his lexicon (Pearson, Fernandez, Lewedeg & Oller, 1997; see below).
However, maybe more importantly, Pearson (1998), based on earlier studies, reviewed the issues surrounding bilingual lexical acquisition when compared with monolingual lexical acquisition. She proposed, amongst other things, that when describing a bilingual child's lexicon, one should take into account the fact that there are not one, not two, but three different lexicons to account for. If we take the example of a French-English bilingual child, the research would need to account for the French lexicon, the English lexicon and a third lexicon composed of items present in both languages. That last lexicon overlaps with both of the others. She proposed several ways of counting lexical items in a bilingual vocabulary. One of them is what she called TCV or total conceptual vocabulary. This includes all of the items that are present in only one language, plus those which have an equivalent in the other language counted as only one concept. Instead of being a score of lexical items, the researcher aims at a concept score. Vocabulary is thus assessed from a different perspective. For example, if a child produces: *voiture* [car], *car*, *apple* and *bébé* [baby], his total TCV will be 3: one for the concept *apple* appearing only once, one for the concept *bébé* appearing only once and one for the concept *car voiture* which is represented by two words in both languages, for which the child knows the two words. By doing so, Pearson (1998) claimed that the lexicon totals would be more adequate especially if wanting to compare them with monolinguals. Pearson was one

of the first to look at lexical development in bilingual children in a less 'typically bilingual' point of view.

Since researchers have accounted for both lexicons, at least, when wanting to compare bilinguals and monolinguals. Junker & Stockman (2002) found that bilingual toddlers at two were not inferior in conceptual vocabulary size and verb diversity when words in both languages were accounted for. Yet, their study used a German 'home-made' translation of the Language Development Survey. The test was not normed. Finally, they did not account for language dominance. They recommended a longitudinal study in order to confirm these results. Another recent study worth mentioning here is by Allman (2002). She also looked at bilingual lexical development from a more monolingual perspective too. She assessed monolingual and bilingual receptive and expressive vocabularies from Spanish-English bilingual pre-schoolers. She found that the English monolingual group had greater receptive and productive vocabulary accuracy rates than the bilinguals but that the results were not statistically different. However, there was no difference with the Spanish monolingual group. Therefore she concluded that bilinguals and monolinguals were similar. However, the language dominance of the bilingual group was not taken into consideration when studying them.

At least one study has looked at the age at which bilingual children reach major milestones of lexical development from a more monolingual perspective (Holowka *et al.*, 2002). According to them, bilingual children across two modalities reached their first word at around 1 year old. They found that the bilinguals in their study reached the 50-word milestone around 1;8 (when adding both languages together). No conclusive evidence can be drawn if looking at each language separately as only two of the children reached the 50-word stage in both languages separately (respectively at 2;0 in both languages for the first child and at 2;0 and 1;10 for the second child). It is important to note that their study followed a relatively small number of children.

Based on results from new assessment tools, recent research in bilingual lexical acquisition has started to be used against the bilingualism deficit hypothesis outlined above. However, studies of lexical development in bilingual children are very few and they have mostly focused on specific bilingual features such as translation equivalents. In order to argue that bilingual children's language development is by and large similar to monolinguals', then lexical development needs to be examined from a more

monolingual point of view. This does not mean that one should compare the raw number of words produced by a bilingual child with that of a monolingual. It means that issues such as the distribution of the lexical categories in both languages, which have been studied in monolingual acquisition, should also be considered for bilinguals. This is what chapter 3 will deal with. I will also focus on those issues specific to bilinguals and especially the issue of translation equivalents which has dominated the literature on bilingual lexical acquisition up to now (see chapter 4). The issue of translation equivalents has been a very important one from very early on in bilingual language acquisition research as it has been seen as evidence for the one-system hypothesis described above.

1.6.3. The critical mass hypothesis

An important issue in lexical acquisition is the transition period between the one-word stage and the emergence of grammar or early word combination stages. Some theories like Pinker's dual-mechanism view (Pinker, 1991) claim that the change in performance in children's production can be attributed to the maturation of a separable rule system. Therefore, that there is not close relationship between different components. For example, morphosyntactic acquisition is guided by separate and distinct mechanisms from phonological acquisition or lexical/semantic development. Recent evidence, such as connectionist theories, suggests continuity between lexical acquisition and morphosyntax (e.g. Bates, Bretherton & Snyder, 1988; Marchman & Bates, 1994). One theory which has attracted our attention here is Locke's critical mass theory. Locke (1997) has argued that children's linguistic capabilities develop through a number of phases. Each phase has a clear purpose and each is dependent on the completion of the previous one. For example, in order to develop a phonological system, the child must possess a critical mass of sounds (as seen in babbling). Locke's model was divided into four phases. The first one is "Vocal learning", where the child learns about prosody and sound segments. Between 0;5 and 0;7 months, the child enters the "Utterance acquisition" phase. Basically during this very important phase, the child acquires his/ her lexicon. It is described as being an "acquisitive or storage" phase. The child stores frozen phrases like *time to go to bed* or *oh, dear*. The set of utterances acquired is used in restricted contexts but it contains words that develop

later (e.g. pronouns). Each utterance is an idiom. It is an irreducible figure of speech but it allows children to participate in adult-like interactions. More relevant to this study is the claim about the Grammatical Analysis Module (GAM). In this third phase called "Analysis and Computation", the child learns the rules by which utterances are built, looking at recurring elements, and can therefore create an infinite number of new ones. The activation of this module is experience dependent. This means that it can only be activated if, and only if, the child has stored enough utterances or words in the earlier phase. A shortage of stored lexical items would prevent the analytical mechanism from working. This phase starts when the child is between 20 and 37 months according to Locke (1997). He claimed that there is a vocabulary burst at around 50 words and that this triggers the activation of the GAM. Another point that Locke explained and which is worth mentioning for this study are the factors affecting the different stages. If phase 2 is affected by external factors (e.g. stimulation, input...), phase 3 is only affected by internal elements (elements perceived, stored and submitted for analysis are important). Finally, during the fourth phase (3;0 onwards) called "Integration and elaboration", the operations are automised and the lexicon expands gradually. One of the issues arising with this theory is that Locke did not give any evidence regarding bilingual language acquisition. What happens then if a bilingual child's lexicon is composed of 25 words in English and 25 words in French or 40 words in French and 10 words in English? More interestingly perhaps, what happens if the 25 (or 40) words in French are content and system morphemes and the 25 (or 10) words in English are all content morphemes? Does this mean that the bilingual child will automatically have to code-switch to be able to activate the GAM? Or does this mean that s/he will have to wait until each separate lexicon reaches the 50-word mark?

Data has shown that monolingual children reach the 50-word stage around 1;5 (Tamis-Lemonda, Bornstein, Kahana-Kalman, Baumwell & Cyphers, 1998). Goldfield & Reznick (1990) reported the same stage appearing at approximately 1;7 but that age ranged from 1;3 to 1;10. Holowka *et al.* (2002) reported the attainment of the 50-word stage by bilingual children to be around 1;8 (see above).

I will focus on these questions with the present data (see chapter 6 for a discussion).

1.7. FACTORS INFLUENCING LANGUAGE ACQUISITION

Most researchers in child language acquisition, bilingual or not, would agree that there is great variability in language development. All else being equal, some factors have been put forward as possibilities affecting that variability amongst children. Most of these factors have been studied in monolingual language acquisition, however, there is no reason to believe that they would not apply to bilingual children too (i.e. gender, siblings, socio-economic background). Other factors are maybe more specific to bilingual children (e.g. input and language attitudes).

1.7.1. Siblings and peers

When considering input as a factor for differences in language development, parents are not the only source for a child. Birth order is not often reported in language acquisition studies (Bennett-Kastor, 1988). However, the greater amount of time that first-borns spend with parents in interaction, as opposed to younger siblings, may enhance development. Monolingual language acquisition research has shown that birth order can be a factor influencing lexical acquisition. Bates, Marchman, Thal, Dale, Reznick, Reilly & Hartung (1994) found a small "but reliable correlation" between lexical acquisition of nouns and birth order. Later-born children had a smaller vocabulary than first-born children. Maital, Dromi, Sagi & Bornstein (2000) and Kern (2001) have both reported similar results favouring the older child of the family in the number of words produced. All of these studies used parental checklists as the basis for their claims. Bates, Bretherton & Snyder (1988) have found no such correlation between first-borns and the results of different language tests.

However, when it comes to bilingual language acquisition, most studies seem to have focused on parents as the only or main source of input. There is no systematic study of the role of siblings or peers in bilingual language acquisition, to my knowledge. As Zhu & Li (in press) noted, studies reporting on children growing up in situations where one language (the minority one) is spoken at home and the other one in the community (majority language), usually report that peers and siblings mainly adopt the majority language as their primary language of interaction thus contrasting with that of their parents. Children born in families with older siblings usually form a

separate network among themselves. That is bound to have some kind of impact on the language development of the bilingual child and should be considered as a possible and significant factor.

1.7.2. Gender

Gender is the single most important variable after age reported in language acquisition studies (Benett-Kastor, 1988). It is widely believed that girls are better at acquiring language than boys. Nevertheless some studies have reported different and sometimes opposite results. Differences between boys and girls have been found in lexical acquisition for children under the age of two (Le Normand, Parisse & Cohen, 2002). Tamis-Lemonda *et al.* (1998) did find a significant difference between girls and boys achieving different language milestones in their second year. Similarly Bornstein, Haynes & Painter (1998) reported the influence of gender (favouring girls) on measures of child vocabulary competence. All of these studies noted the superiority of girls. Studies using parental checklists to assess production and comprehension have found a very slight difference favouring girls but all recommend caution regarding the difference which is minimal and that could be due to a methodological issue (Eriksson, 2001; Maital *et a.l*, 2000; Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick & Reilly, 1993). Kern (2001) found a statistically significant difference in favour of girls in the analysis of the French Communicative Development Inventories. However, she noted that the trend inverts at around 30 months old. Various reasons have been brought forward for this apparent superiority of girls. One of them is the fact that mothers do tend to speak more and in a different manner to girls (see Le Normand *et al.*, 2002). Studies comparing MLU of boys and girls did not provide any significant difference (e.g. Schachter, 1979). Other studies did not find significant correlations between vocabulary and gender (Bates *et al.*, 1988; Bates *et al.*, 1994). Others claim that differences may fail to appear before the age of two (Bates *et al.*, 1998). Finally we can say that if the difference exists, it appears rather small and unreliable and may not play a major role in language development.

Studies of bilingual language acquisition have not, to my knowledge, looked at gender as an influential factor.

1.7.3. Parental input

Parental input has been shown to be a major decisive factor on language acquisition in general. In monolingual language acquisition, studies have shown that the quantity and quality of parental input differs immensely as the child grows older. Hart & Risley (1995) claimed that the amount of parent language varied significantly from family to family. According to them, the amount of parent input contributed not only to the children's experience but also to "the richness of quality utterances" (Hart & Risley, 1995: 127). They also noted a rise in speech directly addressed to the child at around the time of the child's first words. Shore (1995) claimed that there were several factors in the parental input that could affect a child's acquisition. The first factor she considered was the variation in the type of input received (e.g. use of formulaic utterances, speed, complexity...). The second one was "parental responsiveness versus directiveness" (Shore, 1995: 75): the more directive mothers, in particular, were, the less referential the vocabulary of their children was likely to be. Finally, the topic or social orientation of the interaction can influence the child. It is well known that if fathers tend to speak less, in general, to their children, their speech is more play-oriented. A study by Bornstein *et al.* (1998) revealed that mothers' own vocabulary directly influenced their children's verbal comprehension and maternal reports of the child's vocabulary. In addition, the quality of the input to which children are exposed does also influence children's production. Saxton (1997) showed that children (aged 5) responded using the correct verb forms when presented with negative evidence more often than when exposed to positive input. Finally, Parisse & Le Normand (2000) showed that about 80% of words (in tokens) produced by a French-speaking child at 2;1 corresponded exactly to those produced by the adults he was interacting with.

1.7.3.1.Quantity of input received

However, input is an even more important issue when it comes to bilingual first language acquisition. Amongst others, Jisa (2000) and LaBelle (2000) underlined the necessity of its study in bilingual language development. Linguistic input is a continuum: from total separation to the lack of separation. Most of actual input

conditions fall somewhere in between the two endpoints of the continuum. Romaine (1995) described six main types of early childhood bilingualism. These have been described above in section 1.2. Although the literature has mainly reported on one-person-one-language families, mixed language families could appear to be more common than first believed (Zhu & Li, in press). What Romaine's typology tells us is that context and parental input are hugely influential in bilingual language acquisition. Parents, the community and immediate carers are critical factors not only in the process of bilingual language acquisition but also in the type of bilingual speaker the child will become. Whether each type of acquisition produces different mental representations of the languages, and whether different patterns of language development can be observed, remains to be argued and is an all different topic. So how does input influence bilingual language acquisition? It is generally assumed that more exposure to one language will mean faster development and a higher level of proficiency in that language. More exposure to one language can be measured at different levels. It can be at the interactional level (i.e. parents or caregivers interactions with the child) or at the community level (i.e. living in a monolingual community). Different studies have approached the question from either a qualitative or quantitative perspective.

Studies that have looked at the quantitative aspects of input are far fewer than those looking at the qualitative aspects. Claims have been made about input without real quantitative measures to support them. For example, we know that children develop faster in the language which is most used in their environment (Romaine, 1995). According to Saunders (1988), the unequal amount of input in each language may affect only part of children language learning (i.e. productive vocabulary vs. receptive vocabulary). Lanza (1997) found that the amount of input received by Thomas and Siri acquiring Norwegian and English was an important factor in their language development. Thomas had a much greater amount of input in Norwegian than Siri and as a result produced Norwegian appropriately with his Norwegian-speaking mother, and in any case much more appropriately than Siri and her own Norwegian-speaking mother. Lanza (1997) wondered whether at equal amounts of exposure to Norwegian, the children would still have behaved in that way. Detailed quantitative analyses and studies controlling exposure and input on a larger number of children are unfortunately too rare to be able to answer that question. One study on lexical acquisition by Pearson, Fernandez, Lewedeg & Oller (1997) found a strong

correlation between language exposure and words known in each language of English-Spanish bilinguals. None of the families in their study did use the One-Person-One-Language strategy, which might have otherwise helped quantifying the amount of language exposure. Nonetheless, they did find that there was a direct effect of time time spent in each language and that "even at reduced levels of exposure to a language, children will still learn its vocabulary" (Pearson *et al.*, 1997: 55). There is also evidence that a change in the balance of the input received can significantly affect the child's language development. Romaine (1995) argued that a bilingual child's linguistic environment is very changeable and exposure to one language may cease suddenly for prolonged periods of time leading to language attrition (i.e. loss). The language loss phenomenon is related to the degree of the lack of input (e.g.: Li & Zhu, 2001). So the minimal input requirements for sustaining active use of language are very important information to be considered. Turian & Altenberg (1991) reported the case of the first author's son who grew up exposed to Russian and English. When aged around 3;6, Joseph's input in Russian decreased dramatically (from eight hours a day to just one during the week). The authors noticed that certain aspects of language the child seemed to have acquired in Russian (e.g. nominative-accusative control) were lost when the child was taped between 4;3 and 4;4. The child also acknowledged "forgetting" some vocabulary items (Turian & Altenberg, 1991: 211). Yukawa (1997) analysed the lexical and grammatical development of two Japanese-English bilinguals. The study showed that the children's abilities to use certain structures and items are closely related to the changing exposure to the two languages (the children moved between Japan and Sweden). Similar results were noted by Lanvers (1999) in a single case study on the acquisition of translation equivalents. She claimed that "changes of input greatly increased equivalent learning" (Lanvers: 1999: 30).

1.7.3.2.Parental Strategy Hypothesis

If quantity of input is important, the quality of parental input, in particular, might be playing an even more crucial role. In the case of monolingual children, Döpke (1992) claimed that quality is even more crucial than quantity when it comes to input parents can provide for their children. Most of the studies considering the quality of input have taken as criteria parental discourse strategies (i.e. the strategies parents use when their child mixes). These are an important feature of the quality of input parents

provide to their children. It has been noted that parents' interactional styles may affect the bilingual acquisition process (e.g. Lanza, 1992).

Early studies on parental discourse strategies include the very complete and detailed account of strategies used by German-English families in Australia by Döpke (1992). She proposed strategies roughly divided into two large categories depending on whether the child is required to respond and give a translation or not: "non-response-eliciting" and "response-eliciting" strategies. She placed those strategies on a continuum representing the increasing constraint on the conversation.

The Parental Strategy Hypothesis is based on Lanza's work (1992, 1997) on two Norwegian-English bilinguals. However, the name was given by Nicoladis & Genesee (1998). Lanza suggested that language choice might be affected by strategies used by parents in particular in response to instances of mixing from the child. Lanza (1992) claimed that Siri's higher frequency of mixing with her Norwegian-speaking father was due to the fact that he was more likely to accept words in English. On the contrary her English-speaking mother used strategies that indicated her non-understanding of Norwegian or her insistence on Siri using English. She placed the strategies on a continuum ranging from monolingual to bilingual context based on her own and previous work by Ochs (1988). The strategies described were as follows:

- Minimal grasp: the parent requests a clarification from the child "through a quizzical expression or through a verbal statement" (Ochs, 1988: 133). Some parents use utterances such as "what does mummy say?" The parent usually conveys a message in which he or she clearly tells the child that this is the context where they speak the other language, so the child has to make their own effort to change the word or sentence. Sometimes parents pretend they do not actually understand even if they do.
- Expressed guess: it is the parent who initiates the repair attempt, not the child as in the previous strategy. Very often this takes the form of a yes-no question that the child can confirm or not. This is a way of forcing the child to monitor his or her own language. In this strategy the parent does not pretend that he does not understand. S/He just asks for clarification.
- Adult repetition: the parent repeats the child's mix in the appropriate language in a non question form. No answer is then required from the child who can then just move on to the next sentence.

- Move-on strategy: with the move on strategy, the parent just continues the conversation. It shows the child that he understood what the child said in the other language.
- Code-switching: it reveals the bilingual identity of the parent as he or she switches language to match the one used by the child. The parent can switch either a single word or a whole sentence.

Strategies encouraging a more monolingual context (i.e. minimal grasp and expressed guess) discourage the child from using the inappropriate language. Whereas, strategies at the other end of the spectrum (i.e. code-switching), convey to the child the message that it is fully accepted to use the inappropriate language, and so the context is completely bilingual.

Based on these five strategies, several studies have tried to quantify their use and further examine the relationship between them and language mixing. Nicoladis and Genesee (1998) classified the strategies giving them a score from 1 (for ultimate monolingual context, so minimal grasp) to 5 (for code-switching, ultimate bilingual context). Unfortunately, their single-case study failed to reveal any kind of correlation between mixing and parental strategies. Deuchar & Muntz (2003) using a similar method also failed to show any correlation. They attributed that to the difficulty of classifying the strategies.

Lanza (2001) argued that had she used a similar method, her data would not have shown a correlation either. She claimed that these strategies should be used as a qualitative measure and not quantitative. This appears to be verified by another study from Juan-Garau & Perez-Vidal (2001) who did not calculate any correlations but attributed the decrease of language mixing of their child to the change of strategies used by the father.

1.7.4. Attitudes towards the language(s)

Another factor bound to influence language acquisition are the general attitudes towards the language(s) being acquired. A positive attitude from mothers towards any aspect of their child's cognitive and linguistic development has been shown to affect the child (Bornstein *et al.*, 1998). Yet, this is probably an issue more specific to bilingual language acquisition. In every relation between two languages, there is

almost always one preferred language or one language viewed as more prestigious or more useful. At earlier stages of language development, parents and their attitudes towards the languages are probably the most important factors for a young developing child as those this study will be dealing with. Saunders (1988) reported the case of his two sons who showed the effect the environment and its general attitudes can have on a child. His son, Thomas, acquiring English and German in Australia, was clearly reluctant to speak any German in front of native English speakers. For example, when at the kindergarten, he refused to speak German to his father. It was only when he found himself in the presence of other bilingual children in primary school that he became less reluctant and aligned himself with them even taking pride into being bilingual (Saunders, 1988: 118). Lyon (1996) accounted for the importance of both parental attitudes towards the languages and the attitudes of the community at large. Parental attitudes towards the language(s) their children speak usually depend on (i) the family's degree of integration into society; (ii) the status of the languages that the parents wish to pass on to their children (e.g. minority versus majority language); and (iii) the reasons that would motivate them to encourage the acquisition of the language(s) by their children (Lyon, 1996). In some communities, bilingualism can appear as the norm or be regarded as a highly positive phenomenon. However, in some other communities, although the child may also become bilingual, the minority language, spoken at home, remains the sole responsibility of the family to maintain. The child has more or less no other exposure to it than from his/her parents. In these two cases, the attitudes and willingness of the parents and the community surrounding them towards both languages is crucial to the way the child will develop both languages and to the kind of bilingual speaker s/he will become. The outcome is very often totally unpredictable. De Houwer (1999b) underlined the importance of the attitudes and beliefs. She argued that although very little empirical data is available on the issue, parental beliefs and attitudes do appear to influence the linguistic choices and interaction strategies done by parents which in turn will have an impact on the children's language development. The attitude factor is a very difficult one to measure and qualify. It is something that should be accounted for and mentioned in relation to each child's own and personal environment.

Other influential factors include the social economic status of the families both in monolingual and bilingual language acquisition (see Oller and Eilers, 2002 and Hart

& Risley, 1995 respectively for bilingual and monolingual children). However, since most of the families in this study will come from the same background, it will not be considered any further. A further factor is the level of education of the parents and especially that of the mother (e.g. Kern, 2001). Again, all of the mothers in this study are at least university graduates and so, it will be impossible to measure the effect of that factor. Finally, the care situation of the child could have an important factor on the language acquisition of the children. Some studies have shown that children who regularly attend a nursery do produce more words than others (Kern, 2001). However, in the case of bilingual children, the day care situation will be mostly reflected in the language that day care is in. Indeed, if a child attends an English-speaking nursery five days a week, then his input in English will be much more important than a child being cared for at home by his French-Speaking parent. So, measuring day care as a variable of language acquisition will not be possible in this study as it will be incorporated in the amount of input received in each language.

After having reviewed the most important studies in bilingual language acquisition research relevant to the topic of the present study, the next section will present the study's main aims and objectives.

1.8. RESEARCH QUESTIONS

The main purpose of this study is to examine the relationship between the two languages in the developing system of bilingual children by providing an insight into the nature of the developing bilingual lexicon and its impact on the emergence of syntax, especially language mixing. Using French/English data from a longitudinal study of 13 children living in the U.K. or in France, the present study is intended to investigate the following points:

1. What are the normal developmental stages of a bilingual child's lexical acquisition?
2. How early can bilingual children produce translation equivalents?
3. To what extent is young bilingual children's language mixing dependent on their lexical development?

4. How does the socio-linguistic environment impact on bilingual children's lexical acquisition and development of language mixing?

In trying to answer those questions, I will address, among others, the following theoretical issues described in the previous sections of this chapter:

- Locke's critical mass hypothesis (1997).
- The Principle of Contrast from Clark (1987).
- The Grammatical Deficiency hypothesis, Meisel (1994b).
- Lanza's parental discourse strategies (1997).

The next chapter will describe the methodology used for this study and the subjects and their environment. Lexical acquisition will be the focus of chapter 3 as I attempt to profile bilingual lexical development in relation with monolingual data. Chapter 4 will deal with an aspect of language acquisition specific to bilinguals: translation equivalents. Both of these chapters are based on data obtained from parental checklists and therefore the validity and reliability of the tools will be assessed in chapter 5. Chapter 6 will focus primarily on the emergence of syntax in young bilinguals, especially in relation with the lexicon. Chapters 7 and 8 review the influence of certain variables on bilingual language acquisition. Finally, in chapter 9, the main results obtained will be discussed.

CHAPTER 2:

METHODOLOGY AND SUBJECTS

2.1. INTRODUCTION

This chapter will outline the methodology used in this study as well as presenting the subjects and their linguistic background. Firstly, I will explain the design of the study and especially the choice of a longitudinal group study. I will, then, present the children who are taking part in the study and their environment. Finally, I will outline the data collection process from general information about the families through to the different tools used to obtain the quantitative and qualitative data.

2.2. DESIGN OF THE STUDY

The field of bilingual language development has largely been contributed to by longitudinal single case studies, often of the researcher's own child/children. A longitudinal study is typically one where a small number of children are observed or recorded at regular intervals within a long time period. If variables such as gender, socio-economic status, etc, are fully described and explained, then this type of studies allows the researcher to compare his/her results with others previously obtained. This type of study is often referred to as 'diary studies'. One of the earliest studies of this type is by Ronjat (1913) who follows his son, Louis, acquiring French and German simultaneously since birth. Leopold's three-volume-study (1939-49) is among the most cited and most complete studies. He follows his daughter Hildegard's bilingual development in English and German. Lisa and Giulia's language development reported in Volterra & Taeschner (1978) is also a longitudinal study. De Houwer (1990) follows a child acquiring Dutch and English between the ages of 2;7 and 3;4. Other case studies include Lanza (1997) who studies two children longitudinally acquiring Norwegian and English; Deuchar & Quay (2000) who study the first author's daughter's language development in Spanish and English. There are numerous examples of this type of study.

There have been very few larger scale longitudinal studies. The DUFDE project is one of them. Meisel (1990) and his team followed several children acquiring French and German longitudinally. Despite this example, larger group longitudinal studies remain scarce.

Cross-sectional studies of bilingual children have also been carried out in the field of bilingual language acquisition. Research on French-English bilingual children done in Canada by Genesee and his colleagues often involved studying different children at different ages (e.g.: Genesee, Nicoladis & Paradis, 1995). More often studies using cross-sectional methods intend to test a specific issue. Those studies usually involve older bilinguals. Amongst others, Oller and Eilers (2002) led a large project aiming at collecting a larger sample of data from school-aged bilingual children. Their research used both longitudinal and cross-sectional methods.

2.2.1. Why a longitudinal study?

Longitudinal studies usually mean diary studies or at least it did until the 1950's. Diary studies rely on notes being taken by the researcher or an observer rather than complete language samples. Diary studies can be merely suggestive and need to be followed up with other research to substantiate a point.

According to Ingram (1989), longitudinal language sampling studies have now replaced diary studies and have become very popular. They involve visiting a child at predetermined intervals (usually not the researcher's child as in diary studies) to collect a representative sample of language over time. It is this type of method that will be used here as opposed to cross-sectional or diary studies.

There are major differences between longitudinal and cross-sectional approaches to language research. First of all, longitudinal data collection allows the researcher to track changes over time. A cross-sectional study would not offer that possibility. They tend to focus more on the differences between groups rather than on change from which development is inferred (Bennett-Kastor, 1988). Since the focus of this study is on development, it is imperative that we focus on the transitions and changes between the children's levels of ability. The ideal study would look continuously at every type of word or utterance produced by the child. However, that would be technically impossible. In order to obtain continual monitoring, the children must be recorded at short intervals (see below for the details of the recording process). Secondly, longitudinal studies allow for a larger amount of language samples to be collected for each child over time. This, in turn, provides a more representative sample of the child's linguistic abilities (Ingram, 1989).

While longitudinal studies are usually carried out on few subjects' language development over time, the larger sample studies tend to be cross-sectional: large number of children studied at different ages. This represents one of its major weaknesses. The main reason behind the smaller number of subjects is that longitudinal studies are very time-consuming and a rather slow process. So, researchers usually tend to focus on one or a couple of children.

2.2.2. Why a group study?

The advantage of single case studies is that they allow the researcher to refute generalisations (Deuchar & Quay, 2001). In language acquisition, any generalisation that claims to be universal is open to being tested by any particular case. If the generalisation, then does not hold, it must be revised or a claim must be made to exclude the case used to test it. For example, Deuchar and Quay (2000) claim that Clark's (1987) Principle of Contrast is not universal, or does not apply to bilingual children, as M, their subject, did produce translation equivalents before the 50-word stage. Another positive feature of the longitudinal case study is that it can show what is possible and therefore what needs to be taken into account in any generalisations. Lanza (1992) takes into account the context of the child's use of language and concludes that this is something that has to be accounted for and explained when studying language mixing of bilingual children. Also, a case study can be a useful source of hypotheses. Volterra & Taeschner (1978) propose a three-stage development to explain a bilingual child's simultaneous language acquisition process. However, it is very difficult to generalise the findings of single case studies given the various individual variations: for example language pairs. Single case studies cannot control variables and are all, in theory, different from one another. So, it is virtually impossible to compare children's data given the inexistence of normative data and the large number of varieties of different languages involved.

Consequently, the choice was made to base the present study on a larger group than just a couple of children. 13 children were followed for almost a two-year period of time. Having a larger number of children will allow eventual participants' attrition (one of the major issues with longitudinal language studies involving children). (The study actually started with 15 children but two of them dropped out at the very

beginning. The 13 remaining families were more or less involved all the way through.) In addition, I aimed at controlling as many variables as possible. All the participants are acquiring the same languages (French and English). The genders are matched. All of the families are from the same socio-economic background (i.e.: middle class). Thus, it will be easier to compare the results and make more general claims. Making universal claims is the aim of this study, however. The results will have to be taken with caution, especially since we are only dealing with one language pair. One should also be aware of, and keep in mind throughout this thesis, the issues raised by having a larger number of children taking part in the study (i.e.: central tendency not fitting any single individual, less flexibility...).

Despite the group study, the number of children is still 'humanly manageable'. The possibility of knowing each family personally will still be available. By keeping the number quite small we will also be able to provide more detailed analyses of individual situations.

2.3. THE SUBJECTS AND THEIR LINGUISTIC ENVIRONMENT

Amongst the six different ways of growing up bilingually (as described by Romaine, 1995), the decision was made to focus on one type of bilingual family: type 1: One Person-One Language. This type of study is the most common one in the literature on bilingual acquisition. By choosing to focus on this type of bilingual family, and bilingual children, it was hoped that their environment would provide a more balanced language input and therefore, the 'ideal' scenario to study bilingual language development. The advantages linked with this family setting are, primarily that the children are exposed to the two languages from birth. In situations where the second language is introduced from the environment, a greater variability would have been observed across families (depending on day care situations for example) and would not have allowed for meaningful comparisons.

The subjects for this study were recruited through an advert in a bilingual newsletter and from word of mouth. The few conditions to take part in this study were, first of all, that the parents use the *une personne – une langue* [one person one language] method (recommended by the linguist Grammont to Ronjat, 1913: 3) as a family

strategy. One of the parents had to be a native English speaker while the other one had to be a French native speaker.

The country of residence was to be either France or England. This choice was made, to have children living in two different countries, in order to try and determine whether the country of residence could be a major influential factor on language development considering that the other variables were controlled and similar.

This being a longitudinal study, families had to commit to what could seem as a time-consuming task, so most of the families had to be and were interested and willing to spare time for it. The result of these conditions is that all of the families are from a middle-class background. And most of the parents hold a degree in higher education. Consequently, most parents also have at least basic knowledge in each other's language. Following the main aim of this study to control variables (in order to make meaningful comparisons), all of the children have a similar linguistic background: similar language history and relationship between the two languages, similar function for the two languages.

There follows a brief description of each child's linguistic environment. A summary table can be found below. All of the names given to the children and their families are pseudonyms.

2.3.1. Thomas (THO)

THO is an only child. He lives in France, with his English mother and French father. The family language is French. THO's mother worked during the study on an irregular basis. From 12 to 17 months old, he was cared for by a French-Speaking nanny, 5 days a week. When the family moved home (within France), THO started going to a childminder 2 days a week. Then from 30 months, he was cared for full time by a French-speaking childminder. THO and his mother travelled to England for a month at the time twice or 3 times a year. The family language is French.

2.3.2. Floriane (FLO)

FLO was an only child at the start of the study. A younger sister was born when she was 21 months old. She lives in England with her French-speaking mother and

English-speaking father. Both parents are fluent in both languages. The family language is French. Her environment is very much multicultural and multilingual. They have many friends and neighbours speaking different languages. FLO was mainly cared for by her mother at home, until she started attending an English-speaking nursery twice a week for half a day from when she was 15 months old. Both parents, especially the mother, were overheard code-switching themselves; on very rare occasions (the odd word would be mixed). A great emphasis is put in the family on language and literacy in general.

2.3.3. Antonia (ANT)

ANT was an only child at the start of the study. She lives in England with her parents. Her father is British and her mother is French. The family language is English and the parents both obey a very strict One Person-One Language (OPOL) rule. ANT's mother was raised as an Italian-French bilingual; however, she does not use any Italian while speaking to her daughter. ANT does not have much contact with English speakers apart from her father. Her mother does not work, so takes care of ANT at home. Her father is away from home most weeks for work. So she only sees him at weekends. She has an older step-brother (her father's son, monolingual English) who visits regularly at weekends. During the visits, ANT's mother speaks in English to him. ANT travels with her mother to visit her grandparents for a couple of months at a time each year.

2.3.4. Emma (EMM)

EMM is the youngest of two children. Her brother is six years older. They live in France. EMM's father is French and her mother is British. The family language is French. EMM goes to a French-speaking childminder 4 days a week. So the only English input she receives is from her mother. The parents obey the OPOL rule, however, it might be of interest to note that EMM's older brother replies to his mother in French even though his English is good. He goes to an international school, where teaching is done in both languages. EMM's mother does not insist on her children

talking to her in English. Consequently, the conversation will go on as if nothing happened if one of her children says something in French.

2.3.5. Duncan (DUN)

DUN is the youngest of two children. His sister is 2 years older. His mother is the French speaker of the family and his father a native English speaker from New-Zealand. The family has lived in many English-speaking countries before DUN's birth and so his father reckons that his English is far from being typical of New-Zealand. They live in France and in a mainly French-speaking environment. Their family language is English though as DUN's father speaks only little French. DUN is mostly cared for by his mother at home.

2.3.6. Pénélope (PEN)

PEN is the youngest of two children. The family lives in France. Her mother is a native French speaker and her father a British English native speaker. Both parents work full time so a nanny takes care of PEN five days a week. The nanny is not a native speaker of French but only speaks French to PEN. The family language is English. The parents obey the OPOL rule quite strictly. They admit code-switching quite a lot but only between themselves. The mother is a trained linguist. PEN's older sister goes to a French-speaking school.

2.3.7. Elisa (ELI)

ELI is the only child in her family. She lives in England with her French-speaking mother and English-speaking father. The family language is English although both parents speak each other's languages. ELI's mother works part-time, so ELI goes to a nursery 2 to 3 days a week since age one. The rest of the time she is cared for by her mother at home with another young English-speaking boy. The parents obey the OPOL rule although it was noticed that the mother switched to English quite often. The family usually travels to France for a few weeks each year. From very early on both parents used a sign language with ELI. They made up their own signs, the same

in both languages for everyday words like *more* or *fish*. But they mentioned on several occasions that they were not consistent enough for ELI to actually pick them up; and she rarely used any of them.

2.3.8. Jack (JAC)

JAC is the youngest of two children. His sister is 3 years older. They lived in France with their English-speaking mother and French-speaking father at the beginning of the study. JAC's mother is not a British-English speaker. She is Australian. However, she claims that having lived for so many years outside of Australia makes her English more British than before. The family language is a mix of English and French. Both parents are fluent in each other's language, and so there was no real rule about who says what to whom, as far as the conversation language between the parents is concerned. However, like all the other families they adopted the OPOL strategy with the children. JAC's older sister used to go to a bilingual school where teaching was equally divided in both languages. Although, they lived in a French speaking country, they also had many contacts with the English-sepaking community. However, the family frequently travelled to Australia for several weeks at a time. When JAC was 24 months old, the family moved to Japan. Although, they continued keeping in touch, the introduction of a third language meant, they could not participate in the study in the same way as the other families did.

2.3.9. Anna (ANN)

ANN is two years younger than her sister. They live in England with their English speaking mother and French speaking father. ANN's father was born in Switzerland of French parents. He grew up hearing French (at home) and Swiss German from age 5. However, he never uses any German when speaking to his daughters. The family language is English. ANN was cared for primarily by her father, who worked from home until she was 32 months old. They used to attend a playgroup together where many bilingual families from different language backgrounds met once or twice a week. When ANN was 18 months old, her older sister started attending school, so she was mostly on her own with her father at home. Shortly before she turned 32 months

old, she started going to an English-speaking day care centre all day, everyday of the week. That proved to be a major change in her language input. The parents obey to the OPOL rule quite strictly.

2.3.10. Oliver R (OLR)

OLR is the oldest of two children. His younger sister was born during the study shortly before he turned 16 months-old. They live in France. OLR's mother is British and his father is French. His father only has very basic knowledge of English, while is mother speaks fluent French even if she uses English for her work. Both parents' work is very flexible and done mostly from home. However, OLR goes to a French-speaking nanny almost every day. The family language is French. OLR's mother is his only source of English, apart from the few occasional days spent with his maternal grandparents.

2.3.11. Liam (LIA)

LIA was an only child during this study. A younger sibling was born after the period of data collection. He lives in France with his English speaking mother and French speaking father. The family language is French, as his father does not speak much English. LIA's mother works in a bilingual environment. Both parents work full time; consequently, LIA goes to a creche five days a week where the staff and the other children speak French. LIA's input in English comes from his mother and the few days in the year where he meets his maternal grandparents.

2.3.12. Rebecca (REB)

REB is the youngest of three children. Her sister is 6 years older and her brother is one and a half-year older. REB's mother is French and her father is British. Their family language is English as the father only has basic knowledge of French. They lived in England throughout the study although they had only recently moved there from France. While in France, the mother spoke English to her children since she thought they would not have enough English input otherwise. She switched to French

when living in England. There have been many changes into REB's life during the length of the study. Overall, her environment has been quite multilingual. Her mother works part-time and she was first cared for by a French-speaking family member at home. Her parents then separated towards the end of the study. It meant less everyday activities with both parents at once.

2.3.13. Oliver B (OLB)

OLB is the youngest of two children. His sister is a year and a half older. His mother is French and his father is British. They live in England. The family language is English as his father has only basic knowledge of French. His mother works part-time from home so he goes to a nursery one afternoon a week. Although the nursery is mostly English speaking, there are many children with different linguistic backgrounds. They also go to a French playgroup half a day per week. The mother also takes care of a younger boy a few hours a week to whom she speaks English. Their environment is quite multilingual (a lot of input from both languages). The parents obey the OPOL rule but the mother would switch easily to English. Chloe, Oliver's younger sister, was born just at the end of the study.

Table 2.1 Children's background information

	Gender	Country of residence	French-speaking parent	English-speaking parent	Family language	Family ranking
THO	Male	France	Father	Mother	French	Only child
FLO	Female	England	Mother	Father	French	Oldest*
ANT	Female	England	Mother	Father	English	Only child
EMM	Female	France	Father	Mother	French	Youngest (out of 2)
DUN	Male	France	Mother	Father	English	Youngest (2)
PEN	Female	France	Mother	Father	English	Youngest (2)
ELI	Female	England	Mother	Father	English	Only child
JAC	Male	France	Father	Mother	English/ French	Youngest (2)
ANN	Female	England	Father	Mother	English	Youngest (2)
OLR	Male	France	Father	Mother	French	Oldest*
LIA	Male	France	Father	Mother	French	Only child*
REB	Female	England	Mother	Father	English	Youngest (3)
OLB	Male	England	Mother	Father	English	Youngest (2)*

*Younger sibling born during or immediately after the study.

2.4. DATA COLLECTION PROCEDURE AND TOOLS

In order to obtain as diverse data as possible, different data collection methods were used: both quantitative and qualitative. The data was collected in both language contexts: French and English. The following data collection methods will be used for this research:

- Family questionnaires and visits (section 2.4.1) will allow for background information and verifications throughout this thesis. Data collected this way will mostly be used in chapter 7 and 8.
- Parental diaries (section 2.4.2.1) will allow for verifications and confirmations but will not form the bulk of the lexical analyses (mostly in chapter 3 and 5). They will only be used in addition to the CDIs.
- MacArthur Communicative Developmental Inventories adapted into French and British English (see section 2.4.2.2) will form the main data collection tool for the lexical data. Data reported in chapters 3, 4, and 5 is largely and mostly based on this method.
- Naturalistic data will also be collected (see section 2.4.3 for more details) in order to obtain language in context and not only words in isolation. This data will also be used for the emergence of syntax (chapter 6), for observing dyadic interactions (chapter 7 and 8) and finally, for reliability analyses of lexical data (chapter 5).

2.4.1. General linguistic data

2.4.1.1.The questionnaires

The first real contacts with the families were questionnaires. The questionnaires were posted to the families prior to the first visit (see below). The decision was made to post them mainly because the families lived at widely dispersed addresses and abroad. At the very beginning of the study each family received a background questionnaire about their child's medical history (in order to eliminate children with major problems or those at risk), motor development and his/her linguistic situation at the time. Each parent also had to answer a few questions about their own linguistic background and the family language choice patterns. Each questionnaire was in the native language of the parent. So there was no unambiguous question and all were able to answer the questions themselves and independently. The questions asked related to several important issues to be considered when working with bilinguals as outlined by Grosjean (1998): language history and stability (e.g.: how and where languages were learnt), function of languages (e.g.: in which contexts are the languages used),

language proficiency (of the parents) as well as biographical data. All of these were useful to establish a profile of the language input received by the children. As De Houwer (1998) suggests, information about input and accounting for sociolinguistic contexts are vital in subject recruitment and should form the basis of any bilingual language acquisition study. Parents were assured that all of the data collected was confidential. The questionnaires were composed of both open and closed questions. Closed questions require little time and are generally easier to process and analyse. These questions allowed for collection of data regarding the family's habits in terms of media in both languages. However, this type of question is also less spontaneous and, in a way, biased as the responses have to be put in categories. Open questions were used to test hypotheses about ideas and awareness (Oppenheim, 1992). For, example, an open question was used to ask parents about their attitudes to code-switching and wrong language choice. An update questionnaire was then sent every six months about different changes that might have occurred in the child's life. Copies of the questionnaires can be found in the appendices. The parents reported any major changes between updates through e-mails. The data from these questionnaires was summarised and entered into a database to provide an identity file for each child and his/her family.

With each CDI checklist, parents were also ask to quantify the language input received by the child during the past month. The questionnaire is based on a similar questionnaire used by Pearson (1995) for bilingual language input in Miami. This questionnaire was very quick and simple for parents to complete. It simply required them to provide information about the past month. Most of the questions simply required a tick or a number as a large amount of information was already being required from parents each month. It was left up to families to determine who was to complete it. At the end of the questionnaires parents were asked to evaluate the percentage of French and English to which the child had been exposed in the past month. These estimates provide the basis for the exposure analysis in chapter 7. However, some of the estimates were slightly modified after discussion with the parents if the researcher felt that the percentage provided by the family was not accurate based on other information gathered (i.e. visits, answers to other questions and emails). All of the questionnaires can be found in the appendix.

The reliability and validity of these questionnaires were assessed by administrating them, first of all, to a couple of bilingual parents not involved in the study who commented on the usefulness and understanding of each question. Moreover, the data reported by the parents in questionnaires was checked by the visits. Thus, each ambiguous point could be verified and clarified.

2.4.1.2.The visits

All of the families (but two) were visited by the researcher, at least once. It was felt that in order to explain to the families what the study was about and what was expected of them, it would be best to meet all of them in person. The visits allowed the parents to feel more involved and interested. During those visits, the researcher was able to understand the family linguistic situation, family dynamics and language choice patterns. The visits were not just a couple of hours talking to a parent. They, in most cases, involved taking part into a daily routine or activity with the whole family (such as a meal or a day out). Even though the questionnaire allowed, in theory, for a large amount of information to be obtained, the visits brought much more concrete information and clarified what was written on paper. The visit was very important to establish how strictly the parents obeyed the OPOL rule. As it was noticed many times, even if they had written on paper that they were quite strict about it, the reality was sometimes different. This was expected as it had been reported by other studies (Goodz, 1989). The meetings with families were a good way of rating the parent's amount of code-switching as well. On this point again, the differences between the answers on paper and the everyday behaviour was different. This is due to the fact that most parents did not actually realise they were code-switching. Most of those visits were very informal and the families were encouraged to be as natural as possible. No recordings were made during those first visits.

During those visits, the language spoken between the researcher and the family was the family language. When only one parent was present with the children, then the researcher would speak the parent's language. So the children had a clear idea that the researcher was also bilingual.

Even if some families were visited only a few times, very regular contact was kept through other means (e-mail and letters mostly) with all of them. Parents often shared comments on their child's language development on a very regular basis.

2.4.2. Lexical data

2.4.2.1.The parental language diaries

Parents were asked to keep a record of each child's new word between the onset of speech and the 50-word stage. Parents were given an explanation of what was meant by a language diary as none of them were familiar with the technique. Some guidelines and examples were given following Braunwald & Brislin (1979) and Dromi (1996). Those examples were adapted to suit the specificity of a bilingual diary (e.g.: language context). Although this can seem to be a rather old-fashioned method, it is still widely used in child language research. This method was chosen so as to provide a more day-to-day account of the child's language development. Due to the large number of subjects, it was impossible for the researcher to account for the child's improvements on a daily basis. So, families were encouraged to write down the first few words of their children as well as any useful comments or remarks they might have. However, diaries have also been criticised to be biased in that the parents tend to note down what they consider to be important and maybe not what the researcher might think is important (Ingram, 1989). This is especially the case in this study as none of the parents were trained linguists. Other researchers believe that parents (usually mothers) are actually the ideal observer as they have a very intimate knowledge of the child's abilities (Bennett-Kastor, 1988). A parent observer will thus be more easily able to spot daily differences. The diary proved to be a rather difficult exercise for many families. The diaries do vary tremendously in their quality and detail. The first few words were quite easily written down, but as the lexicon grew quicker, many families gave up or kept an incomplete diary. This was expected and so most of the analyses in this study are not based on the diary specifically. The diary only allows verifications and general indications. In this case, the diary data will be supplemental (Ingram, 1989: 10). Some others would say that it is not a diary but simply a "longitudinal written record" serving as a "supplementary source of data" (Braunwald & Brislin, 1979). A few families, though, kept a complete diary and those will be included to support claims and analyses. An example of a diary, kept by one family, is provided in appendix IV.

2.4.2.2.The MacArthur Communicative Developmental Inventories (CDI)

Most of the data for the lexical analyses of this study is based on the MacArthur Communicative Developmental Inventory (CDI). There are three main options to assess language development and abilities for toddlers.

The first option would be structured tests. These require a lot of time, trained personnel and the cooperation of the young child. Most instruments show that their reliability is very limited when it comes to very young children (under three years old). In addition, the present study spanning across two different countries, it would have proved almost impossible to carry out such tests. The second possibility is the collection of language samples. I describe the use of this method in section 2.4.3.

Finally, the third option is parent reports. This is the assessment technique chosen to base most of the lexical analyses on. The main advantages of parent reports are that they are more representative than laboratory samples (Fenson *et al.*, 1993). Laboratory samples are only confined to a few situations, whereas parents are able to observe their children in many different types of situations, virtually any situation possible. Also, parental reports do not need to account for the observer's effect as children are used to their parents being around them and would react differently if an outsider was with them. Other benefits include the fact that they are simple and rapid to analyse (compared with language samples) and cost-effective. Consequently, parental reports enable larger scale studies. Another advantage related to this particular study is that due to the country of residence of the children and the frequency of administration (see below), the forms were on several occasions sent to the families through mail. However, the most significant benefit of such parental forms is the fact that they allow meaningful comparisons between children, for example, amongst the present group of bilinguals and even with previously obtained results by other studies on different populations.

When choosing parental checklists, two options are offered. Most adaptations of the CDI have drawn long and short forms. The short forms only contain a smaller number of items and are designed to be easier and quicker to administer. Since this study is a longitudinal one, these could have been a useful tool as they would take less time for parents to complete. Yet, one of the problems with the short forms is that because they contain fewer words, the lexical evolution or development over time is more complex and difficult to observe. So, those short forms were decided against.

One major issue with the CDI and most parental forms is the issue of categorisation. In parental forms, lexical items are classified in semantic categories for easier recognition by parents. However, this type of categorisation does have a number of drawbacks. From a syntactic point of view, pronouns, for example, are put under one single item to be ticked in French regardless of gender or number. Also, auxiliaries in French are not real auxiliaries, they are merely helping verbs. This classification is therefore far from being perfect from a semantic as well as syntactic point of view. This categorisation is arbitrary and subjective as children probably do not recognise those categories and might even have completely different ones. This point will be further raised and discussed in chapters 3 and 5.

2.4.2.2.1. The standard French and British English adaptations

The CDI was chosen amongst other parental reports as it is one of the only checklists that has been adapted into other languages than English and has been normed in these languages. Another advantage of the CDI is that it contains naturalistic productions to a certain extent. Onomatopoeias like *uh oh* or animal sounds are included. Bassano, Maillochon & Eme (1998) underline the importance of these words in early lexical development.

The American original version (Fenson *et al.*, 1993) contains two different forms. An infant version for children aged between 8 and 16 months. And a toddler form aimed at children between 16 and 30 months old that contains 680 lexical items to be ticked and a series of grammatical questions. The original forms have been adapted into many other languages: Spanish (Jackson-Maldonado, Thal, Marchman, Bates & Gutierrez-Clellen, 1993), Italian (Caselli, Casadio & Bates, 1999), Hebrew (Maital, Dromi, Sagi & Bornstein, 2000), Swedish (Eriksson, 2001), etc.

The British English toddler version (BCDI) was developed by Klee, Marr, Robertson & Harrison (1999) for children between 16 and 30 months old. It contains 672 words organised into 22 semantic categories. The French toddler version (FCDI) called *Inventaire Français du Développement Communicatif chez le nourrisson: mots et phrases*, was adapted by Kern (1999). It contains 689 words organised in the same 22 semantic categories.

These two checklists are far from being perfect tools. However, since both are still fairly new and under-development, it is expected that they can be improved. It is

suggested, by the authors of the original American version, that the other language versions should be adaptations rather than literal translations. They should adapt the original American references to cultural values from the corresponding country. Despite that warning, the French vocabulary list contains very 'strange' items like the word *étendage* (which literally means spreading your clothes for drying purposes). The best guess here is that it is a literal translation from the American *dryer*. However, there is doubt about whether this word is ever used by French children if at all by anybody. The problem might lie in the fact that French speakers adapted the CDI. And only one English speaker (it is not mentioned if s/he was American) advised in case of a conflicting opinion. Kern (submitted: 7) states that they "tried to stay as close as possible to the American version, for purposes of cross-linguistic comparisons". Another major issue regarding the French adaptation is the fact that the functional categories have not been adapted to the morphology and grammatical structure of French. For example, for every possessive pronoun in English, French has three (or two) translations: masculine, feminine and plural (i.e.: *mon, ma, mes*). The French CDI puts the three words on one single line with only one box for parents to tick. Consequently, if French children do have a larger proportion of pronouns in their vocabulary (which makes sense since there are more in the language), then that will not be reflected in the analyses based on the French CDI. Other issues include the modals or auxiliaries. The French adaptation of the CDI has used literal translation of words originally called "helping verbs". So *could* becomes *pourrait*, *gotta/got to* becomes *avoir à faire*, etc. These are not auxiliaries in French and not even 'proper' expressions. My regret is that the French CDI is still too close to the original American version and is more a literal translation rather than an adaptation. However, the forms are still being developed and, hopefully, these issues can be improved. Some similar comments and regrets have been made by other researchers using this tool (Bradley and Lorch, 2003).

It was also reported by many of the parents taking part in the study that some words in the British adaptation sounded American English and would never be used by them. These included but were not limited to *pretzel* and *applesauce*.

Despite these drawbacks, the CDI remains one of the most reliable and valid tool that exists today (see below). Creating a new version would have been too time-consuming and would have involved a completely different study.

2.4.2.2.2. Can we trust parents?

The main issue that arises when using parental reports for research purposes is the question of the parents. Can they be trusted? It is widely believed that parents tend to inflate their child's abilities. Their natural pride as well as a lack of specialised training might produce a bias. However, parents have also far more experience than observers when it comes to understanding their own child. Popular books are full of references to language milestones and generally speaking parents have demonstrated much more knowledge than researchers might have credited them with. Generally speaking, parental reports assess current and new emergent behaviours only. There is no retrospective effect as that would prove to be much less reliable and more difficult for parents to assess. The fact that parental reports (and this one is no exception) have a recognition format does require less effort from the parents' part (Dale, Bates, Reznick & Morisset, 1989).

Most studies do assume, rightly or wrongly, that it is the mother who completes the CDI. A recent study by De Houwer & Bornstein (2001) on monolingual children acquiring Dutch compared results from fathers and mothers based on the completion of the CDI. They find that the better the skills of the child, the more disagreement between parents. This appears to be logical as parents need to remember more and more items. They find no specific rater bias. Finally they claim that the parents' of the child agree more together than they do with a third person (usually nannies). Reliability and validity analyses regarding parental reliability for the present study were carried out (see chapter 5 for further details).

2.4.2.2.3. Validity and reliability of the tools

The CDI remains one of the best parent report tool to analyse children's vocabulary (Thal, 2003). Its reliability has been proved many times over and these adaptations are also under going studies to prove their reliability (Robertson 1999, Marr 1999, Harrison 2001, Kern 2001). High internal consistency is usually reported. Correlations (across forms and languages) usually range between r=.60 and r=.83 between parental reports and laboratory tests or language samples.

The French CDI has only undergone two validity tests up to now. The first one is an internal consistency test done while the forms where being developed. However, Kern

(2001) reports that this only consisted in asking parents for their remarks regarding the words themselves. Secondly, they compare the lexical items present in the forms with the sentences parents wrote when asked to provide three typical utterances in the second part of the checklist. Their analysis shows that amongst words that occur more than 6 times in the sentences, 83% are present in the CDI (Kern, 2001). This score and others reported are significantly lower than for other languages (i.e.: Swedish reports 91%). Although they claim that the validity of the FCDI is adequate, not enough tests are available to judge of that at the moment. Other analyses are being carried out but have not yet been published.

The British English CDI is also undergoing validity and reliability analyses. Robertson (1999) and Marr (1999) report significant correlations between the British English long forms adaptations of the CDI and the Reynell Developmental Language Scales (a standardised clinical test of language skills). The correlations are $r=0.47$ and $r=0.67$ when compared with different parts of the Reynell test, both statistically significant ($p<0.05$). Marr (1999) acknowledges that the correlations are slightly lower than those reported by other adaptations of the CDI and the Reynell test. However, she claims that the differences can be due to the children themselves. Indeed, the children came from very different socio-economic backgrounds. Robertson (1999) also reports that parents were very consistent in checking the same items on both the CDI and the Language Development Survey (another parental checklist). 89.9% of lexical items checked were the same (only including the items that occurred on both lists). Once again, other analyses are currently being carried out.

2.4.2.2.4. Using the CDI with bilingual children

In addition to the problems raised by each CDI, further problems arise when wanting to use them with bilingual speakers. These tools are originally designed to be used on monolinguals. A few studies have used them with bilingual children. Among those, Pearson *et al.*, (1995) used the original American CDI together with its Spanish adaptation on bilingual children. The logic wants that when assessing a bilingual child, the two languages need to be assessed as one would not provide a full picture of what the child knows.

One of the issues that arise when asking parents of bilingual children to fill out these forms is that they might be more likely than monolingual parents to inflate their

child's vocabulary (Pearson *et al.*, 1995). Indeed, very often they might attribute a word to one language that was actually produced in the other language. However, it was judged that this problem might mostly arise when trying to judge comprehension. As far as production is concerned, it is easier to be accurate as the parent can refer to something that was more concrete. This can also be avoided by keeping a diary of the child's new words (this solution was adopted in this study for up to 50 words). Another issue when dealing with bilingual children is the problem of undetermined words. It has been acknowledged that bilingual children produce words that could belong to either language, especially early on in their lexical development (Deuchar & Quay, 2000). As far as the CDIs are concerned, if a child says something like *mama*, it is hard to judge from which language the word comes from. One could argue that this problem could also happen in the case of monolingual children. Children use one same word or sound that means many different things in the adult's lexicon.

2.4.2.2.5. Procedure of administration

The procedure followed for the delivery and analyses of the CDI data is similar to the standard procedure as described in the original CDI manual (Fenson *et al.*, 1993). The forms are self-explanatory.

Each parent was asked to fill in the MacArthur CDI monthly. The frequency proved quite time consuming but a close look at the development needed to be kept. The parents were sent or given, each month, a copy of the FCDI and BCDI with a return envelope. They were asked to complete them as they arrived. The next questionnaire was not sent until the previous one was returned in order not to influence their replies. Each parent would fill out the CDI that corresponds to the language s/he is speaking to the child. However, an issue arose that had not appeared in the literature of studies using the CDI with bilingual children because of the different linguistic situations. The first results suggested that there could be a difference between the form filled in by the mother who stays at home all day taking care of the child and the father's who works all day, for example. Most studies usually assume that it is the mother who completes the CDI. In this case, it was impossible for the mother to complete both forms as some had only limited knowledge in their non-native language and did not feel comfortable completing the form in that language. The only possibility was then for each parent to fill in the form corresponding to their native language. So it was

decided that two checklists would be obtained for the same language whenever it was possible. For example, if the mother is French and the father English, they live in England and speak English together, then the father would fill only English and the mother would fill both languages, as it is assumed that she uses English every day, maybe for work and so is quite fluent in English. If in the same situation the parents spoke French together then both parents would be asked from time to time to fill in both languages. The process was not repeated every month, as it was feared that the load would be too important. So the process was random. In the cases where both parents had to complete the same form at the same time, they were advised to do it independently of one another. The data for this study combines both sets (native and non-native parental lists). A brief mention of the difference in results will be made in the appropriate chapter.

The British English CDI was also completed by the two parents who are not native British English speakers. They were both given the British English version and another version (Australian and New Zealand) to compare them. Both estimated that their spoken English was closer to British English and preferred its content. To my knowledge there is no Swiss French adaptation of the CDI to this day.

As only the toddlers' forms were used, the process started at 16 months and stopped at 30 months. Not all the families filled in all of the questionnaires regularly, so the number of checklists is not the same for each child. The parents were warned of the typical mistakes that could be made by parents of bilingual children. They were explicitly explained that if they tick the word *cat* in English, for example, it means that the child actually produced the word *cat* and not *chat* in French. Table 3.2 represents the age (in months) of the child for which CDI data is available.

The forms were scored following the standard procedure. Both forms were scored separately. To obtain a child's total score for vocabulary production, simply add all the items ticked on one form across the 22 categories for a given child at a given age. The grammar questions were completed by parents but not scored or used in this study. To obtain a bilingual total score, simply add the total of the two languages together.

A fourth vocabulary score unique for bilingual children was also created. This is called the TCV or Total Conceptual Vocabulary as described in Pearson (1998) (see chapter 3 for a discussion how this is calculated and why).

Table 2.2 Parental checklists collected

THO	17	20	21	**22**	23	**24**	25	26	27	28					
FLO	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
ANT	16	17	18	19	20	21	**22**	**23**	24	25	28				
EMM	16	18	20	21	23										
DUN	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
PEN	17	18	21	22	**23**	26									
ELI	16	17	18	19	20	**21**	22	23	24	25					
JAC	16	17	18	19	20	21	22	23	24	26	27	29	30		
ANN	16	17	18	19	20	21	22	23	24	25	26	27	28	**29**	30
OLR	16	17	18	19	20	21	22	23	24	25	27	28	29	30	
LIA	16	20	23	27	30										
REB	16	17	19	**20**	21	23	24	26	**29**	30					
OLB	16	17	18	**19**	**20**	21	23	24	25	27	28	29	30		

Numbers in bold mean that only one out of the two CDIs was completed.

2.4.3. Language samples

A wide variety of recordings were made either by the parents themselves or by the researcher. A total of more than 40 hours of naturalistic recording were obtained. The results for each child are very different depending on the parents' willingness. The typical frequency in child language research for recordings is usually at intervals of two to four weeks. Although, ideally, we would have liked the recordings to be regular (once a week), they were quite irregular for most families, depending on their own situation. Some were more regular than others. The recordings are either audio-video or audio only depending on what was available to the parents at the time. The recordings were either made with the parents' own recording device or with one borrowed from the researcher.

In order to obtain as naturalistic speech samples as possible, the presence of the external observer was limited. It is reported in the literature that in order to reduce the effects of the so-called observer's paradox, the presence of the researcher should be limited (Bennett-Kastor, 1988). The parents were advised to record themselves with

the child during everyday activities such as meal times, dressing-up, bath... Each parent was asked to record himself alone with the child in order to obtain single language contexts as well as with the whole family together.

Recordings made by the researcher herself were made in the language of the parent present at the time, or the family language if both were there. These were, almost always, of better acoustic quality and targeted to test a child's specific aspect of language development.

The vast majority of recordings were orthographically transcribed using the CHAT and LIDES format of CHILDES (see LIDES coding manual from the LIPPS Group, 2000). The utterances were especially coded for switches/mixes. Only recordings of a poor sound quality were excluded. The recordings were transcribed by the author, a French-English bilingual. Some recordings were checked for reliability by a couple of bilingual speakers until full agreement was reached.

After having highlighted the main methods to be used in the present study, the next chapter will present the results of the quantitative data: the bilingual lexicon.

SECTION II:

BILINGUAL LEXICAL ACQUISITION

"When I use a word," Humpty Dumpty said, in rather a scornful tone,
"it means just what I choose it to mean—neither more nor less."
"The question is," said Alice, "whether you can make words mean so many different things."
(Lewis Carroll, *Through the Looking-Glass,* chapter VI)

CHAPTER 3:

PROFILING THE LEXICON

3.1. INTRODUCTION

Profiling is a term most commonly used in psychology to mean the recording and analysis of a person's psychological characteristics, (e.g.: personality traits, socio-economic background, etc...) in order to predict or assess their suitability or ability in a given sphere. Speech and language therapists also use that method as part of their assessment. Profiling is presenting a holistic picture of the strengths and weaknesses of what or who is being studied. Here, I will be using lexical profiling in the sense of establishing the nature of the bilingual lexicon, its various strengths and weaknesses or special characteristics. It is therefore my intention to present a clear picture of the developing bilingual lexicon.

In this chapter, having primarily described the aims, the monolingual norms in the languages we are dealing with in this study (English and French) will be presented. Secondly, I describe the data analysis procedure. The largest part of this chapter then defines and compares the results in three different sections. I will first present each lexicon of the bilingual child separately in order to assess its similarities with monolingual data. Then, a comparison of the two languages in question will be drawn and finally, I establish the special features of the bilingual lexicon. Finally, a discussion of the results obtained here will follow.

3.2. AIMS AND OBJECTIVES OF THIS CHAPTER

The aims and objectives of this chapter are:

- Whether each separate language of the bilingual lexicon is similar to the developing monolingual lexicon in the same language;
- To describe and compare the two languages and check for cross-linguistic influences; and
- To establish the special features of the bilingual lexicon.

Although receptive vocabulary was mentioned very briefly earlier, in this chapter, the present study will examine productive vocabulary only.

3.3. MONOLINGUAL NORMS

First I shall present the studies that can be considered as having set monolingual norms for, firstly, language acquisition in English and then consequently, various Romance languages, especially French.

3.3.1. English

In an early study of lexical development, Bloom (1973) reported that "there are more adult-noun words than any other 'part of speech'" (p.110) in her daughter's lexicon at 16 months old. No count was made of the percentage of each category. However, if we look at the 50-word milestone, around 42% of Allison's lexicon was nouns.

Benedict (1979), in a study of early production and comprehension, found that 61% of the first 50 words produced by eight monolingual English-speaking children were nominals. Nominals included words that refer to objects or people, e.g.: *Daddy*, *cat*, etc. Benedict also included pronouns in the nominal category. 19% were action words (social games, locatives, verbs...) and 10% modifiers. Although that categorisation is debatable, the results showed that nouns are clearly the most common category in early lexical development. This is also accurate for comprehension. Benedict (1979) reports that 56% of words comprehended are nominals by the 50-word stage.

The studies mentioned above are however, only small scale studies. Bates *et al.* (1994) conducted one of the first large scale studies of lexical development. Their data comes from the MacArthur Communicative Developmental Inventory. They tested 1803 English-speaking monolingual children. Their results determined three "waves" in lexical acquisition. The first wave, to the 100-word stage, is characterised by an increase in the percentage of common nouns in the child's vocabulary. The second wave sees a decrease in the percentage of nouns and a slow and linear increase in predicates (verbs and adjectives). The second wave typically takes place when the child has between 100 and 400 words in his/her lexicon. Finally, the third wave, which starts when the child reaches the 400-word stage, is a sharp increase in the percentage of closed-class items, which had until then been stable.

At this stage, it is necessary to briefly mention comprehension. We know that early vocabulary growth is very different for comprehension. At the 50-word stage, comprehension is on average five months ahead of production. The rate of word acquisition for comprehension is twice that of production. 22 and 9 new words are acquired per month respectively for comprehension and production (Nelson, 1973; Benedict, 1979; Goldfield & Reznick, 1990). So the 50-word stage for comprehension is said to be reached around 1;1, while for production it seems to be around 1;6.

3.3.2. French and other Romance languages

The 50-word stage is a very important milestone in child language development as it is from that point on that children start combining words (Locke, 1997). Boysson-Bardies (1999) reported that French-speaking children with a lexicon just smaller than 50 words had 68.5% of nouns 21.6% of predicates and 9.9% of other words. Here again, the classification is arguable, but the overall tendency seems to be very similar to that earlier for English-speaking children.

Overall, the same developmental patterns have been observed for monolingual French-speaking children as for monolingual English-speaking ones. These stages have been cross-linguistically recognised as accurate, for monolinguals at least.
Bassano, Maillochon & Eme (1998) reported no developmental progression of the lexical diversity of a French child. All categories are present almost from the very beginning. As expected, they report that nouns largely dominate until 1;8. After that age, the proportion of nouns in the lexicon decreases to make space for predicates and "grammatical words". The first one has a slow and steady progression, while the second category seems to have a more dramatic increase after the age of 2.
A French version of the CDI, based on the original American one, has recently been developed. The normative study is still under way. Preliminary results tend to generally concur with those reported for other languages. Kern (2001) found that the noun category increases until the 100-word stage and then remains quite stable until the lexicon reaches 500 words. Predicates increase regularly as a function of the size of the lexicon. They also found that children with a vocabulary of less than 50 words tend to produce as many closed-class items as predicates.

Moving on from that 3-stage approach, Caselli *et al.* (1999) added a fourth stage, after a study of monolingual Italian children based on the Italian adaptation of the CDI. They proposed a four-stage model of lexical development from their data on Italian and the original American data. Those four stages are *Routines and word games, Reference, Predication and Grammar.* From the onset of speech up to the acquisition of about 10 words in their productive vocabulary, children are said to be in the *Routines and word games* stage. During that time, children acquire social words. What will be called social words are items like *bye bye* or animal sounds. Very often it is quite impossible to classify those words into grammatical categories. These words are referred to as social words as they allow children to have a social interaction with their environment. The second stage (*Reference*) occurs when the child has between 50 and 200 words. That stage corresponds to the rapid acquisition of nouns. Those nouns, most of the time, refer to concrete objects: *table, cat, dress,* etc. The third stage sees the appearance of verbs and adjectives (predicates). This stage may start as early as the 100-word mark, with stages 2 and 3 actually overlapping. Finally, during the last stage, children acquire "grammatical words" (like pronouns, prepositions ...), rare until then. This stage also correlates with the emergence of productive grammar.
Other results for Spanish-speaking children (Jackson-Maldonado, Thal, Bates, Marchman & Gutierrez-Clellen, 1993) have confirmed such stages.

3.3.3. Other (non-Romance) languages

Other types of languages not belonging to the Romance family have also given very similar results. Eriksson (2001) reported the same 4 stages for Swedish monolingual children in a study based on the Swedish adaptation of the CDI. A study of children acquiring Hebrew (Maital, Dromi, Sagi & Bornstein, 2000) has also confirmed these stages. The authors concluded that early lexical development in Hebrew and English followed "remarkably similar developmental patterns" (Maital *et al.*, 2000). These results are even more significant considering the fact that the two target languages are typologically different. Other reports based on Mandarin for example, have shown that the noun-bias hypothesis claiming that nouns are acquired before verbs might not

be consistently verified. Tardif, Shatz & Naigles (1997) found that Mandarin-speaking children produced more verbs than English-speaking children at earlier stages. The main reason behind that was that the speech of English-speaking care givers emphasised nouns over verbs.

3.3.4. Cross–linguistic differences (English vs. Romance languages)

All developmental studies on lexical acquisition seem to agree on the overall distribution of grammatical categories. Some very minor differences have been found between Romance languages (particularly French) and English in studies based on the different adaptations of the CDI or single case studies. A very interesting finding from Bassano *et al.* (1998), is the fact that the closed-class category is proportionally always much larger in French than the same category in English-speaking children. According to Bates *et al.* (1994), the closed-class category never represents more than 15% of the whole lexicon until at least 30 months in English. While, Bassano *et al.* (1998) found that at 1;2 (when she had less than 50 words in her lexicon), Pauline, the child they studied longitudinally, had 20% of closed-class items in her total lexicon. Despite these findings, we know that there is considerable variation amongst children in the early stages of language acquisition. Such results are based on a single case study. In a comparative cross-sectional corpus, Bassano *et al.* (1998) found that at 1;8 and 2;6, between 8% and 45% of the French-speaking children's lexicon were closed-class items (above 20% on average). Parisse & Le Normand (2001) studied different corpora of English and French acquiring children between the ages of 1;6 and 2;6. They established that English acquiring children were more likely to produce nouns, "communicators" and adjectives, whereas, French-speaking children were more likely to produce relative pronouns amongst other categories. Another study of a romance language seems to confirm the results found by Bassano. Caselli *et al.* (1999) compared Italian to English lexical data both based on the Communicative Developmental Inventory. They also found that the proportion of the closed-class category in the lexicon was always slightly higher for Italian children. A possible reason for that difference would be that French, and other Romance languages like Italian, do contain a lot more closed-class items. For example, for a single word like the possessive *my* in English, there are at least three possible translations in French:

mon, ma, mes, which are respectively the masculine, feminine and plural forms. Italian has *mio, mia, miei* and *mie* whilst Spanish has two forms *mi* and *mis*. However, a study by Girouard, Ricard & Décarie (1997) on the acquisition of personal pronouns highlighted a difference. They found that generally speaking, French and English learning children acquired personal pronouns in the same order and at the same rhythm. However, they noticed that in situations where the children were non-addressed listeners, French speaking children were delayed in comprehending the third person pronoun. Girouard *et al.* (1997) suggest that the combination of two factors can explain that native language effect. The two factors are seen as the complexity of the third person pronouns system and the difficulty of the non-addressee system.

Another cross-linguistic difference between French and English worth mentioning here is to be found in the predicates category. The predicates categories also contain slight differences in developments and quantities. English is seen as having a basic SVO word order. Although French also usually has an SVO structure, it sometimes changes to SOV order when the object is pronominalised (e.g.: *Je l'ai mangé*). French also has an extremely rich system of verb morphology and verb agreement plays a crucial role in conveying basic sentence relations compared to English. So, theoretically speaking, French could provide a basis for earlier verb acquisition or a larger proportion in the lexicon. Bassano *et al.* (1998), found a slight difference between English and French. They reported a slightly larger number of predicates initially in French. Predicates in the French-speaking children's lexicon represented on average 18% when around the 50-word stage. However, they only represented 8% for the American children according to Caselli *et al.* (1999). Similarly, Boysson-Bardies (1999) found that French children with a lexicon inferior to 50 words had 21% of verbs, while American children had only 9%. Parisse and Le Normand (2001) also claimed that verbs were significantly fewer in the lexicon of English-speaking children between 1;6 and 2;6. However, a study on another Romance language with a rich verbal morphology does not show the same results at all. Italian is a so-called 'pro-drop' language, in which the verb may be produced without an explicit referential or pronominal subject constituent. Therefore, the subject can be dropped and the first word consequently becomes the verb. In spite of their important role, Caselli *et al.* (1999) found no evidence to prove that verbs would appear earlier in

Italian children. The data shows that up until the 300-word stage, Italian-speaking children had proportionally less predicates in their lexicon, even if the difference was minimal at the later stages.

Although Hebrew is not a Romance language, it is a language where the word order is relatively free and where verb initial constructions are commonly found in everyday speech. However, Maital *et al.* (2000) did not find any conclusive evidence that predicates (or verbs) appeared earlier or were more prominent in Hebrew compared with English (respectively 4% and 8% at the below-50-word stage).

3.3.5. Classifying early words

Many investigators have attempted to classify early words produced by children, with a plethora of differing methodologies having been implemented. The main taxonomies used are: (i) adult parts-of speech (e.g. Dromi, 1996); (ii) grammatical categories that correspond to parts-of-speech (e.g. Benedict, 1979); (iii) semantic relations (e.g. Greenfield & Smith, 1976); (iv) pragmatic categories (e.g. Halliday, 1975). The most common issues related to these taxonomies have been the identification of similarities and/or differences between the child and the adult linguistic systems, determining the proportion of words in each of these categories in the developing lexicon, and the order of emergence of these categories. Some of the researchers using parts-of-speech as the basis for classifying early words have highlighted the issues relating to these formal categories, arguing that early words are semantically distributed. That is why Bloom (1973), amongst others, has proposed semantic or content based categories to avoid this incompatibility. However, her model does not include any mismatch that may have occurred between the meanings of adults' and children's forms. One of the most detailed attempts to classify children's early words is probably Greenfield & Smith's (1976). They proposed a model in which each class was defined according to the semantic relation expressed by the utterance or the non-linguistic context for single-word utterances. The issue with these categories was that they were analysed dependently of adult knowledge of the grammatical categories of words. Therefore, the classification was motivated by adult linguistic intuition and not by the non-linguistic cues or the context. Finally,

pragmatic-based categories focused on how young speakers can convey various communicative intents. Recent lexical development research using parental reports has largely classified early words in content based categories that correspond to parts-of-speech (e.g. Bates *et al.*, 1994).

3.4. THE NATURE OF THE DEVELOPING BILINGUAL LEXICON

To the best of my knowledge, no research has been done with bilingual children to establish the true nature of the bilingual lexicon. It would be interesting to know whether the cross-linguistic differences found in monolingual development are present in the bilingual lexicon, e.g.: whether French-English bilingual children produce more verbs and closed-class items in French.

In terms of pure quantitative analyses, the bilingual lexicon is not easily comparable with a monolingual one due to its composition. Looking at just one language does not provide an appropriate score, as that is only half of the real lexicon (Pearson, 1998). Examining a total of both lexicons includes some overlaps between the lexicons, especially in terms of cognates. The solution devised by Pearson (1998) is a TCV, as described earlier, accounting for concepts rather than lexical items.

Next, will be presented the methods used for this part of the data analyses. This is followed by a discussion of the results obtained following the three main research questions highlighted above. Finally, the results obtained are discussed.

3.5. METHODOLOGY: DATA REDUCTION

Most of the analyses in this chapter are based on the two adaptations of the MacArthur CDI (see chapter 2) completed by the 13 families every month. Both CDIs are organised into 22 semantic categories. These categories needed to be re-classified in order to provide a more grammatical approach to the nature of the bilingual lexicon. In order to determine how the lexicon is organised, it appeared necessary to classify the reported words into content-based grammatical categories, which are mutually exclusive. Since the data is taken from the CDI, there is no contextual information. By classifying the words under grammatical categories, I do not want to say that children

recognise grammatical categories like noun, verb or adverb, as there are some examples where they deviate from the adult pattern even for a same word. However, children's words have largely been classified according to that of adults' part-of-speech (English or French) in the literature. The categories are seen as variables. Due to the input received, children treat nouns differently than verbs. However, it does not imply the child's ability to realise the emergence of such categories either. This methodology only allows the researcher to point out cross-linguistic differences in the vocabulary composition.

The categories used are the same as those used by Bates *et al.* (1994) from data obtained thanks to the CDI. Those categories were also used to compare American and Italian data (Caselli *et al.* 1999). Similar categories were used by Dromi (1996): *object words, action/state words, modifiers, social words.* The only difference being Dromi created an *indeterminant* category to put in anything that did not fit into any of the other categories. The categories are as follows:

Common nouns:

Animals

Toys

Vehicles

Food

Clothing

Body parts

Furniture and rooms

Outside things

Household objects

Predicates:

Action words

Adjectives

Closed-class category:

Helping verbs

Connecting words

Quantifiers and articles

Prepositions and locations
Question words
Pronouns

Social words:
Sound effects
People
Games and routines

Two categories of the CDI are left out: *Places to go* and *Time words*. The reason being, that the words contained in these categories can be adverbs as well as nouns both in French and in English. Overall, this classification can appear arbitrary and arguable, especially for social words or expressions (e.g. *vroom* or *give me five!*). Consequently, cross-linguistic differences that can be over-turned by one or two items should be taken very carefully. These categories almost correspond to the function word/content word classification. Content words would approximately be common nouns and predicates. Although this classification was originally used in English, it was also successfully used in Italian, Swedish, Hebrew and other languages including French. The only issue that might arise in French is the auxiliary category. Modals in English (*can, could...*) are not equivalent to their translated forms in French like *pouvoir, vouloir* and *devoir*. These French verbs behave like normal lexical verbs, especially in terms of the inflections they take (Jones, 1996). The only true auxiliaries are *être* and *avoir* which occur in compound tenses (Parisse, 2002). The FCDI does not have many real auxiliaries in its auxiliary category. However, it was decided that the category would be kept within the closed-class items for several reasons. Very young children produce very few auxiliaries (so the choice will not dramatically change the results). As the categories do not aspire to be real grammatical categories recognised by the children, auxiliary category can be left in the closed-class part. Some verbs such as *aller, avoir* and *faire* are also in the *action words* category (therefore, also in the predicates). So, the forms proposed as auxiliaries are thought to be taken as such: *fait, a fait*, etc. Moreover, these verbs do not carry the main semantic meaning. They are all to be attached to something else. Their role is largely if not wholly grammatical. One can consider auxiliaries as: verbs that help the main verb in expressing inflection (Jones, 1996). Even if some are more modals than

auxiliaries: e.g. *vouloir*, they do not carry the main semantic meaning. This category could actually be renamed inflection verbs. As such, they are to be classified with the closed-class items, being simply non-action verbs (Bassano, 2000). Finally, Kern (2001) made the same choice for French monolingual lexicon analyses, although she did not account for social words in her classification. Another possible choice could have been based on the system and content morpheme distinction (e.g.: Myers-Scotton, 1993) especially since it is a useful notion to then analyse language mixing in bilinguals. Yet, no context information was available. Without contextual information it is rather impossible to make a clear distinction between the two.
The CDI also has a number of grammatical questions at the end of each form. However, I will not consider the grammatical questions here.

Before starting an analysis of the results, it should be emphasised that the children studied are bilingual. It has been recognised that normative guidelines based on monolingual children give inaccurate predictions for bilingual populations. Comparisons of number of words should therefore be avoided (Pearson, 1998). For this reason, this kind of comparison will be avoided. The aim here is not a quantitative study of the composition of the bilingual lexicon, but rather a qualitative one, although quantitative methods shall be used for some of the analysis.

3.6. RESULTS: COMPARING AND DESCRIBING

3.6.1. The French lexicon

Taking into account the high rate of variability among children at the same age, grammatical distribution was considered as a function of vocabulary size and not as a function of age. The entire sample was divided into seven categories according to the size of the produced vocabulary in that language (not the total of both languages): (1) less than 50 words, (2) 51-100 words, (3) 101-200 words, (4) 201-300 words, (5) 301-400 words, (6) 401-500 words and (7) more than 500 words.

The early distribution of the French lexicon of the bilingual children (less than or equal to 50 words), indicates that, as previously found, it is mainly composed of

nouns (29.2%) and social words (57.7%). Closed-class items represent the smallest category (5.8%).

Table 3.1 French lexicon composition (for 50 words or less)

	Mean percentage
Common Nouns	29.2%
Predicates	7.4%
Closed-class items	5.8%
Social words	57.7%

Below is the mean percentage of each lexical category represented as a function of lexicon size (French only) for the 13 children involved in the study.

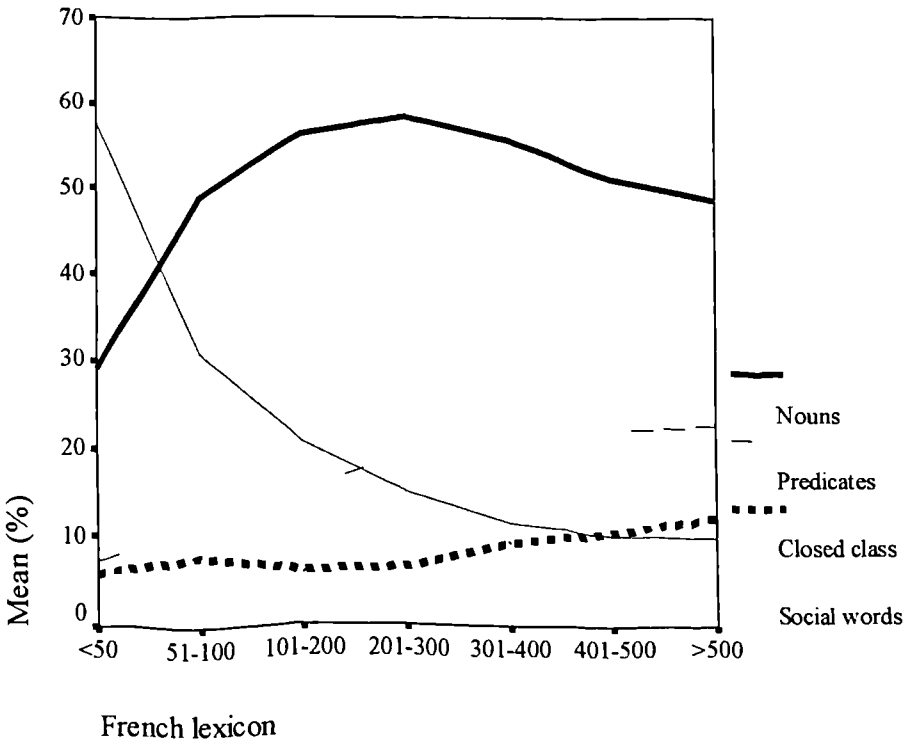


Figure 3.1 Composition and evolution of the French lexicon

Here again, the closed-class category represents most of the items produced for children with a lexicon smaller than 50 words. This category, however, decreases very rapidly, with nouns becoming the biggest category. They increase rapidly until around 100 words and then stabilise. A slight decrease can be observed towards the

end of the study period. As the decrease of the proportion of social words triggers (or is triggered by) the increase of nouns, the slightly smaller proportion of nouns appears to be due to the slow and linear increase of the two other categories: closed-class items and predicates. Attention should be drawn to the fact that there seems to be a slight drop in the proportion of closed-class items between 100 and 300 words.

3.6.2. The English lexicon

The English part of the lexicon is distributed as follows for the children with a vocabulary smaller than 50 words.

Table 3.2 English lexicon composition (for 50 words or less)

	Mean percentage
Common Nouns	28.4%
Predicates	4.9%
Closed-class items	3.4%
Social words	63.4%

As we can see, here again, social words represent the larger category with 63.4%. The predicates and closed-class items are the smallest ones with respectively 4.9% and 3.4%.

The evolution of the different categories (figure 3.2) reveals that nouns constitute the most produced items, with the exception of children with a repertoire fewer than 50 words. Several clear trends are immediately noticeable. The largest category at the below-50-stage is by far the social words category. This category falls quickly to the benefit of nouns. Around the 100-word landmark, the common nouns category stabilises before slowly decreasing. At the same time, the number of predicates and closed-class items start increasing. The first one rises more quickly than the second one.

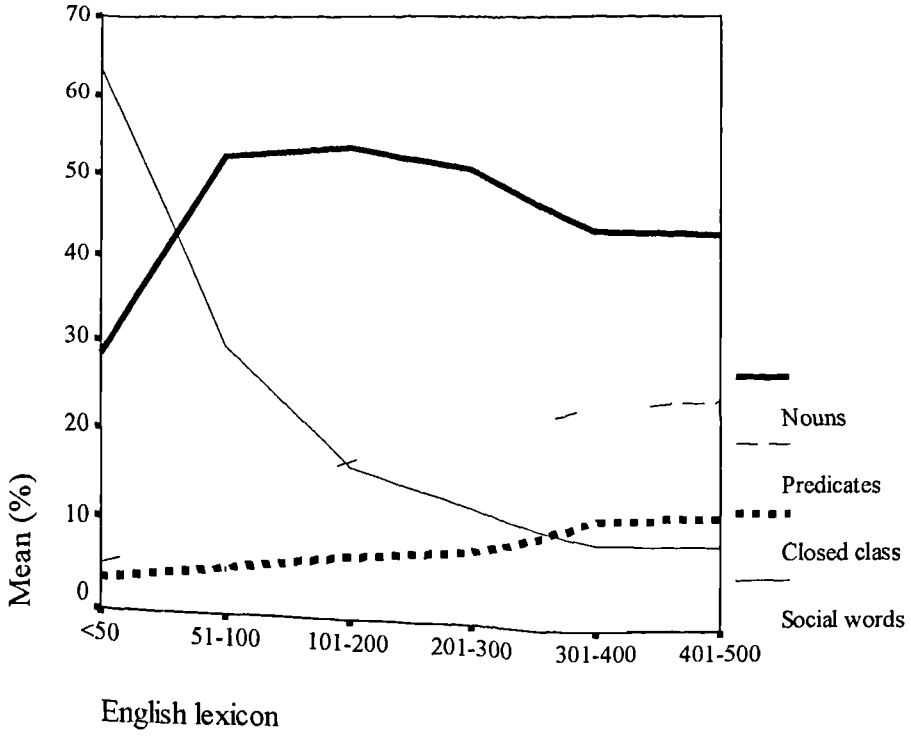


Figure 3.2 Composition and evolution of the English lexicon

3.6.3. Cross-linguistic differences and influences

Having described the results for each language separately, it is now possible to compare the two languages with each other. The fact that we are dealing with bilingual children allows comparisons between the data from one language to that of the other one for the same set of children.

It is important to note at this stage that none of the children reached the last stage (more than 500 words) in their English lexicon. Therefore, the category is not represented on the following graphs. The reason for this is not necessarily that they were all dominant in French. Although this may be true for some children, it indicates that they reached the upper limits of the FCDI more quickly than those of the BCDI. The evolution of each type of lexical category will now be compared for both languages.

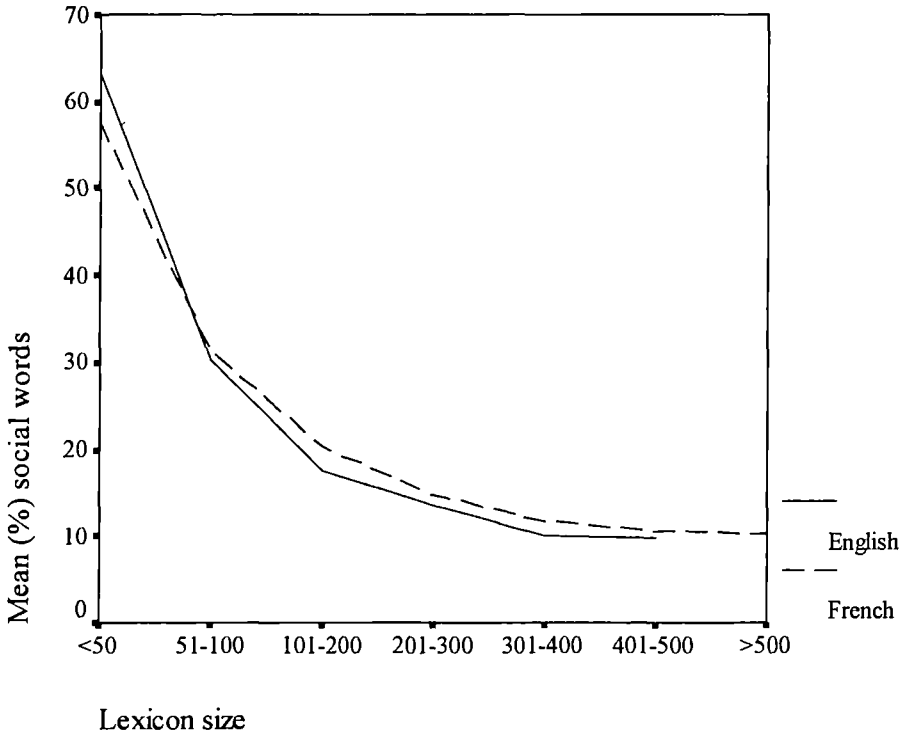


Figure 3.3 Evolution of the social words category

As we can see on figure 3.3, the social words category follows a very similar pattern, if not parallel in each language. This is the most striking example. The two lines are almost exactly the same for the two languages, with the social words representing about 10% of the total vocabulary at the lowest point in both English and French and about 60% at the highest.

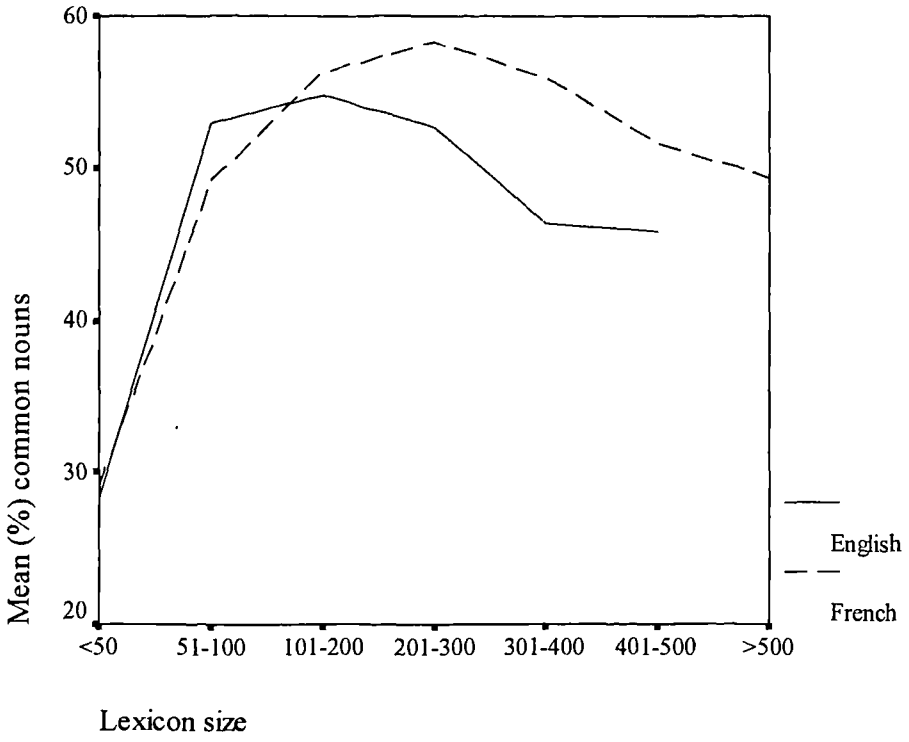


Figure 3.4 Evolution of the common nouns category

The common nouns category also follows a parallel pattern. However, the English category seems to decrease more rapidly than the French one. By the 400-500-word stage, the children have reached the ceiling proportion of nouns in their English lexicon (45%), whilst the French lexicon have not yet caught up by the 500-word stage (ceiling is 47.6%, mean is 49.3%).

As we can see, these two categories do not show any major differences. The development of French and English lexicons appears to be rather similar, to this point, even for bilingual children. The closed-class items and predicates will now be examined closer as those were the categories where differences were found in the literature.

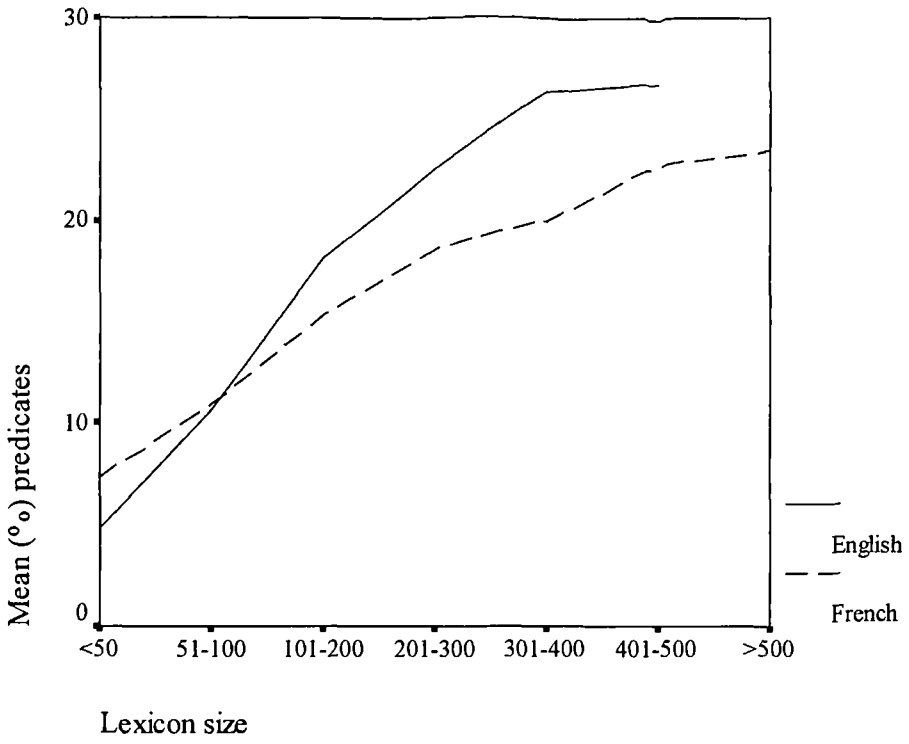


Figure 3.5 Evolution of the predicates category

The only noticeable difference between the two lines on figure 3.5 representing predicates, is their slightly larger number in the French lexicon for children with the smallest vocabulary (BCDI: 4.8%, FCDI: 7.3%). However, this trend fades away as the vocabulary increases and even inverses by 100 words. From 50 words, the number of predicates rises much quicker in English than in French. At their highest point, respectively for a lexicon of 401-500 and above 500 words, predicates represent 26.8% in English and 23.4% in French. It is important to note that the ceilings of the predicates category are 24.7% for the BCDI and 24.2% for the FCDI.

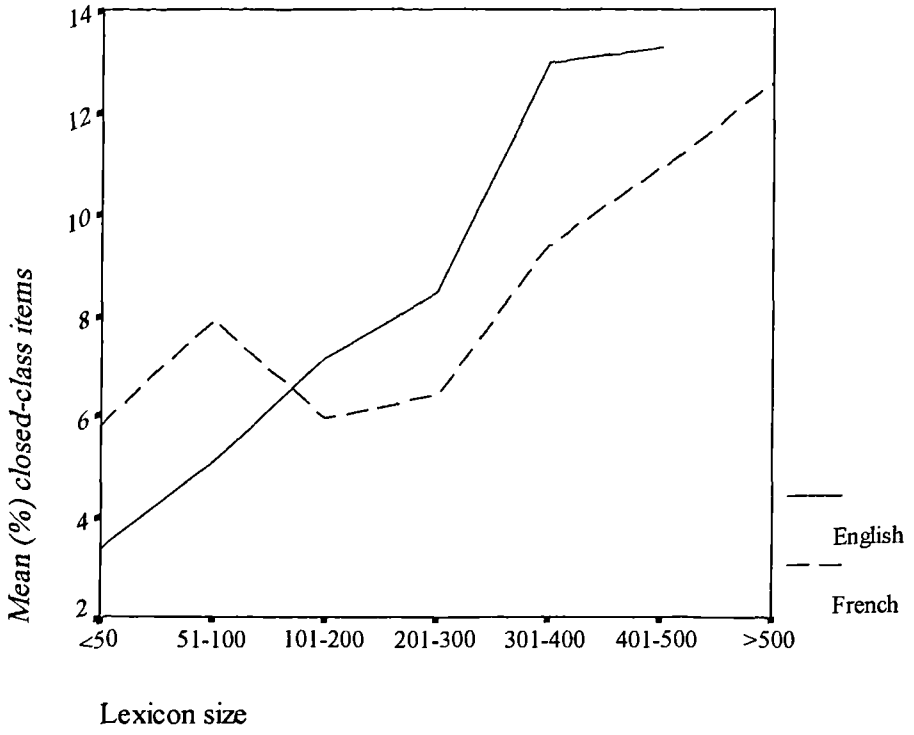


Figure 3.6 Evolution of the closed-class category

Finally, the closed-class category is about the only one that does not follow a parallel pattern in the two languages. Here again, the French lexicon seems to contain more closed-class items for children with the smallest vocabulary. However, this trend inverses from about 100 words. The ceiling percentages for this category are 15.1% and 13.4% respectively for the BCDI and FCDI. The bilingual children by 500 words reaches on average the 12.5% mark in French, while their English lexicon reaches 13.2% of closed-class items. That shows that they almost reached the maximum ceilings, even more so in French than in English. Something that had been noticed earlier, but which is clearer here, is that the proportion of French closed-class items decreases between 50 and 100 words before increasing again.

3.6.4. Special features of the bilingual lexicon

So, bilingual children do appear to be developing similarly to monolingual when comparing each language separately. The debate about bilingual acquisition being different or similar to monolingual acquisition can be looked at in a different view with the present data. Previous research of that kind has compared the means of

vocabulary production between bilinguals and monolinguals (Pearson *et al.*, 1993). Looking at these means for the present data provides an insight into bilinguals' abilities when compared to monolinguals in terms of lexical acquisition. Figure 3.7 below represents the means of production for all 13 children over the whole period of study for both FCDI and BCDI for monolingual children, plus the bilingual scores from the present bilingual study for the two languages separately, (Fre-Biling and Eng-Biling) as well as more 'typical' bilingual measures, the total of the two languages (Biling Fre+Eng) and the TCV score. TCV (total conceptual vocabulary) is a recently developed count (see Pearson, 1998) which has been described earlier in chapter 1. That measure is thought to best represent bilinguals' lexicons, as total scores contain a certain amount of overlap between the two languages.

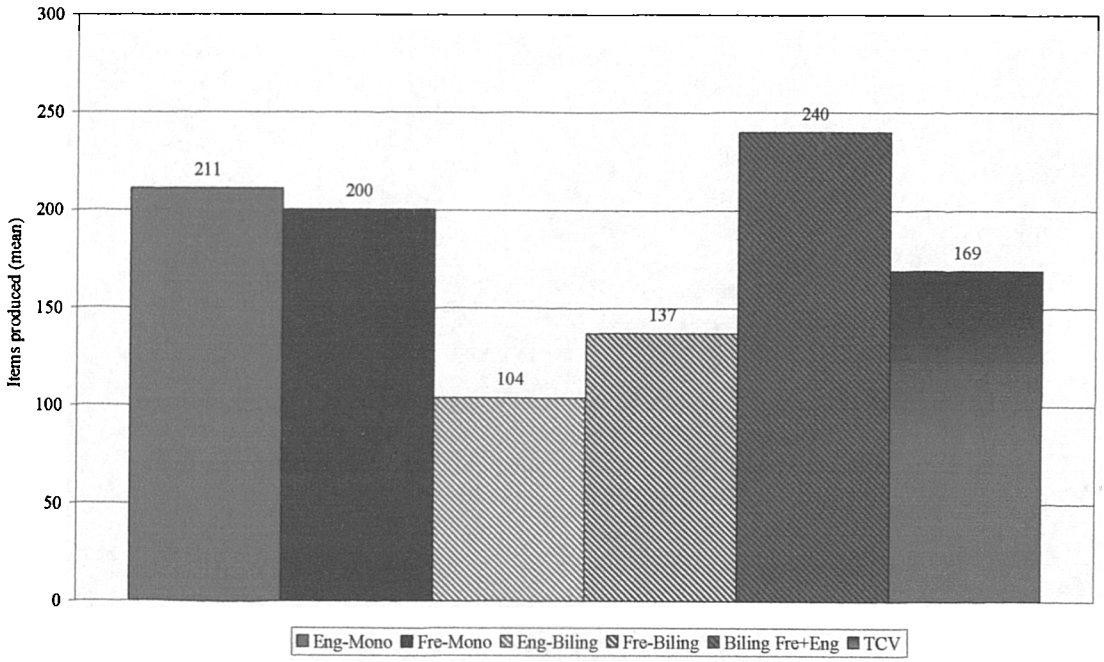


Figure 3.7 Mean number of items produced for monolinguals and bilinguals

Overall then, bilingual children appear to be producing more lexical items than monolingual children if we simply add French + English. The TCV score is slightly lower however, from the monolinguals' average scores. Finally, looking at only one of the languages of the bilingual child would rank that child in the lowest percentiles on a monolingual test as the average mean number of items produced in one language is significantly lower for bilinguals than for monolinguals. These results correlate with previous studies such as Pearson *et al.*'s (1993). It is also interesting to examine

these scores from a developmental perspective, which has not been done by previous studies.

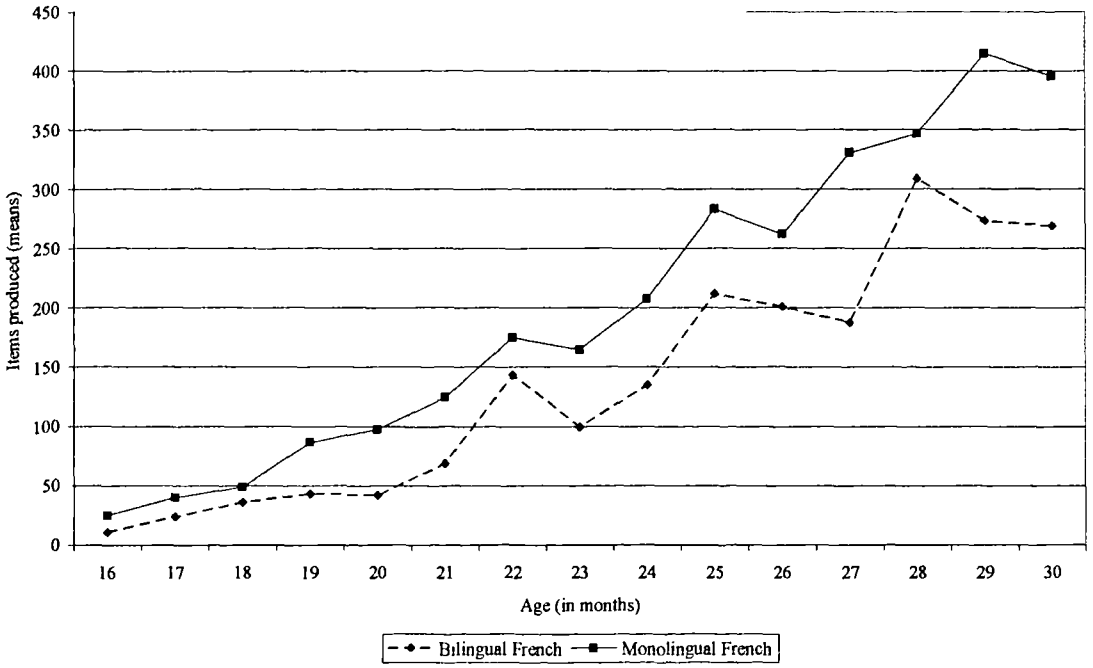


Figure 3.8 Monolingual and bilingual data in French

Figure 3.8 reveals an almost perfect parallel development between monolingual French children based on the FCDI data and the French scores of the bilingual children in this study. We notice that when the lexicon grows quicker in monolinguals (e.g. between 23 and 25 months old), the 13 bilinguals of the study do behave similarly. In addition, when the lexicon slows down or the average even decreases (e.g. between 22 and 23 months old), bilinguals exhibit the same pattern.

Table 3.3 reveals the lexical developmental trends for the same scores comparing monolinguals with bilinguals. First of all, it is possible to say that the rate of acquisition is overall similar. What the table highlights, is the fact that when comparing monolinguals and bilinguals' scores on one language, bilinguals clearly score lower on average. When looking at the total (a simple addition of English + French), bilingual children can be said to have a larger vocabulary than monolinguals except in the first stages, in which the difference is less significant. Finally, if taking into account TCV as their actual lexicon, they score lower than monolinguals. Language dominance was also accounted for based simply on raw lexical scores of the bilingual children. These did not provide any significant difference from the

results provided above. For example, children dominant in French were still clearly below the French CDI average for monolinguals, although to a lesser extent (Bilinguals French dominant mean = 156, monolinguals French = 200).

Table 3.3 Evolution of the total lexicons of bilinguals and monolinguals

Ages (in months)	BCDI	FCDI	BilEng	BilFre	Biltotal (Fre+Eng)	TCV
16-18	36	38	15	24	38	31
19-21	82	103	35	51	85	70
22-24	167	183	89	126	213	154
25-27	346	292	146	200	346	244
28-30	423	386	235	284	515	347
MEAN	211	200	104	137	240	169
STDV	168	140	89	107	195	129

We have, up to now, looked at the two lexicons separately in order to see if bilingual children do behave as monolingual ones respectively in French and English. However, we are not dealing here with monolingual children but with bilingual ones. So the comparisons will not go any further. We should now look at both lexicons combined together and not each one separately. By doing so, it might be possible to shed more light onto some issues, such as the larger number (or not) of predicates in the early lexicon in French. There might be overlap between the two lexicons, but by combining the two lexicons, we do not aim at making any direct comparisons with other populations.

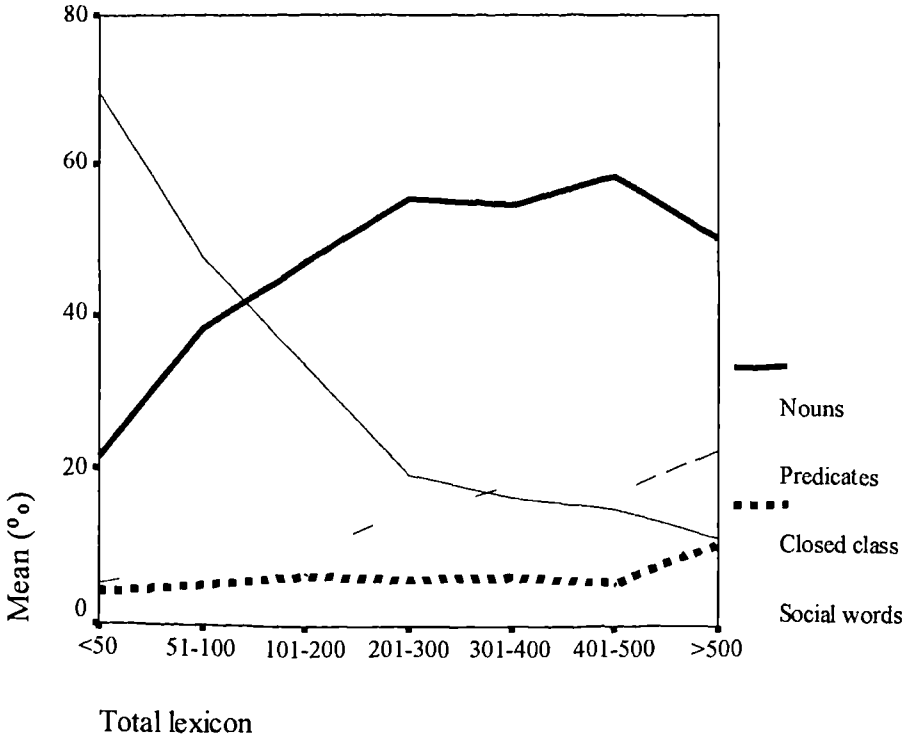


Figure 3.9 Evolution of the categories in the bilingual lexicon (both languages combined)

Firstly, as we can see on figure 3.9, the overall shape of the total bilingual lexicon (French + English) reveals the same pattern as when both languages are separated, although each stage (or wave) seems to be taking longer to be achieved. For example, the common nouns category does not reach its peak until the 300-400-word stage. This is valid for all word categories.

Let us now take a more detailed example of a child whose lexicon is 'unbalanced'. By this, it is meant that ANT has a much larger vocabulary in French than in English. Despite living in England, she had, at the time of the study, little contact with English speakers. She spent most of her days at home with her French-speaking mother. By age 2;0, her productive vocabulary is composed of 310 and 53 words, respectively in French and in English. As table 4 shows, ANT is clearly at stage 1 of lexical development in English but already reaching stage 4 in French. The main argument here is that even though one language seems clearly dominant, English is still developing according to the stages expected as described above.

Table 3.4 ANT's lexicon

	French	English
Common nouns	181 (58.4%)	36 (67.9%)
Predicates	66 (21.3%)	3 (5.7%)
Closed-class	21 (6.8%)	1 (1.9%)
Social words	38 (12.3%)	12 (22.6%)

The bilingual lexicon allows us to bring more evidence into the issue of earlier acquisition of predicates and closed-class items in French. A list was drawn up for ten children, of their first 10 words in order to investigate further the issue of a possible earlier acquisition of predicates in French. This list is based on detailed diaries kept by the parents and regular correspondence between the parents and the researcher. This list (in table 3.5) shows that out of the ten, only two children produced one verb in English. Those verbs (*cuddle* and *tickle*) are counted as predicates. However, they could also be placed in the common nouns category. If the child produces that word in isolation, it is impossible to decide whether it is a noun or a verb. As the present data is based on parental diaries, and therefore does not contain any context information, it is not possible to state positively to which category the word should be attributed. Giving it 'the benefit of the doubt' and assuming it might be a verb was finally decided. On the other hand, six children produced one or two verbs in French. No child produced a verb in each language. One child had an adjective among his first ten words (placed in the predicates category in the CDIs).

Based on these results, it seems that those children acquired predicates earlier in French than in English. This confirms the tendency that was highlighted earlier. It is also consistent with the literature described above (Boysson-Bardies, 1999 and Bassano *et al.*, 1998).

Table 3.5 First 10 words (diary data)

FLO	JAC	ANN	EMM	REB
Dog	Mummy	Papa	Ball	Bye
Door	Papa	Mummy	Grandpa	Shush
Maman	Banana	Book	Mummy	Doudou
Daddy	**Boire**	Byebye	Byebye	Allo
No	Poisson	Hi	Shush	Bravo
Tiens	Bear	No	Thank you	Maman
Donne	There	Yes	**Cuddle**	**Entendre**
Shoe	Pain	Ball	More	Daddy
Crocodile	Aspirateur	Nez	Allo	Merci
Cat	Byebye	Bear	Papa	Non
ELI	**DUN**	**LIA**	**OLB**	**OLR**
Byebye	Maman	Apple	Allo	Au revoir
Daddy	Allo	Daddy	Encore	Daddy
Tiens	Hi	Hot	Cat	Maman
Maman	**Tomber**	**Attend**	Tracteur	**Tickle**
Merci	Banana	Oh dear	Coucou	Non/no
Encore	Dad	Maman	Salut	Bébé
Calin	Yaourt	Door	Shush	Coucou
Ball	**Boire**	Pardon	Caca	Pain
Hello	Canard	Gateau	Lait	Shoe
Baby	Byebye	Ball	Bras	Papi

The same procedure was conducted for the closed-class category, examining the detailed diaries kept by five families. The number of words in each diary (for each child) is not the same as some parents kept a record until 100 words while others stopped at 50. However, looking at the closed-class items or function words present in those diaries, a clearer picture of which ones are acquired first and whether there is a language difference can be provided. The words are ordered chronologically:

Table 3.6 First closed-class items produced

ANN	FLO	JAC	LIA	ELI
There	**There**	**There**	Plus	Encore
This	Encore	**Again**	**More**	Plus
Down	**Up**	**Down**	Dedans	
Up	**More**	**More**		
More	**Down**			
Me	**My**			
Here	**Me**			
Out	**On**			
My	**That**			
Là				
On				
Off				
That				
Ce				
Ma				
Un				
I				
It				
In				
Of				
At				
For				

Surprisingly, most of the first closed-class items produced by bilingual children in this study were in English. Consequently, it seems rather unfounded to claim that French-speaking children do acquire closed-class items quicker than English-speaking ones, at least from this data. This may be due simply to the fact that French has a larger variety of function words and that bilingual children pick up on what they hear more often. Or if they hear *mon, ma* and *mes* on the one hand and simply *my* in the other language, then they might just go for the easier option of *my*.

3.7. DISCUSSION AND CONCLUSION

Our aims in this chapter were, primarily, to determine possible similarities between the French and English lexicons of a bilingual child and the lexicon for a monolingual child. Secondly, I intended to check for cross-linguistic influences. And finally, I wanted to highlight the special features of the bilingual lexicon.

Our results do not provide any major surprise. They are consistent with what has been found for monolingual children in earlier studies. Looking, first, at the case of the French lexicon of these bilingual children and that of monolingual children, the evidence suggests that both predicates and closed-class items do appear later on in lexical development as suggested by Bassano *et al.* (1998), but that for most children almost all categories are present from the very beginning. Similarly to Kern (2001), this study shows that below the 50 word landmark, children do produce almost as many predicates as closed-class items. Kern (2001) also notices this slight fall in the proportion of closed-class items between 100 and 300 words. She does not provide any explanation for that phenomenon. It appears as though the increase of predicates, which is at that time slightly quicker than that of the closed-class items, might be the cause for that small and temporary drop.

Once again the present results, for the English part of the bilingual lexicon, are consistent with what has been found in previous study with monolingual English-speaking children. The results show that nouns are the most prominent type of words (Bloom, 1973, Benedict, 1979). The different stages of lexical development described by Bates *et al.* (1994) do seem to be verified. What is perhaps more interesting is that their predictions as to when each category starts decreasing or increasing are also valid for the bilingual children in this study in each separate language. The first stage of noun increase is true up until the 100-word-landmark. The second stage, characterised by the slow and linear increase of predicates, does happen between 100 and 400 words as validated in their monolingual study (Bates *et al.* 1994).

As it was mentioned earlier, the literature did not find any major difference between the lexical organisation of French and English monolingual children. The four-stage approach to lexical development has been found not only in English-speaking children but also French-speaking one and is also true for other languages (Italian,

Spanish, Hebrew, Swedish, etc...). The only few noticeable differences lie in the amount of predicates and closed-class items produced. The present results are consistent with other studies showing that lexical development is very similar in both languages. No significant cross-linguistic differences can be found as far social words and nouns are concerned.

Regarding the possible larger number of closed-class items in French than in English, no conclusive evidence was found in that direction. Similarly to Caselli *et al.* (1999), the data presented here does not show that children in either language ever have more than 15% of closed-class items in their lexicon. So the 20% found by Bassano (1998) seems to have been an isolated case. The evidence does not suggest either that the French lexicon contained more closed-class items than the English one for these bilingual children contrary to what had been found by Caselli *et al.* (1999). One reason for that could be the way in which the FCDI is written. For example, the French CDI form actually puts *ma/mon/mes* [*my*] which is the first person possessive article respectively for singular feminine, singular masculine, and plural, all as one item to be ticked by parents. So even if French has more closed-class items, the maximum score a child can get is one even if s/he knows two or three forms for the French possessive. This would automatically lower the scores. Another reason could also be simply the complexity and irregularity of categories such as the personal pronouns as mentioned by Girouard *et al.* (1997). However, like Caselli *et al.* (1999) who found that Italian children acquired closed-class items quicker than English ones, the bilingual children in this study seem to have more closed-class items in French before English ones in their lexicon. At the first stage (below 50 words), bilingual children have 5.7% and 3.3% of closed-class items respectively for the FCDI and BCDI. This has been observed in Hebrew (Maital *et al.*, 2000) and Italian (Caselli *et al.*, 1999) also when compared with English. This might be due to the fact that there are more closed-class items in French and so children are more likely to pick up one or two earlier on. The way closed-class items are represented on the FCDI, compared with the BCDI, does not reflect the larger number of closed-class items in the French language. So, even if earlier on children seem to have more closed-class items, the difference might not be maintained simply because of the design of the forms. Another possible reason would be that early on children acquire prepositions like *down* and *up* to mean a lot of different actions involving movements in this case. Those prepositions do not exist in French and French-speaking children would have to

use a more complex structure to express the same idea. Consequently, English-speaking children are more likely to develop those types of closed-class items early on while closed-class items like possessive forms are acquired later by children. As a result, the forms might not reflect the real picture. The slight decrease between 50 and 100 words has also been found in monolingual French-speaking children (see Kern, 2001). The phenomenon remains unexplained even though it might simply be due to the format of the forms.

Finally, looking at comparisons with monolinguals, there is some indication that predicates might appear earlier in French. Little evidence was found in the French lexicon for a larger proportion of verbs earlier than in English based on the CDIs' data. The difference is not as significant as in Boysson-Bardies' (1999) study, for example. The bilingual children seem to have reached the ceilings of the FCDI by the end of the data collection, but not those of the BCDI which might influence the results. It still remains to be seen whether this is just the case for these bilingual children or whether the ceilings of the FCDI are also reached more rapidly by monolingual children. Unfortunately, the French study is still in progress. Here again, the design of the forms might be the cause for the difference. This would need to be further investigated.

Our results do seem to show that bilingual children do acquire the same categories as those acquired by monolingual children around the same ages and in the same developmental pattern. More importantly, these results suggest that both languages follow a parallel development even when the two lexicons are clearly not the same size. It seems rather unlikely that any bilingual child would then be able to develop a lexicon composed only of one type of lexical items in one language and all lexical categories in the other language. A lexicon where only one type of lexical categories would be represented, especially if those were system morphemes (or closed-class items), seems rather unlikely simply because that lexicon would not be able to survive.

Finally, recent literature has widely reported that bilingual children do produce more or similar numbers of lexical items than monolingual children (Pearson *et al.*, 1993; Allman, 2002; Junker & Stockman, 2002). The present results do appear to go in the same direction. However, when looking at the developmental trends, one can observe that at the earliest stages of language production, bilingual children are clearly at the

same level as monolinguals. It appears that it is only from around 22 months that the difference can be noticed. Pearson *et al.* (1993) had found an overall similar lexical production score for monolinguals and bilinguals. Ours reveal a slightly higher number for bilinguals. However, Pearson *et al.* did remove from the total cognates and words that had similar phonological forms, which probably explains such a difference. These results show that a bilingual child's lexicon should be based on two languages and not just one. The fact that the TCV scores are lower than the monolingual scores can be explained by the fact that the monolingual child's score also contains synonyms or words that relate to the same context. Pearson (1998) had suggested TCV as a means of comparing bilingual and monolingual lexical data. However, when compiling the TCV for bilinguals, the researcher has to group items according to concepts. Consequently, items such as *ici* and *là* can be put together as one concept matching with *here*. The monolingual's score on the other hand is not a concept score and still counts *ici* and *là* as two items. Consequently, the TCV score reduces the bilingual score but not the monolingual one. One solution would be to reduce the monolingual child's score with the same conditions to ensure a more adequate comparison.

This chapter has shown that bilinguals follow a very similar lexical development pattern to monolinguals in terms of quantity as well as quality. Special bilingual features will be considered next.

CHAPTER 4:

SYNONYMY IN THE BILINGUAL LEXICON

4.1. INTRODUCTION

The issue of translation equivalents (TEs) and synonymy is an important one in bilingual language acquisition, as it appears to provide evidence for the one-vs.-two-system debate. Volterra & Taeschner (1978) claimed that since the girls they studied had no translation equivalents, they must have had only one lexical system at the beginning of the study. Pearson, Fernandez & Oller (1995) argued that since bilingual children clearly had doublets from the very beginning of lexical development, three solutions were possible: the one system hypothesis without Principle of Contrast or the two systems hypothesis with or without the Principle of Contrast. Deuchar & Quay (2000) came to the same conclusion: translation equivalents by themselves are not enough evidence to claim two separate systems from the start.

In the chapter, I will study the evidence available from the 13 children followed hoping to provide new evidence to the above mentioned debates and continuing in this study's larger purpose of describing the nature of the bilingual lexicon. Firstly, below, the aims and objectives of this chapter are outlined. Secondly, the main body of literature about synonymy in language (as a system) is reviewed together with studies that have looked at synonymy in language acquisition and especially bilingual children. Thirdly, the main methodology of combining and matching the BCDI with FCDI for this chapter's analyses is outlined. Subsequently, the main results will be described and later discussed within the larger framework of the study's purposes.

4.2. AIMS AND OBJECTIVES

Our aims and objectives in this chapter will be to:

- Analyse the bilingual subjects' lexicon for the existence of translation equivalents in order to obtain further evidence that cross-linguistic equivalents exist in the lexicon of bilingual children from a very early age;
- Study possible disparities between children;
- And establish what types of words have synonyms in both languages.

4.3. LANGUAGE AVOIDS SYNONYMY

The issue of synonymy in semantics can be and has been studied from two different points of views. The issue can be examined in terms of the speaker or in terms of the overall language system. I will first look at theories that have described synonymy from the larger language as a system aspect. The Competition Model (MacWhinney, 1989) is a model in which the learner has to adapt his/her behaviour and constantly make decisions between alternatives that are presented to him/her. It is a mental process in which many candidates compete for categorisation. And it is up to the learner to evaluate the possibilities in terms of the cues that are present in the environment. In terms of lexical acquisition and categorisation, the model claims that language avoids synonymy. In other words, two words cannot mean the exact same thing. Even though we are taught that synonyms could be, to take MacWhinney's example, words like *boy* and *lad*, we eventually realise that such words are slightly different. In this case, there is a register difference. The two words are not completely interchangeable in any given context. In a more formal situation *boy* would be preferred to *lad* which is more of a colloquial term. It is cues from the environment that help the learner distinguish and choose between two words. Let's take a further example with the words *ship* and *boat*. These two words have neighbouring semantic properties. Cues like "transportation mode" and "sea travel" place the two words in the same categories. However if we go deeper in the categorisation, there appears to be a size difference. A ship is said to be able to contain boats but a boat can never contain ships. There is therefore competition between meanings. Two words rarely mean exactly the same thing, even in the adult language.

4.4. SYNONYMY IN LANGUAGE ACQUISITION

Many linguists have come to this same conclusion and outline the same type of principle as MacWhinney's, however, looking at the issue from a different perspective. Pinker (1984) calls it the Unique Entry Principle, whilst Markman (1989), the Mutual Exclusivity Principle and Clark (1993), the Principle of Contrast. All of these look at synonyms from the speaker's point of view, analysing the techniques used by speakers and consequences. All of these principles recognise a

constraint against synonymy. The Principle of Contrast and the Mutual Exclusivity Principle interest us the most, here, since they have been formulated to make a clear analysis of lexical development.

Applying this idea of homonymy to language acquisition raises another question as to why children reject synonyms. According to Clark (1993), children assume that adults and people have intentions. So if they choose a word against another word, it is for a special purpose. By doing so, children assume that one word differs in some way to the other word. Thus, it allows children a greater "economy" of effort (Clark, 1988: 324). They do not need to wonder if a word is different from a previous one acquired but in what way, or how different that word is. New words can be acquired more quickly and require less effort from the child.

4.4.1. The Principle of Contrast

"[W]herever there is a difference in form in a language, there is a difference in meaning."(Clark 1987: 1). The Principle of Contrast states that children reject all types of translation equivalents. "Children [...] assign different meanings to words that differ in form." (1993: 90). This allows them to acquire words much more quickly and easily as they do not need to consider whether the new word is different from the new one. It is taken for granted. They only need to focus on how different it is. According to Clark this principle is more of a constraint than a strategy.

There seems to be some confusion in the literature as to what exactly the Principle of Contrast really is. Clark (1987) first claimed it was a semantic constraint. However, more recently, (Clark 1993, 2003) she has classified the Principle of Contrast as a pragmatic principle.

Gathercole (1987, 1989) proposed an alternative to the Principle of Contrast (based on monolingual acquisition). She explained that Contrast can not be an absolute semantic constraint. She claimed that if the Principle of Contrast was a real semantic constraint as it was designed to be, then children would consistently adhere to it. She instead provided other alternatives that would help understand what Contrast was designed to explain. Since there is great variability from context to context, she suggested it was indicative of pragmatic effects.

So the Principle of Contrast might rather be a Contrast Strategy used by some children.

Clark (1993) also claimed that the Principle of Contrast applied to bilingual language acquisition. She stated that bilingual children start building up a lexicon where cross-language synonyms were rare if not completely non-existent. Clark (1987) claimed that bilingual children attribute a single "label" for a concept until they have around 150 words. However, her claim was revised later to less than 50 words (Clark, 1993). Cross-linguistic contrast will apply until the child has realised that s/he is dealing with two different languages. In other words, Contrast will apply cross-linguistically until pragmatic differentiation is achieved. From that point on, Contrast will only be true inside each language. Even if the child, then, has synonyms across languages or across systems, s/he does not have synonyms within a language or a system. It is due to this distinction between the two systems, that Clark delineates how bilingual children acquire equivalents much earlier than monolingual children.

Clark's theory was reinforced by an earlier study from Volterra and Taeschner (1978). They proposed that young simultaneous bilingual children reject cross-language synonyms in their earliest lexicons. According to their argument, children start acquiring lexical synonyms only when they realise that they are dealing with two different languages, or two different systems. Even if the two sisters they followed had a few words like *acqua* [water] and *Wasser* [water] at 1;11 in their lexicon, they claim that these words were not used in the same contexts and therefore not equivalents or synonyms. The awareness of the two systems would, according to them, start around 2;5, although it is not explicitly mentioned. Therefore, the appearance of translation equivalents would also start around that age.
Another study by Nicoladis (1998) found that her subject did not have any synonyms in his vocabulary until after the language differentiation stage.

However, the evidence is mixed and other studies have indeed found translation equivalents in early bilingual language acquisition. Some of the earliest studies of bilingual acquisition like Ronjat's (1913) mention pairs of translation equivalents in their subjects' lexicon. Ronjat found that his son had at least two pairs of equivalents by age 1;5, even if he did not keep track of all of his vocabulary. Leopold (1939-49)

included an analysis on translation equivalents in his study of his daughter, bilingual English-German. He distinguished between three types of equivalents: (1) "Successive bilingual synonyms" in which the second word to appear would replace the first one in the production; (2) "competing bilingual synonyms", which competed with each other until one displaced the other one; and finally, (3) "permanent bilingual synonyms" which were used side-by-side. The last category especially does seem to provide evidence against the Principle of contrast.

Vihman (1985) used both a diary and recordings to keep track of her son's bilingual development in Estonian and English. Vihman (1985) reported that at approximately 1;10, 59% of Raivo's English words had equivalents in Estonian.

Pearson, Fernandez & Oller (1995) used another data collection method to test the Principle of Contrast. They collected a large amount of data using the Spanish and American versions of the CDI. They mapped out the two checklists onto each other. The children were said to have a pair of translation equivalents when parents had ticked two words credited with the same meaning in the adult language. In spite of the limitations of this method (Pearson, 1998), the results (Pearson *et al.*, 1995) were consistent in showing that on average children had 30% of translation equivalents (or "doublets") at any stage of language development (from 2 to 500 words). But their results also showed a wide range of scores. Individual differences between children were very significant. For the younger children (lexicon < 100 words) the minimum percentage of translation equivalents was 0% and the maximum 100%. For the other children (lexicon >100 words) the minimum was 9% and the maximum 76% (Pearson *et al.*, 1995: 356-357). Despite the evidence, Pearson *et al.* (1995) did not totally reject the Principle of Contrast altogether. They ruled out a strong version of it across languages but left an option opened of the Principle of Contrast being valid with a two-system hypothesis. I will come back to that issue in more details later.

Other, more recent studies tend to reject the Principle of Contrast altogether. Johnson and Lancaster (1998) found 16 pairs of unambiguous, or phonetically distinct, translation equivalents in the Norwegian-English bilingual child they studied up until his second birthday. Deuchar and Quay (2000) showed that M (acquiring Spanish and English) produced equivalents from age 0;11. Their data is not only based on a diary but also on video recordings, which allow more detailed analysis due to the availability of the context, semantic and phonetic information. M produced, for example, the word *bye* at 0;10.17 and the word *tatai* at 0;10.25. By age 1;10, she had

146 words with equivalents (out of a total vocabulary of 330), representing 44%. They emphasise the fact that the Principle of Contrast does not hold for bilingual children so it cannot be a universal principle. Petitto *et al.* (2001) found that the French-English bilinguals and the French-Quebec sign language children they followed, also had between 36% and 51% of translation equivalents in their lexicons.

In summary, we can see that there is mixed evidence from the literature about the Principle of Contrast. More specifically, it seems as though not all children apply Contrast as it is explained by Clark (1987, 1993, 2003), probably due to the fact that the characteristics of the principle themselves have slightly changed over time. The idea of Contrast being a strategy rather than a constraint (Gathercole, 1989) was reintroduced by Van der Linden (2000) for bilingual children. However, her study involved a single case study and therefore could not be checked any further than a suggestion.

4.4.2. The Mutual Exclusivity Principle

Before moving on to the present study, it might be useful to say a few words about the Mutual Exclusivity Principle which is another principle, against synonymy, formulated for language acquisition that has recently been used in bilingual language acquisition. By doing that, I hope to highlight further the confusion and mixed evidence available that exist in the literature. Most of the evidence for the Principle of Contrast is also evidence for the Mutual Exclusivity Principle (Markman, 1989). However, the Mutual Exclusivity assumption is stronger that that of Contrast (Markman, 1994).

This principle claims that children map the meaning of novel words onto novel objects. For example, given the choice between an apple and a whisk for an unfamiliar word like "glorp", they are more likely to choose the whisk which is less familiar to them. However, this principle can be overridden. In various situations children are able to apply a novel word to a familiar object. This is seen for example, when the speaker emphasises a part of an object. There are also cases in which children may interpret a novel word as a quality of the object or as a more specific name. Markman (1994) argued that the Mutual exclusivity assumption (as well as

Contrast) is available to children at the time of the "naming explosion" (1994: 223). This principle has been tested in numerous studies on monolingual, as well as bilingual children (Markman, 1989). Some of the most recent research done with bilingual children will now be described.

Frank & Poulin-Dubois (2002) found no differences in the way monolingual and bilingual children (from 27 to 35 months old) adhered to the Mutual Exclusivity Principle. They also found that the bilingual children performed similarly when the constraint was violated across or within language.

Despite those results, earlier studies (e.g.: Au & Glusman, 1990) showed that older children (from three years upwards) were more likely to honour Mutual Exclusivity only within language. Rhemtulla & Nicoladis (2003) also tested the Mutual Exclusivity Principle on monolingual and bilingual children aged between 27 and 46 months old. They replicated earlier results showing that there is no significant difference between bilingual and monolingual children's application of the principle when it comes to simply introducing an unfamiliar word in English (within language). They also replicated the same experiment but with introducing a speaker of an unfamiliar language (Chinese). The Chinese speaker would ask the children to point to "ping guo" when shown a familiar and a non-familiar object. In that case (across languages), there was a significant difference between the two groups of children. Bilingual children pointed at the familiar object about 50% of the time acknowledging that that could be the word for the object in another language and violating Mutual Exclusivity. It would have been interesting to see whether there was a developmental pattern in the children's responses. They concluded that it is a pragmatic principle.

Most studies testing Mutual Exclusivity test it on the basis of nouns. But other types of words can have synonyms. No literature, to the best of my knowledge, highlights the nature of the synonyms children might have. As well as the Principle of Contrast, in theory, synonymy should apply to all lexical categories.

4.5. METHODOLOGY

As in the previous chapter, all of the 13 children will be used for this part of the analyses. Parental reports will form the basis of the analyses for this chapter also.

In order to carry out the analysis of translation equivalents in developing bilingual children, it was necessary, first, to determine what was meant by translation equivalent. Below will be considered TEs words that refer to the same object or that can be used in more or less the same situations and contexts in the two languages, e.g.: *chat cat* or *me moi*. No assumptions will be made that a child will actually use all of these pairs as translation equivalents. But it is assumed that pairs that are close enough in the adult language can be considered equivalents. Translation equivalents are words that refer to the same concept.

4.5.1. Choosing a method

The first step for the analysis of translation equivalents was to actually get translation equivalents available. It would have been possible to simply rely on naturalistic productions hoping that the children would say the same thing in the two language contexts. However, that would have resulted in only a very limited number of items being available to work with and would have forced a control the home recordings of the family up to a certain extent. The aim of the home recordings was to get as naturalistic productions as possible. So, controlling them or determining topics would have taken out the naturalistic dimension off those. The only way was then to look at the lexicon. Since the families were keeping diaries, a choice between diaries and the CDIs was available. The inconvenient of the diaries was that most families only kept it up to the 50-word milestone. Therefore, only a very limited number of items would have been available to work with, if any at all. Deuchar & Quay (2000) used that method for their analysis of translation equivalents. However, the mother of the child is a trained linguist. None of the parents in this study were trained linguists so the diaries they kept were far from being perfect. It was therefore decided to use both adaptations of the MacArthur CDI for the translation equivalent analysis. One advantage of this was that it provided a wider range of lexical items and over a longer period of time (see chapter 2 for a discussion the issues linked with the use of parental reports). Using the CDIs for an analysis of TEs is not without inconvenience. We will highlight some of them in section 4.5 and 4.7 (see also Pearson, 1998 for a discussion).

The data used is only the data obtained for the 'native' parent (see chapter 5 for the differences between native and non-native). In cases where only one of the forms was completed for the same month that form was excluded from the analysis. As a result 131 sets of data to be analysed were obtained, distributed as follows:

Table 4.1 Data available for translation equivalents' analyses

Child	Ages (in months)														
THO	17	20	21	23	25	26	27	28							
FLO	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
ANT	16	17	18	19	20	21	24	25	28						
EMM	16	18	20	21	23										
DUN	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
PEN	17	18	21	22	26										
ELI	16	17	18	19	20	22	23	24	25						
JAC	16	17	18	19	20	21	22	23	24	26	27	29	30		
ANN	16	17	18	19	20	21	22	23	24	25	26	27	28	30	
OLR	16	17	18	19	20	21	22	23	24	25	27	28	29	30	
LIA	16	20	23	27	30										
REB	16	17	19	21	23	24	26	30							
OLB	16	17	18	21	23	24	25	27	28	29	30				

4.5.2. Matching the CDIs

In order to carry out these analyses it was necessary to determine what translation equivalents existed in both forms of the CDI and which ones were to be used. Immediately excluded were all the sound effects from the TEs analyses because they are very/too close phonetically. Since both forms have been adapted from the original American English MacArthur CDI, it was expected that both forms would be close enough to provide a certain number of synonyms. Moreover, early vocabularies in any language address the same type of situations. This was confirmed by the fact that about 94% of both forms were matched. 50 words were left out of the French form and 30 words out of the English form.

Our first goal was to find one-to-one correspondences between words on the two forms. There are many quite straightforward and relatively uncomplicated: e.g.: *fish/poisson*, *beach/plage*, etc... Several reasons, however, prevented finding matching pairs: the first being cultural differences. Foods and other items reveal the cultural differences between countries. For example, French has unpaired *baguette*, *madeleine* (a kind of soft biscuit similar to muffins), while British English has unpaired *muffin* or *pumpkin*. The second reason for unpaired items was linguistic differences. Some adjectives and prepositions translate especially poorly: e.g.: *fine*, *gentle*, *off*, etc... A third reason for not pairing some of the items was gaps on the forms. Some items not particularly characteristic of the language or culture are left on their own: e.g.: *chèvre* [goat], *cœur* [heart], *wardrobe*, etc...

Some of the pairs are not actually pairs but triads or more. Some of the relationships could not simply be one-to-one but many-to-one or vice versa. The first reason for this is that both forms do not contain the same number of words. In addition to that, a word in one language can have two meanings in another. Take for example the words *house* and *home*, both present in the British English form. The French translation for both is the same: *maison*. French does not distinguish between the actual building and the place where one lives irregardless of the fact that it might be an apartment or a studio. Other cases of these 'multi-pairs' include pronouns and prepositions. French has four different ways (present in the CDI) to mean *it*: *il*, *elle*, *ça*, and *on*.

While pairing items, word classes were respected, meaning that a verb and an adjective were never matched together, for example. Words were preferably left unmatched in those cases.

After making those obvious matches, individual matches were examined for less obvious or ambiguous semantic properties: ball, balloon, balle and ballon. It is probably worth giving a short definition of each of those terms for the purpose of these matching decisions. Ball is "any object in the shape of a sphere, especially one used as a toy by children or in various sports such as tennis and football". Balloon is a "small, very thin rubber bag that you blow air into until it is round in shape, used for decoration at parties or as a children's toy"[1]. Balle is: "petite sphère de matière élastique qui sert dans certains jeux" [a small elastic round object used in games].

[1] From the *Cambridge Advanced Learner's Dictionary* (2003).

Ballon is: "grosse balle gonflée d'air dont on se sert pour jouer, pour pratiquer certains sports ou une vessie gonflée d'un gaz plus léger que l'air, qui sert de jouet aux enfants." [Large ball used to play sports or a large rubber bag inflated with gas used as a toy by children.][2]

In such cases, whether there are three or more words involved the following reasoning was made. If we put the four words under one match, and the child says *balloon* and *balle*, then he will be credited for two TEs. That would not be considered correct. *Balloon* is not the same as *balle* in adult language. Parents tick *ballon* in the CDI without specification of the concept or context so they can assume that *ballon* refers either to *balloon* or *ball* or both. So it was decided that matching *balloon* and *ballon*, and *ball* and *balle* matched separately was the best (but still arguable) solution. In this case the *ballon/ball* match is missed but it proved impossible to get every single match in. Since the aim was to show that children had a large number of translation equivalents, going against that was deemed an appropriate solution to emphasise the chapter's main point. A very similar technique was used by Pearson *et al.* (1995) for matching English and Spanish forms of the CDI.

Many resulting matches are approximate and debatable: they could be argued for or against at length. For the majority of them, the symmetry is not an issue but even for those, the context data available is limited. So, without making a list of exact synonyms, I aimed at demonstrating "that there is a lexical representation from a given 'semantic target region'" (Pearson, 1992).

After having matched most of the items on the CDIs, a list of matches was submitted to 20 bilingual French-English adults and parents. Half of them were the parents of the children taking part in the study, the other half were adults fluent in both languages. Each adult received about 100 matches to examine and comment upon if they felt it was necessary. Their comments were taken into account and several changes were done to earlier matching decisions. Ultimately, maybe arbitrarily, the final decisions were taken by the author, a French-English bilingual herself. The final matches can be found in appendix VIII.

[2] From the *Dictionnaire Universel Francophone Hachette* (1997), translations are the author's own.

4.5.3. 'Conservative estimates'

It could be argued that bilingual parents are even more likely (than parents of monolingual children) to inflate the lexical score of their child due to the impossibility of being able to distinguish between the two languages in some occasions. For example, if a child says [oko] for the concept of chocolate, it might be difficult to know if s/he means *chocolate* (in English) or *chocolat* (in French). So, in addition to the 'normal' count of translation equivalent described above, a second, more 'conservative' count was also calculated. In this total, cognates were excluded. Cognates can be defined, here, as words that are related in terms of sound to one another, in other words homophones. However, that criterion was taken very broadly in order to exclude most possibilities where the child would pronounce a part of the word and the adult would be unable to distinguish between the two languages. This prevented counting those words as translation equivalents when in fact parents might have ticked them both even if the child only knew one of the two. Those cognates included pairs like: *tiger tigre*, *eye yeux*, *touch/toucher*, etc... This conservative count was also based on remarks made by some parents for some of the words. About 11% of all concepts (79) were removed because of similarities. Once again, these exclusions are debatable and the list is far from being exhaustive.

4.6. RESULTS

4.6.1. First appearance of translation equivalents

First of all, it is important to note that since there is great variability in the children's lexicons at the same given age, here again, the translation equivalents are represented below as a function of the total lexicon (French + English) and not as a function of age.

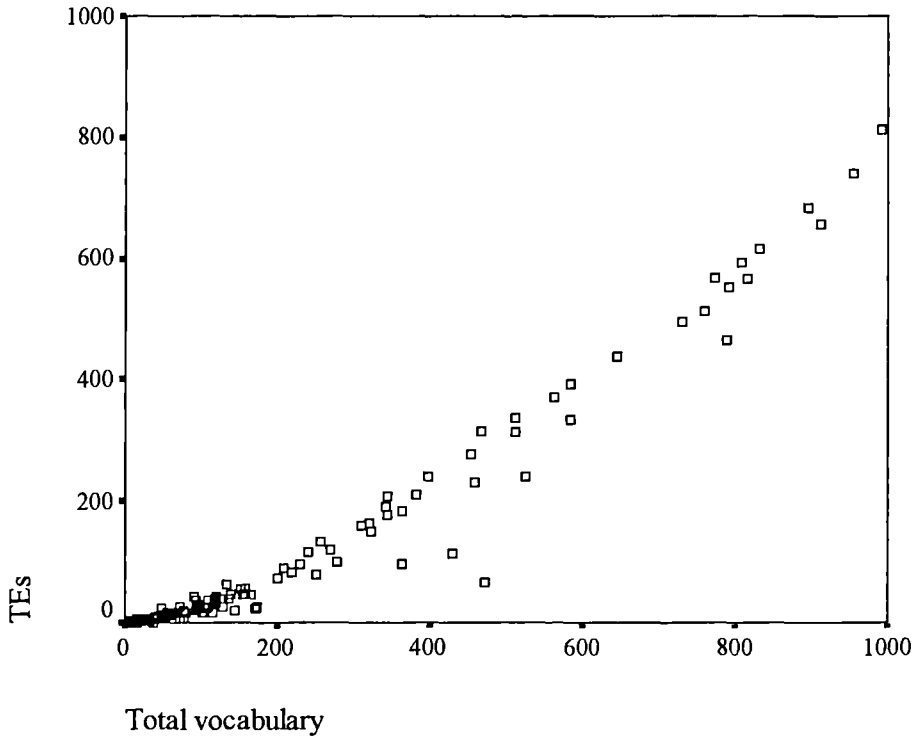


Figure 4.1 Number of translation equivalents as a function of age

As we can see on figure 4.1, bilingual children do have translation equivalents in their bilingual lexicon. The maximum in this study is 813 words with translation equivalents and the minimum is 0.

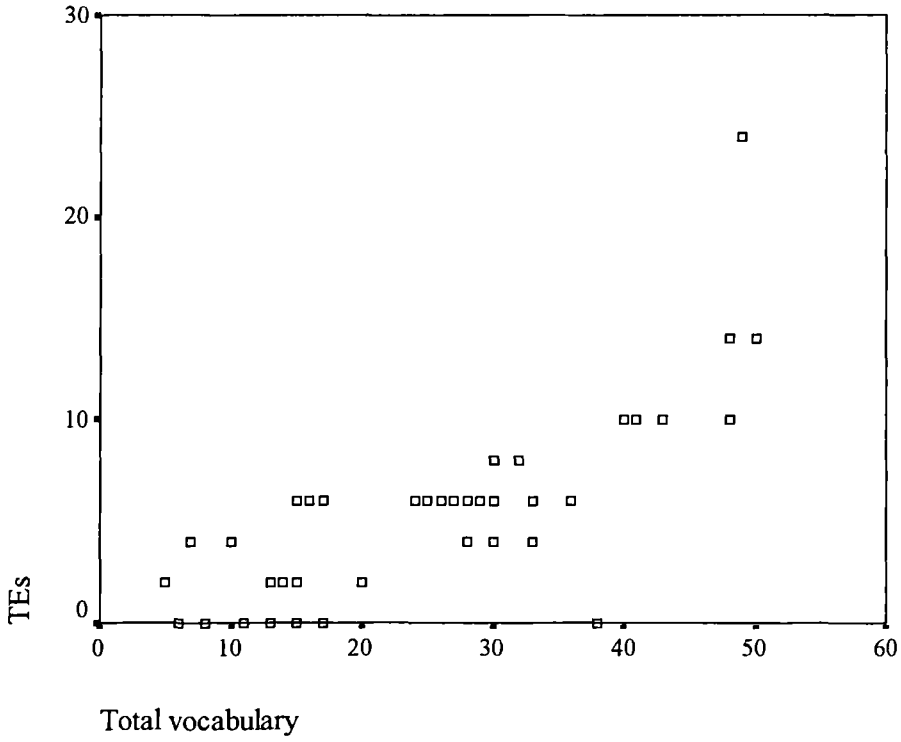


Figure 4.2 Number of translation equivalents for lexicons smaller than 50 words

Figure 4.2 confirms that those translation equivalents do appear from very early on in their lexicon. Even for total lexicons smaller than 50 words, most of the children have TEs. They do not seem to appear only after that 50-word landmark. However, some children do have none at all and that cannot be ignored.

The results examined up to now, are the results obtained from the analyses of all the items that could be matched from the CDIs. The second analysis of the 'conservative' results gives a relatively similar result.

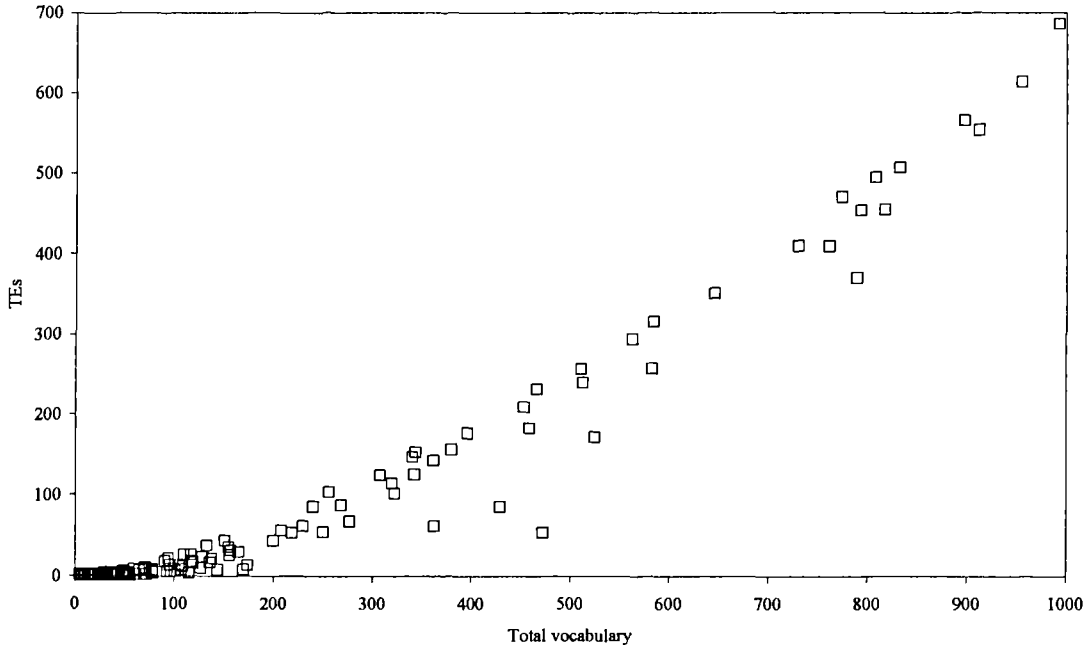


Figure 4.3 Number of translation equivalents: conservative estimates as a function of total vocabulary

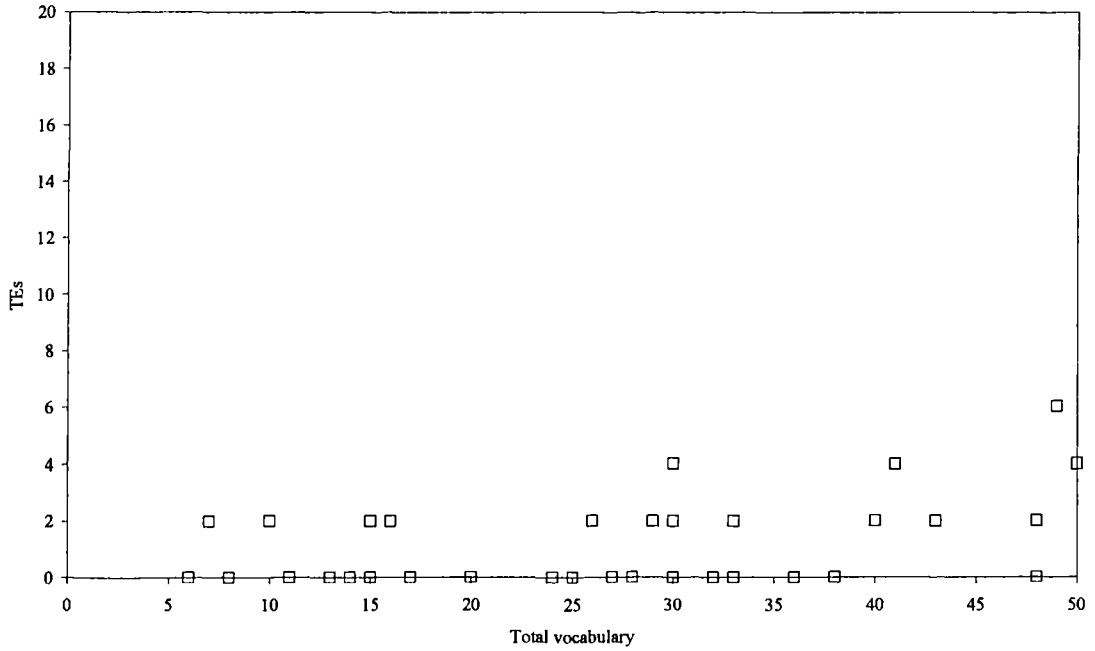


Figure 4.4 Number of translation equivalents for lexicons smaller than 50 words (conservative estimates)

The conservative estimates (figure 4.3 and 4.4) reveal similar charts and tendencies to the previous analysis, so these results seem to be valid and reliable. The only difference is that the maximum total obtained is only around 700 as the maximum to

be achieved was lower. More children get a score of 0 for a total lexicon under 50 words. That appears to be logical, as ambiguous words (or hybrids like [oko] for chocolate/chocolat) are often acquired first. Despite that, most children do seem to acquire translation equivalents before the 50-word milestone. The words counted as translation equivalents in the 'conservative estimates' are unambiguously 'real' and clear TEs as they cannot be words ticked twice by parents because of their similar pronunciation.

Chart 4.5 represents the proportion of translation equivalents present in the total lexicon of the children in function of their age. (The graph represents the total count and not the conservative estimates.) With this graph, we are able to see the great variability amongst the children. The differences are sometimes very significant. For example, at 27 months, the maximum percentage of TEs in the lexicon of a child represents 73.8% while the minimum is 8.3%. Overall, the children do seem to have equivalents for about 35% of their lexicon but the average varies between 11% and 56%. The data shows that there appears to be a development in the mean proportion of TEs. The more lexical items bilingual children have, the bigger the chance of getting translation equivalents. If we look at the conservative estimates, the overall mean is about 22% and we can also notice a developmental pattern.

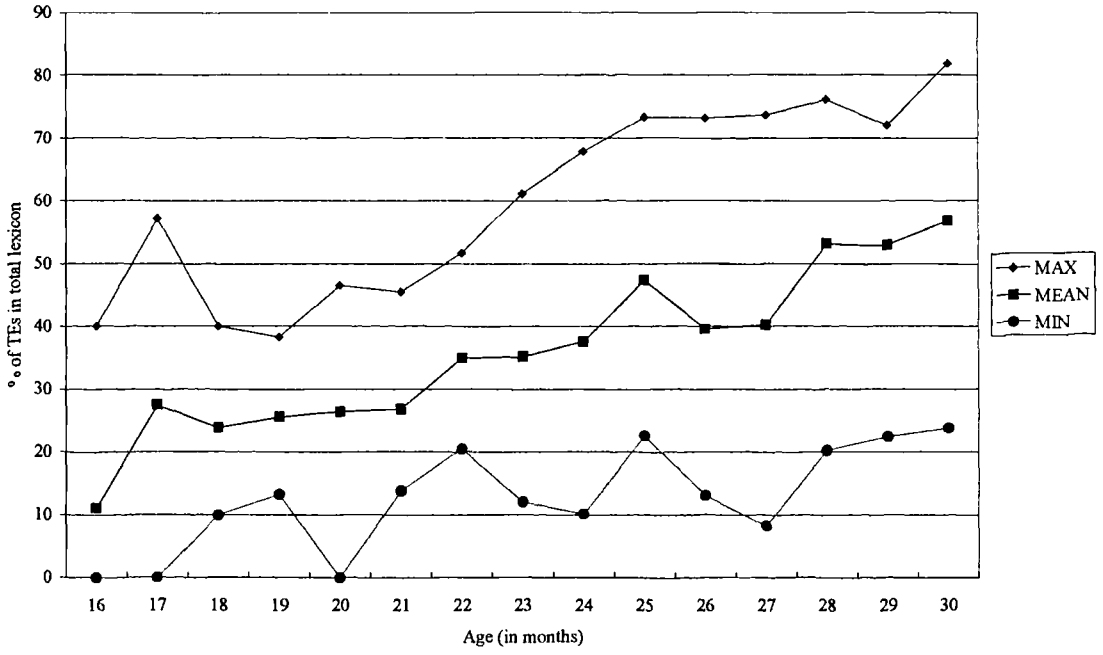


Figure 4.5 Minimum, mean and maximum proportions of TEs in total lexicon (in percentage) as a function of age

The individual differences can be illustrated by the flowing table. This is the data for 11 of the children at 23 months old (total count).

Table 4.2 Number of translation equivalents at 23 months and percentage in total lexicon

Child	Number of TEs	% of TEs in the total lexicon
THO	10	18.5
FLO	205	53.9
EMM	16	22.8
DUN	6	18.1
ELI	206	59.8
JAC	20	21.5
ANN	308	**60**
OLR	30	31.2
LIA	4	**12.1**
REB	24	48.9
OLB	58	37.1

At the same age, there are real important differences between children. However, the data here is presented as a function of age. And we now know that there is great variability in language acquisition, so the number TEs are actually reflecting that variability. What is perhaps more interesting here, is the fact that there seems to be a clear developmental pattern emerging for the acquisition of translation equivalents. Also, we should note that some children do seem to have a lot of translation equivalents when others do not, independently of the age variability.

4.6.2. Types of translation equivalents

Our second aim in this chapter was to determine what types of words those translation equivalents are.

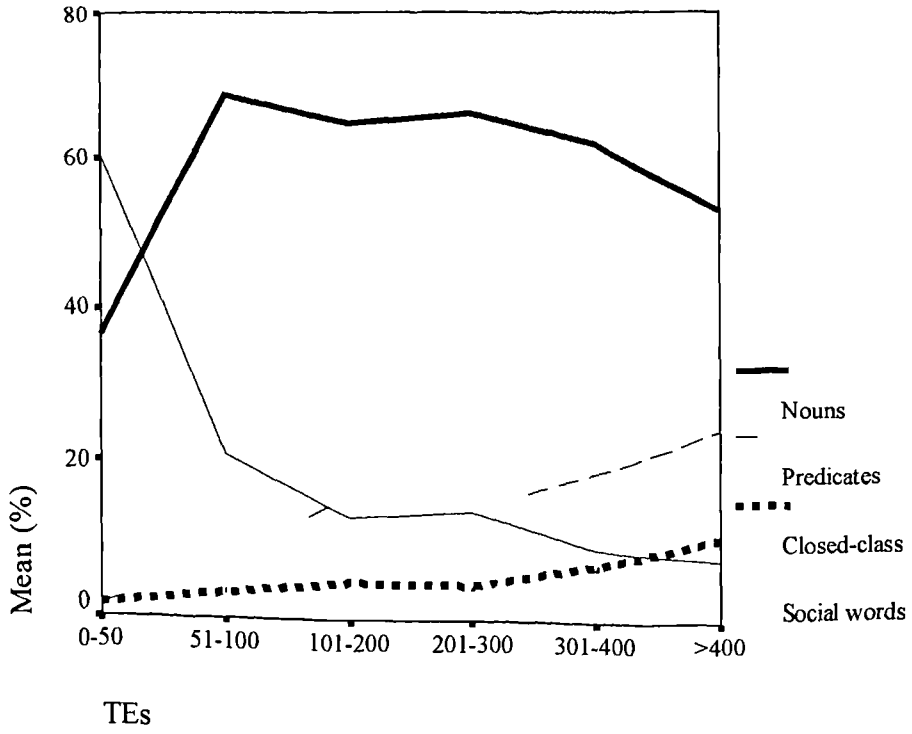


Figure 4.6 Categories of words with a translation equivalent and their evolution as a function of the total number of translation equivalents

Figure 4.6 represents the mean percentage of translation equivalents that are nouns, predicates, closed-class items and social words as a function of the total number of translation equivalents. It reveals that bilingual children obtain equivalents for all types of words in their lexicon. The developmental pattern of the categories, however, follows a very similar pattern to that of their lexicon (see figures 3.1 and 3.2 in chapter 3). This can be seen further on figure 4.7 which shows the mean percentage of TEs as a function of their total lexicon.

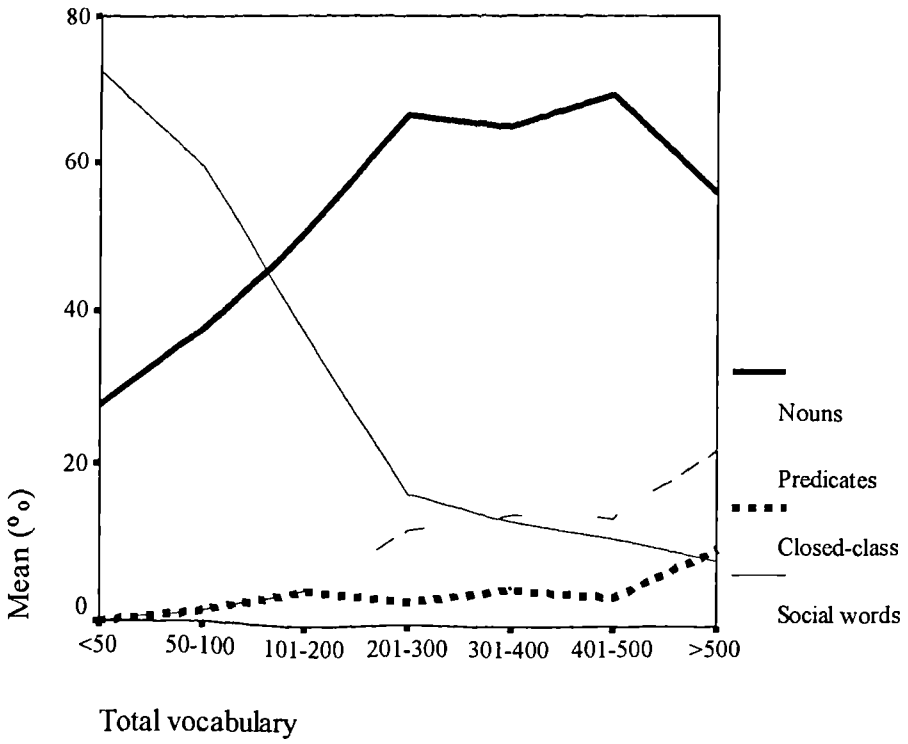


Figure 4.7 Categories of words with a translation equivalent and their evolution as a function of the total lexicon of the child

Those two figures show that children do not have just one type of translation equivalents like nouns, but that as a lexical category grows, the number of translation equivalents in that category increases too. The majority of translation equivalents are nouns, except for a total lexicon smaller than 100 words, where, in that case, social words have more equivalents. Predicates and closed-class items only have synonyms from the 200 words-mark. Two children, EMM and DUN had no equivalent for any of their predicates by 30 months old. Otherwise, all the categories are represented for all the 13 children.

We can use Deuchar & Quay (2000)'s list of M's translation equivalents in order to provide more evidence for the lexical categories developmental pattern. By the end of the study M has 330 words at 22 months old. Nouns represent about 74% of her translation equivalents, 15% are social words, 7% closed-class items and 4% predicates. M fits in the 301-400 words category in the above chart. The main difference is that she has very few equivalents for her predicates. This difference can be due to cross-linguistic influences as M is acquiring Spanish and not French with English. It is also important to note that the authors do match nouns with predicates or predicates with closed-class items in some cases.

4.6.3. Time difference

Finally, in these analyses of bilingual children's translation equivalents, we look at the time difference between the appearance of a word and its equivalent(s) in the other language. On average, the results find that, 1.74 month elapse between the acquisition of a word and its equivalent(s). The maximum time difference is 8.14 months while the minimum is 0. All lexical categories do seem to be acquired with about the same average time difference although predicates do take longer (1.81 month). Nouns have the shortest time difference with 1.75.

Table 4.3 Time difference between the acquisition of a word and its equivalent

	Mean	Minimum	Maximum	Standard deviation
Total	**1.74**	0.00	8.14	1.38
Nouns	**1.75**	0.00	**8.14**	1.36
Predicates	**1.81**	0.00	5.75	1.38
Closed-class	1.77	0.00	5.50	1.34
Social words	1.78	0.06	6.67	1.70

If we look at each child individually, for most children, the mean time difference is around one month. However, one child is significantly different. ANT's mean time difference between the appearance of a word and its equivalent is 4.26 months. The explanation for that important difference might lie in the input amounts she received at that time. Indeed, ANT and her mother spent 3 months in a French-only speaking environment when she was around 21 months. During that time, which is generally when other children's lexicons increased dramatically, she received absolutely no input in English whatsoever. For a couple of months after that, her father reported that

she was using only French, even with him. So, it might have taken about four or more months for her lexicon in English to start growing again. That might be why the difference is so important.

The evidence suggests that on an average of all the children, in 69.8% of all pairs (or triplets) of translation equivalents, it is the French word that was acquired first. This confirms the idea that most of the bilinguals in this study are dominant in French (or as highlighted earlier that the design of the forms influenced that dominance in French). ANN and OLB are the exceptions. For all the words for which they have TEs in the four lexical categories, the English word, on average, always appears first.

4.7. DISCUSSION AND CONCLUSION

Our results are consistent with those reported previously by other studies. Bilingual children do acquire synonyms across languages before the 50-word stage. Deuchar & Quay (2000) also find that the child they studied had equivalents before the 50-word stage. However, the results complement those based on a larger study. On average about 35% of their lexicon consists of translation equivalents. Pearson *et al.* (1995) find a similar result, although they did not find any developmental change over time. The results do show that the proportion of translation equivalents do increase as the child's lexicon increases too (representing from 11 to 56% of the total lexicon). Moreover, children acquire translation equivalents for all types of words they have in their vocabulary at the time. So, as the number of function words increases, children acquire equivalents for those too and not simply for nouns.

Deuchar & Quay (2000) claim that the Principle of Contrast is not a universal constraint as it does not hold for bilingual children. The present results would tend to go in the same direction. However, one can wonder whether the nature of translation equivalents in different languages is the same as synonyms within a language. This question was raised by Li (2002). Part of the issue comes from the fact that there is some confusion as to what the principle of contrast really is: a semantic constraint (Clark, 1993), a semantic strategy (Gathercole, 1989) or a pragmatic principle (Clark, 1993, 2003). We can look at that issue in relation to the decisions speakers (children

or not) have to make in their lexical choices. For a monolingual speaker choosing between the words *boat* and *ship* is primarily a question of semantics. Indeed, although we could consider *boat* and *ship* synonyms they are not really completely interchangeable. A boat is slightly smaller than a ship. There is a slight semantic difference between the two words, even if that difference is minimal and not obvious at first sight or when first acquired by children. There is also some stylistic and register differences. A boat is a small vehicle for travelling on water but can also be used with the same definition (or semantic properties) as a ship in informal contexts. So although pragmatics play a role in the choice of lexical item, we cannot ignore the semantic difference that exists between the two concepts. If we now turn to the choice a bilingual speaker has to make between *bateau* and *boat*, the choice is not the same. A bilingual person (being a child or not) purely makes a choice based on pragmatics. There is no semantic difference, no semantic specificity (or very little in this case) between the two words. Both refer to the concept of a smaller vehicle to travel on water. The French language also has other words (*voilier* [sailing boat], *paquebot* [passenger ship], etc) to mean larger types of boat. So the choice the bilingual speaker has to make is purely as to what is socially appropriate to use - which language to use? Although translation equivalents are a matter of semantic compatibility, how a person uses them is constrained by pragmatics first. One exception may be the case of code-switching where a person would purposely switch to emphasise the fact that there is no semantically compatible word in the other language. However, in that case, the word would not have any equivalent in the lexicon.

It has also been emphasised that bilingual children are better at pragmatic differentiation earlier than monolingual children (Rhemtulla & Nicoladis, 2003). Similarly, bilingual children have been shown to use the mutual exclusivity constraint but not to the extent of monolingual children (Au & Glusman, 1990; Davidson *et al.* 1997). If the Principle of Contrast is a pragmatic principle as Clark later claimed (1993, 2003) and bilingual children perform better at pragmatic tasks, then they learn to override that constraint (or abandon that strategy) earlier. This could also help to explain the important differences in terms of the number of translation equivalents in the lexicon of these bilingual children. Some children reach the pragmatic differentiation stage earlier than others. Nicoladis (1998) suggested that translation equivalents only appear after the pragmatic differentiation stage. There is however, one major issue with that: some children do have translation equivalents from almost

the very beginning of lexical acquisition. Would those children have already achieved pragmatic differentiation by that stage? It is widely acknowledged that pragmatic differentiation is achieved from about 2 years old (Köppe & Meisel, 1995; Nicoladis & Genessee, 1996). By that age, some of the children in the study had already a large number of translation equivalents.

Consequently, the fact that bilingual children do have translation equivalents from an early age does not necessarily mean evidence against the Principle of Contrast. It is evidence against a strong (universal) version of the principle that would claim that all children do apply it. It might simply be that, firstly, bilingual children are better at pragmatic differentiation. Secondly the nature of synonyms across languages is different to the nature of synonyms within a language. So, the Principle of Contrast might be valid for synonyms within a language but not for synonyms across languages. It might simply be that Contrast is a strategy used mostly by monolingual children or that it is a "default option" (Au & Glusman, 1990) honoured only when no other contradicting evidence is available.

Another issue that arises when looking at translation equivalents in bilingual children is the fact that language acquisition is context sensitive and translation equivalents are no exception. Some of the effects that the context can have on language acquisition will be highlighted in the next chapters. The data seems to assume that children do treat these words as synonyms or translation equivalents. However, no evidence points to that. We do not know whether these words are treated as equivalents. Volterra & Taeschner claim that '[t]he children often do not appear to consider such words as exactly corresponding to each other' (1978: 314). They give as an example Lisa, who at stage I uses Italian *là* 'there' 'for things that are not visible at the time of speaking' and German *da* 'there' 'for things that are present and visible to her' (Volterra & Taeschner, 1978: 315). Other studies (e.g. Deuchar & Quay, 2000) have made sure that the words they study are considered as real synonyms as they are using, for example, the same toy in both language contexts. In this study, it is not clear whether the children are actually considering the words as equivalents as no context information is available for most of these words (coming from the CDIs). One conversation with a family revealed the same issue as the one raised by Volterra & Taeschner. One mother revealed that her child (ANN) used the word *poisson* [fish] for the animal swimming in a river, while its synonym *fish* would be used to describe what she has in her plate and eats. Clearly here, the child associates each word with a

different concept. However, a small number of recordings are available where children look at the same picture book with both parents separately. In these cases, the children do seem to be giving two different words for the same animal or toy depending on the context. A possibility would be that children do not realise that these are translation equivalents, but rather that they associate words with separate contexts. For example, *monkey* is used with daddy and *singe* [monkey] with mummy. They associate the word with one parent. A different approach to the issue of translation equivalents and synonymy in bilingual children might need to look at the issue from the child's point of view and not from the adult concepts of what synonyms are.

Our results in this chapter provide more evidence in the one-vs.-two-system debate in bilingual language acquisition. The evidence put forward by Volterra & Taeschner (1978) for their one system claim was partly based on the non-existence of translation equivalents in the early stages of language development. Although they found some translation equivalents, they claimed that those were not real translation equivalents as they were not used in the same context. Their claim (reinforced by Clark, 1993), is that bilingual children would not have any translation equivalents up until they realise they are dealing with two languages. The evidence would concur with that. The evidence seems to show that translation equivalents do provide evidence for the dual system hypothesis. However, Pearson *et al.* (1995) provides an interesting point to this debate. They claim that the fact that bilingual children have translation equivalents in their early lexicon does not automatically provide evidence against the one-system hypothesis. They add that the single system hypothesis with the presence of early doublets make the Principle of Contrast impossible. They argue that the single lexicon hypothesis without a Principle of Contrast is possible and that it has not been proved wrong up to now. In other words, a child would have a single list of words with translation equivalents in it but would not know that they are synonyms as there is no constraint or rule to say otherwise. This would probably only be possible at a very early stage of lexical development. The child would not be aware that there are two languages. So s/he would not be able to tell whether a potential synonym is in one language or the other. Clark (2003) now argues, in order to maintain the validity of the Principle of Contrast, that, in light of recent evidence, children might start with two separate systems as they do have translation equivalents from an early age. In

conclusion, we can say that the presence of translation equivalents in the early bilingual lexicon is not sufficient evidence to decide on the one vs. two system debate. Deuchar & Quay (2000) argue that lexical differentiation cannot simply be defined in terms of translation equivalents. If it is definitely a necessary element, other factors like the use of the right language in the right context are needed to attempt to solve the issue. If we assume that the Principle of Contrast does not hold for bilingual children, or across languages, then we still have two possibilities available to us: one system without Principle of Contrast with a single list and no sufficient linguistic and pragmatic awareness from the child to know whether the words are cross-linguistic synonyms. Secondly, the possibility of two systems from the very beginning without a Principle of Contrast is also possible. In addition, the issue of linking translation equivalents with the one (or two) system hypothesis might be hazardous. Indeed, children do overextend the meaning of words to larger categories (Genesee, 1989), so why could not monolinguals not overextend a word they know in English to a French context? Instead of intra-linguistic overextension, they would also overextend inter-linguistically. This is a possibility which would need to be further examined by experimental methods. Moreover, the issue of translation equivalents might not be directly related to the one-vs-two-system debate. Deuchar & Quay (2000) claimed that input and context might be the reason why Volterra & Taeschner did not find any equivalents in the girls' lexicon. Indeed, the way the recordings were made in Italian with the presence of the German speaker might have inhibited the production of some Italian words. That is a possibility that will be further examined in chapter 7.

This chapter has shown that there appears to be a developmental aspect to the quantity of TEs present in the bilinguals' lexicon and that although most bilingual children produce cross-linguistic synonyms, there is significant variability amongst the children.

CHAPTER 5:

VALIDITY AND RELIABILITY

5.1. INTRODUCTION

The data from the two previous chapters is based mostly on data that has been reported by parents through the CDI checklists. Several methodological issues have been raised by the present study that need to be addressed into more specific details in order to validate and consolidate the present results obtained throughout this thesis. First of all, although, parental reliability has been proved in recent studies (see chapter 2), most of the data supporting the claims made in this study are based on reports from parents. So, it appears to be necessary to examine the extent to which these reports are valid and reliable in this study. Secondly, the lexical data from the study is based on grammatical categories from the CDI checklists. These, again, despite their apparent reliability, need to be validated by, perhaps, another means. Therefore, the aims of this chapter are to validate and consolidate the results obtained. First of all, I will examine parents as assessors comparing the reliability between mothers and fathers (section 5.2) and their reliability over time. Secondly, the validity of the CDIs when compared with other data collection methods will be measured. And finally, I will examine the results obtained from the norming studies of the FCDI and BCDI and outline possible recommendations for future use.

5.2. PARENTS AS ASSESSORS

5.2.1. Inter-observer reliability

As previously explained, parents in this study each completed the form of the CDI that corresponded to the language they spoke to the child, the language being their mother tongue. In some cases, one parent was asked to fill out both forms in the two languages in order to be able to verify their reliability (or inter-rater reliability). In total, 324 forms were completed and returned. For all of the analyses described in the present study and all the results reported previously (unless otherwise stated), these are the forms that were used for reasons to be explained below. Two issues arise when looking at inter-observer reliability. First of all, it is necessary to question the

differences between mothers and fathers, and secondly, to examine the differences between native and non-native parents.

In the present study, both fathers and mothers completed their share of the forms. However, it is generally assumed, rightly or wrongly, that mothers do complete parental reports (De Houwer & Bornstein, 2001). As mentioned previously, De Houwer & Bornstein (2001) did not find any rater bias when it came to comparing mothers and fathers. However, very few studies have looked at this issue. Here, because of the nature of the bilingual situation, in some families, it was not possible to ask the mother to complete both forms as she might not have been proficient or competent enough to do so in the language that was not hers. So, the father had to score the second form. There follows a comparison of fathers' and mothers' scores.

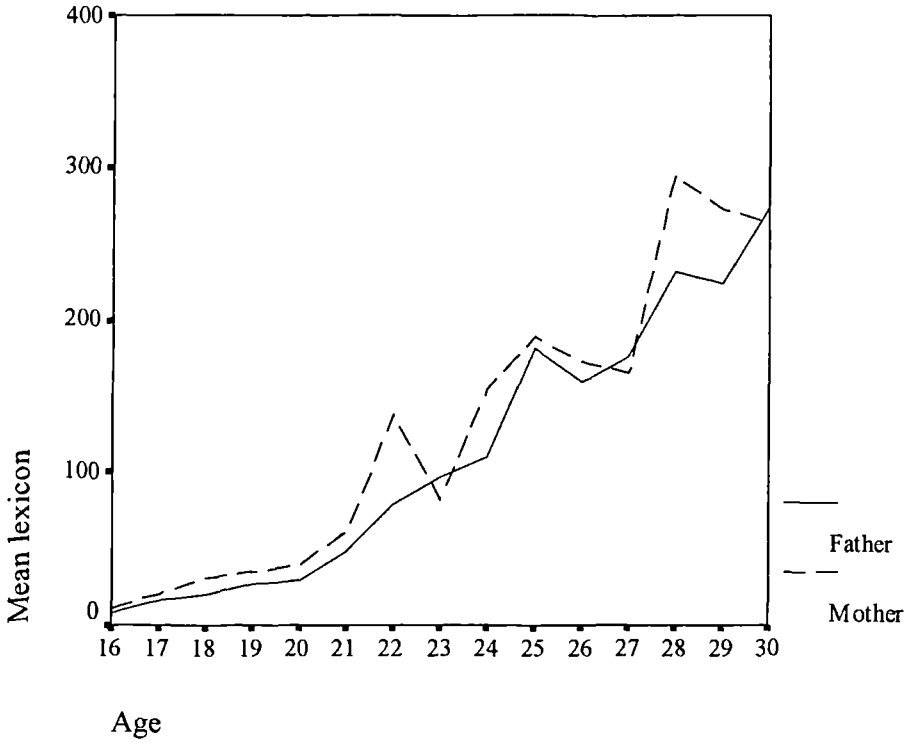


Figure 5.1 Mothers' and fathers' reports of their child's lexicon

Chart 5.1 represents the mean lexicon size (in just one language, not a total of both English and French) by age. As chart 5.1 highlights, a slight difference was found between the mean scores given by the mothers and those from the fathers favouring the first ones. It appears that mothers do tend to give higher scores to their child. However, the difference between the two parents is neither statistically significant if

all scores are taken into consideration (Mann-Whitney U Test, p=.970, two-tailed) nor at each age stage.

In addition to the differences between fathers' and mothers' scores, it was decided that the same parent would complete two forms at the same month in the two different languages. Parents were simply asked whether they felt able to complete the form in their non-native language. 86 of all the checklists representing 10 children for whom 2 forms had been returned at the same age and for the same language will be analysed further for reliability. The analyses of these forms reveal that inter-rater reliability between mothers and fathers are actually extremely high. To determine inter-rater reliability, the number of agreements for each single item in the forms was calculated and divided them by the number of total observations. The results can be seen in table 5.1 below:

Table 5.1 Inter-rater agreement

Mean	93%
Minimum	67%
Maximum	100%
Standard deviation	9%

On average, for about 93% of all single lexical items in the CDIs both parents agree as to whether their child produces it or not. So, even if mothers do tend to give higher scores, the overall agreement is very high. It seems as though the agreement decreases as the child gets older and as s/he acquires more words (table 5.2).

Table 5.2 Inter-rater agreement as a function of the age of the child

Age of the child (in months)	Mean inter-rater agreement
17	98.60%
18	97.51%
19	97.09%
20	97.77%
21	96.84%
24	81.54%
26	81.17%
30	70.35%

Indeed, that impression is confirmed if we look at the correlation between the child's knowledge and the agreement between the parents. This negative correlation is highly significant ($r=-0.96$, $p<0.01$, $N=43$) as illustrated on figure 5.2.

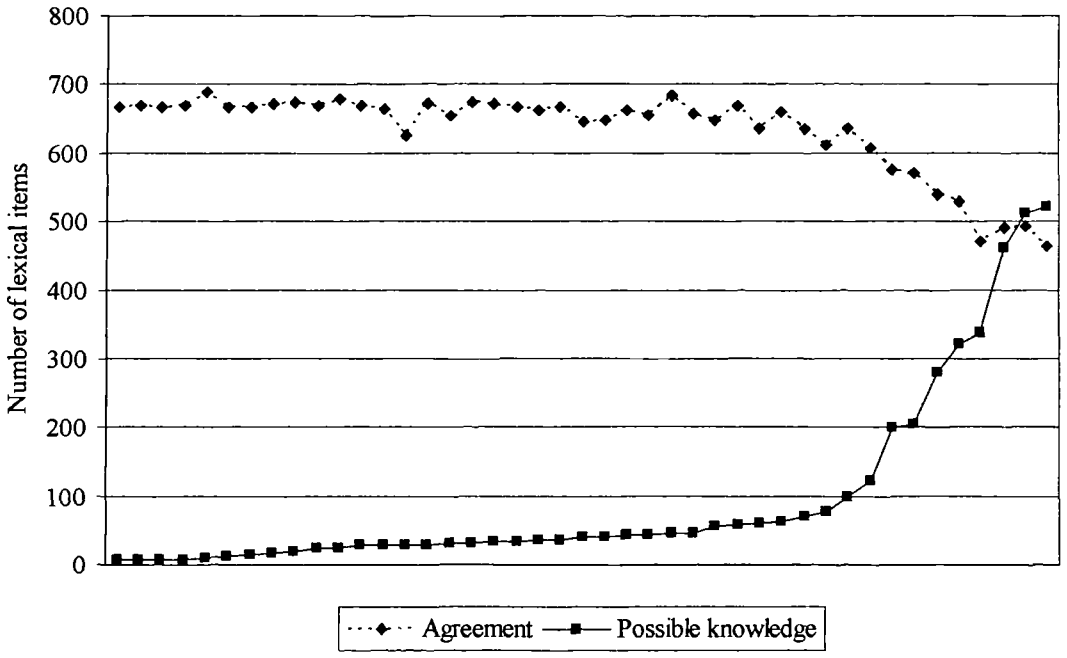


Figure 5.2 Inter-rater agreement as a function of the child's growing abilities

Figure 5.2 represents the number of items on which parents agree with each other and the number of items the child (possibly) knows (i.e. where at least one parent has

ticked it on the form)[1]. This shows that the more lexical items a child knows, the less likely the parents are to agree. That would appear to be logical as parents have to remember more items. These results are consistent with what has been found by other previous studies (e.g. De Houwer & Bornstein, 2001).

However, the present numbers have to be considered with caution. For the children aged 24, 26 and 30 months old, the average scores are calculated only from two different scores. Therefore, a mean score based on five or six children, as in the previous cases, might have provided a very different picture.

Another analysis was carried out to measure inter-observer reliability into more details. These forms obtained at the same age from two different parents also allowed another type of reliability check. It was decided to look at inter-rater reliability based on native or non-native parents. Thereafter the word "native" will mean that the CDI completed was in the native language of the parent. "Non-native" will mean that the CDI was in the other language, not the one usually spoken by the parent to the child.

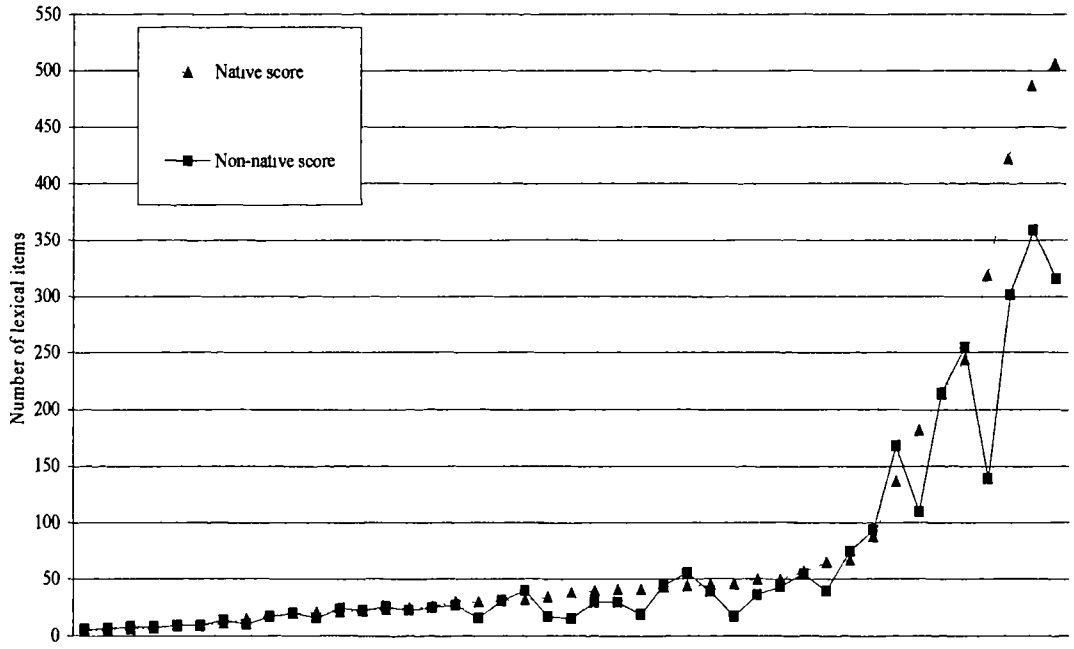


Figure 5.3 Native and non-native scores[2]

[1] The x axis represents each data collection point where two forms were obtained for the same language for the same child (N=43) arranged by increasing number of lexical items known.

[2] Here, again, the x axis represents each data collection point where two forms were obtained for the same language for the same child.

Figure 5.3 represents those 43 instances in which two forms were completed for the same language, the same child at the same age ranked by ascending order of the native score. As shown above, there seems to be a slight difference between native and non-native scores. Parents appear to be more able to estimate their child's abilities in their own language than in the one they do not directly use with them. A statistically significant difference is found between native and non-native scores (Mann-Whitney test, p=.001, two-tailed) overall.

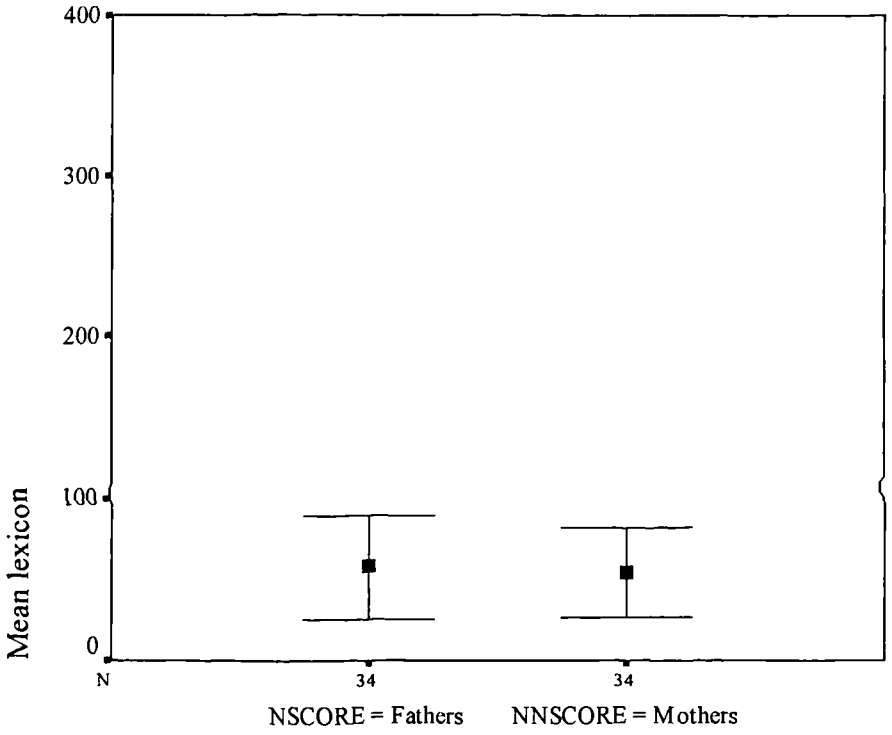


Figure 5.4 Mothers as non-native scorers

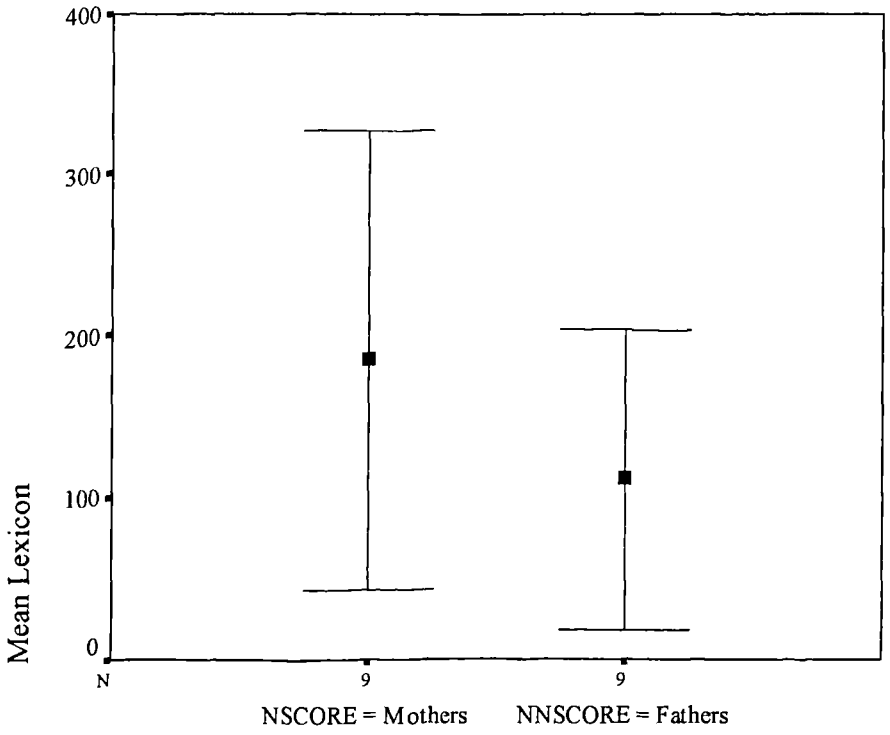


Figure 5.5 Fathers as non-native scorers

Figures 5.4 and 5.5 compare native and non-native scores depending on who the non-native parent is. What is important on these charts is the difference between the two figures. What it shows is that when fathers are the non-native parent, there is a bigger difference between the two parents' scores than when it is the mother who is the non-native speaker. This would show that fathers are more reliable when they are the native speaker of the language they are assessing. This would also indicate that the native parent consistently gives a higher score than the non-native parents, confirming the previous results. Finally, one issue with this result is that fathers as non-native scorers are only representing a very low proportion of the 86 forms (20%). Consequently, the results need to be taken with caution as a couple more forms may have changed the pattern.

To summarise this inter-rater reliability study, the evidence appears to suggest that mothers tend to give higher scores than fathers. Native parents as scorers tend to give higher scores than those who are completing the form in the language they do not speak with the child. And finally, fathers when they are non-native scorers give smaller score. Unlike, De Houwer & Bornstein (2001), a rater bias was found,

although very small. A number of questions arise from this including: Are mothers over-estimating at all times or are they more knowledgeable than fathers? Do they pay more attention? Do mothers know best? Are they more reliable? Or do fathers actually not know? Other types of bilingual family settings would allow to answer some of these questions as well as a larger sample of bilingual families. However, these research questions were not intended to be part of this project and therefore will require a considerable amount of additional research to be answered.

The difference in parental scores, although minimal, poses the question of accuracy when considering that the bilingual child's vocabulary score will be the one given by one parent only, usually the mother. A solution to the problem would be to create a composite score as suggested by De Houwer & Bornstein (2001). That would consist in looking at each individual item and verifying the consistency between raters introducing a third scorer (e.g. nanny or grandparent). However, this is extremely time-consuming. As a conclusion, we can say that taking into the main analyses all of the scores in order to calculate a child's lexicon was the best option. As, if mothers are over-estimating or inflating their child's scores, then the scores of the fathers did counter-balance them. Furthermore, the reality lies probably somewhere in the middle of all of these scores. So, once again, having two raters or observers for the same language and thus three or four forms for each child at the same given age, was the best possible option in the time scale given, considering that the focus of the main study was not reliability of parental forms.

5.2.2. Reliability over time

The issue of reliability of parental reports together with the nature of this study brought forward another methodological issue to be considered here. As the study was a longitudinal one, for reasons explained previously, parents had to complete the CDIs once a month. Parents were only sent the next CDIs when the previous ones had been completed in order to avoid any "copying". This method was adopted in order to avoid parents from just copying the words they had ticked the previous month and add a few extra ones. This method allowed checking for any disparities over time, thus showing parental reliability in general. Looking at the overall curve of the average vocabulary scores for all 13 children does not provide any answer. As some parents

have not completed all of the forms, and as the number of children is relatively small, one form missing one month from one child can make a significant difference in the overall total mean score especially if that child has one of the lowest or highest production scores. Other studies looking at the reliability of the CDI have shown that there is a significant correlation between two tests at different time intervals. Fenson *et al.* (1993) found correlation coefficients between .6 and .9 between two forms completed for the same child at 6 weeks interval. Eriksson (2001) argued for a 70% of words marked on both forms at two months interval. When we consider that children are expected to acquire more words in 6 weeks or two months interval and thus one would expect the scores to increase naturally, the results are said to be showing the high reliability of, respectively, the American and Swedish CDI.

For this analysis, we used each child's longitudinal productive data based on the parental reports of the native parent only. We flagged out each single item, in each language, for which the parent was showing inconsistency. This means that, for example, if the parent marked a word as produced by the child at 18 months, then at 20 months but not at 19 months, we assumed that this was an example of inconsistency. We recognise the limits of that method, as word attrition is a phenomenon widely reported in child lexical development. However, we assumed that if a child had said a word only once, a parent would be less likely to tick it on the CDI thus minimising the evidence for word attrition. Another issue we acknowledge with that method is that parents are not perfect and can involuntarily forget to tick words. The results vary enormously across the families (see table 5.3).

Table 5.3 Parental inconsistencies over time

Mean	12.5%
Maximum	43.8%
Minimum	1.3%
Standard deviation	10.12%

On average, about 12% of the total items in the CDIs (for one child in one language) were reported with inconsistencies. The results vary from 1% to 43%. The reasons behind the discrepancies can be found in the nature of the child's lexicon. Correlation

between the percentage of inconsistent items and the total number of CDIs completed by each parent was significant: the more forms the parents completed, the larger the number of inconsistencies ($r=.461$, $p<0.05$). That appears to be logical as parents are likely to make unintentional mistakes, i.e. forgetting a word. A further reason for the important differences appears to be the average time gap between two forms returned. As mentioned previously, some parents did not return every single CDI. Correlation between the percentage of inconsistent items and the mean time gap between two completed forms is significant ($r=-.445$, $p<0.05$): the longer the gap between the CDIs, the less errors or inconsistencies. That appears to be rather surprising and is not fully understood yet. One suggestion might be that it might not actually be linked to the mean gap between two CDIs, but rather to the parent itself; some parents being more consistent than others, more careful in judging their child's lexical abilities, or more careful in filling out the form. Last but not least, the most significant correlation was found between the percentage of inconsistent items and the final lexical score of the child obtained ($r=.734$, $p<0.01$). This shows that the larger the child's vocabulary, the more likely parents are to make mistakes and forget some items. No evidence of a correlation between mothers or fathers and the number of inconsistencies was found, suggesting that mothers are as reliable as fathers (over time).

Mistakes can happen and word attrition is common in language acquisition, so one other way to look at this issue would be to do a quantitative analysis of the same parent's reports over time, to see if there is any significant variation (increase/decrease). In charts 5.6 and 5.7, we show the overall lexical development of two children, LIA and FLO, chosen as they represent respectively the minimum and maximum percentages of inconsistencies in at least one language.

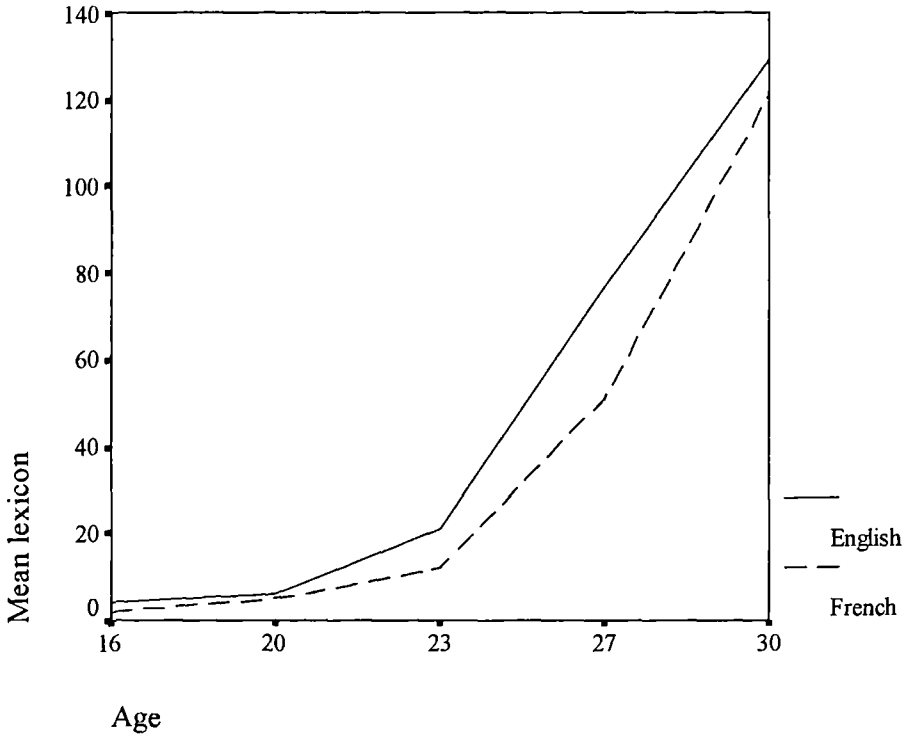


Figure 5.6 LIA's lexical development

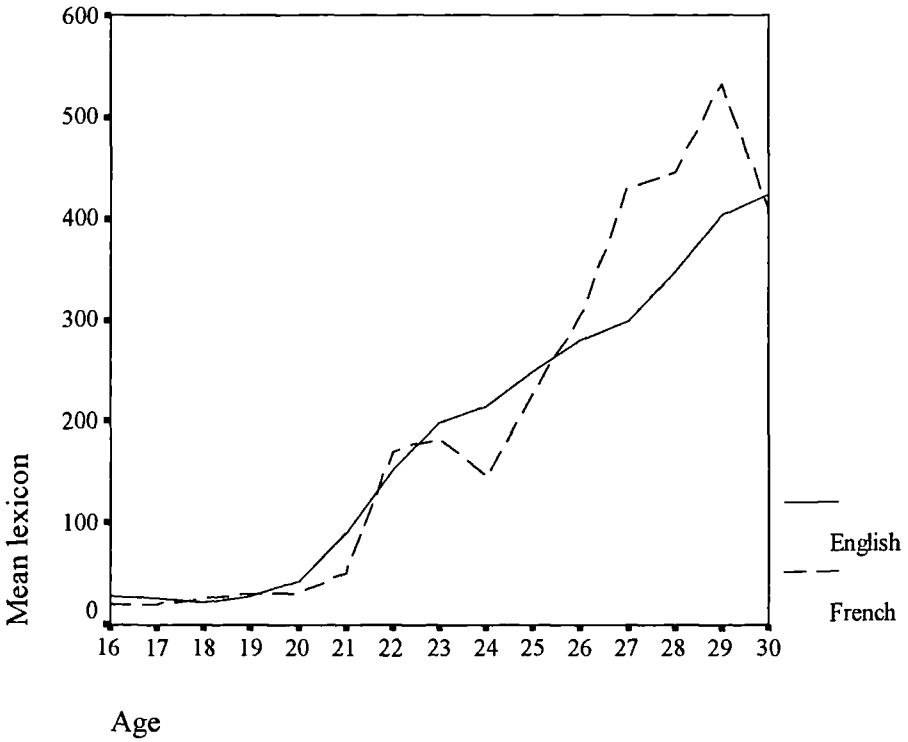


Figure 5.7 FLO's lexical development

As shown on the charts, LIA's lexical development does not show any significant decrease and his parents' reliability appear excellent. He falls into the three criteria mentioned above: the average time gap between two CDIs was longer than average (3.5 months; mean 1.5 months), consequently, the total number of CDIs completed was quite low (5 in each language for native scores; mean 10 in each language), and his final lexicon was lower than average (129 in English and 121 in French at 30 months old; mean 240). That seems to be the best "recipe" for a reliable parental account of a child's lexical development over time. On the opposite side of the reliability scale, FLO's parents (and especially her mother in French) appear to be less reliable as her vocabulary curb shows a certain number of decreases. To their credit, it is worth mentioning that FLO was the girl with the largest lexicon in the study (with ANN) and so her parents would have had to remember more individual items than LIA's for example. Also, they completed absolutely every single CDI that was sent to them (44 in total) so might have spent less time completing each of them every month.

Overall, we can say that parents are generally reliable over time as they provided only a minimal number of inconsistencies (12% of the overall forms), a score which is very similar (even lower) to the one found by Eriksson (2001) for parents who only completed two forms over two months. The families in the study were all volunteers, interested in the topic of the study and from an upper-middle class background. So that might be a reason for that slightly lower score. We can say that completing a CDI or two each month appears to have been a demanding exercise for parents. Those who did not complete all of them appear to have spent more time on each of them and thus might have provided more reliable scores. However, this is only a generalisation. Some families completed many CDIs and were extremely reliable (ANN's) others filled out only a few and showed a large number of inconsistencies.

5.3. THE CDI VS. OTHER LEXICAL ASSESSMENT TOOLS

After having discussed and illustrated the issues arising with the reliability of the CDI forms, we need to justify the choice of that tool. The reasons for choosing the CDI as the main lexical assessment tool have been described earlier in this study.

Nevertheless, we feel it is necessary to verify its concurrent validity as a lexical assessment tool by comparing the results with other methods.

Fenson *et al.* (1993) have verified the concurrent validity of the CDI by comparing its results with other lexical assessment tools and laboratory tests. The correlations they reported range from .33 to .85. Marr (1999) compared the British adaptation of the CDI with another tool, i.e. the Reynell Developmental Language Scales. She found positive correlations between the total number of items marked on each form and the number of "word pair agreement" (i.e. words either marked or not marked on both forms).

In this study, the data available to test validity comes from two other sources: the parental diaries and the audio/video recordings. Consequently, we decided to verify for similarities or differences in the quantity and quality of the data reported. First of all, the diaries kept by parents allowed for a verification of the quantitative side of the question. Unfortunately, not all of the diaries are complete or have lasted for a long enough period to make any meaningful comparisons, with one exception. ANN's diary was kept with very meticulous care up until she was 22 months by both her parents (each kept a separate one in their respective language that they arranged as one in the end of the study). Their data reveals that she reached the 50-word stage (both languages combined) at 1;6.14. If we compare that with the CDI data, adding both French and English, she had 51 words in her lexicon at 17 months. However, the CDI data contains animal sounds that the parents did not report in the diary. The discrepancy between the diary and the CDI can also be explained from the way they compiled their diary. When they were not sure whether a word was in English or in French they entered it as one in the diary as opposed to both ticking it in the CDI therefore counting as two separate items. There were very few of those words though as they usually waited for it to be pronounced in a way that would allow them to distinguish which language it belonged to before writing it down in the diary. Overall, their reports appear to match and be reliable and valid. All of the words reported in the CDIs at 17 months are included in the diary.

Perhaps more important is the quality of what is reported. Again, if we take the diaries (both languages combined), the first 50 words of ANN and ELI are reported below in pie charts 5.8 and 5.9.

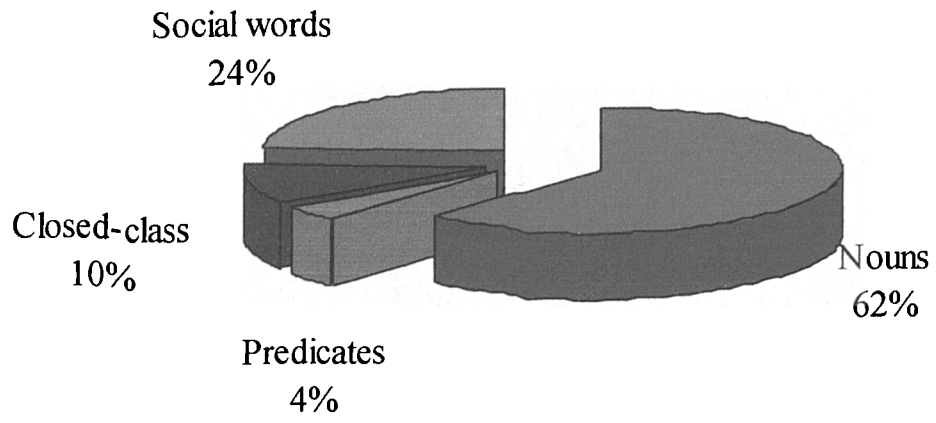


Figure 5.8 Composition of ANN's first 50 words (diary data)

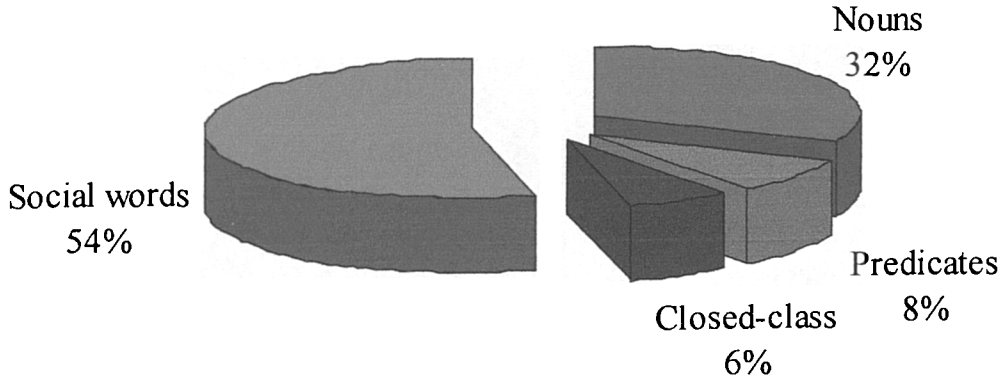


Figure 5.9 Composition of ELI's first 50 words (diary data)

One can expect that with a lexicon smaller than 50 words, the majority of lexical items will be social words. ELI's pie chart matches what was described in chapter 3: social words represent the majority of the items produced for a child with a lexicon smaller than 50 words. Yet, ANN's is different. The majority of the items she appears to know or produce, according to the diary, are nouns. It has already been mentioned that ANN's parents excluded animal sounds from their diary and that is probably what causes the difference within the CDI data.

We also included in these analyses a sample of recordings made at different stages. The analyses reveal two different trends. First of all, some of the recordings as illustrated by FLO's data (in figure 5.10) provide a confirmation of what was found with the CDI.

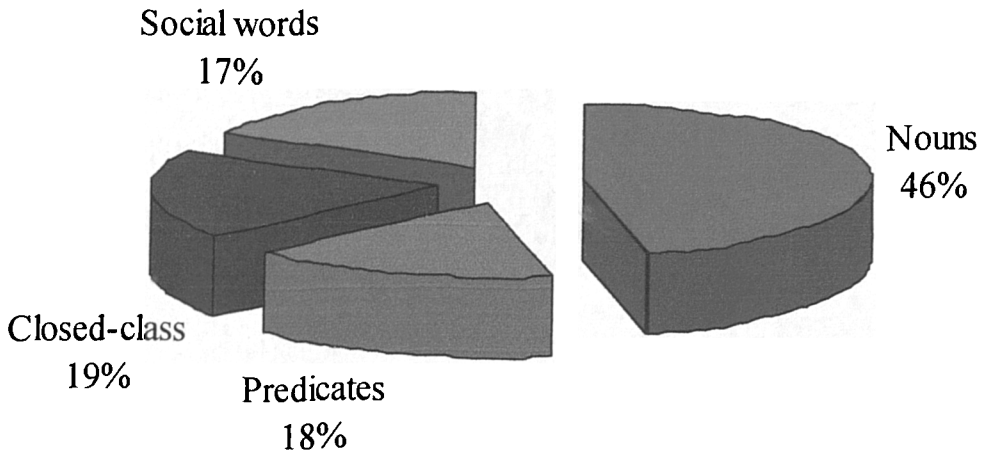


Figure 5.10 Composition of FLO's lexicon at 1;10 (recordings data)

At 1;10 (308 words in her total lexicon, i.e. French and English), FLO produced a majority of nouns (46% of types). The only unusual number here is the high proportion of closed-class items. We will come back to the explanation for that below. The second trend is illustrated by OLB. The pie chart (figure 5.11) reveals an unusually high proportion of closed-class items produced. At 2;3 (458 words in his lexicon), he produced 40% (types) of closed-class items in the recordings, when the pie chart should have looked very similar to FLO's above.

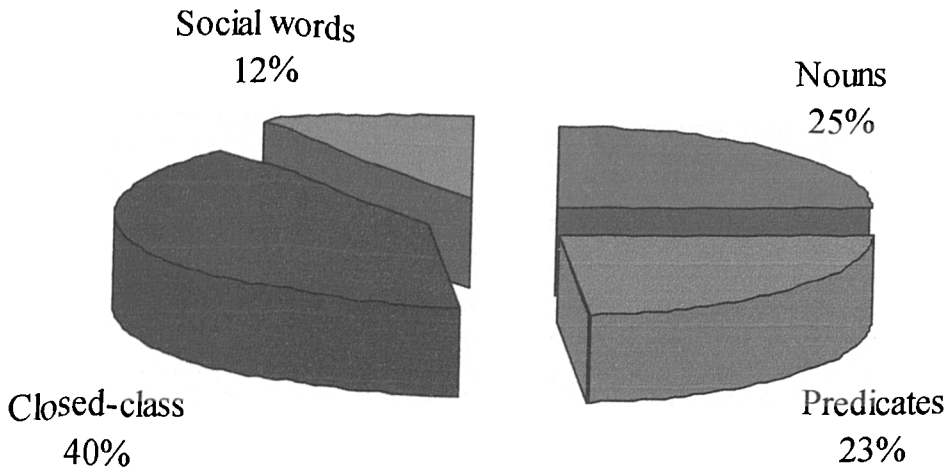


Figure 5.11 Composition of OLB's lexicon at 2;3 (recordings data)

Generally speaking, the recordings data shows the same developmental trend as that described in chapter 3. However, one category which does appear to be produced much more than is reported in the CDI is that of the closed-class items. We have to bear in mind that the data reported in the parental checklists and the diary relates to what the child is able to produce, whereas what is in the recordings is what the child produces more often, the real use that is made of his/her abilities. The CDI represents the children's competence, while the recordings represent *their performances. Other* previous studies have reported that there is a difference between what is known and what is used in child language development.

Overall, we can say that the CDI represents relatively accurately what children know even though there might be a slight difference between what they know and what they actually use.

5.4. About the French and British English adaptations

Finally, one last issue to be considered deals with the particular adaptations of the CDI used in this research. As mentioned previously, the forms of the CDI used are still under development. However, it appears necessary to note a few problems encountered while using them. Parents using these forms in the present study have noted and reported several issues to be raised here. First of all, the BCDI, according to

the families, still contains too many words which are visibly direct translations from American English and would not be used by them (e.g. applesauce, pretzel...). This is also true, perhaps to a larger extent for the FCDI. We have already mentioned the word *étendage*. However, it is felt that perhaps the way some of the categories were built inadequately represent the French language. We have raised the issues of the possessive pronouns being under one single box. Also, the words put under the auxiliaries or helping verb category are not real auxiliaries in French and have been literally translated so that they do not reflect the specificity of the French language in that area. There are, in French, only two real auxiliaries: *être* and *avoir*. Although this is debatable, we feel that further use of these forms should ensure a better organisation and better adaptation of some cultural items.

Finally, after careful examination of the results, we had examined the possibility that the children reached the ceilings of the FCDI too quickly. There is a possibility that the French forms contain a larger number of items that children acquire at an earlier stage. The figure (5.12) comparing the scores for monolingual children on the BCDI and FCDI provides a rather puzzling result.

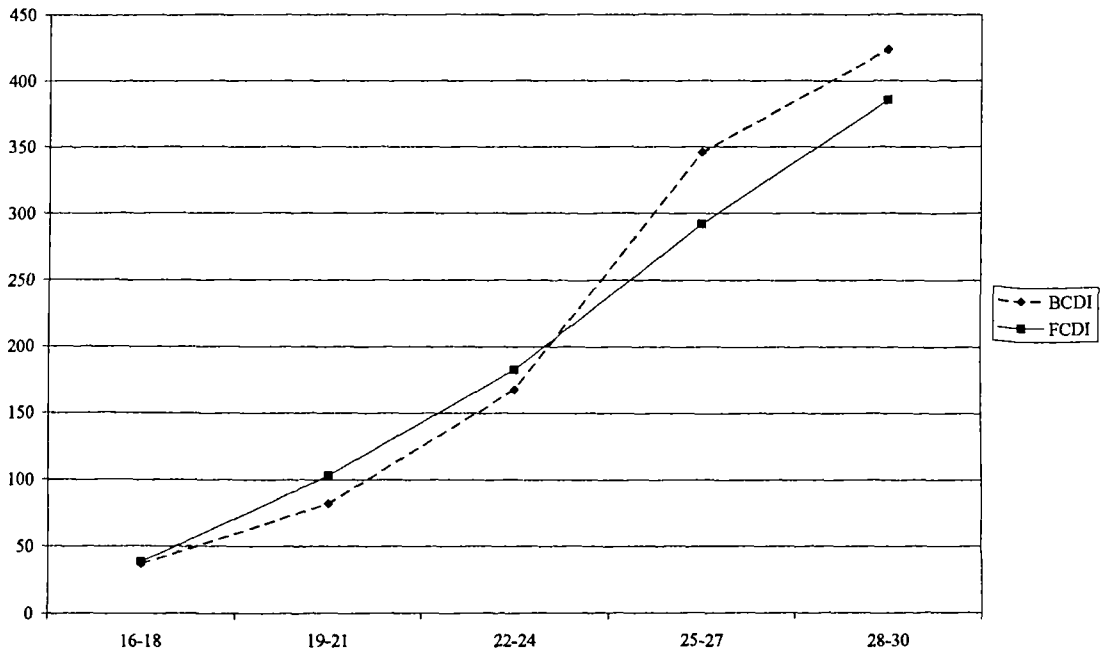


Figure 5.12 Norming (monolingual) data of the two CDIs compared

Indeed, it is difficult to explain why at around 24 or 25 months old, the data for the BCDI increases more rapidly than the FDCI normed data. One reason could be that

the data on which the BCDI is based only contains a relatively small number of children (112 vs. 663). In addition to that, the different ages are not equally represented in the British study. Only 13 children are part of the 22-24 months data while the other stages contain data from at least 20 children. This issue is raised simply to caution further use of the CDIs with bilingual children. Researchers should ensure that the CDIs are normed and provide a similar and accurate picture of the lexical development of the children across languages.

By raising these methodological and reliability issues, we hope to have shown that parents are reliable when it comes to reporting their child's vocabulary even when they have two different languages to consider. In addition, despite their excellent validity, this chapter has provided caution with regards to the use of some of the adaptations of parental reports.

SECTION III:

FROM WORDS TO SENTENCES

BÉLISE

Ton esprit, je l'avoue, est bien matériel.

« J e » n'est qu'un singulier, « avons » est pluriel.

Veux-tu toute ta vie offenser la grammaire?

MARTINE

Qui parle d'offenser grand'mère ni grand-père?

Molière, *Les Femmes Savantes*, Acte II, Scène VI.

CHAPTER 6:

THE EMERGENCE OF SYNTAX

6.1. INTRODUCTION

The previous chapters have dealt with bilingual lexical acquisition. We have established that bilingual children develop in a similar manner as monolingual children. Lexical acquisition cannot be studied without linking it with the emergence of grammar. This is one aspect in which bilinguals differ from monolinguals as they need to coordinate two grammars. The onset of syntax and language mixing will therefore be the focus of this chapter. It will be studied in relation with lexical development. Research on language mixing in recent years has focused on issues such as the grammatical constraints and pragmatic conditions that characterise this unique phenomenon in bilingual discourse. However, one question which has hitherto received little attention is that of the developmental aspects of language mixing in very young bilingual children.

In this chapter, having described the aims, background information will be primarily provided about the 50 word stage and the developmental patterns observed for language mixing in young bilinguals. Some of the analyses used in this chapter will then be further explained as they differ from previous chapters. Subsequently, the main results will be described and later discussed.

6.2. AIMS AND OBJECTIVES OF THIS CHAPTER

The wider aim of the present study in studying lexical development in relation to the early utterances of bilingual children is to examine the validity of Locke's critical mass hypothesis (1997). In order to do that, we need to be able to answer several other questions related to early utterances and language mixing. Consequently, in this chapter, I will seek to determine to what extent language mixing might be dependent on the child's lexical development asking:

- whether there may be a developmental threshold for switching;
- to what extent switches might be a response to lexical gaps; and,
- whether a developmental pattern exists in the quantity and the nature of mixing observed.

6.3. THE RELATIONSHIP BETWEEN THE LEXICON AND FIRST WORDS COMBINATIONS

In order to be able to analyse the language mixing of bilingual children in more details, it is necessary to focus, first, on the transition period between single lexical items and the early multi-word utterances children produce. It has been mentioned earlier (see chapter 1) that Locke (1997) argued that in order for children to activate their GAM (Grammatical Analysis Module), which will allow them to go from producing lexical items in isolation to two-word utterances, they need a critical mass of lexical items to have been acquired and stored during the previous stages. Locke (1997) mentioned (based on Benedict's work, 1979) that the vocabulary burst that happens around the 50-word stage is the trigger for the activation of the GAM. Therefore, a child needs at least 50-words in his/her vocabulary before being able to produce multi-word utterances (other than frozen or formulaic phrases). That theory, as outlined by Locke, only accounted for monolingual children.

Independently, it has been reported in the literature that monolingual children produce their first two-word utterances between 1;5 and 1;10 with great variability among children (e.g. Tamis-Lemonda *et al.*, 1998; Goldfield *et al.*, 1990). Other studies have established the 50-word stage between 1;3 and 1;10 (e.g. Benedict, 1979). Further studies on monolingual children have shown the relationship between grammar and the lexicon: correlations between lexicon type and MLU (e.g. Bates *et al.*, 1988), or the correlation between total lexicon and grammatical complexity (e.g. Bates & Goodman, 1997).

Furthermore, we know that around the 50-word stage, the composition of the lexicon changes dramatically, in monolingual children, at least. Based on studies that used the MacArthur CDIs, we can see that, cross-linguistically, up to the 50-word stage children produce a majority of social words. From about 50-words, the proportion of social words, decreases to the benefit of nouns (e.g. Caselli *et al.*, 1999; see chapter 3 for more details). Therefore, the 50-word stage appears as a crucial landmark in monolingual acquisition for two reasons: the changing composition of the lexicon and the first word combinations.

These results raise a certain number of issues when dealing with bilingual children. Are we to count 50-words in one language or 50-words in both? What is the

composition of the lexicon of bilinguals at each of these stages and will that affect the emergence of multi-word utterances? Can a child activate the GAM with 40 content morphemes in one language and 10 system morphemes (or closed-class items) in the other language? There are relatively few studies that have looked at the critical mass issue in bilingual children. Petitto *et al.* (2001) have shown that bilingual children do compare with monolingual children when it comes to the age they reach the 50-word (or signs) milestone. The six bilingual children involved in the study produced their first two-word utterances between 1;5 and 1;8. Holowka *et al.* (2002) revealed similar results. Their bilingual subjects reached the 50-word landmark at around 1;8 when adding both languages together, so within monolingual norms.

It is expected that similar results will be found for the bilingual children involved in this study looking into more detail at the relationship between the lexicon and the early multi-word utterances.

It is important to be able to determine from what point on bilingual children develop their grammar, or when the onset of syntax is, as that will provide significant information to establish the beginning of language mixing and its nature.

6.4. LANGUAGE MIXING IN YOUNG BILINGUALS

If we are able to compare monolinguals to bilinguals when it comes to the emergence of syntax, there is one aspect in which they clearly differ: language mixing. Mixing is important as it is part of a bilingual's grammatical abilities. The early emergence of grammar in bilingual children contains this added aspect that does not exist in monolingual language acquisition. Language mixing in very young bilingual children has rarely been studied from a developmental perspective. Previous research has largely focused on grammatical constraints at different given stages rather than focusing on the developmental process.

6.4.1. Developmental perspectives on language mixing

The developmental pattern of language mixing in very young bilingual children is a topic that needs further investigation in bilingual language acquisition. Most studies on this topic start when the child is around the age of three. Only very few deal with

the earlier stages of language development. One study by Meisel (1994b) provides a good starting point to the development of language mixing. Meisel (1994b) predicted a "U-shaped" developmental pattern in the frequency of mixes during bilingual language acquisition. Based on analyses of German-French bilinguals, what he called the "grammatical deficiency hypothesis", will predict the following three stages. First, the child starts off with a very high mixing frequency due to the lack of lexical and grammatical knowledge and competence. As the child acquires sufficient knowledge, mixes become less frequent. The decline in the amount of mixing happens around the age of two to three. Finally, the frequency of mixes increases again as the child starts to produce "adult like code-switching". Meisel (1994b) also noted that the second half of the U-shape is only a possibility and might not happen for every individual as some bilingual adults do switch frequently and others do not. Other studies' data seem to confirm this U-shaped development, at least partly (e.g. Volterra & Taeschner, 1978). Redlinger & Park (1980) reported a decrease of mixing rate with "advancing linguistic development" (1980: 351) by German-English and French-English bilinguals. Vihman (1985) also found a drop in her son's language mixing around age two that she explained as "the dawning of metalinguistic awareness" (1985: 317). She also suggests a later rise of code-switching (around age 5) as a "step forward in metalinguistic [...] sophistication" (1985: 317). Köppe (1996) argued for a decline of mixing around the age of two. Deuchar & Quay (2000) also reported that drop in the amount of mixing around age 2 but did not trace the second half of the U in the bilingual child they followed.

6.4.2. Mixing and the developing lexicon

In addition to the frequency of early mixing, it is important to study the type/quality of mixing children produce from a developmental perspective. Most studies have analysed the grammatical properties of language mixing. Very few have looked at the issue from the perspective of the relationship between the child's mixing and the child's lexicon. Studies which have analysed data in relation to a larger grammatical framework have often deplored the lack of lexical data available to them to confirm or disconfirm hypotheses and claims (e.g. Paradis, Nicoladis & Genesee, 2000).

One of the first issues to address when considering the lexicon and language mixing as a relationship is the lexical gap hypothesis. In an early study, Lindholm & Padilla (1978) argued that the mixing they observed from five bilingual children was due to lexical gaps: they "lack[ed] the lexical entry in the appropriate language" (1978: 334). However, no quantitative data is given about the frequency of these mixes and the composition of the children's lexicon at the time. Quay (1995) also put forward a similar argument. She hypothesised that children might mix words from the other language simply because they do not have the appropriate vocabulary and therefore need to fill a gap. Nicoladis & Secco (1998) did find that about 90% of a single child's early mixes could be accounted for by the lack of translation equivalent in his vocabulary. Several other researchers have tried to verify that hypothesis but found opposite results. Vihman (1985) argued that by 1;11, the majority (61 %, types only) of her child's mixes could not be explained by the lack of equivalent in the 'right' language. Köppe (1996) argued that mixing could not be simply due to a lexical need, during the first stage of language mixing, as the child knew both words in either language in most cases. Lanza (1997) also investigated language mixing in relation to the lexical gap issue also. However, she acknowledged the fact that she could not find evidence in the data available (recordings and diary) of whether Siri had equivalents for many of her mixes (Lanza, 1997: 233) since no detailed and precise account of lexical data was kept. Van der Linden (2000) claimed that only a very small minority of the Dutch words switched by Anouk (at the age of 2) in a French matrix language situation could be explained by the lexical gap hypothesis. Most of the studies mentioned above did not have at their disposition a very detailed account of the child's lexicon (perhaps with the exception of Vihman's). However, the relationship between lexical gap and mixing is especially important in the earlier stages of language development. With clear and detailed lexical data as well as examples of language mixing, we will be able to test that hypothesis.

A further issue regarding the quality of mixes produced by bilingual children that will be dealt with in the present study involves the exact nature of those early mixes. There is a consensus in most studies that single lexical items and especially nouns are the most frequently inserted items in early mixes by young bilinguals (e.g. Lindholm & Padilla, 1978). Amongst others, Redlinger & Park (1980) found that, on average, 40% of mixes were single nouns between two and three years of age. This has also been

observed for bilingual adults (see Poplack, 1980). These studies also agreed that phrasal mixes were extremely rare.

Vihman (1985) did appear to provide slightly different results from those mentioned above. In her study of her son, Raivo, she found a large proportion of what she called function words being mixed. She reported that 42% of types recorded in an Estonian context consisted of an English function word with an Estonian word. The function words, as she defined them, included determiners, auxiliaries, quantifiers, etc. She concluded that these were most often mixed. Lanza (1997) made a similar claim for Siri's data. She found that Siri mixed mostly Norwegian function words with English content words but not vice-versa. She interpreted that as a sign of language dominance or "directionality of mixing". Deuchar (1999) found similar results with M who did mix a large number of function words (defined as closed-class items). She did not argue for a directionality of mixing nor language dominance, however, since M did mix both Spanish and English function words. She interpreted the data as indicating the possibility that children do treat content and function words differently: function words being less language-specific than content ones. Jisa (2000) reported that the vast majority of mixed elements in Tiffany's English at 2;5 were "fonctors" (Jisa, 2000: 1373). By "fonctors", she means unbound grammatical morphemes such as pronouns, prepositions, articles, connectors, etc. English was the weaker language of the child at the time, however, she was being exposed to English only for the past two months in unusual circumstances by the time she reached 2;5. French was her dominant language.

Earlier studies had reported that mixing and/or switching mostly consisted in inserting elements from the stronger language in the weaker one (Poplack, 1980). Poplack also claimed that balanced bilinguals with a greater competence in both languages who learnt their second language early (2 to 6 years old) do code-switch to a larger extent than other bilinguals. The topic of language dominance and directionality of mixing had also been raised by Meisel (1994b). He predicted that mixing was related to language proficiency. Meisel claimed that balanced bilinguals "would tend to use code-switching more frequently" (1994b: 417). Similar claims were made by Lanza (1997) who argued that the bilingual children in her study did mix elements from their dominant language in the non-dominant one. The present data will allow for a detailed analysis of what type of words these bilinguals did mix most.

6.5.METHODOLOGY

6.5.1. Subject selection

The data for this chapter are mostly taken from two of the 13 bilingual children: ANN and OLB (with the exception of section 6.6.1). Those two children were selected as a larger number of data and recordings were available. In addition, those children appeared to mix languages in a larger quantity than some of the other children (in the recordings at least). A profile of both of them was given in chapter 2. The two children share a number of similarities. They live in England and were both cared for, at home, by their French-speaking parent (ANN's father and OLB's mother) during the study. They both have an older sibling. Both of the children had a balanced lexicon that developed in a parallel manner in both languages.

6.5.2. Data collection and analyses

In this chapter, data used are based on the audio recordings of the children in two different language contexts: with their French-speaking parent and with their English-speaking parent. Some recordings were excluded when the child's sibling was on the tape as it was rather difficult to distinguish between them at times. All others were transcribed with the CLAN program in the CHAT format (as described in chapter 2). Lexical items were coded for language.

All utterances in this chapter come from either the recordings, or diary notes taken by the parents (in the case of ANN). The diary examples of language mixing were included if they contained a date and a specific description of the context. However, these were not included in the quantitative analyses. All utterances that contained mixed elements were selected. We will distinguish later on between multi-word utterances with mixed elements and one-word turns with single lexical items in the 'wrong' language. Following Meisel (1994a), capital letters will be used for all switched or mixed items for parents as well as for children.

The mixes were classified in two ways. First, the words mixed were classified according to their lexical nature using the same word categories (i.e. closed-class items, social words, etc) as the ones used in the previous chapters in order to be able

to make comparisons and establish relationships between the lexicon and language mixing. Secondly, the mixes were classified according to their type. To this end, four categories were created. Firstly lexical mixes (see 1a) are mixes where the child uses one or several lexical items or content words in the wrong language (i.e. verb, noun, adjective, etc). The second type of mixing is syntactic mixes (as in 1b). The child mixes the subject and the verb of the sentence or a verb and object. This type of language mixing can affect the structure of the sentence (i.e. word order). Thirdly, mixes were found which were set phrases or formulaic switches, where the child uses an expression or a string of words learnt as such, without, in most cases, being able to separate or use the words individually. As in example (1c), ANN uses the expression *un peu chaud* [a little bit hot] as a single unit in several utterances on different separate occasions. Although she knows the word *hot*, she does not appear to be using the others separately or independently at that time. Finally, the fourth category is composed of mixes made up of one content word and a modifier. This modifier category includes adjectives as well as closed-class items (prepositions, quantifiers, etc.). There are very few of those mixes (see 1d.) Those four categories are not exhaustive. Other types of language mixing might exist. However, these are the only types of mixing found in this data at that point of the child's bilingual development. The vast literature on the subject has classified mixes in various other different ways. Some utterances can be composed of several types of mixes. Lexical mixes (type 1) usually involve one single lexical item (not always), whereas the other types involve a couple or more items.

(1) a. OLB, 2;6.13, French
Sur le HAT une fleur.
'On the HAT a flower.'

b. ANN, 1;11.7, French
ME GOT bavette.
'ME GOT bib.'

c. ANN, 1;10.7, English
Not UN PEU CHAUD.
'Not A BIT HOT.'

d. ANN, 1;10.27, English
Look DANS LA POCHE.
'Look IN THE POCKET.'

In the examples above, and all subsequent ones throughout the thesis, the name of the child is given, together with the age s/he was at the time of the production as well as the ML. It is sometimes difficult to determine the matrix language (ML) or base language (Myers-Scotton, 1993) of a two-word utterance containing mixed elements as they are produced by young children. To do so, the language the child was 'supposed' to be speaking with the parent s/he was interacting was used. That method is common in the bilingual language acquisition research (e.g. Paradis *et al.*, 2000). So the ML will refer to the language which the parent is speaking to the child and in which the child is expected to reply.

In order to test the lexical gap hypothesis, the mapping of the two CDIs (as described in the previous chapter) was used as the basis of the analysis. The equivalent word was looked up, first in the same recording or previous ones. If the equivalent word did not appear in the recording, it was looked up in the nearest previous CDI filled out by the parents of the child. It was not always possible to find a definitive answer as not all items are in the CDIs, so some remained with a question mark. Therefore, some mixes remained with a question mark as far as the lexical gap hypothesis is concerned.

6.6. RESULTS

6.6.1. First words combinations and the 50-word stage

Before turning to the developmental aspects of language mixing, it is interesting to describe the first word combination produced by bilingual children. (For this section, all 13 children of the study are taken into account). In order to examine whether Locke's theory (1997) of the 50-word threshold for word combination is applicable to bilingual children, we need to come back, first of all, to lexical data obtained (in chapter 3). The 50-word stage is seen as a crucial one. However, no details are given about bilingual children's 50-word stage. Should we consider each language

separately? Or both languages together? As the lexicon evolves, the proportion of each category is different at different stages. If we claim that from the 50-word stage, a child can start putting two words together and construct utterances, then, it is vital to explain and determine what is meant by the 50-word stage in the case of a bilingual child.

Table 6.1 Proportion of each word category in the 50 words of bilingual children

50 words in ...	Common nouns	Predicates	Closed-class items	Social words
Total lexicon	32.2%	4.8%	4.9%	**58%**
Each language (mean)	**50.9%**	9.3%	6.6%	33.2%
French	**46.5%**	10.6%	9.3%	33.6%
English	**54.7%**	8.2%	4.2%	32.9%

As shown on table 6.1, the percentage of each category is very different depending on which 50 words are taken into account. Taking into account the total lexicon, bilingual children have at their disposal a very different variety of words and lexical composition than if one looks at each language separately. It is therefore necessary to examine when bilingual children do reach the 50-word stage. However, it appears unlikely that bilingual children will be able to produce multi-word utterances when their total lexicon has reached the 50-word stage as they do not have the adequate lexical resources to do so. Table 6.2 represents the mean age at which the bilingual children from this study reached the 50-word milestone, in each language separately and in both combined. The reader should note that these are averages and that the variability between children is very high. For example, FLO reached the 50 word

milestone (in both language combined) at 16 months, while DUN reached it at 25 months. The table shows that, on average, children reached the 50-word stage around 21 months in each language separately. And it is also at that time that the first word combinations were reported by parents (through the CDIs). It appears that for bilingual children to start combining words, they need 50 words in each language. It is important to note that no indication was given to parents as to what counted as two-word utterances (words could have been all in one language or mixed).

Table 6.2 Age at which the 50 word stage is reached

	50 words			First word combinations	
	Total	English	French	English	French
Mean age (in months)	17.5	21	20.5	21.5 (STDV=2.5)	21.5 (STDV=3.7)

At that crucial point in time where bilingual children start to combine words, their lexicon is then mostly composed of nouns and not social words any more (see chapter 3).

6.6.2. A developmental threshold for mixing?

If bilingual children do start combining words when they reach the 50-word milestone in each language, then one can wonder what triggers language mixing. Is that 50-word milestone activating the GAM as well as language mixing? Do children need a minimum level of competence in each language to able to mix? Would children produce mixed utterances if they had reached the 50-word milestone in one of the languages only?

Let us turn to the two children to be studied in more detail in this chapter. ANN's first word combination appeared shortly before she reached 1;5.0. During that first stage of word combination, she mainly used one type of construction. She used the pronoun *this* with usually a noun attached to it, meaning *I want this* or *this is...* as in (2a).

(2) a. ANN 1;4.30, English
This bubble.

b. ANN 1;4.29, French
THIS à boire.
'THIS to drink.'

c. ANN 1;4.30, English
This MAIN.
'This HAND.'

This type of construction was not exclusively used in an English context. She also used this structure with her father (as in 2b) in a French ML context and with her mother but with a French content word (noun), producing a mixed utterance (2c). One could argue that this item was indeed behaving as a filler (see Veneziano & Sinclair, 2000). This lexical item was pronounced differently almost each time it was used but most often was pronounced [ɪσ˥]. However, as most of the instances were reported by parents, it is difficult to tell whether there was any kind of consistency or logic behind its pronunciation. An interesting hypothesis would be that this filler was pronounced differently according to the language context it was used in. However, that would require data that was not available during this study.

ANN began constructing different types of two-word utterances only a few months later (see examples 3a and 3b), around 19/20 months old (when her parents reported her beginning to combine words on the CDI form). By that time, both languages had reached the 50-word stage on their own. English reached the 50-word landmark earlier than French and her MLU in French (see table 6.3), although slightly lower than in English at the very beginning, caught up very quickly.

(3) a. ANN 1;7.15, English
More butter.

b. ANN 1;7.15, French
Bébé dodo.
'Baby sleep.'

Table 6.3 ANN's MLU

	French		English	
Age	MLU	Lexicon	MLU	Lexicon
1;7	1.128	36	1.211	65
1;8	1.111	48	*	107
1;9	*	90	*	178
1;10	*	147	2.266	215
1;11	2.270	237	2.045	276

Similarly to ANN, ELI's parents reported her combining words from 20 months old. And although, they gave examples of single language utterances at that time, a couple of mixed two-word utterances had been noticed by the researcher a couple of months earlier.

Let us now consider the second child, OLB. His MLU data (as well as parental reports) reveals that he started combining words around 23 months old. If we compare this data to the one reported by his parents, we can see that at that age, OLB's lexicon was composed of more than 50 words in each language.

Table 6.4 OLB's MLU

	French		English	
Age	MLU	Lexicon	MLU	Lexicon
1;11	1.04	80	1.24	76
2;0	1.20	111	1.29	119
2;1	1.40	166	*	175

As soon as he was able to combine words together (after the 50-word stage in each separate language) OLB started building utterances with both languages together:

(4) a. OLB 2;0.7, French
CUT encore.

* Data not available at that age.

'CUT more.'

b. OLB 2;0.16, French
Non LOOK.
'No LOOK.'

It is important to say at this point that no intra-sentential mixing in all of OLB's recordings with his father (in English) before the age of 30 months was found. There were very few examples of wrong language choice and those were isolated words (single nouns). The possible reasons for that are outlined in section 6.6.4.

6.6.3. The lexical gap hypothesis

In addition to considering the emergence of two-word utterances and language mixing, a further aim for this chapter was to test the possibility that lexical gaps were responsible for language mixing. Many studies do mention it as a reason for mixing without actually having at their disposal a systematic account of the lexical data. For this analysis, lexical mixes mostly were taken into account for reasons to be explained. Lexical mixes are often nouns (5a) or predicates (5b), or even a single closed-class item (5c). Some syntactic mixes (6) and formulaic ones were not counted as it was harder to tell whether the child actually had the equivalent structure available rather than looking at each individual item. Also, mixes where it was not clear to what the child was referring or how s/he could have used the appropriate structure in the other language, were not counted in this analysis.

(5) a. ANN 1;10.16, English
I got a POCHE.
'I got a POCKET.'

b. ANN 1;10.20, English
MANGE a tissue.
'EAT a tissue.'

c. OLB 2;6.16, French
Pourquoi aller THERE?

'Why go THERE?'

(6) ANN 2;1.13, English

Careful Y A ARETES DEDANS.

'Careful THERE ARE BONES IN.'

Table 6.5 Proportion of mixes explained by the lexical gap hypothesis

	OLB	ANN
Yes	22%*	18.6%
No	**61%**	**61.4%**
Don't know	12.2%	15.8%
Borrowing	4.8%	4.2%

*The percentages given are per token.

Table 6.5 shows the results for that lexical gap analysis for ANN and OLB. The borrowing category is composed of words like *pancake* that have no translation equivalents in the other language for socio-linguistic and/or cultural reasons. Also included in that category are words like *bobo*. This word was used in ANN's family by both parents and is usually used in French to refer to the visible result of a child hurting himself, for example a small scratch on a leg. There is no equivalent word in English and ANN's mother adopted that word into her vocabulary.

As can be seen on table 6.5, only a minority of elements mixed by both children can be attributed to the lexical gap hypothesis. Even if we add the category for which we cannot say for sure whether the child had a translation equivalent (TE) at that time or not to the possibility that he did not have a TE, then the lexical gap explanation still only accounts for a smaller percentage of mixes. It is also worth noting that the numbers are very similar in both situations, for both children. This does not mean that all children will behave in this way even though it does provide strong evidence for a typical pattern for this age range at least.

6.6.4. A developmental perspective

Below, I will be describing and highlighting the phenomenon of language mixing from a developmental perspective for the two children mentioned above. First I will

focus on the quantity of language mixing and its development over time. Secondly, I shall examine the nature of mixing and its changes.

6.6.4.1.Quantity of language mixing

In order to study the amount of mixing and a possible developmental pattern, it was necessary to take into account any lexical item that the child uses with one parent that is in the 'wrong' language. Here, taken into account are not only multi-word utterances but also one word utterances.

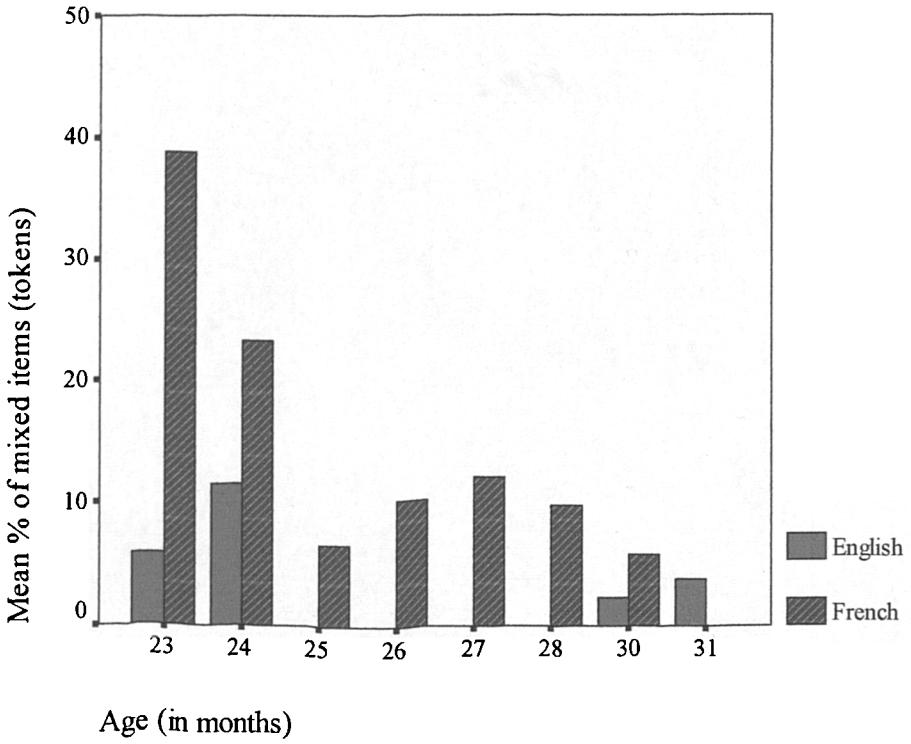


Figure 6.1 OLB's mixing over time

As can be seen on figure 6.1, OLB's language mixing had a tendency to decrease with time. Furthermore, he mixed languages mostly in French language contexts. One explanation for this is the fact that his mother's code-switching was quite significant compared to his father's and to other parents in the study. Although she was the minority-language speaker, his mother would code-switch up to 60% of the tokens in one recording. In addition to that she would not always provide OLB with the right word in the right language when OLB would mix. She would just go on with the

conversation and OLB would clearly understand that speaking with his mother was very much a bilingual context. The relation between OLB's mixing and his mother's is analysed further in the next chapter. Still, OLB's language mixing, in one- or multi-word utterances seemed to be decreasing as he got older and acquired more linguistic skills. Taking only multi-word utterances into account, the proportions of mixed utterances is much smaller as many of these mixes are single items like the word *yes*. They, then, do not represent more than 10% (tokens) at any point in time. However, the same developmental pattern is apparent.

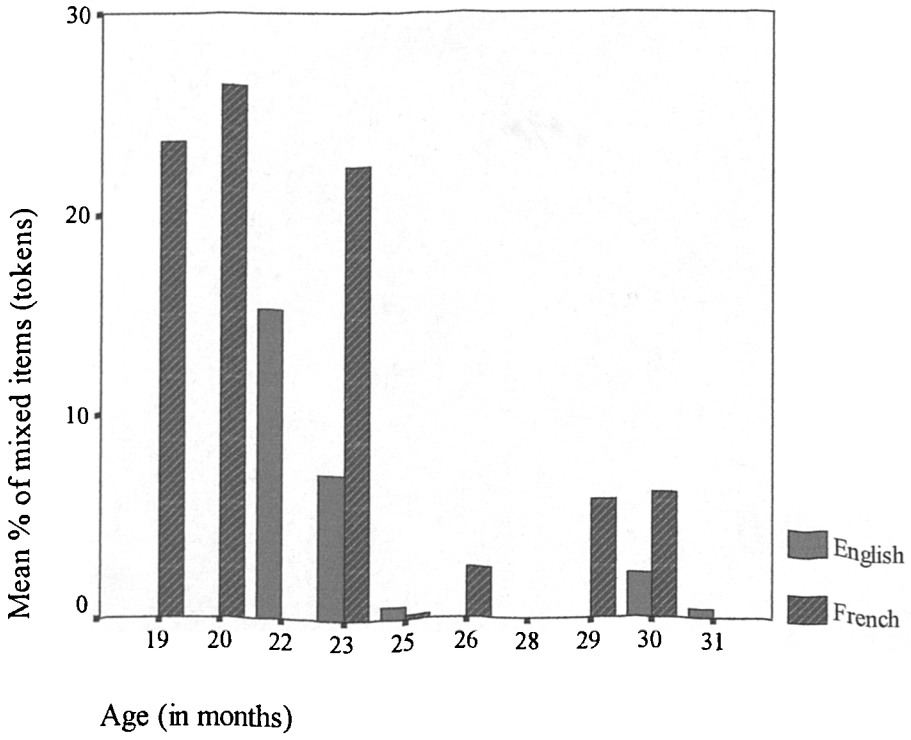


Figure 6.2 ANN's mixing over time

ANN's language mixing represents the same clear decreasing trend as OLB's. Further data would have been extremely useful as the amount of tokens mixed does appear to be increasing again from 29 months old. Again, ANN mixed mostly in French language contexts with her father.

Both children clearly exhibit the same pattern of mixing: a very large amount of mixing or wrong language choice at earlier stages followed by a significant decrease (between 25 and 30 months) as their linguistic abilities increase. Furthermore, ANN's data would appear to suggest that an increase in the amount of language mixing might

appear again towards the end of the period of study. However, it is my opinion that ANNs data would have automatically shown an increase towards 32 months as she started attending a day-care centre five days a week in a mainly English-speaking environment. Although the recordings stopped at that time because of time constraints, her parents reported an increase in her using English items with her French-speaking father. On the other hand, the data available at 29, 30 and 31 months do show an increase without her linguistic situation having changed at that time yet. More data is needed in order to investigate that pattern.

6.6.4.2.Quality of mixing

Next, we examine the type of language mixing produced by the two children. For this analysis only intra-sentential mixing (i.e. from multi-word utterances) is taken into account. As can be seen on table 6.6, OLB's mixing was composed of 70% of common nouns, as in example (7), during the period of study. And a staggering 91% of his mixes were lexical. OLB's mixing patterns did not reveal any developmental change. The pattern was the same throughout the period of study. OLB would mostly mix English single items into French ML structures, as seen previously. As mentioned earlier, mixes in ML English were rare and did not appear until later on.

Table 6.6 Categories of items mixed for OLB

	Total
Common Nouns	**70.2%**
Predicates	10.6%
Closed-class items	19.1%
Social words	0%

(7) OLB, 2;2.19, French
Mange CHOCOLATE.
'Eat CHOCOLATE.'

Even though, overall, lexical mixes represent ANN's largest category of types of switches (81%) and always at least 60% at any given point in time, the type of words mixed is very different from OLB and merits closer examination. On average, ANN

mixed about the same amount of closed-class items as common nouns (respectively 38% and 39%) during the period of study (between 16 and 30 months old). However, changes can be observed throughout the period (figure 6.3).

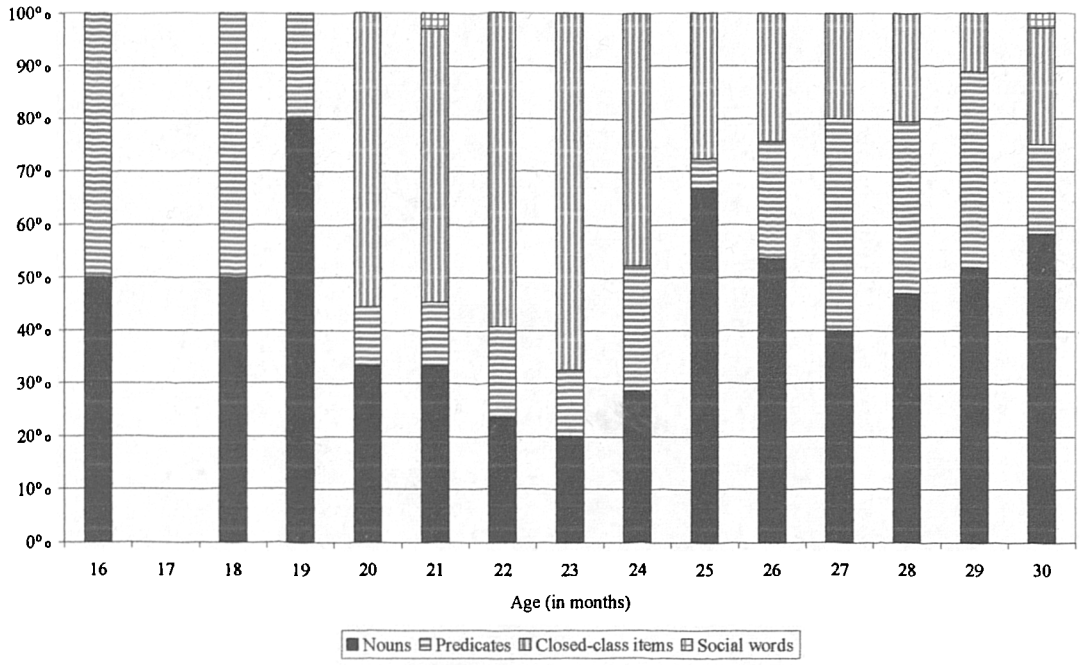


Figure 6.3 Categories of items mixed for ANN over time

When language mixing was at its peak (between 22 and 23 months old), ANN's language mixing was more diversified (as shown in a typical syntactic switch in (8a)) and closed-class items, or function words as they are sometimes called (Deuchar, 1999), represented the majority of items switched (8b). It is important to note that in cases where the child mixed a verb with a preposition (9), these were classified as predicates. The reason for this is that the child did not separate the verb from its preposition and would probably not have mixed the preposition on its own in such cases. However, one could argue that a closed-class item is also mixed. The few instances similar to (9) occurred around 30 months old.

(8) a. ANN, 1;9.22, French
ME WEAR ça.
'ME WEAR this.

b. ANN, 1;9.23, English
Yes AVEC chair.

'Yes WITH chair.'

(9) ANN, 2;6.25, French

Moi veux toi LIFT moi UP.

'I want you to LIFT me UP.'

Adopting Lanza's classification of mixing (1997), we could say that instances where closed-class items are mixed do represent (more or less) what she called grammatical mixing. During the period where ANN mixed most, grammatical mixes were the most common ones. That can be further illustrated by figure 6.4.

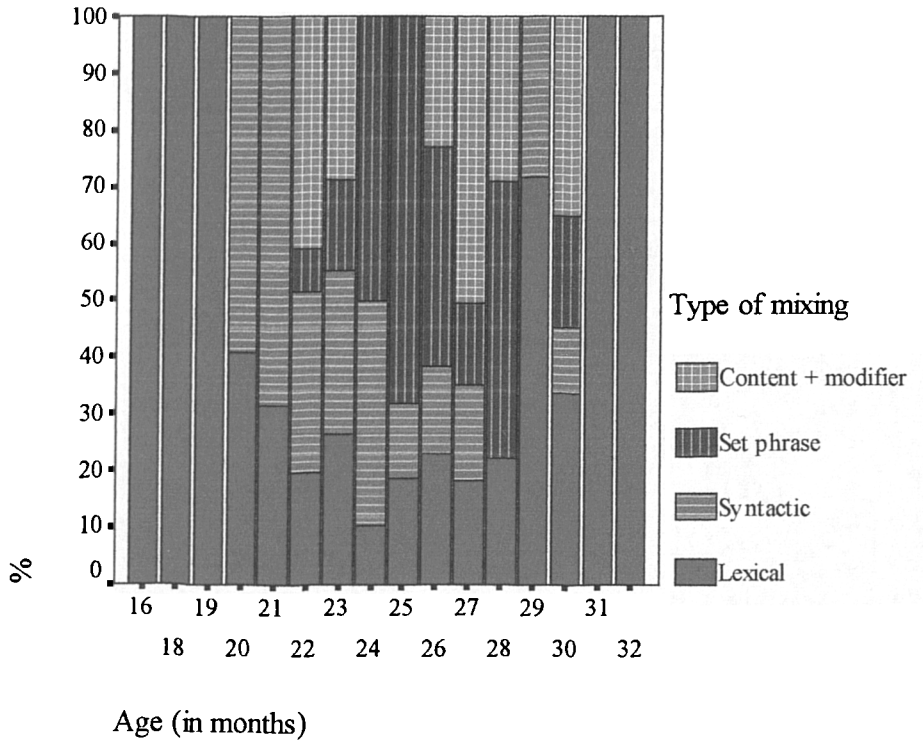


Figure 6.4 Types of mixing for ANN over time

This shows that at periods of intense mixing, ANN did produce a wide variety of mixing. Looking even more closely at the items mixed (table 6.7), we can see that the distribution per language is not exactly the same, despite ANN's earlier apparent balanced lexicon (see appendix IX).

Table 6.7 Categories of items mixed for ANN

	ML = English	ML = French
Common nouns	**58%**	33%
Predicates	15%	25%
Closed-class items	26%	**41.7%**
Social words	1%	0.3%

The table represents the overall mean proportion of mixed elements for the whole of the study period for ANN. It is quite clear that overall most elements mixed when English is the matrix language are nouns. In other words, French nouns were most often inserted into English structures and English closed-class items were most often (41% of the time) inserted into French utterances. The explanation for the pattern of mixing can actually be found into ANN's lexicon. If we look at the proportions of each category of words in her lexicon at 24 months (the end of the mixing peak period), it appears that we can find the answer for her mixing patterns.

Table 6.8 ANN's lexicon at 24 months old

	English	French
Common nouns	**46%**	**58.4%**
Predicates	28%	22.5%
Closed-class items	**15.9%**	**8.7%**
Social words	10.1%	10.4%

ANN had overall a fairly balanced lexicon. What is most striking is that the two categories where there is a noticeable difference between French and English are those she mixed elements in. As previously mentioned, mixes in English ML structures consisted mostly of French nouns. Table 6.8, confirms that at 24 months, ANN's lexicon was composed of more French nouns than English ones. This was also true throughout the period in which language mixing intensified. Consequently, a wider range of items were available to her in those cases where the English equivalent might not have been available. We also found that ANN inserted mostly English closed-class items or function words into her French structures, at that time. Here again, the English closed-class category was the largest and so contained more items to be mixed in.

This period of intense language mixing corresponded to the time when her English vocabulary had developed more rapidly than the French one and the latter was trying to catch up. It eventually did catch up around 24/25 months. That was also when mixing decreased.

6.7. DISCUSSION AND CONCLUSION

The results obtained in this chapter are in accordance with both previous results reported in earlier studies and what was expected. First of all, regarding the early word combination stage and its relationship with the 50-word stage, we can conclude that bilingual children do need 50 words in each language and 50 words in total appear not to be enough in order to create two-word utterances. This is similar to what Petitto *et al.* (2001) found although they did not investigate the relationship between the two milestones. The ages for attainment of 50-words and first two-word utterances are slightly higher than those found by Petitto *et al.* (2001) but slightly below those found by Holowka *et al.* (2002). Regarding the first appearance of the two-word utterances, again the present results show that the children in this study might have been slightly slower than those reported by Petitto at al (2001). One of the reasons for the different results could be the method used. Indeed, Petitto *et al.* (2001) and Holowka *et al.* (2002) used video recordings to measure the children's attainment of the different milestones (as well as the CDI). It has been noticed on different occasions examples where the parents would report the child starting combining

words at 21 months when in fact a couple of two-word utterances had been written down by the researcher during visits a few weeks previously (e.g. ANN and ELI). Parents probably waited that the child clearly did produce a certain number of two-word utterances before actually reporting it as they were made aware of the importance of that stage for the project. In Petitto's study, what was counted as the first two-word combination was the first recorded example. Notwithstanding this technical issue, the children in the study do fall within monolingual range for early word combinations and for attainment of the 50-word milestone. Holowka *et al.* (2002) claimed that it was not appropriate to consider bilinguals' attainment of the 50-word milestone for each language separately as that was not "an accurate index of the maturational time course". It might be true that both lexicons should be considered together, however, children in this study did reach the 50-word milestone in both languages and in each language separately within monolingual range. This is similar to what Petitto *et al.* (2001) claimed. Finally, in view of Locke's theory (1997) for the activation of the grammatical analysis module, bilingual children do appear to need 50 words in each language in order to be able to activate that module. The present results show that as soon as they have attained 50 words in one language, they will start combining words in that language. A number of parents reported that their child was combining words in one language but not in the other for a subsequent period of time. That would suggest that children do need that critical mass of lexical item in each language and that mixing is not an option there (at least in this sample). The original question based on Locke's (1997) critical mass theory was whether it is possible for a child to develop both types of morphemes separately in two different languages (see chapter 1). No child, in the present study, had a lexicon composed of exactly half French half English items. It seems very unlikely that a child could have that "perfect" scenario. Although, different types of lexical categories were used, it is possible to state that no child had only content morphemes in one language and system morphemes in the other. It seems rather unlikely that a child would be able to develop only system morphemes in one language simply because these cannot survive on their own. In a case where the child has both types of morphemes in one language and only one in the other, the language with both system and content morphemes would be the matrix language and would be the one laying down the constraints. In other words, that language would be the dominant one. More longitudinal studies are needed looking at not only lexical but also syntactic development in order to be able

to address this question with more certainty. There is also a need for longitudinal studies of children growing up in different family settings.

Secondly, we turn to the possibility of a developmental threshold for language mixing. The present data suggests that as soon as the child is able to combine languages, s/he is also able to mix, whatever the composition of his/her lexicon. The children in this study started combining words with items from both languages very early on. However, both of the children studied here had a quite balanced lexicon in terms of lexical categories and number of words in each language. As mentioned above, some children were reported to combine words only in one language for a few months. Although no detailed specific data is available for these children, parents did not report language mixing for them at that time. One other issue arising from the parents reporting the two-word utterances is the fact that they did so on separate, language-biased forms. For example, the French speaker was asked to state the child's three longest sentences heard in the last month on the French form of the CDI. Parents might have felt obliged to report French or English utterances only. So mixed utterances might have been left out. Children might have been able to produce a few mixed utterances before they actually started combining words in one language. That is certainly the case for ANN. A few mixed utterances were noted a few months before her parents reported her combining words. So, a threshold for language mixing might involve the child actually reaching 50 words in the total of the two languages. It seems as though language mixing peaks a few months after the beginning of word combination. That would show that children need a certain linguistic ability to be able to mix languages, being able to combine words might not in itself be enough. An imbalance in the lexicon sizes in each language might trigger language mixing too. More data from more children is needed, however, to address that question in more detail.

Thirdly, the possibility of the lexical gap hypothesis to justify or explain the mixes of young bilingual children was examined. The availability of lexical data allowed a systematic analysis of the lexical gap hypothesis. The results correlate with what has been found previously (e.g. Vihman, 1985) but contradict the single case study of Nicoladis & Secco (1998). The lexical gap hypothesis does not appear to hold for the majority of cases. Nicoladis & Secco's single case (1998) may therefore have been an

isolated case. There appears to be no developmental pattern for the lexical gap explanation. So the lexicon might not be the only factor responsible for the mixes produced by the child. Some more possible explanations will be outlined in the next chapter when factors for individual differences in the production of cross-linguistic synonyms are analysed.

Finally, this chapter aimed at a developmental perspective on language mixing, especially focusing on the relationship with the lexicon. As found previously (e.g. Redlinger & Park, 1980; Deuchar & Quay, 2000), the present data shows a fall in the amount of language mixing between the ages of 2 and 3. However, the second half of the U shape put forward by Meisel (1994b) was not traced. That is due to the fact that the children were not studied after 31 months old. ANN appeared to show a slight increase of the amount of mixes towards the end of the study period.

As previously reported by others (e.g. Meisel, 1994b), phrasal mixes only occurred on rare occasions for both children. Mixes involving other elements than a single noun only appeared at later stages of the development. That confirms what Meisel (1994b) claimed when he argued that sentence internal mixing "involving more than single words, require[d] grammatical knowledge". The type of mixes produced by the child is very much dependent on what is available to them at the time in the lexicon. The evidence showed that ANN did appear to mix lexical items from the categories in which a difference was found between English and French. Also ANN did mix a lot of function words (as in Deuchar, 1999), but OLB did not; his mixes were almost only single nouns (as in Redlinger & Park, 1980). So, it appears that the two types of mixing reported in the literature by different studies are both valid. Both types of language mixing appear to be present in young bilingual children. The reasons behind these two types of mixing are not, however, clear. If we take into account Deuchar's explanation (1999) that function words are not language-specific, then why did OLB not mix them too? As Deuchar (1999) noted for M, both ANN and OLB did have translation equivalents for closed-class items. The explanation may rest in the fact that ANN used the articles *this* and *that* on numerous occasions in early word utterances and she mixed them with French content words to a very large extent in the earlier stages of mixing. This probably inflated the percentage of closed-class items mixed, as the percentage of closed-class items mixed decreased towards the end of the study

period. Vihman's claim (1985) that language mixing at earlier stages in characterised by mixed function words is true for ANN but not for OLB.

One last point made by Meisel (1994b) suggesting that balanced bilinguals tended to mix languages more than less proficient or non-balanced speakers appears to be verified in this study. The reason for choosing ANN and OLB as the focus of this chapter was because they mixed to a much larger extent than other children in the study. It is also true to say that they were two of the children for whom the two lexicons developed almost at the same rate and speed.

We can thus say that the lexicon plays a very important role on the development of language mixing. However, the lexicon is probably not the only factor that plays a role. The next chapter investigates further influential factors in language mixing and lexical development.

SECTION IV:

FACTORS INFLUENCING BILINGUAL LEXICAL ACQUISITION

What children say, they have heard at home.

Wolof proverb.

CHAPTER 7:

SOCIO-LINGUISTIC VARIABLES

7.1. INTRODUCTION

As mentioned throughout this study, bilingual language (lexical) acquisition does not differ from monolingual acquisition in the sense that important differences are found amongst children who, at first sight, appear to have a similar linguistic environment. This chapter will first highlight and then try and explain some of these differences by examining different available variables affecting lexical acquisition and, to a lesser extent, language mixing in bilingual children. Much of the literature relevant to this chapter has been reviewed earlier.

Firstly, after having outlined the aims of this chapter, I will establish in more details the type of individual differences found. Secondly, the methods and analyses used in this part of the study will be described in order to account for these differences. Next, results will be presented regarding different variables that have been identified as influential in monolingual language acquisition. Later, possible influences will be studied on the quality and quantity of (i) the lexicon and the emergence of word combination; (ii) cross-linguistic synonyms; and (iii) language mixing. Finally a discussion of the present results will follow.

7.2. AIMS AND OBJECTIVES

Here, the main research question is: how does the socio-linguistic environment impact on bilingual children's lexical acquisition and development of language mixing? In order to answer this question this chapter will focus on several issues:

- bringing more evidence to the debated issue of gender and its influence on language acquisition;
- testing the extent to which the amount of exposure to a language is responsible for the lexicon size of a child;
- studying the impact of the quality of the input provided by parents on different aspects of the children's language;
- examining a possible link between the directionality of mixing and parental strategies.

7.3. ACCOUNTING FOR INDIVIDUAL DIFFERENCES

One of the main limitations of group studies is the fact that individual differences fade away as results are represented as means. By voluntarily keeping the number of subjects to a humanly manageable size, this study provides a great opportunity to look in more detail at individual differences in group studies and individual differences in bilingual lexical acquisition in general. Throughout the previous results chapters, I have tried to convey the great variability that exists amongst the children taking part in the study. More details on individual children will be given here comparing them with means and other individual children.

In chapter 3, the very similar lexical development of bilingual children and monolingual children was examined in terms of number of words produced and also in terms of the balance and development of lexical categories. Despite these apparent similarities, among the group of bilinguals followed for this study, lexical development varies to a great extent.

Figure 7.1 and 7.2 respectively represent OLR and PEN's lexical development as measured by the parental reports. It was reported earlier that lexical acquisition of bilingual children was very similar to that of monolinguals. However, that does not mean that the two languages develop in the same way and/or at the same speed. As illustrated on figure 7.1, OLR's two languages follow a very parallel development.

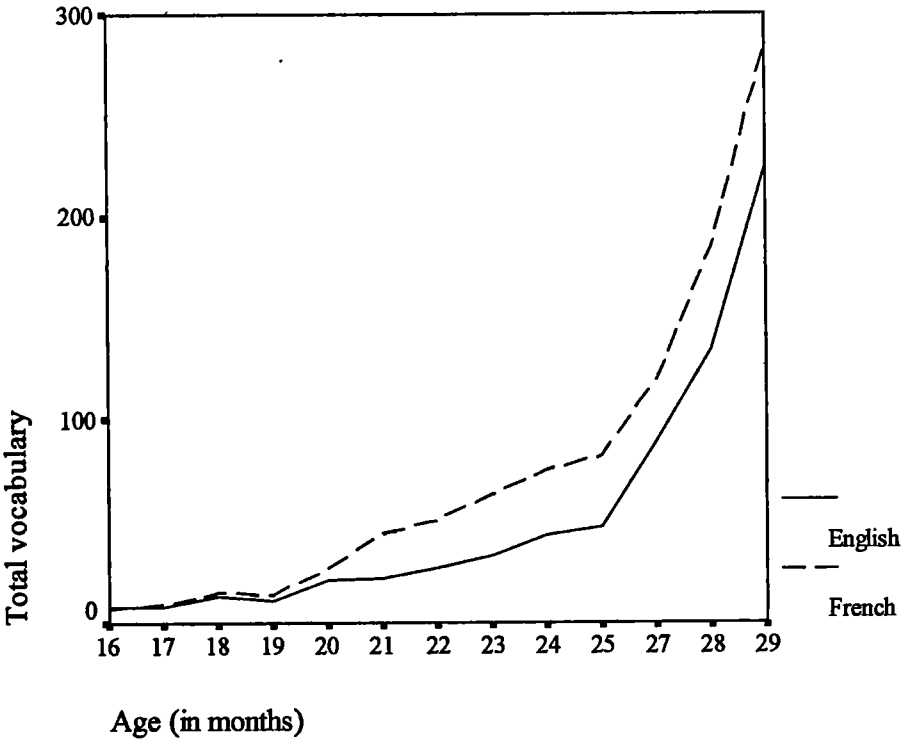


Figure 7.1 OLR's lexical development

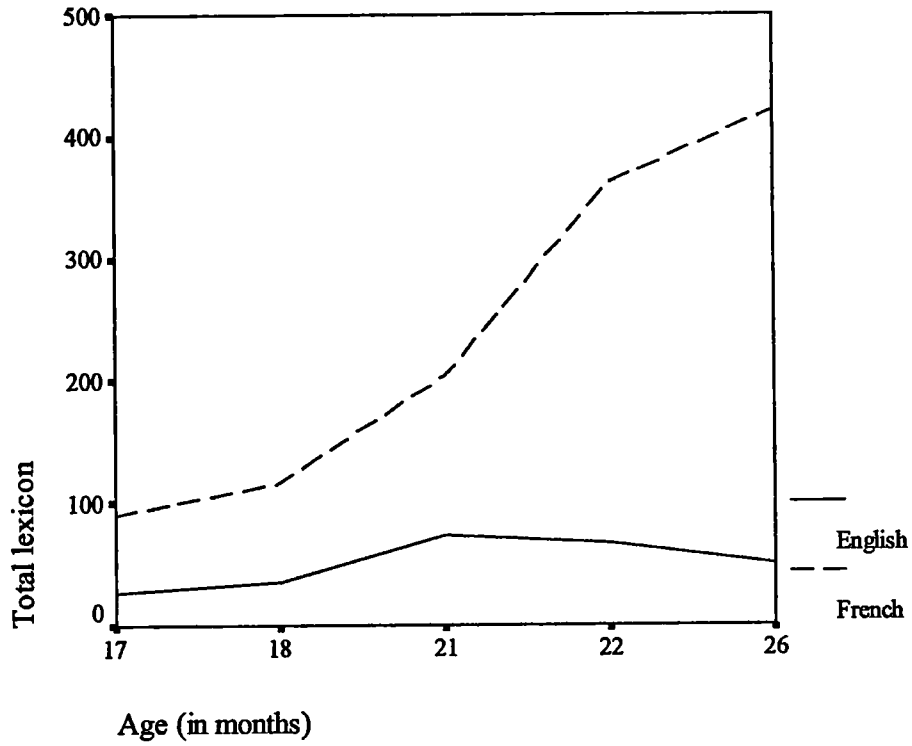


Figure 7.2 PEN's lexical development

In contrast, PEN's English does not develop at the same rate as her French at all. Although the two languages do appear to develop at the same rate in the earliest stages (until 18 months old), later, French increases much more rapidly than English. The main reason behind that significant difference between the two languages is the input PEN receives. She was exposed to French between 80 and 90% of the time. In addition, her father (the minority language speaker) was not insistent on PEN using English with him (he did not always correct her or provide feedback for her in English whenever she used French).

The second type of difference observed comes from the number of words produced. Although, on average, the children in this study compared to monolinguals (see chapter 3), the variability observed amongst the children (as is the case in monolingual acquisition, e.g. Bates *et al.*, 1994) is highly significant. Table 7.1 gives an overview of the lexical data available for eleven of the children at 23 and 30 months old. It is obvious from that table that at the same given age, bilingual children do vary to a very large extent in the number of lexical items they produce. At 30 months old, for example, FLO produced 993 lexical items, according to her parents, while DUN produced nine times fewer words than FLO (109).

Table 7.1 Lexicons at 23 and 30 months old

	23 months	30 months
THO	54	N/A
FLO	380	**993**
EMM	70	N/A
DUN	33	**109**
ELI	344	N/A
JAC	93	323
ANN	**513**	**955**
OLR	96	466
LIA	33	250
REB	49	256
OLB	**156**	**817**

This table also shows that the rate and speed of lexical acquisition varies from child to child. For example, OLB only had 156 words at 23 months but caught up with ANN by 30 months despite the fact that he had three times fewer words than her at 23 months. This can be further illustrated by the figure below.

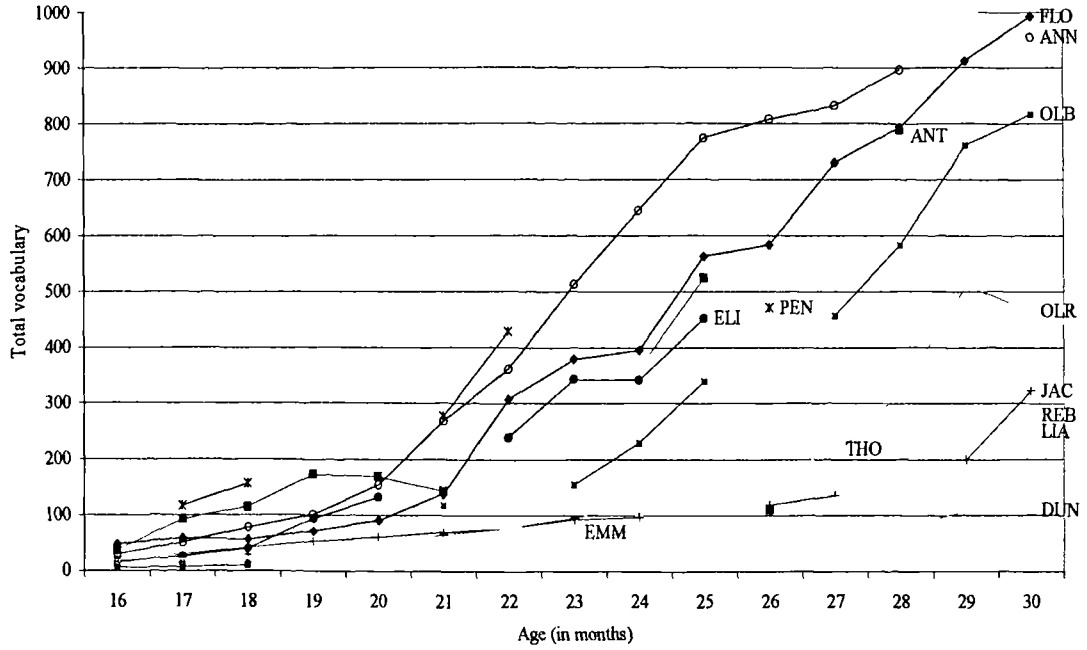


Figure 7.3 Lexical development of the 13 children

This chart represents each individual child and their total lexical production throughout the study period. Two groups appear to emerge out of this chart. Up until 19/20 months old, there is very little to differentiate the children. However, one group of children (all girls) appear to have their vocabulary burst shortly before their second birthday. By 23/24 months old, we can clearly distinguish a group of children above the 200 word-stage. Their vocabulary expands more rapidly than the others. The children in the other group only appear to have a vocabulary burst towards the end of the study period (e.g. JAC, OLR). Overall, their lexical development is slower. OLB appears to be somewhere in between the two groups. His vocabulary burst happened slightly later than that of the girls in the first group (around 2 years old).

Individual differences in the acquisition of cross-linguistic equivalents have been illustrated in chapter 4.

Individual differences also exist both in the presence and absence of language mixing. First of all, not all children do mix languages or at least, not to the same extent. Here, we take into account all instances of wrong language choice, including single word utterances in the wrong language and multi-word utterances comprised of at least one item in the wrong language.

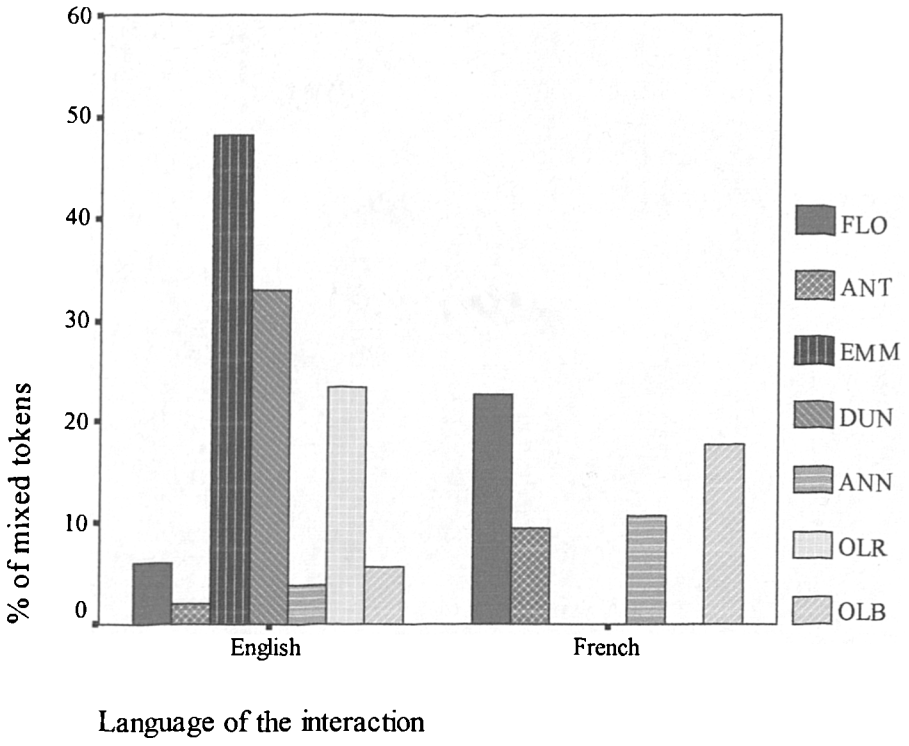


Figure 7.4 Mean percentage of tokens produced in the wrong language

Figure 7.4 represents the percentage of tokens in the wrong language for seven of the children in the study. It shows that not all children do make the wrong language choice to the same extent. For example, EMM uses a large number of French lexical items with her English-speaking mother. One of the main reason behind that is due to the number of hours she is exposed to French and also her mother's strategies as will be explained later in this chapter. Perhaps more importantly, not only is there variation between children but there are also individual differences between languages. Figure 7.5 shows the percentage of wrong language choice by OLB. One can clearly notice that the majority of his mixes are done when French is the matrix language (with his mother). Again, the reasons behind this appear to be language exposure and parental strategies.

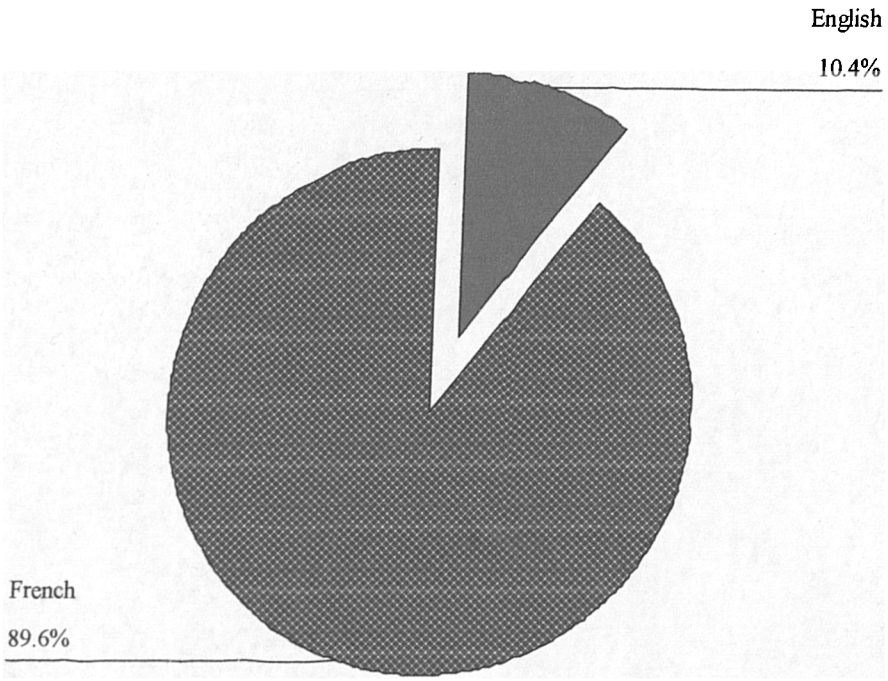


Figure 7.5 Wrong language tokens for OLB

Individual differences in bilingual lexical acquisition are very important and although group studies do not usually include them, it was necessary to highlight the fact that the results represented in the previous chapters were averages across children. Individual children show different rates of acquisition from one another. Bilingual lexical development is very similar to monolingual however, this does not mean that the two languages develop at a similar rate. Based on these significant differences, possible explanations will be brought forward.

7.4. DATA ANALYSES

7.4.1. Lexical data

In order to account for the factors influencing lexical acquisition and the factors responsible for the individual differences, different methods were used. The data presented in this chapter is based on several methods previously described in chapter

2. First of all, lexical measures are based on the CDIs. I will not explain the process here as it has been described above in more detail. The CDIs, in turn, allowed compiling data about the quantity and nature of translation equivalents in the children's lexicon. The reader should note that for the analyses below, all of the data collected will be taken into account (native and non-native scores; see chapter 5 for the differences between those scores). In addition, the reader should be aware that the total vocabulary of each child will be used, meaning their bilingual lexicon composed of French and English data (unless otherwise stated). The possible overlap in some cases between the two lexicons (see the distinction between TCV and total vocabulary in chapter 1) is acknowledged. However, since no comparison will be made between the lexicon and monolingual norms quantitatively, it will not influence the results. The total vocabulary score is therefore only a guide. Furthermore for all of the analyses below, TEs numbers given are not the 'conservative estimates' but the count including cognates.

7.4.2. Language input

Secondly, as mentioned previously, parents were asked to quantify the language input received by their child every month at the same time as they were completing the CDI. Some parents chose not to do so every month but only when a change occurred in the child's life or when an unusual event occurred (e.g. visit of the grandparents). From that data, rough monthly estimate of the child's language exposure to each language (in percentage) was calculated. An effort was made to stay as close as possible to what really went on in the child's life. Nonetheless, estimates like these are quite subjective and should be taken with caution as they mostly rely on the parents' idea of "being exposed to a language", though an explanation of what this meant was given to the parents. And for example, the estimates included the time children overheard a language even though they might not be directly included in the activity happening (e.g. the television being switched on in the background). The visits were a great help as I was able to observe the child's input first hand and not rely on parental reports only. Consequently, what will be referred to as input or the amount of input a child received at a given period directly refers to these percentages.

7.4.3. Naturalistic data

Thirdly, in this chapter will also be used data based on the audio recordings made of the children interacting with each of their parent separately in two different language contexts. The recordings were transcribed with the CLAN program. Analyses made included TTR (type-token ratio), D (a measure of vocabulary diversity) and MLU (mean length of utterance) of both parents and children. MLU (words) was calculated on the basis of all utterances produced in one single recording episode. Utterances containing mixed elements (or switched items for the parents) were also included in the analyses as they are an integral part of bilingual discourse. Omitting utterances with mixed elements from the children would not have allowed a large enough sample of utterances solely in one language. It is necessary to give more details about TTR and D. When it comes to measuring vocabulary diversity in child language research, measures have been largely based on the ratio of different words (types) to the total number of words (tokens), also known as the type-token ration or TTR. Recently, a new measure called D has been introduced calculating lexical diversity based on a probabilistic model yielding type-token characteristic curves (McKee, Malvern & Richards, 2000). D has been validated by different studies on different populations (see Malvern & Richards, 2004). TTR has been criticised for providing biased results based on the number of tokens in the language sample – "samples containing larger numbers of tokens give lower values for TTR and vice-versa" (MacWhinney, 2004). Malvern & Richards (2004) claimed that the TTR results have distorted many research findings. In this study, although parents were encouraged to keep the recordings lasting for around 30 minutes, they varied enormously in length. Consequently, the D index of lexical diversity seems rather more appropriate as it is less sensitive to samples' length. Both of these measures are available, and were calculated, using the FREQ and VOCD programs of the CLAN software. Differences yielded by the two methods will be briefly highlighted below.

It is important here to note that the use of MLU and TTR measures is widely used in child language acquisition with relatively reliable results up until a certain stage (see Brown, 1973). However, the use of MLU for older children and adults poses a certain number of issues that have lead Dewaele (2000), for example, to claim that it is unreliable for these types of population. Indeed, very short sentences (e.g. *Oh yes!*)

tend to bring the average of MLU, for example, to a very low score which in reality does not reflect the adult's abilities. Taking these issues into consideration, I would like to point out that MLU and TTR (or D) scores for parents in this study are not, in any case, an exact measurement of their capacities. It is merely a readily available tool that enables a brief analysis of the parents' language as it is not the aim of this research to analyse in depth parents' language.

7.4.4. Mixed utterances and parental strategies

Finally, from the transcripts were counted the number of tokens produced by both parents and children in the 'wrong' language. This means that if a French-speaking mother said to her child "YES, c'est le chapeau", the token *yes* would have been counted as a wrong language item. The word 'wrong' does not mean that this is a wrong behaviour, rather that it is the inappropriate language given the language rule parents imposed. In addition to that, for each utterance containing at least one mixed token from the child, the strategy of the parent used in response to it was examined. In order to examine parental strategies, Lanza's (1992) five parental discourse strategies as discussed in chapter 1, were used as classification system. These strategies were then classified into two main categories: feedback-providing (example 1) and non-feedback-providing (example 2). In example (1), OLR's mother uses two different strategies in response to OLR's two mixes. First of all, she makes an expressed guess, providing at the same time the correct word. Secondly, she simply repeats the right word. In both cases these are feedback-providing strategies as she clearly shows that *pain* was not the right word to use there. In example (2), EMM's mother does not provide any feedback to EMM's French utterance thus causing EMM to repeat that utterance once again in the wrong language (French). Had she been providing with feedback and the right structure in English she might have used it the second time round. This type of conversation actually goes on for more than five minutes of EMM asking in French *C'est quoi ça?* and her mother answering the question without any translation of the original question.

(1) OLR, 1;10.24, English

OLR: PAIN.

[BREAD].

MUM: Some bread ?

OLR: PAIN.

OLR: Maman.

MUM: Here is some bread.

(2) EMM, 1;11, English

EMM: C'EST QUOI ÇA ?

[WHAT IS THIS?]

MUM: That's a sheep .

EMM: bêê.

MUM: bêê.

MUM: The sheep goes bêê.

EMM: C'EST QUOI ÇA ?

[WHAT IS THIS?]

MUM: It's a dog.

We felt that the minimal grasp, expressed guess and repetition strategies had the added benefit of providing feedback to the child and offer the 'right' alternative for what s/he had just produced in the inappropriate language. Döpke (1992) had classified parental discourse strategies according to whether they elicited a response from the child. Following this categorisation one would have included only minimal grasp and expressed guess. However, it was felt that the repetition strategy offered by Lanza (1992) would also be beneficial to the child as it provided the right alternative. Even though the parent clearly demonstrated his knowledge of the other language and did not require the child to reply, at least the child would hear that word and hopefully take it into account for the next time s/he might need it. In the minimal grasp strategy, the parent does not immediately give the right alternative. It is up to the child to offer it. So the child is, in a way, his own feedback-provider. In cases where the child does not find the alternative, the parent usually ends up providing it (at least in the vast majority of the cases observed in this study).

7.5. DIFFERENT VARIABLES AFFECTING LEXICAL ACQUISITION

Before we turn to input and discourse strategies, let us examine well-known variables that have influenced monolingual language development. The first goal in this chapter is to establish the impact of the well-known variables such as gender or birth order on bilingual children's developing lexicon. As outlined earlier (in chapter 1), these factors have been shown to influence language acquisition in general despite there being contradictory results from different studies apparently using apparently similar methods. The literature has generally favoured girls and first-borns where any differences to be found.

7.5.1. Gender

Figure 7.6 represents the children's lexicon in function of their age. Lines are defined by gender.

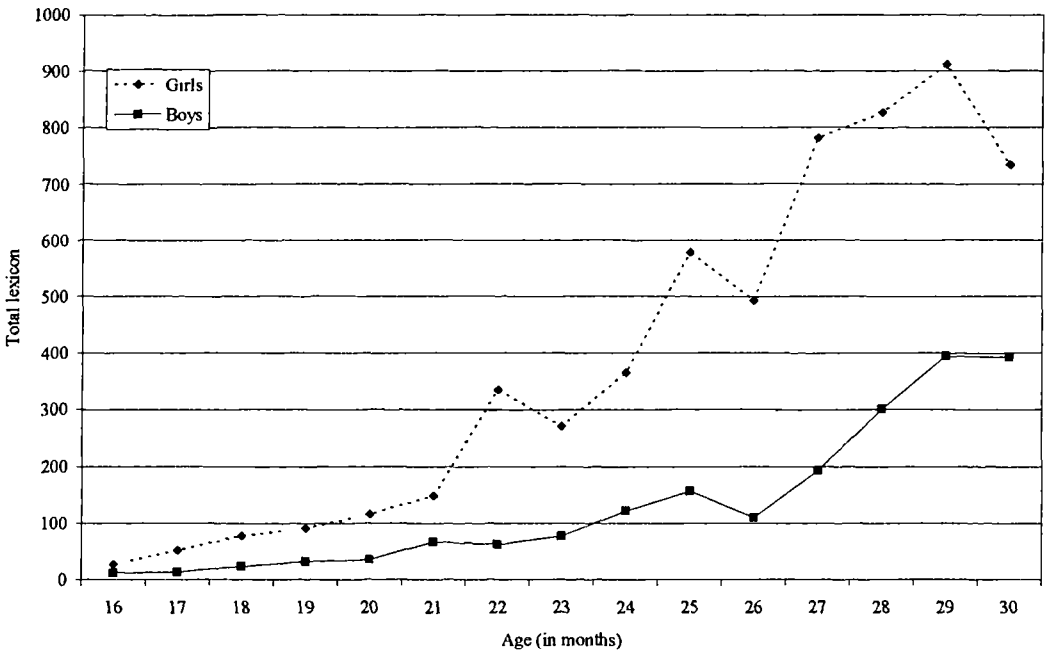


Figure 7.6 Lexical acquisition as a function of gender

We can see clearly that there is a significant difference between the boys and the girls in this study. The evidence suggests that the girls have a much larger vocabulary than the boys and that the gap increases as they grow older. The present data only represent

the means for each month. However, looking at the different scores for each child, it is essential to notice that the variations within each group are very important. Towards the end of the period of study, the chart would tend to show that the gap closes and that boys might catch up.

Table 7.2 Differences in items produced as a function of gender

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Total vocabulary	Male	157	63.57	83.451	6.660
	Female	167	135.62	147.086	11.382

The average mean of words produced in one language over the whole time period by the girls is higher than the boys' (135 vs. 63; see table 7.2). There is a small significant correlation between vocabulary (native scores only) and gender ($r=.297$, $p<0.01$, $N=312$) favouring girls. It is also important to note that these correlations would be higher if calculated on the basis of the native parents' scores only.

The results are consistent with previously reported results for monolingual acquisition. Kern (2000) reported a significant difference between the boys and the girls' production data from the FCDI favouring the girls. Le Normand *et al.* (2002) found similar differences for French acquisition. Fenson *et al.* (1993) also noted the higher score of the girls overall in the American CDI. However, Kern (2001) mentioned that at 30 months old, the boys have caught up with the girls and produce on average slightly more words than the girls. The reason behind this is not clear though. Le Normand *et al.* (2002) interpret their results as a memory issue. They advance the possibility that girls might have a better declarative memory than boys. This part of the memory is responsible for the mental lexicon and the set of grammar rules children need to build utterances. The fact that the bilingual girls and boys in this study appear to be behaving in the same way as monolingual children is not in itself surprising. Bilingual children would then appear to have and use the same type of memory as monolingual children to process their lexicon and utterances regardless of whether they have one or two languages to deal with.

7.5.2. Birth order

The second factor which has been seen as an important variable affecting language acquisition in monolingual children is birth order. No correlation was found between vocabulary and birth order. Despite the balance between first-borns and children with older siblings, the number of subjects might be simply too small to find a significant difference between the two groups.

7.5.3. Impact of the quantity of input...

We will now turn to the question of input as this has been reported as probably the most influential factor in language acquisition especially that of bilingual children (see chapter 1). It has been highlighted earlier the fact that both the quantity and the quality of the input received by the children is important. I will, first of all, focus on the quantity of input and its effects on lexical development. Secondly, I will examine the apparent impact of the parental input's quality.

7.5.3.1.... On the lexicon

One of the first aims for this section is to examine the impact of language exposure on the lexicon. Pearson *et al.* (1997) have shown a strong correlation between language exposure and the amount of words known. It was expected that these results would be replicated with the 13 children in this study.

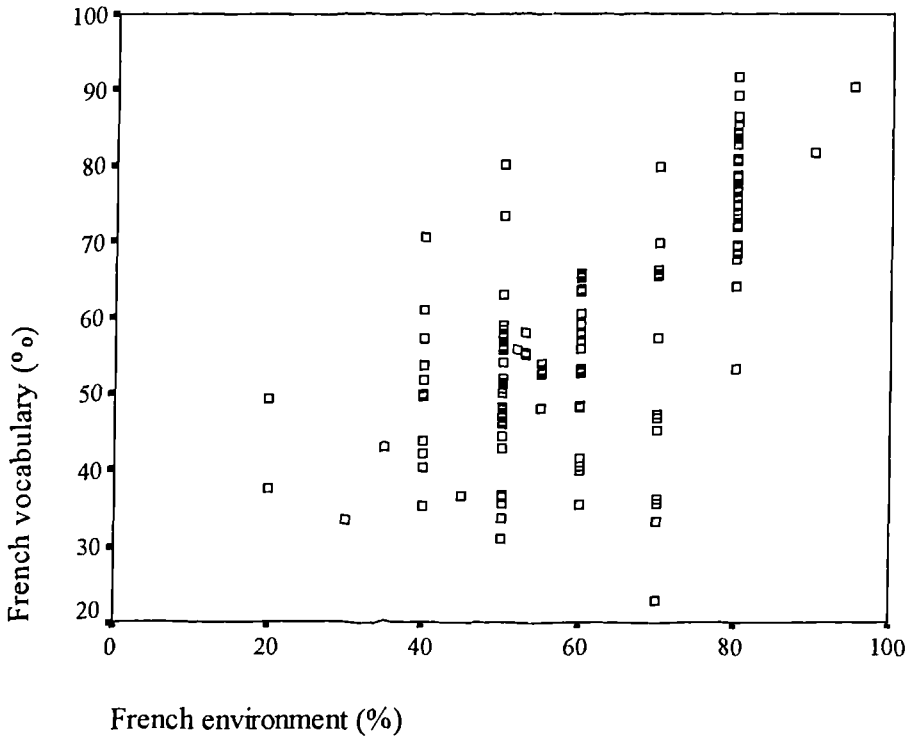


Figure 7.7 Proportion of the French lexicon as a function of the exposure to French

Figure 7.7 shows that the more a child is exposed to French, the more likely s/he is to acquire more lexical items from it. The correlation between exposure and vocabulary is significant ($r=.650$, $p<0.01$, $N=131$). The number 131 represents, here, the number of data collection points taken into account for this analysis. Children with a balanced input (40%-60%) do also appear to have the most balanced lexicon. It is also important to note that even children whose exposure to one language is limited will acquire words in that language. However, no significant correlation was found between exposure and vocabulary for the cases where the exposure to French was lower than 50% probably due to the fact that the number of children in that situation was quite small. It would have been interesting to have access to a wider range of situations as it turned out that most of the children were predominantly exposed to French in their early years.

One more question to be asked in this section relates to whether major changes in the input to a child could influence their lexicon size. And the results showed that there is a small significant correlation between changing environments and lexicon size ($r=.246$, $p<0.01$, $N=131$). Changing environment means that the change in language exposure was greater than, or equal to, 20% from one month to the next. One could

argue that that figure of 20% is rather subjective. Nonetheless, this figure does not include children who just spent two weeks holidays in the presence of speakers of only one language (and it did not). So, the data seems to favour children with a more stable input pattern. It is important to note that in the stable input category were the two children with the largest vocabulary and the two children with the lowest too. So, other factors must play a role, some of which will be accounted for further below. The effects of a changing environment can be seen on figure 7.8 that represents REB's changing exposure to French and the proportion of French lexical items in her total lexicon.

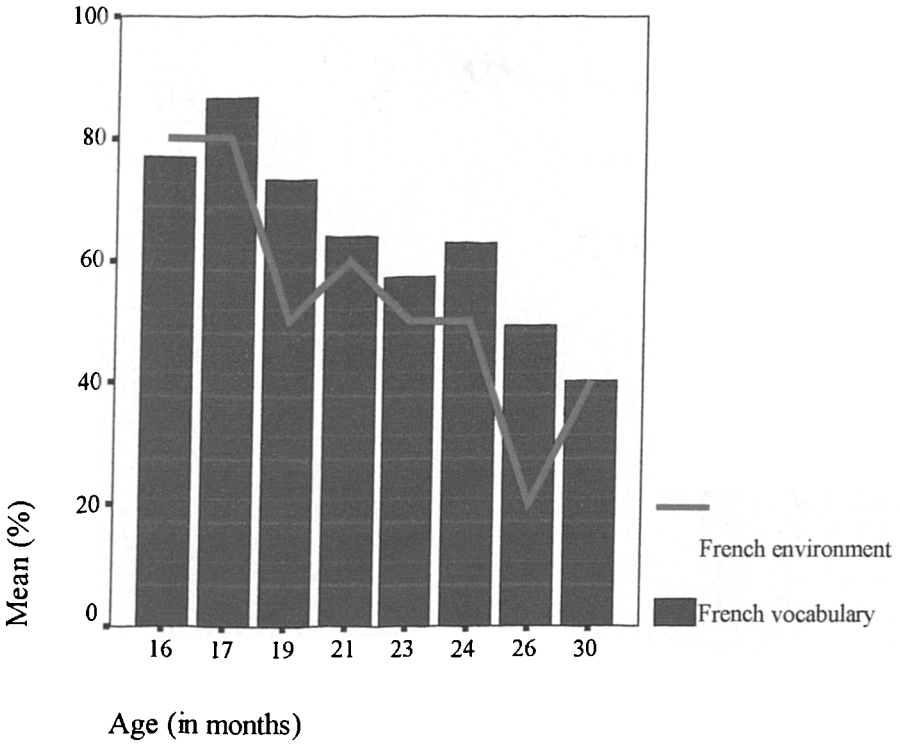


Figure 7.8 Evolution of REB's French lexicon as a function of her exposure to French

We can see that as REB's exposure to French diminishes the proportion of her French lexicon in the total lexicon decreases too. REB lives in an English speaking environment (England). At the beginning of the study, REB was cared for at home by a native French speaker (a family relative) and her French speaking mother and by the end of the study was going to an English-Speaking nursery while her mother worked gradually more often. Between 19 and 24 months, the stabilisation and even slight increase in the amount of exposure to French seems to result in the slower decrease

(or even slight increase) in the proportion of her French vocabulary at 24 months. The sharp decrease in her exposure to French at 26 months is due to a visit to her English-speaking grand-mother for three weeks on her own, without her parents.

Indeed, these results match the earlier study of Spanish-English bilinguals by Pearson *et al.* (1997). They reported the smallest correlation for children less exposed to Spanish. That parallels the present results and shows the very important role that language exposure plays in language acquisition, and not only bilingual acquisition.

7.5.3.2.... On the presence of translation equivalents

We have illustrated the fact that the amount of exposure to a language a child receives affects his/her knowledge of lexical items in that language. Translation equivalents (TEs) or cross-linguistic synonyms are subset of that lexicon and therefore must be influenced too. Moreover, in chapter 4, significant disparities between the numbers of translation equivalents present in the children's lexicon were highlighted. One can wonder what might have been the factors influencing these significant differences between children. Pearson *et al.* (1995), who also found differences in the number of TEs in Spanish-English bilinguals, did not offer any explanation. They only mentioned that balanced bilinguals were rare and that most of the children in the study were clearly dominant in one language. So, one hypothesis could be that the number of translation equivalents depends on the language dominance situation of the child. However, the authors did not go deeper into this possibility. No studies have, to the best of my knowledge, looked at the link between translation equivalents and language exposure.

Table 7.3 describes the correlations between the translation equivalents and language exposure on the first instance and the correlation between translation equivalents and the lexicon in the second instance. The table reveals that the correlations are significant in both cases. Only taken into account are cases where language exposure in French is superior or equals to 50% and where the French lexicon represents 50% or above of the total lexicon. Only the cases for which the French lexicon's percentage was superior to 50% (or the exposure to French was more than half) are

represented as most of the children in this study turned out to be balanced or dominant in French. Yet, the correlations are also significant in the other direction although to a lesser extent probably due to the smaller number of cases.

Table 7.3 Correlations between translation equivalents and language exposure and lexicon

	r	p	N
TEs & language exposure	-.517	<0.01	89
TEs & lexicon	-.632	<0.01	89

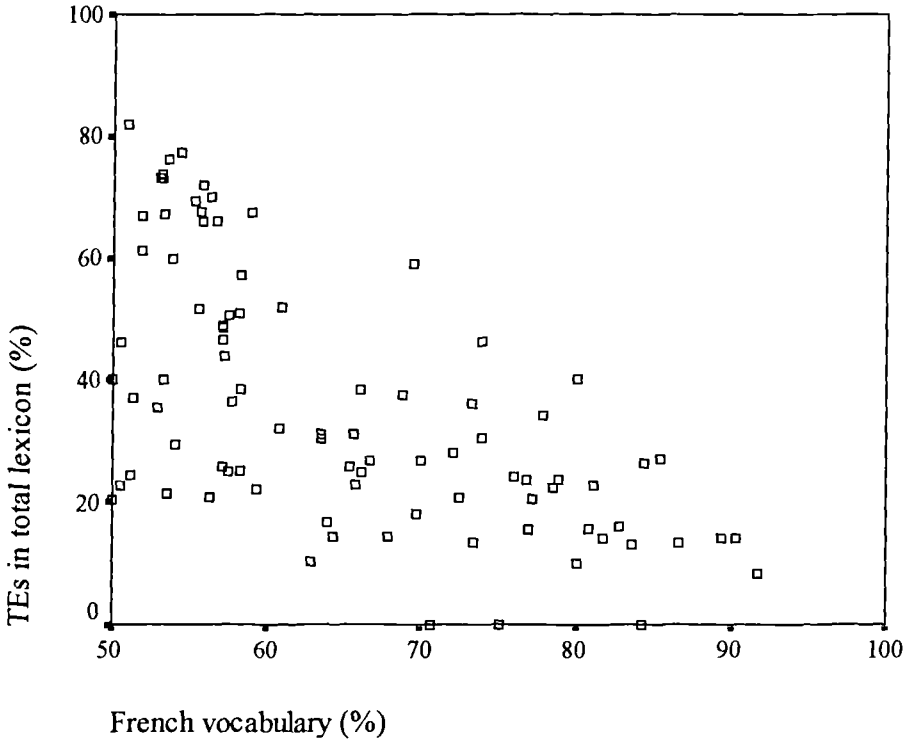


Figure 7.9 Percentage of translation equivalents in the total lexicon as a function of the proportion of the French vocabulary

Figure 7.9 represents the percentage of translation equivalents in the child's lexicon as a function of the percentage of the French lexicon in the total lexicon (English + French). That chart illustrates that relationship even further revealing that the less balanced a child's lexicon is (or the more dominant a child is in one language), the less likely it is to contain a large number of cross-linguistic synonyms.

The more linguistically balanced the child's environment is, the more likely the child is to have cross-linguistic synonyms in his/her lexicon. And the more a child's lexicon is balanced, the more likely s/he is to have translation equivalents in his/her overall

lexicon. The three relationships are linked with each other. Consequently, the quantity of input a child receives in a language impacts both the lexicon and the presence of TEs. These results are key for two reasons. Firstly, they demonstrate that not only does the input influence the quantity of words acquired in each language but it also influences the nature of these words. Secondly, the results described probably explain, at least partly, some of the disparities between the children found in chapter 4 on the presence or not of translation equivalents.

One important issue to be raised here is that although the amount of exposure to a language can impact greatly a child's lexical development, not only in its quantity but also in its quality, we have assumed all along that amount of exposure translates into amount of input. It has been assumed that the number of hours a child spends with speakers of a given language will automatically provide him or her with language input. That might not always be the case. The literature has shown that the range of utterances addressed to a child vary greatly in their nature and quantity as the child's abilities develop (e.g. Hart & Risley, 1995). So this study assumes a relative measure. It is obvious that the nature and quantity will vary depending on the situations and speakers. The method used only takes into account a proportion of the total heard rather than the actual number of words.

7.5.4. Impact of the quality of input...

The quantity of the input received clearly appears to impact the lexical development of bilingual children. But the literature has revealed that quantity is not a sufficient factor. Quality of input is at least as important. Quality of input has been widely measured in studies with monolingual children, using two related features: the vocabulary and the sentence complexity parents use (Hart & Risley, 1995). We shall now examine to what extent the quality of input received by bilingual children influences (i) the lexicon; (ii) translation equivalents; (iii) word combinations and (iv) language mixing.

7.5.4.1.... On the lexicon and the emergence of word combinations

Table 7.4 Correlations between parents and children's language (MLU, D and lexicon)

	Child's Lexicon [a]	MLU Child
MLU Parent	**.630****	**.621****
D Parent	**.485****	**.515****

[a] lexicon in one language only (the parent's language) **significant at p<0.01

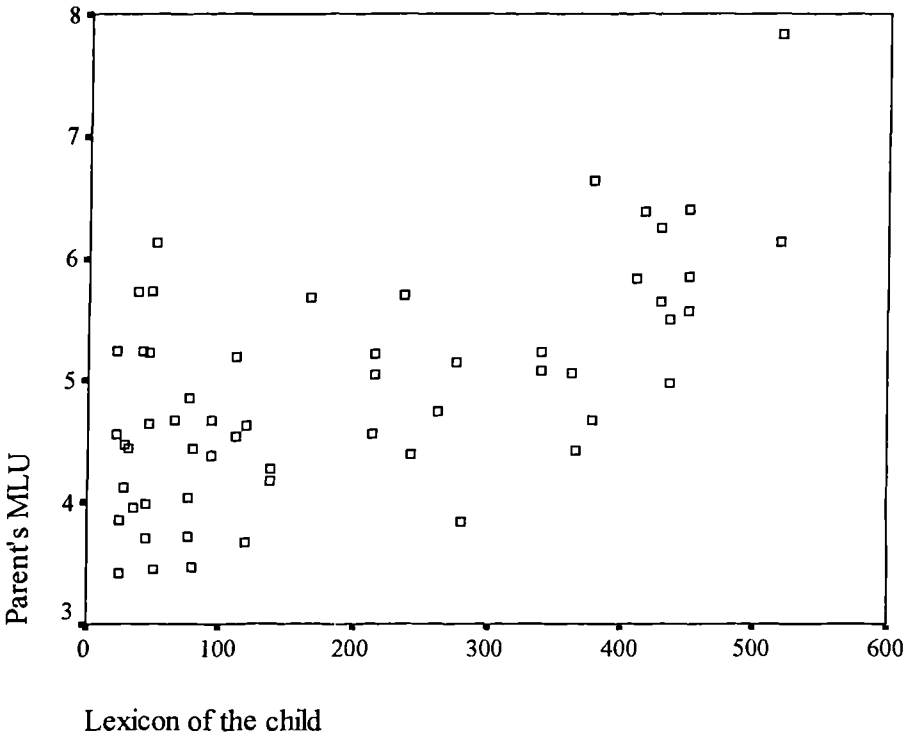


Figure 7.10 Parents' MLU as a function of children's lexicon

Table 7.4 and figure 7.10 highlight the relationship between the quality of the input from parents on the lexicon of the child. The most significant result is the fact that the mean length of utterances from parents and parental lexical diversity do impact on the number of lexical items known by the child. It appears that, as the child's abilities increase, parents use longer and more complex utterances. This result is not at all surprising as parents do talk differently to younger children. Not only does the length and lexical diversity of the parents' utterances influence the lexicon of the child, but

they also impact the MLU of the children. Table 7.4 also illustrates the fact that the correlation between parents' MLU or parents' lexical diversity and the child's MLU are statistically significant in both cases. Which factor influences which is rather unclear. One could hypothesise that parents notice the child's growing abilities and therefore realise that they can use more complex and longer sentences. Or maybe it is a case that the child's abilities impact on the parents' speech. It is likely that it is a 'virtuous circle'. Both influence each other and the child benefits from the whole process. It allows him/her to improve his/her language skills.

No statistical differences were found between mothers' and fathers' indexes of lexical diversity or mean length of utterances. Correlation between parental and children D is not significant.

Although, it was previously emphasised that TTR might not be the best way of measuring lexical diversity (see section 7.4.3), it was felt that using TTR might provide a larger sample of data to work with. D "is calculated from averaging the TTRs of 100 trials on sub-samples consisting of the number of tokens for that point, drawn at random from throughout the transcripts" (MacWhinney, 2004), therefore D does not yield any result if the language samples are too small and consequently limits the analyses. However, the TTR ratio gives very different if not completely opposite results to D. For example, the correlation between parents' TTR and the child's lexicon yields a significant negative correlation. That result is extremely surprising. It would mean that the more children know, the less diverse the speech of the parents become. This result seems to confirm the fact that TTR is function of the total number of tokens and that its results distort research findings as previously anticipated.

These results are consistent with previous studies on monolingual children that demonstrated that parents talk differently to younger children not only in the quantity of input given but also in the nature of input (e.g. Hart & Risley, 1995).

7.5.4.2.... On the presence of cross-linguistic synonyms

As explained above, the results regarding the number of translation equivalents in the bilingual lexicon yielded very significant differences amongst the 13 children. It has been shown that the quantity of language exposure plays a role in how many TEs a child acquires. A second possible factor would be the nature and the quality of the input. Döpke (1992) found that the way in which parents apply the one person-one

language strategy affect the way the children use their languages. So, one could hypothesise that if the parents apply a strict rule, then the child has clearer cues to understand the bilingual situation s/he is in and is therefore helped to develop cross-linguistic synonyms. Thus, we need to take into account the quality of the parents' input in the issue of the presence of TEs.

First of all, the results suggest that the more diverse the parents' vocabulary is, the more TEs the child has in his/her lexicon. This is no surprise as earlier results showed that as the child grows older, and his/her lexicon grows, parents use longer and more diverse utterances. Since the evidence in chapter 7.4 suggested that the proportion of TEs in the lexicon increases with time, then the correlation between TEs and parental D was likely to be, and is, significant (r=.504, p,0.01, N=57).

We did not find that the amount of code-switching done by the parents influenced the number of TEs produced by the child nor that the type of parental strategies, in case of mixing from the child, correlated with the number of TEs. The correlations were not significant at the individual level either. This can be illustrated with the case of OLB. He had a large number of TEs (above average) and did also produce a large amount of mixing as seen in the previous chapter. His parents used similar parental strategies to other parents (see below) and did not always provide an equivalent by using feedback-providing strategies, thus providing no link between the two. I shall come back to that issue in the following chapter. That is not what might have been expected in light of Döpke's claims (1992). Nonetheless, the sample might be too limited to have shown a significant correlation. Another reason might be that when parents are recorded, they are clearly more aware of their language behaviour. So, the actual amount of code-switching or responses to it, might be different from the one observed. It appeared, however, from the visits that the behaviour displayed during the visits was fairly consistent with recordings. Given that the visits lasted for sometimes more than a day, it was hoped that parents would have not been able to modify their behaviour in a dramatic way during that substantial amount of time. Still, this possibility cannot be completely excluded.

7.5.4.3.... On language mixing

Finally, the previous chapter has highlighted the fact that some children did mix languages while some others did not. In addition, among the children who did mix, the extent to which they did, in each language, varied immensely. The reasons behind this are unclear. Studies have advanced the fact that parents have a great role to play in the amount of language mixing children produce (e.g. Goodz, 1989).

The present results provide further evidence for the highly important role of parents in language mixing. First of all, a small correlation is found between strategies offering an alternative word (or feedback-providing) and the amount of language mixing produced by the child ($r=-.285$, $p<0.05$, $N=58$). That correlation is even greater if FLO is excluded from the analyses as she did produce mixed utterances but mostly at her father's request (see next chapter). However, once again great disparities between parental styles are noticed (as can be seen on figure 7.11). The percentage of feedback-providing strategies as opposed to the non-feedback-providing ones can vary from 0% to 100%. These differences are mainly due to the large differences in the length of the recordings and the amount of language mixing from the child. Nevertheless, the overall trend seems to indicate that children tend to mix less when parents use more strategies like minimal grasp, expressed guess or repetition. The correlation was not found when classifying strategies according to (non) response-eliciting strategies as suggested by Döpke (1992).

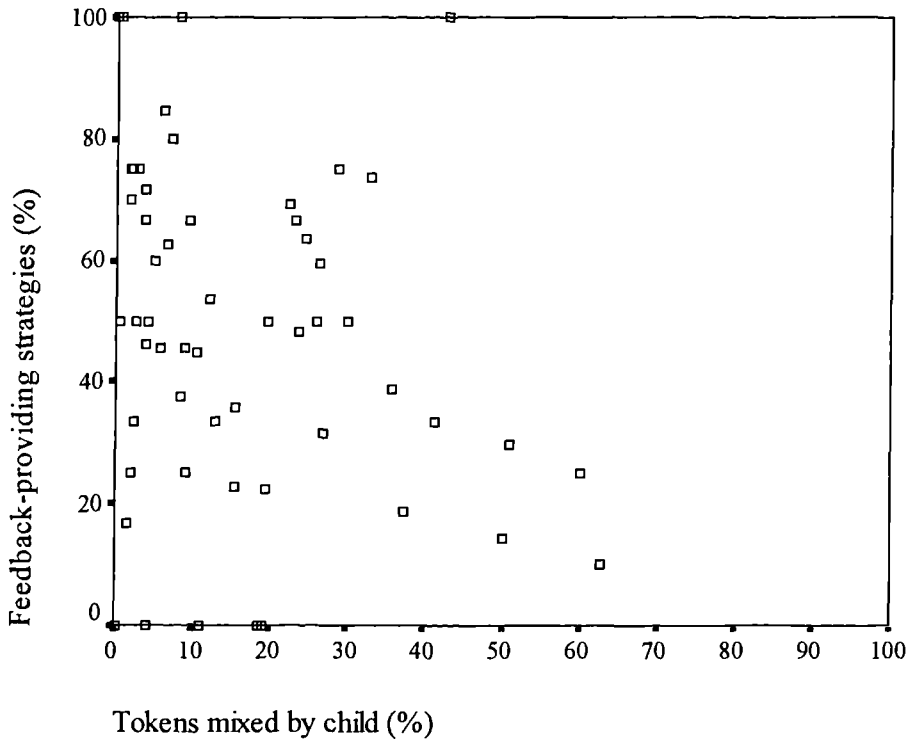


Figure 7.11 Percentage of tokens mixed by children as a function of parental strategies

Perhaps more importantly, the correlation between the parents' code-switching and the amount of mixing done by the child is significant. The correlation is .374 ($p<0.01$; N=58) between the percentage of tokens switched out of the total number of tokens produce in one recording episode from the child and the parent. For some dyads, that relationship is even more significant: for example, OLB and his mother ($r=.816$ significant at 0.01 level; N=22). The more the parent code-switches, the more the child does too. However, it could also be true in the other direction: the more the child mixes, the more the parent does too. Which one influences the other is not yet clear. From OLB's data, about half of all of his mother's switches were also initiated by her and were not a response to his. So, the other half is in response to his mixing. In this case, we have a 'vicious' circle where both interlocutors influence each other's behaviour. It is also possible that the mother's switching is influenced by her own environment. Indeed, she is the minority language speaker (French) and uses English on a daily basis with her husband and the community at large as they live in England. It is important to note that OLB's older sister did appear to switch languages to a large extent too.

Lanza (1997) had raised an important issue regarding directionality of mixing and parental strategies (see chapter 1). She claimed that children mixed more elements with the parent who was the most likely to accept them. A possible link between directionality of mixing and lexicon (chapter 6) was previously highlighted. Both OLB and ANN inserted more English items when using French. For both of them, French was indeed the minority language and that is probably one of the main factors. Minority language parents have to work harder at their child's maintenance of that language as it is not supported by their environment.

Table 7.5 OLB's and ANN's parents' use of parental strategies

	OLB's mother*	OLB's father	ANN's mother	ANN's father*
Minimal grasp	2.3%	5%	34.6%	10%
Expressed guess	8.2%	15%	19.2%	18%
Repetition	21.3%	25%	26.9%	28.7%
Move-on	31.1%	20%	19.2%	26.7%
Code-switching	37%	35%	0%	16.7%

*minority language speaker.

It appears from the data above (table 7.5) that both minority language parents used more non-feedback-providing strategies than the parent speaking the majority language. In addition to that both minority speakers (OLB's mother and ANN's father) used less response-eliciting strategies than their spouse. All in all, this certainly plays a role into the child's language mixing as well as the all important language exposure highlighted earlier. Consequently, one can hint that directionality of mixing is probably linked with parental strategies used. However, should the case of a family where the child mixes extensively despite parents using mainly feedback-providing or response-eliciting strategies be found, that hypothesis would be invalidated. No such case was found here. Yet, that does not mean it cannot exist.

7.6. DISCUSSION AND CONCLUSION

In this chapter, I have highlighted some of the possible reasons for the significant differences in lexicon sizes, number of cross-linguistic synonyms and amount of language mixing.

First of all, results show a significant difference between mean vocabulary sizes of boys and girls. Girls appear to have a larger vocabulary to boys and develop quicker. However, that difference does seem to fade away by 30 months old. This is consistent with some studies which reported a superiority of girls in early lexical development (e.g. Le Normand *et al.*, 2002). The reason behind this is not clear. Le Normand *et al.* (2002) interpret their results as a memory issue. They advance the possibility that girls might have a better declarative memory than boys. This part of the memory is responsible for the mental lexicon and the set of grammar rules children need to build utterances. The fact that the bilingual girls and boys in this study appear to be behaving in the same way as monolingual children is not in itself surprising. Bilingual children would then appear to have and use the same type of memory as monolingual children to process their lexicon and utterances regardless of whether they have one or two languages to deal with. Birth order did not provide any statistically significant differences.

Most importantly, parental input plays a crucial role into lexical development and the emergence of grammar. The development of the lexicon appears to be linked with the amount of exposure the child receives in each language. Even children exposed for a minimal amount of time to a language do acquire its vocabulary. In addition, the way parents talk to their child influences the size of their lexicon. That is consistent with studies of monolingual development which claimed that quality and quantity of input do matter in lexical acquisition (e.g. Hart & Risley, 1995).

The significant differences in terms of cross-linguistic synonyms in the bilingual lexicon have been explained by the richness of the vocabulary employed by the parents and also the language dominance of the child: the more balanced a child, the more likely s/he is to have a larger number of TEs. However, surprisingly, no evidence for parental strategies influencing the acquisition of TEs was observed. Döpke (1992) and Lanza (1997) had outlined the possibility that the more consistent a parent was in maintaining a One Parent- One Language strategy; the more proficient

the child was likely to be in that language. However, that proficiency does not appear to be reflected in that presence of TEs for the children studied here. A caution must be issued here as the language samples obtained for this study were relatively small. This issue will be dealt with further in the next chapter.

Finally, a very strong link between parental strategies and mixing was found. The two children who mixed the most in this study mixed in the minority language and did so because their parents did not use a majority of feedback-providing strategies. This confirms the hypothesis that the minority language speaker and his/her efforts to maintain that language are extremely important. On the issue of direction of mixing, those two children mixed mostly elements from the majority language into the minority one. That might be a sign of language dominance (as suggested by Lanza, 1997) but it is also probably partly due to the use of the parental strategies. Deuchar & Muntz's (2003) lack of correlations between parental strategies and child mixing on the one hand and parental mixing and child mixing on the other hand might have been an isolated case. However, it is important to bear in mind that although the correlation was significant, it was still relatively low and, looking at each child separately only provided significant results for some of them, not all.

Despite the evident influence of the parental strategies on the lexicon and mixing, the correlation with the TEs was not significant. Moreover, some of the correlations with mixing were not very strong. One issue which may be the cause of the lack of correlation here, is the coding of the parental strategies. This is an issue that was recently also raised by Deuchar & Muntz (2003). They claimed that it was difficult at times throughout the coding process to attribute one specific strategy to one category. A similar problem arose in the present analyses. Parents can use more than one strategy in reply to one single mixing from the child. Moreover, at times, the classification of the strategy is debatable. In some cases, parents repeat the word the child has just said with a rising intonation indicating a question as in example (3) below.

(3) MUM: Qu'est ce que c'est ça?
[What is this ?]
OLB: A SWING SWING.
MUM: SWING SWING?
OLB: Oui.
MUM: C'est la balançoire?
[Is it the swing ?]

Thus they do code-switch but also require a response from the child. Deuchar & Muntz (2003) had analysed the data by giving a score to each strategy (1 to 5). By choosing to classify the responses into feedback-providing or not, the aim was to minimise this issue of strategy coding. An analysis was also done using similar scoring than Nicoladis & Genesee (1998) and Deuchar & Muntz (2003) and no significant correlations were found. Lanza (2001) had mentioned that strategies were to be seen as a more qualitative way of measuring input. However, it appears difficult to do so without any kind of quantitative measures at all. A possible solution would be to classify the strategies into wider categories such as the one used in the present study: feedback or non-feedback providing.

This chapter has shown that numerous factors do influence lexical and language development in bilingual children. These factors are very similar to those influencing monolingual children. Nonetheless, the question of input in bilingual children is extremely important and it probably is one of the bases for a child becoming dominant in one language or for the presence of mixed elements in early utterances. This chapter highlights the importance of taking into account these factors when studying and working with bilingual children. Further research dealing with bilingual children does need to make a note of such factors in order for bilingual research to be more comprehensible and results to be comparable and verifiable. The next chapter deals further with the issue of parental strategies and its relationship with lexical development.

CHAPTER 8:

LANGUAGE 'STEERING' AND BILINGUAL IDENTITY

8.1. INTRODUCTION

It has been widely acknowledged in the literature that input and parental discourse structures have a role to play into the bilingual child's language development and especially into language mixing. In the larger study of 13 French-English bilingual children, I have looked at the effects of the environment and the input on the child's lexical development (see previous chapter). Moving on from a quantitative approach to parental discourse strategies and analysing them from a more qualitative point of view, a new parental strategy used by one family in particular emerged. In this chapter, I will describe that strategy and examine the relationship and its effect on the lexicon of the child and language mixing.

8.2. AIMS OF THIS CHAPTER

Most of these studies have shown the influence of parental strategies on language mixing. However, the aim of this chapter is, once again, to show that parental strategies influence both the lexicon and language mixing using a more qualitative approach to the issue of parental strategies. The aims for this chapter are therefore twofold: firstly, although no evidence was found in the previous chapter for a correlation between TEs and parental strategies, I hope to show that this type of strategies can encourage the acquisition of cross-linguistic equivalents. Secondly, I also wish to argue that negotiating a bilingual context might not be as negative as had been outlined by earlier studies which claimed that, by encouraging a bilingual context, the child would be less likely use the appropriate language.

8.3. THE PARENTAL STRATEGY HYPOTHESIS

We have previously outlined the role of parental strategies as found in the literature (for more details see chapter 1). Döpke (1992) and Lanza (1997) have reported that parents who used more child centred speech and were more consistent in language choice were more successful at increasing their child's proficiency in that language they use with them. Juan-Garau & Perez-Vidal (2001) examined the parental

strategies used by the family of a Catalan-English bilingual child. They claimed that the English father's use of strategies negotiating a more monolingual context favoured the child's use of English (the minority language in that case). It consequently required a decline in the amount of mixing form the child's part. However, a more recent study by Deuchar & Muntz (2003) did not find any significant correlation between M's mixing and the parental strategies used. Neither did they demonstrate any correlation between the degree of the parents' mixing and that of the child. They explained that these lacks of correlations were due to two main factors: the inadequacy of the coding of parental strategies and the unusual situations of some of the recordings.

8.4. DESCRIPTION OF THE LANGUAGE STEERING STRATEGY

Most of the families in this study used the five strategies as defined by Lanza (1997), although it was sometimes difficult to actually classify each use (see previous chapter). Families used these strategies in very different ways. However, when studying the 13 French-English bilingual children, a different strategy used by one family mainly was found, which, to my knowledge, had not appeared before in the literature as such (illustrated by example 1). The strategy was not present in any of the other 12 families' discourse strategies.

(1) FLO, 2;2.22, English

FLO: Look.

FLO: There.

DAD: Daddy says look and maman says ?

FLO: Put it there.

DAD: Maman says ?

FLO: REGARDE.

[LOOK.]

DAD: REGARDE.

[LOOK.]

DAD: Good.

The family in question in this chapter is that of FLO. It is worth reminding the reader that the child lives in England with her French-speaking mother and her English-speaking father. The parents obey the OPOL rule. The family language is French, meaning that the mother and father speak French together when not addressing the children directly. Both parents are fluent in each other's language. FLO's younger sister was born during the study when she was 21 months old. FLO started going to nursery when she was 15 months old for two afternoons a week at first. Although the nursery is an English-speaking setting, many of FLO's friends there are actually bilinguals. My first impression of the family was that they are very 'language- and literacy-oriented' generally speaking. They place a great emphasis on reading, talking, and languages in general. They also have a large number of friends and acquaintances from different linguistic backgrounds. Of possible relevance, here, is the fact that FLO's father is a language teacher (French, Spanish, and German).

The strategy highlighted is what I will call *language steering* or *steering strategy.* Both FLO's father and, to a lesser extent, her mother used that strategy. Note the following exchange between FLO and her English- speaking father:

(2) FLO 2;2.22, English
DAD: And what is that?
FLO: A frog.
DAD: Good.
DAD: And maman says?
FLO: GRENOUILLE.
[FROG.]
DAD: Good girl.

The strategy consists in encouraging the child to providing the adult with translation equivalents of words she does already know in the 'right' language. In other words, it 'steers' the child towards the right word or words in the right language. It also shows the child the fact that two labels exist for one single concept. Döpke (1992) uses the expression "teaching techniques" as a mean of describing corrections and elaborations made by the parents of the bilingual children in her study. Here, it is a response-eliciting strategy in the sense that the child has to make an effort to provide an

adequate translation for a word s/he already knows. A content response is expected from the child, similar to the *minimal grasp from Lanza (1992), and not a simple* "yes" or "no" as with the expressed guess strategy. However, the conversation will not stop if the child does not know the word, the parent is simply adding to the child's knowledge. The right word is not necessary to the on-going conversation (as in (3)).

(3) FLO 2;2.22, English

DAD: Can you give me another one?
DAD: Please.
FLO: Please daddy.
DAD: And daddy says please.
DAD: And maman says +/.
FLO: Please.
FLO: Oh.
DAD: What's that?
DAD: That's daddy's alarm ringing.

In some rare cases, the father does not require a response from FLO and simply provides her with the right word as in (4).

(4) FLO 1;10.17, English

FLO: Daddy TIENS.
[Daddy HERE.]
DAD: Thank you.
DAD: TIENS.
[HERE.]
DAD: Say TIENS to maman.
[Say HERE to maman.]
DAD: Daddy says here.
DAD: Thank you.
DAD: Here.
DAD: Here you are.

If classifying the strategy in terms of bilingual or monolingual context (following Lanza, 1997), one can claim that the parent is initiating a real bilingual context. S/He does clearly show that both words are right and that s/he understands both languages. Even if the child has to make a content response as in the minimal grasp strategy, the parent, here, does not pretend to be monolingual. However, as opposed to "moving-on" or "code-switching" (Lanza, 1997), the parent sets the boundaries of each word use very clearly by explaining that one is in one language and the other in another language. Or in terms, in which younger children can understand: one is to be used by or with the mother and the other one by or with the father. So, it still compels the child to adhere to the use of one language in the given context. Contributing to the bilingual context is also the fact that the child is required to code-switch herself and to show her bilingual capabilities.

(5) FLO 1;10.22, English
DAD: What's that?
FLO: ROUGE.
[RED.]
DAD: Maman says ROUGE.
[Maman says *RED*].
DAD: Daddy says red.
FLO: ROUGE.
[RED.]

It is very important at this point to mention, if it is not already obvious from the examples given, that the strategy currently being described is not necessarily a response to the child using the wrong language with the wrong person (as in example (5)) or to the child code-switching. It is merely an additional tool for parents to make her aware of her bilingual identity and teach her vocabulary or grammar. It could be argued that example (5) was rather a case of the father using the repetition strategy (Lanza, 1992).Or maybe a double strategy using the adult repetition and the code-switching strategies. However, one could also argue that as opposed to a repetition, the father here not only gives the child the right word but also explains to the child that the wrong word (in this context) is the one used by her mother. And as a benefit

to the traditional code-switching strategy, the adult provides the child with the right word.

Lanza (1997) acknowledged something very similar used by the parents of Siri and especially her father. She gives the following example between Siri and her father:

(6)

DAD: Har du vært noen andre steder i dag? Vært på butikken? Hm?
[Have you been anywhere else today? Been to the store?]
SIRI: Ja.
[Yes.]
DAD: Ja? Kjørte bil?
[Yes? Drove the car?]
SIRI: CARRIAGE!
DAD: Kjørte du CARRIAGE?
[Did you drive a CARRIAGE?]
SIRI: Ja.
[Yes.]
DAD: Barnevogn.
[Carriage.]
SIRI: Ja.
[Yes.]
DAD: Ja.
[Yes.]
Adapted from Lanza (1997: 297-298)

In this example, however, the child is not required to give a full response. The father does not request one from her. A mere "yes" is enough. Lanza (1997) explains that those "language-teaching episodes" are mostly found in triadic interactions where both of Siri's parents are present as in the example below.

(7) DAD: Skal du ta av deg smekken først kanskje?
[Are you going to take off your bib first maybe?]
SIRI: Skitten, skitten.
[Dirty, dirty.]
DAD: Er den skitten?
[Is it dirty?]
SIRI: Ja.
[Yes.]
MUM: What does Mama say for 'skitten'?
SIRI: Dirty
MUM: Dirty. Good!
Adapted from Lanza (1997: 297)

In these cases, no code-switching is required from the child's or the parent's part since both parents are present and the child is addressing herself to the mother in English and to her father in Norwegian. In the present case, the parent is purposefully asking the child to code-switch when only one of them is present.
As a result, it seems as though the child is aware of the two labels for the same word, or so it seems as in example (8).

(8) FLO, 2;0.11, English
DAD: Yes, it is a monkey with a dummy in its mouth.
FLO: Dummy.
FLO: TÉTINE.
[DUMMY.]
DAD: Maman says *TÉTINE*.
[Maman says DUMMY.]
DAD: Daddy says dummy.

The father clearly explains that *tétine* has to be used with the mother and dummy with the father. Consequently, it wouldn't be appropriate to use dummy with her mother or *tétine* with her father. It also conveys to the child that it is right to use that word but not in this context. In this example, we can see that the child knows the two words for

the same single concept and acknowledges that by herself. Consider also the following example.

(9) FLO 1;8.7, English

DAD: Do you want some more?

FLO: OUI.

[YES.]

FLO: Yes .

DAD: Is that yes or OUI?

[Is that yes or YES?]

FLO: Yes.

DAD: Yes or OUI?

[Yes or YES?]

FLO: Yes.

In this utterance, the father again, conveys to the child the message that there are two different words for one concept. He asks the child to make the choice between the two words. He even repeats the question asking her to choose as if to check she is sure of herself. The child seems to be aware of which one is the right word as she chooses the word *yes* not once but twice correctly.

The technique or strategy is used not only for single lexical items but also for more complex structures:

(10) FLO 2;2.22, English

FLO: C'EST LA.

[IT'S THERE.]

DAD: Maman says C'EST LÀ.

[Maman says IT'S THERE.]

DAD: Daddy says it's there.

FLO: It's here.

This steering strategy is mostly used by the father. The fact that he is a language teacher might have has influenced his use of this strategy. He claims that it was not a deliberate technique and that he did not realise its use until mentioned to him.

However, he also underlines the fact that he is a language teacher and so that must have an affect on his behaviour with his daughter even if not voluntary. He also claims that both him and his wife are language learners themselves and so realise the benefits of such techniques.

As we can see the description of this discourse strategy is rather a complex one and this strategy can take very different forms. It is described as a *steering technique* as the father steers the child into the right path or the right language to be used while acknowledging the alternative. This is used by this family to enrich the child's linguistic abilities and her bilingual identity as well as that of her parents. The strategy was used during the whole period of study but with increased frequency as FLO's linguistic abilities improve.

Table 8.1 MLU and use of the steering strategy in FLO's data

Age	MLU	Steering strategy
1;7.9	1.476	0
1;8.7	1.276	1
1;8.10	1.304	1
1;10.17	1.434	2
1;10.22	1.316	4
2;0.11	1.722	2
2;2.22	1.817	10

8.5. APPARENT EFFECTS OF THE STRATEGY

The family completed all the CDIs from 16 to 30 months without exception. The parents also kept a very concise summary of their child's first few words. Relatively few hours of audio-recordings were obtained: 270 minutes between 1;7.9 and 2;2.22 in three different language contexts: both parents separately and both parents together.

8.5.1. Language steering and lexical development

The visible results on FLO's lexical development are quite significant. In the larger study, FLO was one of the girls with the largest lexicon from very early on in the study. Table 8.2 shows her lexical data based on the CDIs. The scores given are the native parent's scores only. FLO's lexicon is balanced throughout the study period. Each language overtakes the other one on a certain number of occasions, however; the difference between the two languages is minimal compared to other children in the study. Of course, the fact that her input was balanced certainly plays a non-negligible role in that also. Table 8.2 also shows the number of TEs present in her lexicon and the percentage they represent in her total vocabulary. The mean percentage of TEs in her total lexicon is 52.96% over the 14 months of study while the average is 35.94%.

Table 8.2 FLO's lexical data

Age	BCDI	FCDI	**Total**	Number of TEs	***% of* TEs in total CDI**
16	28	20	**48**	14	**29.17**
17	38	21	**59**	14	**23.73**
18	22	34	**56**	18	**32.14**
19	30	41	**71**	26	**36.62**
20	45	46	**91**	42	**46.15**
21	88	50	**138**	48	**34.78**
22	137	171	**308**	159	**51.62**
23	197	183	**380**	211	**55.53**
24	214	182	**396**	241	**60.86**
25	244	319	**563**	370	**65.72**
26	281	303	**584**	390	**66.78**
27	300	430	**730**	494	**67.67**
28	347	446	**793**	553	**69.74**
29	403	509	**912**	657	**72.04**
30	487	506	**993**	813	**81.87**

Further illustration can be obtained by the two figures below showing FLO's lexicon in comparison with the average scores for the 13 children (figure 8.1). The figure shows that FLO's lexicon is consistently and significantly above average and on several occasions the higher score of the 13 children in the study.

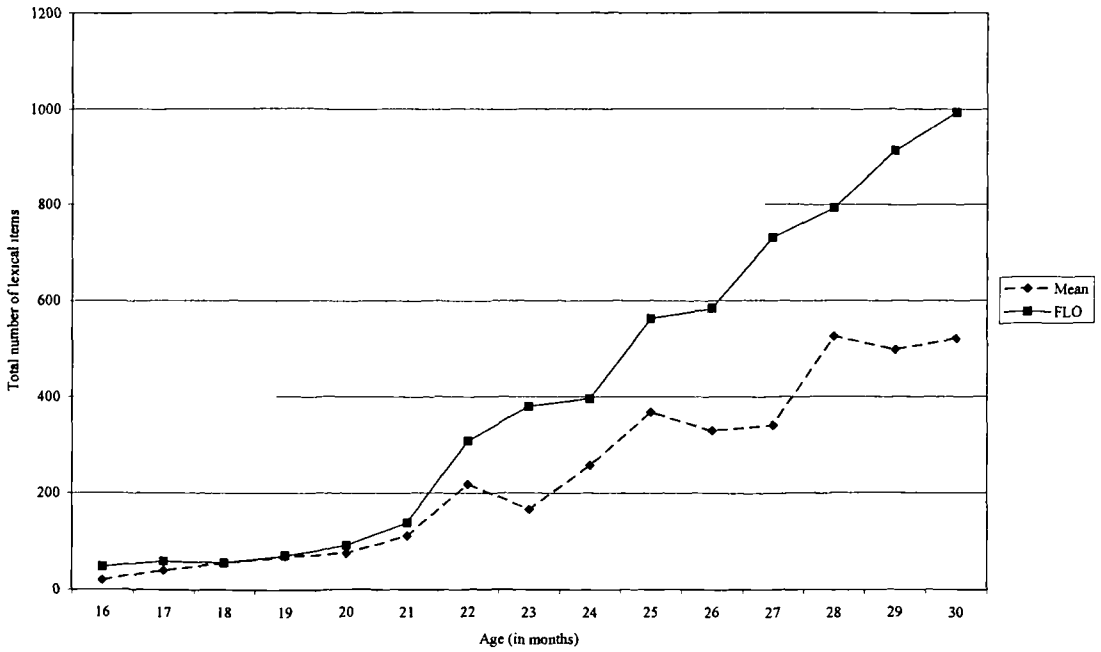


Figure 8.1 FLO's lexical development compared with the study's means.

In addition, the percentage of cross-linguistic synonyms acquired by FLO is also consistently above average (with one exception at 17 months old) and on several occasions the maximum of all of the children.

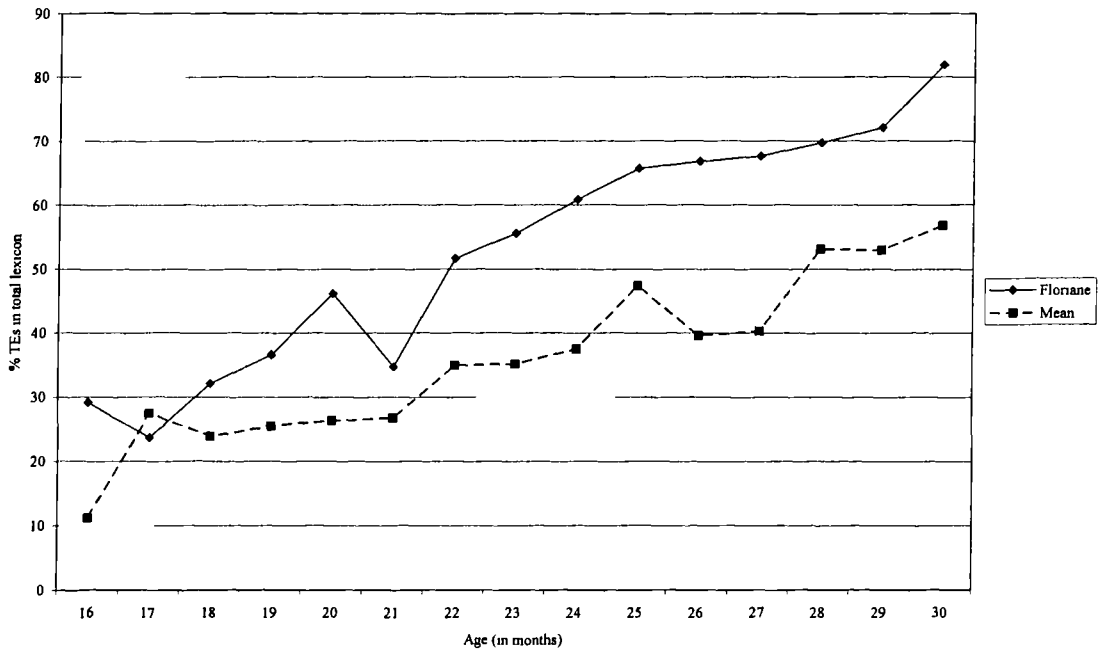


Figure 8.2 FLO's translation equivalents compared with the study's means.

8.5.2. Language steering and language mixing

One other fact worth mentioning is that FLO only rarely mixed compared to other children in the study with a large lexicon. FLO rarely used elements of two languages in the same sentence and rarely used the wrong word in a given language context. Compare her amount of wrong language choice with other children's (in figure 8.3). Overall, FLO mixed languages to a lesser extent than ANN and OLB. Both of these children had a very similar lexical development to FLO (very large number of TEs and large balanced lexicon). The only factor differentiating them is the fact that FLO did rarely mix languages. The unusually large number of mixes for FLO at 21 months can be explained by the fact that the data is based on only one recording where she only produced about 10 tokens and mixed about 4 of them with her mother.

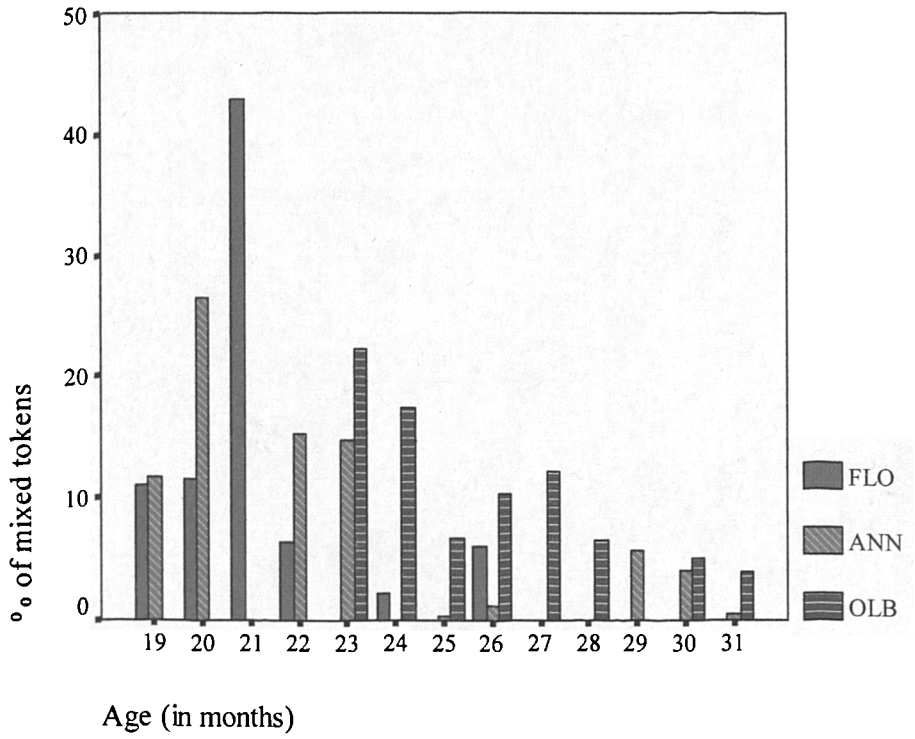


Figure 8.3 FLO's, ANN's and OLB's mixed tokens by age

8.6. DISCUSSION AND CONCLUSION

The single case study of FLO and her parents' use of a steering strategy have shown that translation equivalents and parental strategies could be linked despite the lack of correlation in the previous chapter. Secondly, although this steering strategy is clearly negotiating a bilingual context as the father shows his understanding of French and even his abilities to use it, that does not appear to translate into the child mixing languages more than any other child, quite the opposite. Caution must be issued to the reader here. No claim is made here that the fact that FLO had a large lexicon and a large number of TEs in her lexicon are solely due to her father using that strategy, especially since he started using it more consistently towards the end of the study period. However, the fact that the family as a whole unit did emphasise the bilingual identity of the child by means of that strategy or any other means that have not been studied but might have been at their disposal might have played a very important role on the child's growing lexicon. In addition, by emphasising that bilingual identity, they showed her that one concept could be represented by two different words in two different languages very early on to the child. Although no significant correlation was

found in the previous chapter between parental strategies and TEs, evidence from this child does appear to show that there could be a link. This needs to be further investigated by means of qualitative rather than quantitative methods or a combination of both as quantitative methods do not appear to relate the special features of each strategy used by parents. The suggestion by Lanza (2001) that parental strategies are not to be analysed quantitatively but qualitatively appears from this chapter to be founded. One other possibility would be to provide a different classification system to be able to analyse them quantitatively. Further research will need to elaborate on this matter. This chapter also highlights the issue arising when coding parental strategies. Counting most of FLO's father's switches as such would have meant scoring 5s for each of them according to Nicoladis & Genesee's classification (1998). His total score would have therefore been close to those of parents who do code-switch regularly or do not provide any alternative to their child. That would have provided a wrong overview of the fact that this score is supposed to represent the acceptance of the child's use of the inappropriate language. FLO's father very strongly explains that each word should be used in the right context. Both are right but they are context-dependent. Therefore, we can see how the scoring of parental strategies as such might cause problems. The strategy used by FLO's father was thus classified as a feedback providing strategy as in most cases he requires or provides some kind of information/explanation from the child or gives it himself.

To come back to the differences between FLO, ANN & OLB, the major difference lies, from what has been observed from the families, in the attitude of the parents towards mixing and bilingualism in general. As previously explained, FLO's father and mother encouraged FLO to show off her bilingual abilities. She was made aware from an early stage of the positive aspects of bilingualism and encouraged to give two words for one concept as outlined by the strategy used by her father. FLO lives in a very multicultural environment and frequents a local nursery where she has bilingual French-English friends as well as friends from other language pairs. This can be further illustrated by the fact that shortly after the end of the study; FLO's parents introduced German as an extra language. I do not mean, here, that ANN's and OLB's parents did not encourage their child's bilingualism. However, their attitude towards it and especially towards mixing was different. ANN's father who is the minority speaker accepted from very early on that ANN would eventually become dominant in

English and admitted switching at night when tired and talking to his children. Both of ANN's parents felt that English was going to win over at some point but that the basis they had given her would be a good one for the future as illustrated by this quote from her mother: "ANN's French has not developed for a long time - she is at an English speaking nursery now and she makes too many mistakes to correct. It's a pity, but we hope it is a basis to learn later." (December 2003). A similar view exists in OLB's family. Although OLB and his sister attend a French playgroup where they meet other French-speaking children living in England, OLB's mother is only too keen to accept mixes from OLB as seen in previous chapters. In contrast the view in FLO's family is quite different. In response to my mentioning of the steering strategy once the study was finished here is what her father replied: "I am not making conscious decisions to correct FLO by using the right word...but I am a teacher by profession and I don't think you can avoid doing the same in your private life, especially when you are a language learner yourself [...] and your wife is also a language learner [...]. It's more a way of life. FLO is bilingual and very confident in both languages. [...] We are involved with the local French community and there is a Saturday school with 6 classes for bilingual children. Each case is slightly different, but I think we have got the right balance in terms of aiming towards bilingualism." (January 2004). This is very much in-line with what had been proposed by De Houwer (1999b) claiming that parental beliefs can influence children's linguistic development through their linguistic choices and strategies. Future studies could investigate the impact of these strategies on the degree of bilingualism of the children, for example.

The emphasis made in this chapter has been that parental discourse strategies are not limited to the ones often reported in the literature and that there are many ways parents respond to their child's mixing or are able to emphasise their bilingual identity. As seen in the previous chapter, parental strategies, and perhaps more generally their attitudes towards bilingualism and the bilingual identity of the child, play a significant, non-negligible role in the language acquisition process and can be (partly) responsible for the size or rather the composition of a child's bilingual lexicon and the use they make of it. Negotiating a bilingual context can prove very positive for the child.

SECTION V:

DISCUSSION AND

CONCLUSION

CHAPTER 9:

GENERAL DISCUSSION & CONCLUSION

9.1. INTRODUCTION

In chapter 1, four main aims and objectives were set out in order to investigate the relationship between the two languages of the developing system of a bilingual child. Thus, this thesis has provided an insight into the developing bilingual lexicon and its impact on the emergence of grammar (especially language mixing). This chapter will summarise the factual, theoretical and methodological contributions this thesis brings to the field of bilingual language acquisition. First, I review the main findings. Secondly, this chapter will present a general discussion of the major results obtained in the present piece of research in relation to the wider topics that had been outlined in chapter 1 and 2, i.e. mainly the debate on the differences between monolingual and bilingual language acquisition. Thirdly, I briefly discuss methodological issues by establishing the limitations and issues raised by the study (section 9.4) and suggest some further research (section 9.5). Finally, section 9.6 presents a general conclusion for the present study.

9.2. MAIN FINDINGS

The first aim in this study was to provide an accurate pattern of development for bilingual lexical acquisition. The evidence suggests that the normal developmental stages of a bilingual child's lexicon follow the four stages: Routines and Word Games, Reference, Predication and Grammar. Regardless of the differences in exposure or the differences in rate of acquisition, these four stages were observed for all of the 13 children studied. Despite, overall developing according to the four stages established for monolingual children, the two languages of these bilinguals showed cross-linguistic differences. These differences were similar to those observed for monolingual children acquiring French or English. For example, it was shown that predicates were acquired earlier in French than in English. Chapter 4 of this study asked how early bilingual children were able to use cross-linguistic synonyms. It was established that even in lexicons smaller than 10 words, some bilinguals had translation equivalents. The evidence suggested that an overall mean of about 30% of a child's lexicon was composed of synonyms although great variability was observed.

Most results in chapters 3 & 4 were based on the parental checklists. Therefore, in chapter 5, I checked for the validity and reliability of the results asking whether parents were accurate raters. Results showed that mothers tended to be more reliable than fathers especially when fathers were rating the language they did not use with the child (i.e. their non-native language).

Chapter 6 examined another aspect of bilingual language acquisition. The question was asked to what extent the emergence of syntax, especially language mixing, was influenced by lexical development. The results showed that, although, not all of the children mixed, most of the mixing was not due to the lexicon (or to lexical gaps). However, as their grammatical abilities increased, children mixed less. As to what extent different words were mixed, both nouns and closed-class items were mixed by the children, some preferring the first category, others mixing both. Also in that chapter, the critical mass hypothesis from Locke (1997) was evaluated on the bilingual population. The present results complement previous ones (e.g. Petitto *et al.*, 2001) thus providing a significant number of children acquiring different language combinations to form the basis for establishing norms of bilingual lexical development. Based on these, one could suggest that bilingual children do reach the 50-word stage by around 1;8 ranging from 1;4 to 2;1 for the children described in this study.

Bilingual children in this study did show appreciable variations in their lexical development when comparing both languages of one child and when comparing two children together. It has been explained throughout this thesis (and especially in chapters 7 and 8) that one of the main explanations for these variations is to be found in the quantity and quality of the input received by the children. A child, who receives less exposure to one language, will acquire the vocabulary in that language at a slower pace than others who are exposed to it in larger quantities. However, even with limited exposure to one language, these bilingual children were able to acquire its vocabulary, perhaps with more difficulty. Furthermore, parents who are more consistent in their language choices and code-switch less will influence the language choices of their child. These findings are predictable on the assumption that vocabulary is based on learning, as opposed to syntax, for example, which is based on triggering (see chapter 1 and Paradis & Genesee, 1996). The variability in the age of emergence of two-word utterances (which is one of the first stages of grammar with

the acquisition of relevant word categories) is due to the variability in lexical acquisition.

These were the main results obtained from this longitudinal group study of bilingual lexical acquisition. Having reviewed the main facts and results of this piece of research, we will now turn to the more theoretical contributions this thesis brings to the field of bilingual language acquisition.

9.3. MONOLINGUAL VS. BILINGUAL LANGUAGE ACQUISITION

In chapter 1, the debate about bilinguals having one or two systems was explored. However, due to several inherent issues linked with it, the focus was put on a new, maybe more interesting issue. The new emerging debate in the field of bilingual language acquisition, moving away from the one-vs.-two-system dichotomy, focuses on whether bilingual language acquisition resembles monolingual language acquisition in any shape or form. This debate is largely based on the assumption that bilinguals are born with two distinct systems and that monolingualism is not the default setting. Researchers have focused on different aspects of language and tried to determine similarities and differences between monolingual and bilingual. This new focus is partly based on, and motivated by, the common assumption made by many parents and professionals that bilinguals lag behind monolinguals in terms of language development (and lexical acquisition in particular). A large number of recent studies have focused on cross-linguistic influences especially in terms of syntactic structures (e.g. Döpke, 2000; Hulk & Müller, 2000) but none have looked at it from the lexicon's point of view. The debate focuses on whether language development in bilinguals is autonomous (i.e. the two languages develop separately or as a monolingual's) or interdependent (i.e. the two languages influence each other and bilingual development is different from monolingual; see chapter 1 for more details). The present study fits into this debate of monolingual versus bilingual language acquisition bringing new evidence based on lexical development. Lexical data from bilingual children does appear to indicate that they do resemble monolingual children on a certain number of points. First of all, the order and overall rate of acquisition of the different lexical categories, as well as the overall size of the lexicon is very alike

between monolingual and bilingual children, based on results from this study. Thus the present results have shown that bilingual lexical development appears to be autonomous. Secondly, cross-linguistic differences found for monolinguals between English and French (i.e. predicates and closed-class items; Parisse & Le Normand, 2001) were also found in the corresponding lexicon of the bilingual child. The evidence suggests here that there is little or no cross-linguistic influence in terms of the proportion of the lexical categories in the bilingual lexicon. One clear example of that is the case of the predicates. The bilingual children did produce a larger number of predicates in French than in English. If one had argued that bilinguals were developing differently from monolinguals, then one would have expected the children to produce a larger amount of predicates in English when compared with monolinguals due to the influence of French. However, no such phenomenon has been observed. So, the fact that there are slight differences between the French and English lexicons of these bilinguals would appear to show that the two languages do develop independently of each other in relation to the language-specific patterns of language acquisition. Similarly to Paradis & Genesee (1996) for syntax, the present results allow the claim that bilingual children do develop their two languages autonomously. Subsequently, it was shown that the rate of acquisition was similar to monolinguals. Significant milestones such as the 50-word stage, first word combinations fell within the range of variation observed for monolingual children (see chapter 6). The present findings support the claim that bilingual children are not consistently slower to develop than monolinguals (Pearson *et al.*, 1993; Petitto *et al.*, 2001; Holowka *et al.*, 2002). Considering that bilinguals receive less input than monolinguals in each language that is quite a remarkable achievement. Furthermore, regardless of the amount of input received in each language, the children always developed in the same recurring order: social words, common nouns, predicates and closed-class items. Even a child with significantly more exposure to one language (e.g. see ANT in chapter 3) followed that developmental pattern.

However, caution must be taken in claiming that the bilingual lexical development is autonomous. First of all, this might be true for French-English bilinguals but might not for other language pairs. Indeed, we have seen from other studies on other language pairs that the influences between languages in syntactic acquisition, for example, vary to a large extent (e.g. Yip & Matthews, 2000). Secondly, while the

lexicons of a bilingual child might develop autonomously with little cross-linguistic influence that does not imply that other aspects of language might behave in a similar manner. Genesee (2003) argued that different aspects of language could develop autonomously while other interdependently. However, both syntax and lexicon do appear to develop autonomously based on the present results.

In addition, claiming that the two languages are autonomous does not automatically include the claim that the bilingual child is simply two monolinguals in one. The bilingual child does not have a switch that s/he can turn on or off depending on which type of monolingual s/he wants to be (Grosjean, 1999). The other language can never be totally ignored. It remains there even if it is not active. Even though the two languages develop in a similar way, the bilingual child has additional aspects to manage. S/He has to acquire other abilities and capabilities that the monolingual child does not need to acquire. Chapter 6 provided evidence for the fact that the language mixing was linked with lexical development, to a certain extent. Language mixing is one example of an ability the bilingual child has to master. The processing and correct use of translation equivalents, in accordance to the relevant pragmatic rules is another example of additional abilities the bilingual child needs to learn. For a bilingual child, matching a lexical item with the interlocutor is one extra task to perform. Another example of differences raised is the effect of cross-linguistic synonyms on the overall size of the lexicon. I mentioned previously that the total conceptual vocabulary of the French-English bilinguals was slightly inferior to monolingual norms. This suggests that there are more aspects than just quantity to take into account with the developing bilingual lexicon (i.e. quality) and that the bilingual child is not simply two monolinguals in one. This is a further issue to be examined in more details.

Although by and large, bilingual lexical development is very similar to monolingual, one can wonder why many parents and professionals still have the impression that bilingual children lag behind in terms of acquisition rate when compared with monolinguals. Some of the answers can be found in this research. First of all, differences between vocabularies of different children are susceptible to environmental factors. A child who is mainly exposed to French, for example, by spending all day in a French-speaking nursery will acquire more French words than English ones. Parents might then worry that the child has a clear imbalance and appears not to acquire English as quickly. However, as mentioned above, the child

will still achieve the main linguistic milestones (e.g. 50-word stage) at roughly the same age as monolinguals and balanced bilinguals. One other key factor for the impression that bilinguals do not develop as quickly as monolinguals is due to the nature of the lexicon itself. As mentioned throughout the present research, bilinguals produce some words which cannot be attributed to any of the two languages and which are ambiguous. These words are often not recognised as such by parents and therefore not taken into account in assessing the child's abilities (Petitto *et al.*, 2001). This is especially true in the case of two languages such as English and French that are typologically and phonologically similar in many respects. The fact that the TCV appears, in this case, to be slightly inferior to the total lexicon of monolingual children, might be another reason for that impression. Therefore, although it is widely believed that bilinguals do develop slower than monolinguals, lexical acquisition data does not support this view.

9.4. LIMITATIONS OF THE STUDY

After having reviewed the factual and theoretical aspects of this thesis, and by examining the limitations of this study, we shall now focus on the methodological contribution this thesis makes to the field. The present study raised a number of practical issues and questions relating to the use of parental checklists as the main method of data collection. Although the CDIs allowed for a systematic and regular assessment of the children's lexical development, that tool is not without drawbacks. We have already outlined the differences between parents in chapter 6. In addition to that, despite their validity and reliability tests proving successful, the CDIs used here proved to be more unreliable than expected. This is especially true of the Standard French adaptation of the CDI. At the start of the present study the FCDI was still under development. However, it is my impression that the FCDI is (or was) too close to a literal translation of the original American form rather than an adaptation as is recommended. The form does not reflect the nature and structure of the French language accurately. Examples have been given throughout the previous results chapters.

Furthermore, using the CDIs with children bilingual in two similar languages, as French and English are phonetically, did cause a problem when matching lexical

items for a cross-linguistic equivalents analysis. Cognates or items that are pronounced similarly in French and English (especially by a child) have been excluded from some of the analyses. However, it would have been interesting to examine the possibility that cognates trigger the appearance of synonyms. The hypothesis would have been that there would be more cognates with synonyms than other words with synonyms at least in the first few stages of lexical development. It was unfortunately not possible to test that hypothesis as it appeared that parents tended to tick both words when they were unsure to which language they should attribute the word. Consequently, although the CDI is one of the only ways of obtaining systematic developmental lexical data, that could have been complemented by recordings of the children starting at an earlier time than what had been originally planned. Another solution could have been psycholinguistic experiments (practicality issues would have then arisen due to the subjects' country of residence).

Finally, one last issue raised by using the CDIs as the basis of the data analyses is that of categorisation. As it has been mentioned previously, the CDIs are very good lexical assessment tools, however, they do categorise words in a manner that might not be appropriate for all types of analyses. For example, in chapter 3 was mentioned the fact that possessive pronouns were all put under one item in the French forms regardless of gender or number. Therefore, should the focus of this study have been pronouns, the tool would have given flawed results for the French part of the acquisition at least. Categorisation is an important issue with all CDIs and not just the French adaptation. We have seen that in balance it is a better tool than any other as the study dealt with a group and it allowed for meaningful comparisons. However, collecting only naturalistic data would have been ideal. That would have eliminated all the limitations due to the CDI. Yet, using only naturalistic data would not have allowed comparisons between the children's developmental patterns. It would have been extremely difficult to make sure that all of the children used all of the words they knew at any given point in time, for example. Naturalistic data would have represented what they used rather than what they knew. In addition, from a semantic and pragmatic point of view, the data collected in this study provides the basis for further similar studies that will be able to compare results.

Other limitations include participants' involvement in the study. Most of the families did not provide as much as information as was originally planned (e.g. recordings). However, that is an issue that arises with most longitudinal studies. Despite these

drawbacks, the group longitudinal study is one of the innovations of this study and brings a major contribution to the field by allowing for a larger amount of data to be collected, thus providing grounds for comparisons and further developments.

9.5. SUGGESTIONS FOR FURTHER RESEARCH

Several suggestions for further research can be drawn from the present study regarding study design, population choices for bilingual lexical acquisition studies and further topics related to the present one.

One of the major innovations and originalities of this study is its design. The longitudinal group study approach is certainly not the most common method used in bilingual language acquisition research. One of the main issues, previously outlined, in single case studies is the fact that they make comparisons more complex as each is unique in one way or another. Longitudinal group studies are quite rare in the field. However, such studies contribute to larger amounts of quantitative and qualitative data from children with similar backgrounds. Thus, they would allow researchers to be able to compare results on different populations, with different language pairs or different acquisition settings. The multi-disciplinary field of bilingual acquisition is in need of more longitudinal group studies of this kind as Genesee (2004) pointed out. Despite the limitations, the data collection method will allow future research to be compared with the present one.

The second aspect of further research that could be suggested from the actual study relates to the population(s) to be studied. Most of the recent research on bilingual language development has focused on middle-class families using the one parent-one language strategy. Potential further research on childhood bilingualism needs to enlarge the type of children being studied. Firstly, the role of different contexts of acquisition in bilingual language development has rarely been raised. In chapter 1, I highlighted the fact that all of the children studied here would belong to what Romaine called Type 1 of her bilingual family typology (1995). Above, it was outlined that the children in this study had relatively balanced lexicons as they were exposed to a relatively balanced amount of each language. Longitudinal group studies of the developing lexicon of children being brought-up bilingually but not necessarily in a One-Person-One Language family context would be interesting to be compared to

the present one. The results could prove different if one studied families where both parents used the same language at home and the child acquired his/her second language only when outside home. In strict One Person- One Language contexts, the outcome tends to be more balanced, at least lexically speaking. However, children being brought up in different contexts of acquisition (e.g.: where both parents code-switch) might develop an unseen scenario not previously envisaged. Namely, they might acquire only system or only content morphemes in one language. Such studies should examine the role of the environment and the input and look at the relationship between process and outcome in bilingual language acquisition. To the best of my knowledge, no study has looked at that possibility yet.

In addition to different settings of acquisition, different language pairs should be studied in bilingual lexical acquisition. A comparative analysis of bilingual acquisition involving different language pairs would be needed. French and English are relatively close languages when it comes to lexical and grammatical structures. And all of the aspects of lexical acquisition highlighted in this piece of research would need to be confirmed by other bilinguals learning typologically different language pairs (e.g.: Chinese-English, Arabic-French which). Furthermore, a certain number of issues were raised relating to the fact that French and English were phonetically too similar languages (e.g. cross-linguistic synonyms). These could be re-examined on the basis of different language pairs.

9.6. CONCLUSION

The present innovative research provided an account of early lexical development by 13 French-English bilingual children. The study has outlined four main conclusions.

- The normal developmental stages of a bilingual child are similar to that of a monolingual. Bilingual children's lexical categories in each language develop in a parallel manner whether or not the children are dominant in a language. However, crucially, that does not mean that the two languages do evolve at the same speed. Cross-linguistic influences between French and English revealed that bilingual children acquire French predicates earlier than English ones. That is in line with monolingual research.

- Bilingual children do have translation equivalents before the 50-word stage in their lexicon, which seems to contradict the Principle of Contrast. However, the idea that translation equivalents across languages are not the same as synonyms within a language is brought forward. Therefore, comparisons between the two should be avoided. Cross-linguistic synonyms are one aspect in which bilinguals differ from monolinguals.
- Thirdly, naturalistic production data combined with systematic lexical reports suggest that the bilingual child also falls within monolingual norms when it comes to early syntactic milestones (e.g. first word combination stage). Moreover, as soon as the child is able to combine languages, s/he is also able to mix language, whatever the composition of his/her lexicon. Bilingual children do appear to need 50-words in each of their two languages in order to be able to produce multi-word utterances in that language. No evidence is found for a lexical gap hypothesis.
- Finally, as far as the role of the input on the lexicon and language mixing is concerned, the evidence suggests that although the children are growing up in a similar linguistic situation, differences in lexical categories, number of words, number of translation equivalents are very important. Possible explanations from the input are brought forward. For, example, a strong correlation was found between the number of translation equivalents produced and language (lexical) dominance or between parental strategies and language mixing.

The thesis as a whole contributes to the expanding body of literature on cross-linguistic studies of bilingual language development as well as to the current debate on bilingual language development, i.e. whether it is the same or different from monolingual language acquisition. These findings provide further support for the argument that bilingual children are not two monolinguals in one as they have added abilities and capacities (cross-linguistic synonyms and language mixing) but that their overall language development is similar to that of monolinguals (in terms of lexical acquisition at least).

BIBLIOGRAPHY

Allman, B. (2002). *Differences in the lexical development of monolingual and bilingual preschool children.* Paper presented at the IXth IASCL/SRCLD, Madison, WI, USA.

Au, T. K. & Glusman, M. (1990). The Principle of Mutual Exclusivity in word learning: to honor or not to honor? *Child Development,* 61, 1474-1490.

Bassano, D. (2000). Early development of nouns and verbs in French: exploring the interface between lexicon and grammar. *Journal of Child Language,* 27, 521-559.

Bassano, D., Maillochon, I. & Eme, E. (1998). Developmental changes and variability in the early lexicon: a study of French children's naturalistic productions. *Journal of Child Language,* 25, 493-531.

Bates, E., Bretherton, I. & Snyder, L. (1988). *From first words to grammar: individual differences and dissociable mechanisms.* Cambridge: Cambridge University Press.

Bates, E. & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes,* 12(5/6), 507-584.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P. S., Reznick, J. S., Reilly, J. & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language,* 21, 85-123.

Benedict, H. (1979). Early lexical development: comprehension and production. *Journal of Child Language,* 6, 183-200.

Bennett-Kastor, T. (1988). *Analyzing children's language.* Oxford: Blackwell.

Bialystok, E. (2001). *Bilingualism in development: language, literacy, & cognition.* Cambridge: Cambridge University Press.

Bloom, L. (1973). *One word at a time: the use of single word utterances before syntax.* The Hague: Mouton.

Bornstein, M. H., Haynes, M. O. & Painter, K. M. (1998). Sources of child vocabulary competence: a multivariate model. *Journal of Child Language,* 25, 367-393.

Bosch, L. & Sebastián-Gallés, N. (1997). Native language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition,* 65, 33-69.

Bosch, L. & Sebastián-Gallés, N. (2001). Early language differentiation in bilingual infants. In J. Cenoz & F. Genesee (Eds.), *Trends in bilingual acquisition* (pp. 71-94). Amsterdam: John Benjamins.

Boysson-Bardies, B. (1999). *How language comes to children: from birth to two years.* Cambridge, MA.: MIT Press.

Bradley, A. & Lorch, M. (2003). *Lexical acquisition in bilingual children.* Poster presented at the Child Language Seminar, University of Newcastle upon Tyne.

Braunwald, S. R. & Brislin, R. W. (1979). The diary method updated. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental pragmatics.* New York: Academic press.

Brown, R. (1973). *A first language: the early stages.* London: George Allen & Unwin.

Caselli, C., Casadio, P. & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language,* 26, 69-111.

Clark, E. (1987). The principle of contrast: a constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1-33). Hillsdale, NJ: Lawrence Erlbaum Associates.

Clark, E. (1988). On the logic of Contrast. *Journal of Child Language,* 15, 317-335.

Clark, E. (1993). *The lexicon in acquisition.* Cambridge: Cambridge University Press.

Clark, E. (2003). *First language acquisition.* Cambridge: Cambridge University Press.

Dale, P. S., Bates, E., Reznick, J. S. & Morisset, C. (1989). The validity of a parent report instrument of child language at twenty months. *Journal of Child Language,* 16, 239-249.

Davidson, D., Jergovic, D., Imami, Z. & Theodos, V. (1997). Monolingual and bilingual children's use of the mutual exclusivity constraint. *Journal of Child Language,* 24, 3-24.

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 P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language* (pp. 219-250). Oxford: Blackwell.

De Houwer, A. (1998). By way of introduction: methods in studies of bilingual first language acquisition. *International Journal of Bilingualism,* 2(3), 249-263.

De Houwer, A. (1999). Language Acquisition in Children Raised with two Languages from Birth: an Update. *Revue PArole,* 9-10, 63-87.

De Houwer, A. (1999b). Environmental factors in early bilingual development: the role of parental beliefs and attitudes. In G. Extra & L. Verhoeven (Eds.), *Bilingualism and migration.* New York: Mouton de Gruyter.

De Houwer, A. & Bornstein, M. H. (2001). *Lexical knowledge at 13 months as measured by the MacArthur CDI: does it matter how many raters there are?* Poster presented at Early Lexicon Acquisition, Lyon, France.

De Leòn, L. (2001). Recent cross-linguistic studies in early lexical development: reassessing factors of input, interaction, and cognitive constraints. In M. Almgren, A. Barreña, M.-J. Ezeizabarrena, I. Idiazabal & B. MacWhinney (Eds.), *Research on Child Language Acquisition: Proceedings of the 8th Conference of the International Association for the Study of Child Language* (pp. 923-927). Somerville, MA.: Cascadilla Press.

Deuchar, M. (1999). Are function words non-language-specific in early bilingual two-word utterances? *Bilingualism: Language and Cognition,* 2(1), 23-34.

Deuchar, M. & Quay, S. (1998). One vs. two systems in early bilingual syntax: Two versions of the question. *Bilingualism: Language and Cognition,* 1(3), 231-243.

Deuchar, M. & Quay, S. (2000). *Bilingual Acquisition (Theoretical implications of a case study).* Oxford: Oxford University Press.

Deuchar, M., & Quay, S. (2001). What use is a case study of bilingual acquisition? In M. Almgren, A. Barreña, M.-J. Ezeizabarrena, I. Idiazabal & B. MacWhinney (Eds.), *Research on Child Language Acquisition: Proceedings of the 8th Conference of the International Association for the Study of Child Language* (pp. 128-136). Somerville, MA.: Cascadilla Press.

Deuchar, M. & Muntz, R. (2003). Factors accounting for code-mixing in an early developing bilingual. In N. Müller (Ed.), *(In)vulnerable Domains in Multilingualism* (pp. 161-190). Amsterdam: John Benjamins Publishing Company.

Dewaele, J.M. (2000). Saisir l'insaisissable? Les mesures de longueur d'énoncés en linguistique appliquée. *International Review of Applied Linguistics*, 38(1), 17-33.

Döpke, S. (1992). *One parent, one language: an interactional approach.* Amsterdam: John Benjamins Publishing Company.

Döpke, S. (2000). Generation of and retraction from cross-linguistically motivated structures in bilingual first language acquisition. *Bilingualism: Language and Cognition,* 3(3), 209-226.

Dromi, E. (1996). *Early lexical development.* San Diego: Singular Publishing Group.

Eriksson, M. (2001). *Children's words in inventories and tests: what questions do they answer?* Paper presented at the Early Lexicon Acquisition seminar, Lyon, France.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J., Pethick, S. & Reilly, J. (1993). *MacArthur Communicative Development Inventories: User's Guide and Technical Manual.* San Diego, CA: Singular Publising.

Frank, I. & Poulin-Dubois, D. (2002). Young monolingual and bilingual children's responses to violation of the mutual exclusivity principle. *International Journal of Bilingualism,* 6(2), 125-146.

Garcia, E. (1983). *Early childhood bilingualism.* Albuquerque: University of New Mexico Press.

Gathercole, V. (1987). The contrastive hypothesis for the acquisition of word meaning: a reconsideration of the theory. *Journal of Child Language,* 14, 493-531.

Gathercole, V. (1989). Contrast: a semantic constraint? *Journal of Child Language,* 16, 685-702.

Genesee, F. (1989). Early bilingual development: one language or two? *Journal of Child Language,* 16, 161-179.

Genesee, F. (2003). Rethinking Bilingual Acquisition. In J.-M. Dewaele, A. Housen & W. Li (Eds.), *Bilingualism: Beyond Basic Principles.* Clevedon: Multilingual Matters.

Genesee, F. (2004). *The role of dominance in bilingual first language acquisition.* Paper presented at the 2nd Lisbon Meeting on Language Acquisition, Lisbon, Portugal.

Genesee, F., Boivin, I. & Nicoladis, E. (1996). Talking with strangers: a study of bilingual children's communicative competence. *Applied Psycholinguistics,* 17, 427-442.

Genesee, F., Nicoladis, E. & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language,* 22, 611-631.

Girouard, P., Ricard, M. & Gouin Décarie, T. (1997). The acquisition of personal pronouns in French-speaking and English-speaking children. *Journal of Child Language,* 24, 311-326.

Goldfield, B. A. & Reznick, J. S. (1990). Early lexical acquisition: rate, content and the vocabulary spurt. *Journal of Child Language,* 17, 171-183.

Goodz, N. S. (1989). Parental language mixing in bilingual families. *Infant Mental Health Journal,* 10(1), 25-44.

Goodz, N. S. (1994). Interactions between parents and children in bilingual families. In F. Genesee (Ed.), *Educating second language children* (pp. 61-81). Cambridge: Cambridge University Press.

Greenfield, P. M. & Smith, J. H. (1976). *The structure of communication in early language development.* New York: Academic Press.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition,* 1(2), 131-149.

Grosjean, F. (1999). The bilingual's language modes. In J. L. Nicol (Ed.), *One Mind, Two Languages: Bilingual Language Processing.* Oxford: Blackwell.

Halliday, M. A. K. (1975). *Learning how to mean: explorations in the development of language.* London: Edward Arnold.

Harrison, C. (2001). *Validity and preliminary norms for a British English adaptation of the MacArthur Communicative Development Inventories: Toddlers.* Unpublished dissertation, University of Newcastle-upon-Tyne.

Hart, B. & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore: Paul H. Brookes Publishing Co.

Hickmann, M. (1997). The acquisition of French as a native language: structural and functional determinants in a cross-linguistic perspective. *Journal of Speech-Language Pathology and Audiology,* 21(4), 236-257.

Holowka, S., Brosseau-Lapré, F. & Petitto, L. A. (2002). Semantic and conceptual knowledge underlying bilingual babies' first signs and words. *Language Learning,* 52(2), 205-262.

Hulk, A. & Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition,* 3(3), 227-244.

Ingram, D. (1989). *First language acquisition: method, description and explanation.* Cambridge: Cambridge University Press.

Jackson-Maldonado, D., Thal, D., Marchman, V., Bates, E. & Gutierrez-Clellen, V. (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language,* 20, 523-549.

Jisa, H. (2000). Language mixing in the weak language: evidence from two children. *Journal of Pragmatics,* 32, 1363-1386.

Johnson, C. E. & Lancaster, P. (1998). The development of more than one language phonology: a case study of a Norwegian-English bilingual child. *International Journal of Bilingualism,* 2(3), 265-300.

Jones, M. A. (1996). *Foundations of French syntax.* Cambridge: Cambridge University Press.

Juan-Garau, M. & Perez-Vidal, C. (2001). Mixing and pragmatic parental strategies in early bilingual acquisition. *Journal of Child Language*, 28, 59-86.

Junker, D. A. & Stockman, I. J. (2002). Expressive vocabulary of German-English bilingual toddlers. *American Journal of Speech-Language Pathology,* 11, 381-394.

Kern, S. (1999). *Inventaire Français du Dèveloppement Communicatif chez le nourisson: mots et phrases.* Lyon: Institut des Sciences de l'Homme.

Kern, S. (2001). *Développement des premiers comportements linguistiques chez le jeune enfant* (Rapport d'activités APN 98-2000). Unpublished report.

Kern, S. (submitted). Lexicon development in French-speaking infants. *Applied Psycholinguistics.*

Khattab, G. (2002). VOT production in English and Arabic bilingual and monolingual children. In D. B. Parkinson & E. Benmamoun (Eds.), *Perspectives on Arabic Linguistics XIII-XIV* (pp. 1-38). Amsterdam: John Benjamins.

Klee, T., Marr, C., Robertson, E. & Harrison, C. (1999). *The MacArthur Communicative Development Inventory: Toddlers.* University of Newcastle upon Tyne.

Klee, T. & Harrison, C. (2001). *CDI Words and Sentences: validity and preliminary norms for British English.* Paper presented at the Child Language Seminar, University of Hertfordshire, UK.

Köppe, R. (1996). Language differentiation in bilingual children: the development of grammatical and pragmatic competence. *Linguistics,* 34, 927-954.

Köppe, R. & Meisel, J. (1995). Code-switching in bilingual first language acquisition. In L. Milroy & P. Muysken (Eds.), *One speaker, two languages: cross-disciplinary perspectives on code-switching.* Cambridge: Cambridge University Press.

LaBelle, C. (2000). A longitudinal study of lexical and prosodic differentiation by a simultaneous French/English bilingual child. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University Conference on Language Development.* Somerville, MA.: Cascadilla Press.

Lanvers, U. (1999). Lexical growth patterns in a bilingual infant: the occurrence and significance of equivalents in the bilingual lexicon. *International Journal of Bilingual Education and Bilingualism,* 2(1), 30-52.

Lanza, E. (1992). Can bilingual two-year-olds code-switch? *Journal of Child Language,* 19, 633-658.

Lanza, E. (1997). *Language mixing in infant bilingualism.* Oxford: Oxford University Press.

Lanza, E. (2001). Bilingual first language acquisition: a discourse perspective on language contact in parent-child interaction. In F. Genesee & J. Cenoz (Eds.), *Trends in bilingual acquisition* (pp. 201-230). Amsterdam: John Benjamins.

Le Normand, M.-T., Parisse, C. & Cohen, H. (2002). *Developmental norms in the acquisition of lexical categories: trends and variability in normally developing French children aged 2 to 4.* Unpublished manuscript.

Leopold, W. (1939-49). *Speech development of a bilingual child: a linguist's record,* 4 volumes. Evanston, Illinois: Northwestern University Press.

Li, W. (2000). *The bilingualism reader.* London: Routledge.

Li, W. & Zhu, H. (2001). Development of code-switching and L1 attrition in L2 setting. In M. Almgren, A. Barreña, M.-J. Ezeizabarrena, I. Idiazabal & B. MacWhinney (Eds.), *Research on Child Language Acquisition: Proceedings of the 8th Conference of the International Association for the Study of Child Language* (pp. 174-187). Somerville, MA.: Cascadilla Press.

Li, W. (2002). Review of M. Deuchar & S. Quay, Bilingual acquisition: theoretical implications of a case study. *Journal of Child Language,* 29, 205-223.

Lindholm, K. J. & Padilla, A. M. (1978). Language mixing in bilingual children. *Journal of Child Language,* 5, 327-335.

LIPPS Group (2000). The LIDES coding manual: A document for preparing and analyzing language interaction data. *International Journal of Bilingualism,* 4(2).

Locke, J. (1997). A theory of Neurolinguistic Development. *Brain and Language,* 58, 265-326.

Lyon, J. (1996). *Becoming bilingual (Language acquisition in a bilingual community)*: Clevedon: Multilingual Matters Ltd.

MacLaughlin, B. (1984). Early bilingualism. In M. Paradis & Y. Lebrun (Eds.), *Early bilingualism and child development.* Lisse: Swets & Zeitlinger.

MacWhinney, B. (1989). Competition and lexical categorization. In R. Corrigan, F. Eckman & M. Noonan (Eds.), *Linguistic categorization* (pp. 195-242). New York: John Benjamins.

MacWhinney, B. (2004). *CLAN Manual.* Available: http://childes.psy.cmu.edu.

Maital, S. L., Dromi, E., Sagi, A. & Bornstein, M. H. (2000). The Hebrew Communicative Development Inventory: language specific properties and cross-linguistic generalizations. *Journal of Child Language,* 27, 43-67.

Malvern, D. & Richards, B. (2004). *A unified approach to the measurement of vocabulary diversity, morphological development and lexical style.* Paper presented at the BAAL/CUP Colloquium: Vocabulary knowledge and use: measurements and applications, Bristol, UK.

Marchman, V., & Bates, E. (1994). Continuity in lexical and morphological development: a test of the critical mass hypothesis. *Journal of Child Language,* 21, 339-366.

Markman, E. M. (1989). *Categorization and naming in children: problems of induction.* Cambridge, MA.: MIT Press.

Markman, E. M. (1994). Constraints on word meaning in early language acquisition. *Lingua,* 92, 199-227.

Marr, C. (1999). *The concurrent validity of British English adaptations of the long and short forms of the MacArthur Communicative Development Inventories (toddlers).* Unpublished dissertation, University of Newcastle-upon-Tyne.

McKee, G., Malvern, D. & Richards, B. (2000). Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing,* 15(3), 323-337.

Meisel, J. (1989). Early differentiation of languages in bilingual children. In K. Hyltenstam & K. Obler (Eds.), *Bilingualism across the lifespan* (pp. 13-40). Cambridge: Cambridge University Press.

Meisel, J. (1990). *Two first languages: early grammatical development in bilingual children.* Dordrecht: Foris.

Meisel, J. (1994a). *Bilingual first language acquisition: French and German grammatical development.* Philadelphia: John Benjamins.

Meisel, J. (1994b). Code-switching in young bilingual children: The acquisition of grammatical constraints. *Studies in Second Language Acquisition,* 16, 413-439.

Müller, N. (2004). *Wait and see: determiner omissions and object clitics in the acquisition of Italian by bilingual children.* Paper presented at the 2nd Lisbon Meeting on Language Acquisition, Lisbon, Portugal.

Myers-Scotton, C. (1993). *Duelling Languages.* Oxford: Clarendon Press.

Myers-Scotton, C. & Lake, J. L. (2000). Four types of morphemes: evidence from aphasia, codeswitching and second language acquisition. *Linguistics,* 38(6), 1053-1100.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development,* 38.

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), 105-116.

Nicoladis, E. & Genesee, F. (1996). A longitudinal study of pragmatic differentitation in young bilingual children. *Language Learning,* 46(3), 439-464.

Nicoladis, E. & Genesee, F. (1997). The role of parental input and language dominance in bilingual children's code-mixing. In E. Hughes, M. Hughes & A.

Greenhill (Eds.), *Proceedings of the 21st annual Boston University Conference on Language Development.* Somerville, MA.: Cascadilla Press.

Nicoladis, E. & Genesee, F. (1998). Parental discourse and codemixing. *International Journal of Bilingualism,* 2(1), 85-99.

Nicoladis, E. & Secco, G. (1998). The role of translation equivalents in a bilingual family's code-mixing. In A. Greenhill, M. Hughes, H. Littlefield & H. Walsh (Eds.), *Proceedings of the 22nd annual Boston University Conference on Language Development.* Somerville, MA.: Cascadilla Press.

Ochs, E. (1988). *Culture and language development: language acquisition and language socialization in a Samoan village.* Cambridge: Cambridge University Press.

Oller, D. K., Eilers, R. E., Urbano, R. & Cobo-Lewis, B. (1997). Development of precursors to speech in infants exposed to two languages. *Journal of Child Language,* 24, 407-425.

Oller, D. K. & Eilers, R. E. (Eds.). (2002). *Language and literacy in bilingual children.* Clevedon: Multilingual Matters.

Oppenheim, A. (1992). *Questionnaire design, interviewing and attitude measurement.* London: Pinter Publishers.

Paradis, J. (1996). Phonological differentiation in a bilingual child: Hildegard revisited. In A. Stringfellow, D. Cahana-Amitay, E. Hughes & A. Zukowski (Eds.), *Proceedings of the 20th annual Boston University Conference on Language Development* (pp. 528-539). Somerville, MA: Cascadilla Press.

Paradis, J. (2001). Do bilingual two-year olds have separate phonological systems? *International Journal of Bilingualism,* 5, 19-38.

Paradis, J. & Genesee, F. (1996). Syntactic acquisition in bilingual children: autonomous or interdependent? *Studies in Second Language Acquisition,* 18(1), 1-25.

Paradis, J., Nicoladis, E. & Genesee, F. (2000). Early emergence of structural constraints on code-mixing: evidence from French-English bilingual children. *Bilingualism: Language and Cognition,* 3(3), 245-261.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism.* Amsterdam: John Benjamins Publishing.

Parisse, C. (2002). Personal communication.

Parisse, C. & Le Normand, M.-T. (2000). How children build their morphosyntax: the case of French. *Journal of Child Language,* 27, 267-292.

Parisse, C. & Le Normand, M.-T. (2001). *Early lexical development in English and French children aged 1;6 to 2;6.* Paper presented at the Early Lexicon Acquisition seminar, Lyon, France.

Pearson, B. Z. (1992). *Rationale for English-Spanish CDI mapping.* Unpublished manuscript.

Pearson, B. Z. (1995). *A temporary mini-manual for the associated forms of the MacArthur CDIs.* Unpublished manuscript.

Pearson, B. Z. (1998). Assessing lexical development in bilingual babies and toddlers. *International Journal of Bilingualism,* 2(3), 347-372.

Pearson, B. Z., Fernandez, S., Lewedeg, V. & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics,* 18, 41-58.

Pearson, B. Z., Fernandez, S. & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: comparison to monolingual norms. *Language Learning,* 43(1), 93-120.

Pearson, B. Z., Fernandez, S. & Oller, D. K. (1995). Cross-language synonyms in the lexicon of bilingual infants: one language or two? *Journal of Child Language,* 22, 345-368.

Petitto, L. A., Katerelos, M., Levy, B. G., Gauna, K., Tétreault, K. & Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language,* 28, 453-496.

Pinker, S. (1984). *Language learnability and language development.* Cambridge,MA: Harvard University Press.

Pinker, S. (1991). Rules of language. *Science,* 253, 530-535.

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en español: toward a typology of code-switching. *Linguistics,* 18, 581-618.

Quay, S. (1995). The bilingual lexicon: implications for studies of language choice. *Journal of Child Language*, 22, 369-387.

Redlinger, W. E. & Park, T.-Z. (1980). Language mixing in young bilinguals. *Journal of Child Language,* 7, 337-352.

Rhemtulla, M., & Nicoladis, E. (2003). *Is Mutual exclusivity a pragmatic principle? Evidence from bilingual children.* Paper presented at the 4th International Symposium on Bilingualism, Arizona State University, USA.

Robertson, E. (1999). *The effect of list length on parent-report measure of children's early vocabulary production.* Unpublished dissertation, University of Newcastle-upon-Tyne.

Romaine, S. (1995). *Bilingualism.* Oxford: Blackwell.

Ronjat, J. (1913). *Le développement du langage observe chez un enfant bilingue.* Paris: Librairie Ancienne H. Campion.

Saunders, G. (1988). *Bilingual children: from birth to teens.* Clevedon: Multilingual Matters.

Saxton, M. (1997). The Contrast Theory of negative input. *Journal of Child Language,* 24, 139-161.

Schachter, F. F. (1979). *Everyday Mother Talk to Toddlers: Early Intervention.* New York: Academic Press.

Schlyter, S. (1993). The weaker language in bilingual Swedish-French children. In K. Hyltenstam & A. Viberg (Eds.), *Progression and regression in language: sociocultural, neuropsychological, & linguistic perspectives* (pp. 289-308). Cambridge: Cambridge University Press.

Shore, C. M. (1995). *Individual differences in language development.* London: Sage Publications.

Tamis-Lemonda, C., Bornstein, M. H., Kahana-Kalman, R., Baumwell, L. & Cyphers, L. (1998). Predicting variation in the timing of language milestones in the second year: an events history approach. *Journal of Child Language,* 25, 675-700.

Tardif, T., Shatz, M. & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: a comparison of English, Italian and Mandarin. *Journal of Child Language,* 24, 535-565.

Thal, D. (2003). *Early identification of specific language impairment: present status and future directions.* Paper presented at the Child Language Seminar, University of Newcastle upon Tyne, UK.

Turian, D. & Altenberg, P. (1991). Compensatory strategies of child first language attrition. In H. W. Seliger & R. M. Vago (Eds.), *First language attrition* (pp. 207-226). Cambridge: Cambridge University Press.

Van der Linden, E. (2000). Non-selective access and activation in child bilingualism: the lexicon. In S. Döpke (Ed.), *Cross-linguistic structures in simultaneous bilingualism* (pp. 37-56). Amsterdam: John Benjamins Publishing Company.

Veneziano, E. & Sinclair, H. (2000). The changing status of 'filler syllables' on the way to grammatical morphemes. *Journal of Child Language,* 27, 461-500.

Vihman, M. M. (1985). Language differentiation by the bilingual infant. *Journal of Child Language,* 12, 297-324.

Vihman, M. M. (1998). A developmental perspective on codeswitching: Conversations between a Pair of Bilingual Siblings. *International Journal of Bilingualism,* 2(1), 45-84.

Volterra, V. & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language,* 5, 311-326.

Yip, V. & Matthews, S. (2000). Syntactic transfer in a Cantonese-English bilingual child. *Bilingualism: Language and Cognition,* 3(3), 193-208.

Yukawa, E. (1997). *L1 Japanese attrition and regaining.* Stockholm: Centre for Research on Bilingualism.

Zhu, H. & Dodd, B. J. (In press). *Cross-linguistic perspectives on child phonology.* Clevedon: Multilingual Matters.

Zhu, H. & Li, W. (In press). Bi- and multi-lingual acquisition. In M. Ball (Ed.), *Handbook of clinical linguistics.* Oxford: Blackwell.

Does the child have a history of middle ear infections? (If yes, how many?)

II The child's motor development

At what age did your child…?

	Age
Sit	
Crawl	
Walk	
Babble	
Say his first word	
Combine 2 words together	

Is the child talkative?

Does he/she repeat syllables? (dada)

Does he/she repeat different syllables? (dida)

III The child's linguistic background

What language does the child hear from…?

	Language heard
You	
Your partner	
His playmates	
Your parents	
His other grandparents	
His brothers or sisters	

How is the child cared for? How many times a week? For how long? (Please be precise)

What language(s) does he/she hear there?

If you are using the one person-one language method, when your children use the 'wrong' language how do you react?

How does your partner react?

Are you satisfied with the way your child is developing?

You can add any comment that you have in mind and that might be useful to the study.

APPENDIX II

Linguistic questionnaire:

LANGUAGE QUESTIONNAIRE

Please be as precise as possible in answering this questionnaire.
Only the researcher will have access to the information given and your identity will be kept confidential at all times.

Name:

Profession:

Ethnic background:

Age group: under 25 □ 25-29 □ 30-34 □ 35-39 □ 40-44 □ 45 and above □

Baby's name:

Baby's date of birth:

Languages most often spoken at home:

I Language Background

What is your mother tongue (the first language(s) you learnt as a child)?

What other languages do you know? (List all of them)

How did you acquire each of them and at what age?

	Age	Method of acquisition
_____		
_____		
_____		

How often do you use each of them?

	Mother tongue: _____	Other: _____	Other: _____
Everyday			
Very often			
Often			
Rarely			
Never			

Are you learning one or several languages at the moment?

If yes, which one(s) and how?

Have you always lived in the country you are living now?

If no, where else have you lived? When? For how long? What language did you use or hear there? (Give as many details as you wish).

How well can you perform in each of these tasks in each language?

	Writing			Reading			Speaking			Understanding the speech of others		
	Eng.	Fre.	Other:	Eng.	Fre.	Other:	Eng.	Fre.	Other:	Eng.	Fre.	Other:
Native												
Near native												
Quite well												
Well												
Not so well												
Ignore												

II Language Behaviour

In what language do you…?

	Language
Keep a diary or notes	
Calculate	
Remember telephone numbers	

Which language do you use with…?

	Language used
The baby?	
Your partner?	
Your other children?	
Close friends?	
Neighbours?	
Your parents?	
Your partner's parents?	
People at work?	

Do you have access to TV, radio or newspapers in another language than the country's majority language? (Please be as precise as possible)

How often do you…?

	Everyday	Twice a week	Once a week	Once a month	Never
Watch TV or a film in **English**?					
Watch TV or a film in **French**?					

	Everyday	Twice a week	Once a week	Once a month	Never
Watch TV or a film in another language: _____					
Read a book in **French**?					
Read a book in **English**?					
Read a book in another language: _____					
Read a newspaper in **English**?					
Read a newspaper in **French**?					
Read a newspaper in another language: _____					
Listen to the radio in **English**?					
Listen to the radio in **French**?					

Listen to the radio in another language: _____					

Estimate how long you used each language? For example, if you used as much French as English, put English 50%, French 50%.

Yesterday?

Last week?

Over the last month?

Over the last six months?

Do you ever use two languages in a conversation?

Never □ rarely □ sometimes □ often □ very often □

When you do, with who do you mostly do?

Do you find yourself saying something in one language while you should use the other one?

Never □ rarely □ sometimes □ often □ very often □

What kind of things are you talking about when this happens?

Would you say you are bilingual?

Do you think code-switching is good? *(Code-switching is when people use two or more languages in a sentence or a conversation)* Why?

Add any comments about your language background or experience that might be useful to the research.

APPENDIX III

Language input:

LANGUAGE EXPOSURE

Today's date: _____ Name of the child: _____

1. What language has the mother been speaking to your child in the past month?

English ____% French ____% Other _____

2. What language has the father been speaking to your child in the past month?

English _____% French _____ % Other _____

3. If the grandparents are able to visit or live with your child, please indicate what language they speak to your child?

Maternal grandmother: English _____ French _____ Other _____
Paternal grandmother: English _____ French _____ Other _____
Maternal grandfather: English _____ French _____ Other _____
Paternal grandfather: English _____ French _____ Other _____

4. Approximately how much time do the grandparents spend with your child?

Maternal grandmother: _____

Paternal grandmother: _____

Maternal grandfather: _____

Paternal grandfather: _____

5. How do the grandparents spend this time with your child? Do they, for example, read or play with your child or is the child sleeping during this time?

Maternal grandmother: _____

Paternal grandmother: _____

Maternal grandfather: _____

Paternal grandfather: _____

6. Are there any other people who may play a significant role in your child's life, for example, daycare personnel, a nanny, siblings...?

What is the relationship of that person to your child? _____

What language does s/he speak to your child? English _____ French _____ Other ___

How much time does s/he spend with your child? _____

How is the time spent with your child (ie. playing, sleeping)?

_____

If there is another significant person in your child's life, please indicate:
What is the relationship of that person to your child? _____

What language does s/he speak to your child? English _____ French _____ Other ___

How much time does s/he spend with your child? _____

How is the time spent with your child (ie. playing, sleeping)?

_____

If there is another significant person in your child's life, please indicate:
What is the relationship of that person to your child? _____

What language does s/he speak to your child? English _____ French _____ Other ___

How much time does s/he spend with your child? _____

How is the time spent with your child (ie. playing, sleeping)?

_____

7. Have there been any significant recent changes in the family, which might have affected the child's amount of exposure to the different languages? An example would be a situation where the mother goes back to work and the child is spending more time with another caretaker. Please describe it briefly.

_____

_____

_____

8. As a very rough estimate, what percentage of time do you believe that your child hears each language?

English _____ % French _____ % Other _____ %

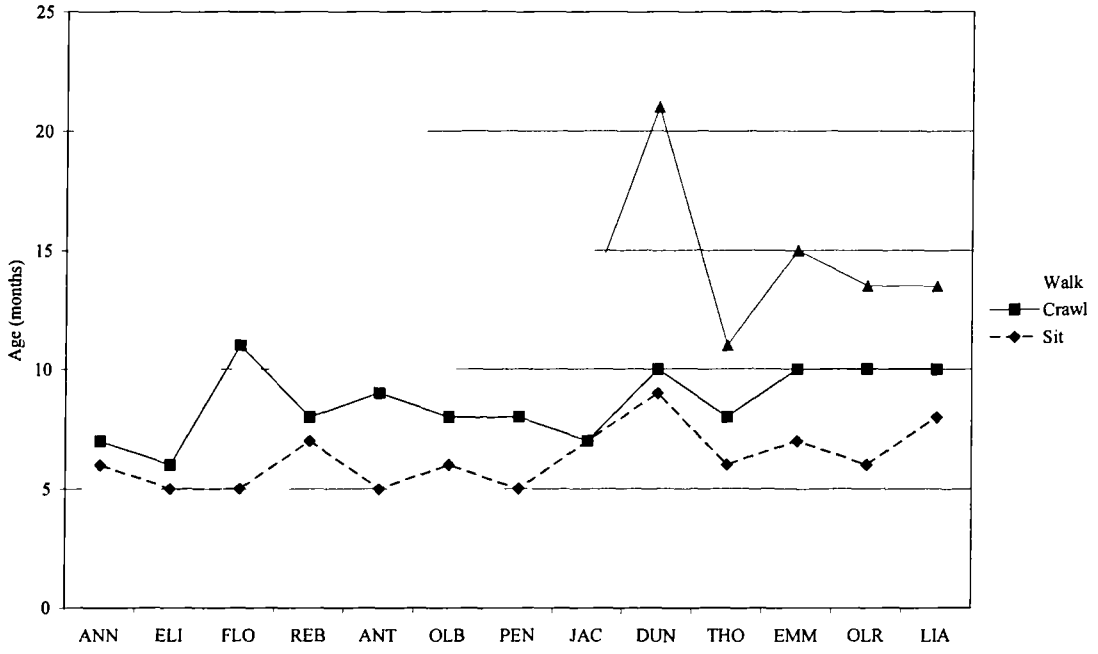
APPENDIX IV

Diary sample:

Date	*Word/Phrase said*	*Pronunciation*	*Who was he talking to?*	*Topic/Situation/ Question asked*	*Who else was there?*	*Gestures or reactions/word repeated ? times*	*Special events?*
04/10/01	Apple	Abble	Mummy & Grandma	Eating apple compote for lunch. Mummy & Grandma were saying " Isn't this nice apple"	Grandad	Said 3 times	In the UK
20/10/01	Daddy	Dadd-ee	Daddy	Daddy came in the door and Mummy said 'Look its Daddy'.	Mummy	Said once	
20/10/01	Hot	'ot	Mummy & Daddy	In the kitchen. Pointing to the oven, Daddy said "chaud", Mummy said 'Hot". Liam said 'ot'.		Said once	
Dec 01	Attend	Att	Daddy	When Daddy says 'attend', he repeats 'att'.		Only repeats after Daddy says it.	
Dec 01	Oh dear	Oh de de	Mummy	When Mummy said ''oh dear'' when Liam dropped something.		Repeats after Mummy but also says on his own when drops something etc.	

APPENDIX V

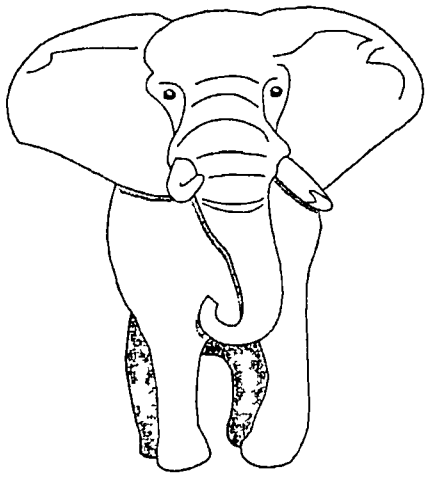
Motor development:



APPENDIX VI

Child's Name: _____ Sex: _____

Date of Birth: _____ Age: _____ Today's Date: _____



The MacArthur Communicative Development Inventory: Toddlers

British English Adaptation

A. VOCABULARY CHECKLIST

Ch ldren understand many more words than they say. We are particularly interested in the words your chi d SAYS. Please go through the ist and mark the words you have heard your ch d use. If your ch ld uses a different pronunc ation of a word (for example, "raffe" instead of "g raffe" or "skett for "spaghett "), mark the word anyway. Remember that this is a "catalogue" of a the words that are used by many different children. Don't worry if your child on y knows a few of these right now.

1. SOUND EFFECTS AND ANIMAL SOUNDS

baa baa	O	meow	O	uh oh	O
choo choo	O	moo	O	vroom	O
co kadoodledoo	O	ouch	O	woof woof	O
grr	O	quack quack	O	yum yum	O

2. ANIMALS (Real or Toy)

animal	O	duck	O	owl	O
ant	O	e ephant	O	penguin	O
bear	O	f sh	O	pig	O
bee	O	frog	O	pony	O
b rd	O	g raffe	O	puppy	O
bunny	O	goose	O	sheep	O
butterfly	O	hen	O	squirrel	O
cat	O	horse	O	teddybear	O
chicken	O	insect/fly	O	t'ger	O
cockerel	O	kitty	O	turkey	O
cow	O	lamb	O	turt e	O
crocod'le	O	lion	O	wolf	O
deer	O	monkey	O	zebra	O
dog	O	moose	O		
donkey	O	mouse	O		

3. VEHICLES (Real or Toy)					
aeroplane	O	fire-engine	O	sledge	O
bicycle	O	helicopter	O	tractor	O
boat	O	lorry	O	train	O
bus	O	motorbike	O	tricycle	O
car	O	pram	O		

4. TOYS					
ball	O	cha k	O	pen	O
balloon	O	crayon	O	pencil	O
bat	O	doll	O	play dough	O
block	O	game	O	present	O
book	O	glue	O	story	O
bubbles	O	jigsaw	O	toy	O

5. FOOD AND DRINK					
apple	O	fish	O	pickle	O
applesauce	O	food	O	pizza	O
banana	O	grapes	O	popcorn	O
beans	O	green beans	O	potato	O
biscu t	O	hamburger	O	pretzel	O
bread	O	ice	O	pudding	O
butter	O	ice cream	O	pumpkin	O
cake	O	icepop/icelolly	O	raisin	O
carrots	O	jam	O	salt	O
cereal	O	je ly	O	sandwich	O
cheese	O	juice fizzy ju'ce	O	sauce	O
chew ng gum	O	lol ipop	O	soup	O
ch cken	O	meat	O	spaghetti	O
ch ps	O	melon	O	strawberry	O
ch c ate	O	milk	O	sweetcorn	O
c ffee	O	muff n	O	sweets	O
coke	O	nood es	O	toast	O
cracker	O	nuts	O	tuna	O
crisps	O	orange	O	vanilla	O
doughnut	O	pancake	O	vitamins	O
drink	O	peanut butter	O	water	O
egg	O	peas	O	yogurt	O

6. CLOTHING					
beads	O	jeans	O	slipper	O
belt	O	jumper	O	sock	O
bib	O	mittens	O	sweatshirt	O
boots	O	nappy	O	t'ghts	O
button	O	neck ace	O	trainers	O
coat	O	pyjamas	O	trousers	O
dress	O	scarf	O	underpants	O
g oves	O	shirt	O	zip	O
hat	O	shoe	O		
jacket	O	shorts	O		

7. BODY PARTS

ankle	O	feet	O	nose	O
arm	O	finger	O	penis*	O
belly button	O	hair	O	shoulder	O
buttocks/bottom*	O	hand	O	tooth	O
cheek	O	head	O	toe	O
chin	O	knee	O	tongue	O
ear	O	leg	O	tummy	O
eye	O	lips	O	vagina*	O
face	O	mouth	O		

*or word used in your family

8. SMALL HOUSEHOLD ITEMS

basket	O	hammer	O	plant	O
b anket	O	hoover	O	plate	O
bott e	O	jar	O	purse	O
bowl	O	keys	O	radio	O
box	O	kn fe	O	rubbish	O
brush	O	lamp	O	scissors	O
bu ket	O	l ght	O	soap	O
camera	O	medicine	O	spoon	O
can	O	money	O	tape	O
c ock	O	mop	O	telephone	O
comb	O	nail	O	tissue/kleenex	O
cup	O	napkin	O	toothbrush	O
dish	O	paper	O	towel	O
fork	O	penny	O	watch	O
g ass	O	picture	O		
glasses	O	p llow	O		

9. FURNITURE AND ROOMS

bath	O	garage	O	shower	O
bathroom	O	h gh chair	O	sink	O
bed	O	kitchen	O	sofa	O
bedroom	O	ving room	O	stairs	O
bench	O	oven	O	stove	O
cellar	O	playpen	O	table	O
cha r	O	porch	O	TV	O
cot	O	potty	O	wardrobe	O
door	O	refr gerator	O	washing machine	O
drawer	O	rocking chair	O	window	O
dryer	O	room	O		

10. OUTSIDE THINGS

backyard	O	pool	O	star	O
cloud	O	ra n	O	st ck	O
flag	O	rock	O	stone	O
flower	O	roof	O	street	O
garden	O	sandp t	O	sun	O
grass	O	sky	O	swing	O
hose	O	s de	O	tree	O
ladder	O	snow	O	water	O
lawn mower	O	snowman	O	wind	O
moon	O	spade	O		
pavement	O	sprink er	O		

11. PLACES TO GO

beach	O	home	O	school	O
camping	O	house	O	shop	O
church*	O	outside	O	woods	O
c rcus	O	park	O	work	O
city centre, town	O	party	O	yard	O
country	O	petrol station	O	zoo	O
farm	O	picnic	O		
film	O	playground	O		

*or word used in your family

12. PEOPLE

aunt	O	doctor	O	child's own name	O
baby	O	fireman	O	people	O
babysitter	O	friend	O	person	O
babysitter's name	O	g rl	O	pet's name	O
boy	O	grandma*	O	police	O
brother	O	grandpa*	O	postman	O
ch d	O	ady	O	sister	O
clown	O	man	O	teacher	O
cowboy	O	mummy*	O	uncle	O
daddy*	O	nurse	O		

*or word used n your fami y

13. GAMES AND ROUTINES

bath	O	lunch	O	shopping	O
breakfast	O	nap	O	snack	O
bye	O	night night	C	so big!	O
d nner	O	no	O	thank you	O
g ve me f ve!	O	patty cake	O	this little piggy	O
gonna get you!	O	peekaboo	O	turn around	O
go potty	O	phone	O	yes	O
hi	O	please	O		
hello	O	shh shush/hush	O		

14. ACTION WORDS

b te	O	drink	O	hold	O	read	O	swim	O
blow	O	drive	O	hurry	O	ride	O	swing	O
break	O	drop	O	jump	O	rip	O	take	O
bring	O	dry	O	k ck	O	run	O	talk	O
bu ld	O	dump	O	k ss	O	say	O	taste	O
bump	O	eat	O	knock	O	see	O	tear	O
buy	O	fall	O	ick	O	shake	O	think	O
carry	O	feed	O	ke	O	share	O	throw	O
catch	O	f nd	O	l sten	O	show	O	tickle	O
chase	O	finish	O	look	O	sing	O	touch	O
clap	O	f t	O	ove	O	sit	O	wait	O
clean	O	f x	O	make	O	skate	O	wake	O
cl mb	O	get	O	open	O	sleep	O	walk	O
close	O	give	O	paint	O	sl de	O	wash	O
cook	O	go	O	pick	O	smile	O	watch	O
cover	O	hate	O	play	O	spill	O	wipe	O
cry	O	have	O	pour	O	splash	O	wish	O
cuddle	O	hear	O	pretend	O	stand	O	work	O
cut	O	help	O	pull	O	stay	O	write	O
dance	O	hide	O	push	O	stop	O		
draw	O	hit	O	put	O	sweep	O		

15. DESCRIPTIVE WORDS

allgone	O	full	O	orange	O
asleep	O	gentle	O	poor	O
awake	O	good	O	pretty	O
bad	O	green	O	quiet	O
better	O	happy	O	red	O
b'g	O	hard	O	sad	O
black	O	heavy	O	scared	O
blue	O	high	O	sick	O
broken	O	hot	O	sleepy	O
brown	O	hungry	O	slow	O
careful	O	hurt	O	soft	O
clean	O	last	O	sticky	O
cold	O	little	O	stuck	O
cute	O	long	O	thirsty	O
dark	O	loud	O	tiny	O
dirty	O	mad	O	tired	O
dry	O	naughty	O	wet	O
empty	O	new	O	white	O
fast	O	nice	O	windy	O
fine	O	noisy	O	yellow	O
f rst	O	old	O	yucky	O

16. WORDS ABOUT TIME

after	O	morning	O	today	O
before	O	night	O	tomorrow	O
day	O	now	O	tonight	O
later	O	time	O	yesterday	O

17. PRONOUNS

he	O	me	O	their	O	we	O
her	O	mine	O	them	O	you	O
hers	O	my	O	these	O	your	O
him	O	myself	O	they	O	yourself	O
h s	O	our	O	th s	O		
I	O	she	O	those	O		
it	O	that	O	us	O		

18. QUESTION WORDS

how	O	when	O	which	O	why	O
what	O	where	O	who	O		

19. PREPOSITIONS AND LOCATIONS

about	O	down	O	on top of	O
above	O	for	O	out	O
around	O	here	O	over	O
at	O	inside/in	O	there	O
away	O	into	O	to	O
back	O	next to	O	under	O
behind	O	of	O	up	O
beside	O	off	O	with	O
by	O	on	O		

20. QUANTIFIERS AND ARTICLES					
a	O	each	O	other	O
all	O	every	O	same	O
a lot	O	more	O	some	O
an	O	much	O	the	O
another	O	not	O	too	O
any	O	none	O		

21. HELPING VERBS					
am	O	does	O	need/need to	O
are	O	don't	O	try/try to	O
be	O	gonna/going to	O	want to	O
can	O	gotta/got to	O	was	O
could	O	hafta/have to	O	were	O
did/did you	O	is	O	will	O
do	O	lemme/let me	O	would	O

22. CONNECTING WORDS					
and	O	but	O	so	O
because	O	if	O	then	O

B. HOW CHILDREN USE WORDS	Not Yet	Sometimes	Often
1. D es your ch d ever ta k about past events or peop e who are not present? For example a chi d who went to the circus last week might ater say c rcus, c own or band.	O	O	O
2 Does your ch d ever talk about someth ng that s going to happen n the future for example, say ng "choo choo" or "aerop ane" before you leave the house for a tr p, or saying "sw ng" when you are going to the park?	O	O	O
3 D es your ch d talk about objects that are not present such as a k ng about a m ss ng or absent toy, referring to a pet out of view or ask ng about someone not present?	O	O	O
4 D es your ch d understand if you ask for someth ng that is not n the room for examp e by going to the bedroom to get a teddy bear when you say "where's the bear?"	O	O	O
5. Does your ch d ever p ck up or point to an object and name an absent person to whom the object be ongs? For examp e, a ch ld m ght point to mammy's shoe and say "mammy".	O	O	O

PART II – SENTENCES AND GRAMMAR

A. WORD ENDINGS/PART I	Not Yet	Sometimes	Often
1. To ta k about mo e than one th ng, we add an s to many words. Examp es inc ude cars (for more than one car), shoes, dogs and keys Has your ch ld begun to do th s?	O	O	O
2. To ta k about ownersh p, we add an s', for example, Daddy's key, kitten s d sh and baby s bottle. Has your child begun to do this?	O	O	O
3. To ta k about activ ties, we sometimes add 'ing' to verbs. Examples inc ude looking, running and crying. Has your child begun to do this?	O	O	O
4. To talk about things that happened in the past, we often add 'ed' to the verb. Examples inc ude kissed, opened and pushed. Has your ch ld begun to do this?	O	O	O

THE MACARTHUR COMMUNICATIVE DEVELOPMENT INVENTORY: TODDLERS
British English Adaptation

B. WORD FORMS

Following are some other words children learn. Please mark any of these words that your child uses.

NOUNS

children	O	men	O	teeth	O
feet	O	mice	O		

VERBS

ate	O	fell	O	made	O
blew	O	flew	O	ran	O
bought	O	got	O	sat	O
broke	O	had	O	saw	O
came	O	heard	O	took	O
drank	O	held	O	went	O
drove	O	lost	O		

C. WORD ENDINGS/PART 2

Young children often place the wrong endings on words, for example, a child might say "Aunt e goed home". Mistakes like this are often a sign of progress in language. In the following lists please mark all the mistakes of this kind you have heard your child say recently.

NOUNS

b ockses	O	mans	O	sockses	O
childrens	O	mens	O	teeths	O
childs	O	m ces	O	toeses	O
feets	O	mouses	O	tooths	O
foots	O	shoeses	O		

VERBS

ated	O	comed	O	goed	O	ranned	O
b ewed	O	doed	O	gotted	O	runned	O
b owed	O	dranked	O	haved	O	seed	O
bringed	O	drinked	O	heared	O	satted	O
buyed	O	eated	O	holded	O	sitted	O
breaked	O	falled	O	losed	O	taked	O
broked	O	flied	O	losted	O	wented	O
camed	O	getted	O	maked	O		

HAS YOUR CHILD BEGUN TO COMBINE WORDS YET, SUCH AS "NOTHER BISCUIT", OR "DOGGIE BITE"?

Not Yet O Sometimes O Often O

IF YOU ANSWERED NOT YET, PLEASE STOP HERE. IF YOU ANSWERED SOMETIMES OR OFTEN, PLEASE CONTINUE.

D. EXAMPLES: Please list three of the longest sentences you have heard your child say recent y.

1. _____
2. _____
3. _____

E. COMPLEXITY

In each of the following pairs, please mark the one that sounds MOST like the way your child talks right now. If your child is saying sentences even longer or more complicated than the two provided, just pick the second one.

Two shoe.	O	Baby blanket.	O	Read me story, ma/ummy.	O
Two shoes.	O	Baby s blanket.	O	Read me a story, ma/ummy.	O
Two foot.	O	Doggie table.	O	No wash dolly.	O
Two feet.	O	Doggie on table.	O	Don't wash dolly.	O
Daddy car.	O	That my lorry.	O	Want more juice.	O
Daddy's car	O	That's my lorry.	O	Want juice in there.	O
(Talking about something happening right now)		(Ta king about something that already happened)			
Kitten s eep.	O	Daddy pick me up.	O	There a kitten.	O
Kitten sleeping.	O	Daddy picked me up.	O	There s a kitten.	O
(Talking about something happen ng r ght now		(Ta king about something that already happened)			
I make tower.	O	Kitten go away.	O	Go bye-bye.	O
I making tower.	O	Kitten went away.	O	Want go bye-bye.	O
(Talking about something that already happened)		(Ta king about something that already happened)			
I fall down.	O	Doggie kiss me.	O	Where ma/ummy go?	O
I fell down.	O	Doggie k ssed me.	O	Where did ma/ummy go?	O
More biscuit!	O	Baby crying.	O	Coffee hot.	O
More biscu'ts!	O	Baby s crying.	O	That coffee hot.	O
These my tooth	O	You fix it.	O	I no do it.	O
These my teeth.	O	Can you fix it?	O	I can't do it.	O
I want that.	O	Baby want eat.	O	I l'ke read stories.	O
I want that one you got.	O	Baby want to eat.	O	I like to read stories.	O
Turn on l ght	O	Where s my dolly?	O	Don't read book.	O
Turn on the ight so I can see.	O	Where s my dolly name Sam?	O	Don't want you read that book.	O
Want biscuits.	O	Lookit me!	O	We made this.	O
Want biscuits and m lk.	O	Lookit me dancing!	O	Me and Paul made this.	O
Biscuit ma ummy.	O	Lookit!	O		
Biscuit for ma ummy.	O	Lookit what I got!	O		
Baby crying.	O	I sing song.	O		
Baby crying coz she's sad.	O	I sing song for you.	O		

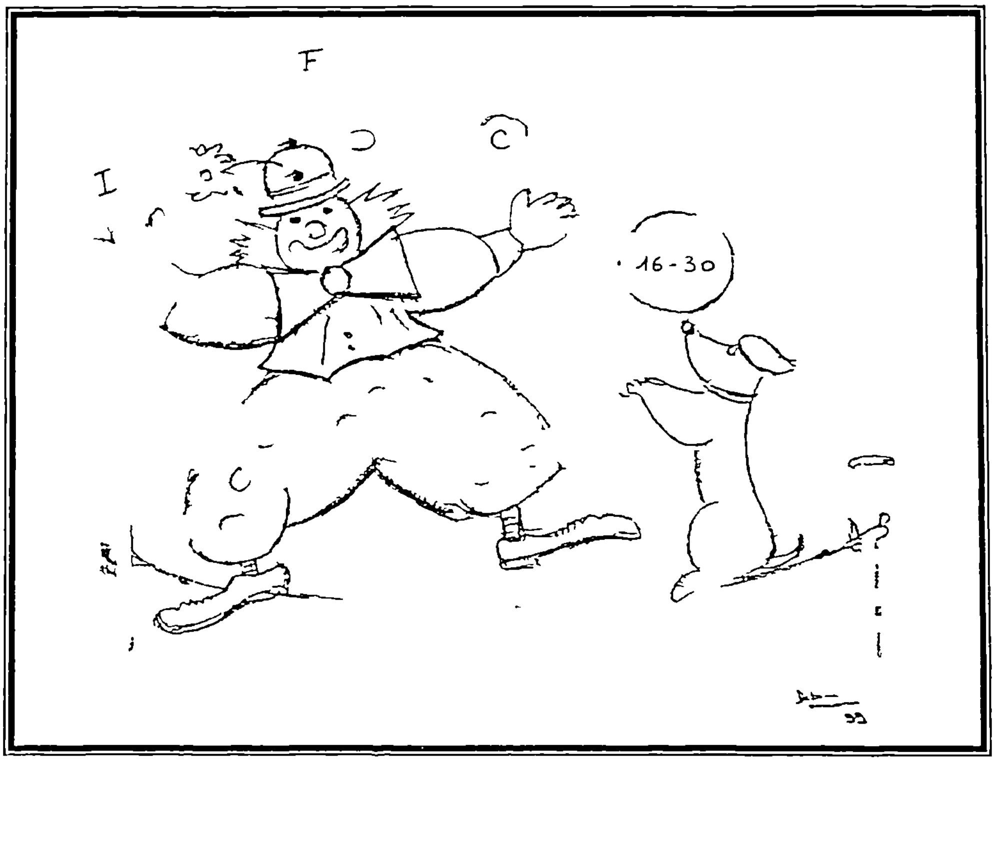
OTHER COMMENTS:

Name of the person who completed this questionnaire: _____

APPENDIX VII

Inventaire Français du Développement Communicatif chez le nourrisson : mots et phrases

C [illegible] S K n, L r e D n m qu du L
I t u I [illegible] H 4, a e u B r h t, 6 3C3 Ly n c d 7 F ance



Vous avez accepté de participer à ce projet. Nous vous en remercions.

Informations sur l'enfant:

Prénom: _____

Date de Naissance: _____

Date d'Aujourd'hui: _____

Nom de la personne qui remplit ce questionnaire : _____

VOCABULAIRE

Regardez SVP la liste suivante et cochez les mots que votre enfant utilise en ce moment.

Cris d'animaux et sons

1	bêê bêê	☐
2	tchou tchou	☐
3	cocorico	☐
4	grrrr	☐
5	miaou	☐
6	meuh	☐
7	allô	☐
8	aïe	☐
9	coin-coin	☐
10	oh oh	☐
11	vroum	☐
12	ouaf-ouaf	☐
13	miam-miam	☐

Véhicules (vrais ou jouets)

14	avion	☐
15	vélo	☐
16	bus	☐
17	voiture	☐
18	camion de pompier	☐
19	moto	☐
20	poussette	☐
21	train	☐
22	camion	☐
23	bateau	☐
24	hélicoptère	☐
25	traîneau	☐
26	tracteur	☐
27	tricycle	☐

Jouets

28	balle	☐
29	ballon	☐
30	cube	☐
31	livre	☐
32	bulles	☐
33	poupée	☐
34	crayon	☐
36	stylo	☐
37	jouet	☐
38	raquette	☐
39	craie	☐
40	jeu	☐
41	colle	☐
42	pâte à modeler	☐
43	cadeau	☐
44	puzzle	☐
45	histoire	☐
46	feutre	☐

Noms d'animaux (vrais ou jouets)

47	animal	☐
48	ours	☐
49	abeille	☐
50	oiseau	☐
51	petite bête	☐
52	lapin	☐
53	papillon	☐
54	chat	☐
55	poule	☐
56	vache	☐
57	biche	☐
58	chien	☐
59	âne	☐
60	canard	☐
61	éléphant	☐
62	poisson	☐
63	grenouille	☐
64	girafe	☐
65	chèvre	☐
66	oie	☐
67	cheval	☐
68	bébé chat	☐
69	agneau	☐
70	lion	☐
71	singe	☐
72	souris	☐
73	hibou	☐
74	pingouin	☐
75	cochon	☐
76	poney	☐
77	bébé chien	☐
78	mouton	☐
79	écureuil	☐
80	nounours	☐
81	tigre	☐
82	dindon	☐
83	tortue	☐
84	crocodile	☐
85	fourmi	☐
87	renne	☐
88	coq	☐
89	loup	☐
90	zèbre	☐

Nourriture et boisson

91	lait	☐
92	céréales	☐
93	cornflakes	☐
94	orange	☐
95	fromage	☐
96	petits pots	☐
97	poulet	☐
98	pizza	☐
99	café	☐
100	raisin	☐
101	petits gâteaux	☐
102	spaghetti	☐
103	gâteaux apéro	☐
104	tartine	☐
105	pomme	☐
106	œuf	☐
107	banane	☐
108	poisson	☐
109	pain	☐
110	nourriture	☐
111	beurre	☐
112	glace	☐
113	gâteau	☐
114	jus de fruit	☐
115	sucre	☐
116	viande	☐
117	carotte	☐
118	boisson	☐
119	eau	☐
120	clémentine	☐
121	purée	☐
122	petits pois	☐
123	baguette	☐
124	soupe	☐
125	pâtes	☐
126	thon	☐
127	kiwi	☐
128	compote	☐
129	haricot	☐
130	chocolat	☐
131	coca	☐
132	maïs	☐
133	beignet	☐
134	frites	☐
135	haricots verts	☐
136	chewing-gum	☐
137	hamburger	☐
138	glaçon	☐
139	vitamines	☐
140	bonbons	☐
141	confiture	☐
142	yaourt	☐
143	sucette	☐
144	melon	☐
145	madeleine	☐
146	noisettes	☐
147	crêpe	☐
148	nutella	☐
149	mayonnaise	☐
150	pop-corn	☐
151	esquimau	☐
152	pomme de terre	☐
153	chips	☐
154	bretzel	☐
155	flan	☐
156	courge	☐
157	raisins secs	☐
158	sel	☐
159	sandwich	☐
160	sauce	☐
161	limonade	☐
162	fraise	☐
163	vanille	☐

Vêtements

164 perles
165 collier
166 bavoir/bavette
167 pyjama
168 bottes
169 pantalon
170 boutons
171 chemise
172 manteau
173 chaussure
174 couche
175 body
176 grenouillère
177 short
178 robe
179 chaussettes
180 chapeau
181 sweet
182 veste
183 tee-shirt
184 jeans
185 salopette
186 chausson/pantoufle
187 pull
188 ceinture
189 gants
190 moufles
191 écharpe
192 basket
193 combinaison de ski
194 collants
195 culotte/slip

Parties du corps

196 bras
197 nombril
198 joue
199 oreille
200 yeux
201 figure/visage
202 pied
203 doigt
204 cheveux
205 main
206 tête
207 genou
208 jambe
209 bouche
210 nez
211 aïe bobo
212 dent
213 doigt de pied
214 langue
215 ventre
216 cheville
217 fesses
218 menton
219 lèvre
220 pénis/zizi...
221 cœur
222 vagin/zezette...
223 pouce

Meubles et pièces

224 salle de bain
225 baignoire
226 lit
227 chambre
228 chaise
229 canapé
230 berceau
231 porte
232 tiroir
233 garage
234 chaise haute
235 cuisine
236 salon
237 four
238 parc
239 pot
240 frigo
241 fauteuil
242 lavabo
243 escalier
244 cuisinière
245 table
246 télé
247 fenêtre
248 étendage
249 entrée
250 douche
251 pièce
252 cave
253 banc
254 wc
255 machine à laver
256 évier

Jeux et routines

257 bain
258 petit déjeuner
259 au revoir
260 dîner
261 ne fais pas
262 bonjour
263 salut
264 déjeuner
265 sieste
266 bonne nuit
267 non
268 ainsi font font
269 coucou
270 s'il te plaît
271 chut
272 merci
273 oui
274 bravo
275 faire les courses
276 goûter
277 coup de fil
278 top là
279 je vais t'attraper
280 va sur le pot
281 tourne-toi
282 ce petit cochon

Endroits où aller

283 plage
284 camping
285 église
286 cirque
287 campagne
288 centre-ville
289 ferme
290 station service
291 maison
292 cinéma
293 dehors
294 parc
295 fête
296 pique-nique
297 terrain de jeux
298 école
299 magasin
300 forêt
301 travail
302 cour
303 zoo
304 garderie
305 crèche

Prépositions et localisations

306 au sujet de
307 au dessus de
308 autour de
309 à
310 loin
311 derrière
312 à côté de
313 chez
314 en bas
315 pour
316 ici
317 à l'intérieur de
318 dans
319 près de
320 de
321 au loin
322 sur
323 au sommet de
325 par dessus
326 là-bas
327 vers
328 sous
329 en haut
330 avec
331 là

Interrogatifs

- 332 comment
- 333 quoi
- 334 quand
- 335 où
- 336 qui
- 337 pourquoi
- 338 le/la/les/quel(les)

Quantificateurs et articles

- 339 tous/tout
- 340 un autre
- 341 encore
- 342 aucun/ne
- 343 pas
- 344 le/la même
- 345 un peu
- 347 un/une
- 348 plein/beaucoup
- 349 du/de la/des
- 350 chaque
- 351 autre
- 352 le/la/les
- 353 aussi

Pronoms

- 354 à elle/sa
- 355 à lui/son
- 356 je
- 357 ça
- 358 moi
- 359 à moi
- 360 ma/mon/mes
- 361 vous/tu
- 362 votre/ta/ton
- 363 il
- 364 ses
- 365 lui
- 366 moi-même
- 367 notre
- 368 elle
- 369 leur
- 370 eux
- 371 ces
- 372 ils/elles
- 373 ceux
- 374 nous
- 375 on
- 376 toi-même

Mots sur le temps

- 377 jour
- 378 après
- 379 matin
- 380 nuit
- 381 maintenant
- 382 aujourd'hui
- 383 demain
- 384 ce soir
- 385 avant
- 386 heure
- 387 hier

Connecteurs

- 388 et
- 389 parce que
- 390 mais
- 391 si
- 392 donc
- 393 alors

Objets d'extérieurs

- 394 rocher
- 395 nuage
- 396 drapeau
- 397 fleur
- 398 jardin
- 399 herbe
- 400 lune
- 401 tuyau
- 402 échelle
- 403 tondeuse à gazon
- 404 piscine
- 405 pluie
- 406 caillou
- 407 toit
- 408 pelle
- 409 bac à sable
- 410 toboggan
- 411 neige
- 412 étoile
- 413 trottoir
- 414 soleil
- 415 balançoire
- 416 arbre
- 417 eau
- 418 ciel
- 419 bonhomme de neige
- 420 arrosoir
- 421 bâton
- 422 pierre
- 423 rue/route
- 424 vent

Personnes

- 425 tante
- 426 bébé
- 427 nounou
- 428 nom de la nounou
- 429 garçon
- 430 frère
- 431 enfant
- 432 papa
- 433 fille
- 434 grand-mère
- 435 grand-père
- 436 dame
- 437 maman
- 438 nom de l'enfant
- 439 gens
- 440 personne
- 441 sœur
- 442 maître/sse
- 443 oncle
- 444 monsieur
- 445 clown
- 446 docteur
- 447 pompier
- 448 copain/ine
- 449 facteur
- 450 infirmière
- 451 police
- 452 nom de l'animal domestique

Auxiliaires

- 453 suis
- 454 sont
- 455 être
- 456 est
- 457 peux
- 458 pourrait
- 459 a fait
- 460 faire
- 461 fait
- 462 ne pas
- 463 aller
- 464 devoir faire
- 465 avoir à faire
- 466 laisse-moi
- 467 avoir besoin de
- 468 essayer de
- 469 vouloir
- 470 était

Petits objets ménagers

471 couverture
472 bouteille
473 bol
474 boite
475 balai
476 brosse
477 horloge
478 peigne
479 tasse
480 plat
481 fourchette
482 verre
483 lunettes
484 marteau
485 clefs
486 lampe
487 lumière
488 médicaments
489 argent
490 papier
491 sous/pièces
492 photo
493 oreiller
494 plante
495 assiette
496 porte-monnaie
497 radio
498 ciseaux
499 savon
500 cuillère
501 téléphone
502 brosse à dent
503 serviette
504 poubelle
505 aspirateur
506 montre
507 feuille
508 musique
509 sirop
510 biberon
511 télécommande
512 sucette
513 panier
514 seau
515 appareil photo
516 ordures
517 pot
518 couteau
519 serpillière
520 clou
521 serviette de table
522 cassette
523 mouchoir
524 plateau
525 trotteur
526 coussins

Mots descriptifs

527 parti
528 endormi
529 pas bon
530 grand
531 bleu
532 cassé
533 attention
534 propre
535 froid
536 mignon/ne
537 sombre
538 sale
539 sec/che
540 vide
541 vite
542 bien
543 doux/ce
544 bon/ne
545 content/te
546 dur
547 chaud/e
548 avoir faim
549 blessé
550 petit/e
551 vilain/e
552 gentil/le
553 vieux/vieille
554 joli
555 rouge
556 avoir peur
557 malade
558 avoir sommeil
559 tendre
560 avoir soif
561 fatigué
562 mouillé
563 dégoûtant/e
564 beau/belle
565 méchant/e
566 être réveillé
567 mieux
568 noir
569 marron
570 premier/ère
571 plein/ne
572 vert/e
573 lourd/e
574 haut/e
575 dernier/e
576 long/ue
577 fort/e
578 fou/folle
579 neuf/ve
580 bruyant/e
581 orange
582 tranquille
583 triste
584 lent/e
586 collant/e
587 minuscule
588 blanc/he
589 venteux
590 jaune
591 coquin/e
592 coincé

Mots d'action

593 mordre
594 souffler
595 casser
596 apporter
597 se cogner
598 nettoyer
599 fermer
600 pleurer
601 danser
602 dessiner
603 boire
604 conduire
605 manger
606 tomber
607 nourrir
608 finir
609 recevoir
610 donner
611 aller
612 aider
613 taper
614 prendre dans ses bras
615 se dépêcher
616 sauter
617 donner un coup
618 faire un bisou
619 regarder
620 laver
621 aimer
622 ouvrir
623 jouer
624 tirer
625 pousser
626 mettre
627 lire
628 faire du vélo/moto
629 courir
630 dire
631 voir
632 montrer
633 chanter
634 dormir
635 sourire
636 éclabousser
637 arrêter
638 nager
639 balancer
640 prendre
641 jeter
642 chatouiller
643 toucher
644 marcher
645 essuyer
646 écrire
647 construire
648 acheter
649 porter
650 attraper
651 courir après
652 faire bravo
653 sécher
654 déposer
655 trouver
656 aller bien avec
657 réparer
658 détester
659 avoir
660 entendre
661 cacher
662 tenir
663 couper
664 frapper à la porte
665 lécher
666 aimer bien
667 écouter
668 faire
669 ramasser
670 faire de la peinture
671 renverser
672 verser
673 faire semblant
675 déchirer
676 secouer
677 partager
678 s'asseoir
679 faire du patin
680 glisser
681 travailler
682 souhaiter
683 être debout
684 rester
685 balayer
686 parler
687 goûter
688 arracher
689 penser
690 attendre
691 se réveiller
692 couvrir
693 grimper
694 cuisiner

695 Est-ce qu'il arrive à votre enfant de parler d'événements passés ou de personnes absentes ? Par exemple un enfant qui a vu un défilé la semaine précédente peut dire plus tard "défilé", "orchestre", "musiciens".

oui ☐ non ☐

696 Est-ce qu'il arrive à votre enfant de parler de choses qui vont se produire dans le futur ? Par exemple, dire "tchou tchou" ou "avion" avant de quitter la maison pour voyager ou dire "balançoire" quand vous allez au parc.

oui ☐ non ☐

697 Est-ce qu'il arrive à votre enfant de parler d'objets qui ne sont pas présents comme parler d'un jouet manquant ou absent, se référer à un animal domestique hors de vue, ou poser des questions à propos d'une personne absente ?

oui ☐ non ☐

698 Est-ce qu'il arrive à votre enfant de comprendre si vous lui demandez quelque chose qui n'est pas dans la pièce, par exemple est-ce qu'il va dans la chambre à coucher chercher son ours en peluche si vous lui dites "où est l'ours ?"

oui ☐ non ☐

Grammaire

Parmi les phrases proposées, veuillez cocher la phrase la plus fréquemment produite par votre enfant en ce moment :

699 pour dire : la chaussure de maman
- a) chaussure maman ☐
- b) chaussure (d)e maman ☐
- c) chaussure à maman ☐
- d) chaussure de maman ☐

700 pour parler d'une chaise ou d'un frigidaire
- a) chaise / frigidaire ☐
- b) (l)a chaise / (l)e frigidaire ☐
- c) la chaise / le frigidaire ☐
- d) une chaise / un frigidaire ☐

701 pour dire : je veux de la confiture/du chocolat
- a) confiture / chocolat ☐
- b) je veux de la confiture / du chocolat ☐

702 pour dire : il est beau
- a) "yé" beau ☐
- b) l'est beau ☐
- c) il est beau ☐

703 pour dire : elle monte sur la chaise
- a) elle monte chaise ☐
- b) elle monte sur la chaise ☐

704 pour dire : le robot de Grégoire (en parlant de lui-même)
- a) robot Grégoire ☐
- b) robot à Grégoire ☐
- c) mon robot à moi ☐
- d) robot à moi ☐
- e) mon robot ☐

705 pour dire : il / elle cherche
- a) cherche ☐
- b) é cherche ☐
- c) il / elle cherche ☐

706 pour dire : je veux la chaise
- a) veux la chaise ☐
- b) veux la chaise Grégoire ☐
- c) moi veux la chaise ☐
- d) moi je veux la chaise ☐
- e) je veux la chaise ☐

707 pour dire : c'est une voiture
- a) ça voiture ☐
- b) ça c'est voiture ☐
- c) c'est une voiture ☐

708 pour dire : aller à l'école
- a) aller école ☐
- b) aller à école ☐

709 pour dire : c'est dans la valise
- a) c'est dedans la valise ☐
- b) c'est dans la valise ☐

710 pour dire : est-ce que c'est un chien?
- a) chien ? ☐
- b) ça chien ? ☐
- c) est chien ? ☐
- d) c'est un chien ? ☐
- e) est-ce que c'est un chien ? ☐

711 pour dire : les chiens sont là
- a) les chiens est là ☐
- b) les chiens sont là ☐

712 pour dire : je ne joue pas
- a) jouer pas ☐
- b) pas jouer ☐
- c) non pas jouer ☐
- d) non jouer ☐
- e) je joue pas ☐

713 pour dire : je veux descendre
- a) descendre ☐
- b) veux descendre ☐
- c) je veux descendre ☐

714 pour dire : je vais t'attaquer
- a) je vais s'attaquer ☐
- b) je vais t'attaquer ☐

Cochez les cases suivantes et donnez des exemples si votre enfant utilise des verbes à :

715 infinitif (*prendre*) ☐ _____

716 présent (*prends*) ☐ _____

717 passé composé (*ai pris*) ☐ _____

718 imparfait (*prenais*) ☐ _____

719 futur (*prendrai*) ☐ _____

720 impératif (*prends !*) ☐ _____

721 Donnez les trois phrases les plus longues que votre enfant produit en ce moment :

1) _____
_____
_____

2) _____
_____
_____

3) _____
_____
_____

MERCI POUR VOTRE PARTICIPATION

APPENDIX VIII

Matching decisions for the cross-linguistic equivalents analyses:

	English items	French items
1		allô
2	baa baa	bêê bêê
3	choo choo	tchou tchou
4	cockadoodledoo	cocorico
5	grr	grrrr
6	meow	miaou
7	moo	meuh
8	ouch	aïe
9	quack quack	coin-coin
10	uh oh	oh oh
11	vroom	vroum
12	woof woof	ouaf-ouaf
13	yum yum	miam-miam
14	animal	animal
15	ant	fourmi
16	bear	ours
17	bee	abeille
18		tendre
19	bird	oiseau
20	bunny / rabbit	lapin
21	butterfly	papillon
22	cat	chat
23		chèvre
24	chicken	poule
24	hen	
25	cockerel	coq
26	cow	vache
27	crocodile	crocodile
28	deer	biche
29	dog	chien
30	donkey	âne
31	duck	canard
32	elephant	éléphant
33	fish	poisson
34	frog	grenouille
35	giraffe	girafe
36	goose	oie
37	horse	cheval
38	insect/fly	petite bête
39	kitty	bébé chat
40	lamb	agneau
41	lion	lion
42	monkey	singe
43	moose	
44	mouse	souris
45	owl	hibou
46	penguin	pingouin
47	pig	cochon
48	pony	poney
49	puppy	bébé chien
50		renne
51	sheep	mouton
52	squirrel	écureuil
53	teddy bear	nounours
54	tiger	tigre
55	turkey	dindon
56	turtle	tortue
57	wolf	loup
58	zebra	zèbre
59	aeroplane	avion
60	bicycle	vélo
61	boat	bateau
62	bus	bus
63	car	voiture
64	fire engine	camion de pompier
65	helicopter	hélicoptère
66	lorry	camion
67	motorbike	moto
68		poussette
69	pram	
70	sledge	traîneau
71	tractor	tracteur
72	train	train
73	tricycle	tricycle
74	ball	balle
75	balloon	ballon
76	bat	
77	block	cube
78	book	livre
79	bubbles	bulles
80	chalk	craie
81	crayon	crayon
81	pencil	
82	doll	poupée
83	game	jeu
84	glue	colle
85	jigsaw	puzzle
86	pen	stylo
87		feutre
88	play dough	pâte à modeler

89	present	cadeau
90		raquette
91	story	histoire
92	toy	jouet
93	apple	pomme
94	applesauce	
95		baguette
96	banana	banane
97	beans	haricot
98	biscuit	petits gâteaux
99	bread	pain
100	butter	beurre
101	cake	gâteau
102	carrot	carotte
103	cereal	céréales
103		corn flakes
104	cheese	fromage
105	chewing gum	chewing-gum
106	chicken	poulet
107	chips	frites
108	chocolate	chocolat
109		clémentine
110	coffee	café
111	coke	coca
112		compote
114		courge
115	cracker	gâteaux apéro
116	crisps	chips
117	doughnut	beignet
118	drink	boisson
119	egg	œuf
120	fish	poisson
121		flan
122	food	nourriture
123	grapes	raisin
124	green beans	haricots verts
125	hamburger	hamburger
126	ice	glaçon
127	ice cream	glace
128	Icepop/icelolly	esquimau
129	jam	confiture
130	jelly	
131	juice/fizzy juice	jus de fruit
132		kiwi
133		limonade
134	lollipop	sucette
135		madeleine
136		mayonnaise
137	meat	viande
138	melon	melon
139	milk	lait

140	muffin	
141	noodles	pâtes
142		nutella
143	nuts	noisettes
144	orange	orange
145	pancake	crêpe
146	peanut butter	
147	peas	petits pois
148		petits pots
149	pickle	
150	pizza	pizza
151	popcorn	pop-corn
152	potato	pomme de terre
153	pretzel	bretzel
154	pudding	
155	pumpkin	
156		purée
157	raisin	raisins secs
158	salt	sel
159	sandwich	sandwich
160	sauce	sauce
161	soup	soupe
162	spaghetti	spaghetti
163	strawberry	fraise
164		sucre
165	sweet corn	maïs
166	sweets	bonbons
167	toast	tartine
168	tuna	thon
169	vanilla	vanille
170	vitamins	vitamines
171	water	eau
172	yogurt	yaourt
173	beads	perles
174	belt	ceinture
175	bib	bavoir
176		body
177	boot(s)	bottes
178	button	boutons
179	coat	manteau
180		combinaison de ski
181	dress	robe
182	gloves	gants
183		grenouillère
184	hat	chapeau
185	jacket	veste
186	jeans	jeans
187	jumper	pull
188	mittens	moufles
189	nappy	couche
190	necklace	collier

191	pyjamas	pyjama
192		salopette
193	scarf	écharpe
194	shirt	chemise
195	shoe	chaussure
196	shorts	short
197	slipper	chausson
198	sock	chaussettes
199	sweatshirt	sweat
200		tee-shirt
201	tights	collants
202	trainers	basket
203	trousers	pantalon
204	underpants	culotte/slip
205	zip	
206		aïe bobo
207	ankle	cheville
208	arm	bras
209	belly button	nombril
210	buttocks/bottom	fesses
211	cheek	joue
212	chin	menton
213		cœur
214	ear	oreille
215	eye	yeux
216	face	figure/visage
217	feet	pied
218	finger	doigt
219	hair	cheveux
220	hand	main
221	head	tête
222	knee	genou
223	leg	jambe
224	lips	lèvre
225	mouth	bouche
226	nose	nez
227	penis	pénis/zizi
228		pouce
229	shoulder	
230	tooth	dent
231	toe	doigt de pied
232	tongue	langue
233	tummy	ventre
234	vagina	vagin/zezette
235		balai
236	basket	panier
237	blanket	couverture
238	bottle	biberon
238		bouteille
239	bowl	bol
240	box	boite
241	brush	brosse

242	bucket	seau
243	camera	appareil photo
244	clock	horloge
245	comb	peigne
246	cup	tasse
247	dish	plat
248		feuille
249	fork	fourchette
250	glass	verre
251	glasses	lunettes
252	hammer	marteau
253	hoover	aspirateur
254	can	pot
254	jar	
255	keys	clefs
256	knife	couteau
257	lamp	lampe
258	light	lumière
259	medicine	médicaments
260	money	argent
261	mop	serpillière
262		musique
263	nail	clou
264	napkin	serviette de table
265	paper	papier
266	penny	sous/pièces
267	picture	photo
268	pillow	coussins
268		oreiller
269	plant	plante
270	plate	assiette
271		plateau
272		poubelle
273	purse	porte-monnaie
274	radio	radio
275	rubbish	ordures
276	scissors	ciseaux
277		sirop
278	soap	savon
279	spoon	cuillère
280		sucette
281	tape	cassette
282		télécommande
283	telephone	téléphone
284	tissue/kleenex	mouchoir
285	toothbrush	brosse a dent
286	towel	serviette
287		trotteur
288	watch	montre
289	bath	baignoire
290	bathroom	salle de bain
291	bed	lit

292	bedroom	chambre
293	bench	banc
294	cellar	cave
295	chair	chaise
296	cot	berceau
297	door	porte
298	drawer	tiroir
299	dryer	étendage
300	garage	garage
301	high chair	chaise haute
302	kitchen	cuisine
303		
304	living room	salon
305	oven	four
306	play pen	parc
307	porch	
308	potty	pot
309	refrigerator	frigo
310	rocking chair	fauteuil
311	room	pièce
312	shower	douche
313	sink	lavabo
313		évier
314	sofa	canapé
315	stairs	escalier
316	stove	cuisinière
317	table	table
318	TV / television	télé
319	wardrobe	
320	washing machine	machine à laver
321		wc
322	window	fenêtre
323		arrosoir
324	backyard	
325	cloud	nuage
326		
327	flag	drapeau
328	flower	fleur
329	garden	jardin
330	grass	herbe
331	hose	tuyau
332	ladder	échelle
333	lawn mower	tondeuse a gazon
334	moon	lune
335	pavement	trottoir
336	pool	piscine
337	rain	pluie
338	rock	rocher
339	roof	toit
340	sandpit	bac a sable
341	sky	ciel
342	slide	toboggan

343	snow	neige
344	snowman	bonhomme de neige
345	spade	pelle
346	sprinkler	
347	star	étoile
348	stick	bâton
349	stone	caillou
349		pierre
350	street	rue/route
351	sun	soleil
352	swing	balançoire
353	tree	arbre
354	water	eau
355	wind	vent
356	beach	plage
357	camping	camping
358	church	église
359	circus	cirque
360	city centre, town	centre ville
361	country	campagne
362	farm	ferme
363	film	cinéma
364	home	maison
364	house	
365	outside	dehors
366	park	parc
367	party	fête
368	petrol station	station service
369	picnic	pique-nique
370	playground	terrain de jeux
371	school	école
372	shop	magasin
373	woods	forêt
374	work	travail
375	yard	cour
376	zoo	zoo
377	aunt	tante
378	baby	bébé
379	babysitter	
380	babysitter's name	
381	boy	garçon
382	brother	frère
383	child	enfant
384	clown	clown
385	cowboy	
386	daddy	papa
387	doctor	docteur
388	fireman	pompier
389	friend	copain/ine
390	girl	fille
391	grandma	grand-mère

392	grandpa	grand-père
393	lady	dame
394	man	monsieur
395	mummy	maman
396		nounou
397		nom de la nounou
398	nurse	infirmière
399	child's own name	nom de l'enfant
400	people	gens
401	person	personne
402	pet's name	nom de l'animal domestique
403	police	police
404	postman	facteur
405	sister	soeur
406	teacher	maître/sse
407	uncle	oncle
408	bath	bain
409		bravo
410	breakfast	petit déjeuner
411	bye	au revoir
412	dinner	dîner
413	give me five!	top là
414	gonna get you!	va sur le pot
415	go potty	je vais t'attraper
416	hello	bonjour
417	hi	salut
418	lunch	déjeuner
419	nap	sieste
420		ne fais pas
421	night night	bonne nuit
422	no	non
423	patty cake	ainsi font font
424	peek-a-boo	coucou
425	phone	coup de fil
426	please	s'il te plaît
427	shh / hush / shush	chut
428	shopping	faire les courses
429	snack	goûter
430	so big!	
431	thank you	merci
432	this little piggy	ce petit cochon
433	turn around	tourne-toi
434	yes	oui
435	bite	mordre
436	blow	souffler
437	break	casser
438	bring	apporter
439	build	construire
440	bump	se cogner
441	buy	acheter
442	carry	porter

443	catch	attraper
444	chase	courir après
445	clap	faire bravo
446	clean	nettoyer
447	climb	grimper
448	close	fermer
449	cook	cuisiner
450	cover	couvrir
451	cry	pleurer
452	cuddle	
453	cut	couper
454	dance	danser
455	draw	dessiner
456	drink	boire
457	drive	conduire
458	drop	
459	dry	sécher
460	dump	déposer
461	eat	manger
462	fall	tomber
463	feed	nourrir
464	find	trouver
465	finish	finir
466	fit	aller bien avec
467	fix	réparer
468	get	recevoir
469	give	donner
470	go	aller
471		goutter
472	hate	détester
473	have	avoir
474	hear	entendre
475	help	aider
476	hide	cacher
477	hit	taper
477	kick	donner un coup
478	hold	prendre dans ses bras
478		tenir
479	hurry	se dépêcher
480	jump	sauter
481		
482	kiss	faire un bisou
483	knock	frapper a la porte
484	lick	lécher
485	like	aimer bien
486	listen	écouter
487	look	regarder
487	watch	
488	love	aimer
489	make	faire
490	open	ouvrir

491	paint	faire de la peinture
492	pick	ramasser
493	play	jouer
494	pour	verser
495		
496	pretend	faire semblant
497	pull	tirer
498	push	pousser
499	put	mettre
500	read	lire
501	ride	faire du vélo/moto
502	rip	déchirer
502	tear	arracher
503	run	courir
504	say	dire
505	see	voir
506	shake	secouer
507	share	partager
508	show	montrer
509	sing	chanter
510	sit	s'asseoir
511	skate	faire du patin
512	sleep	dormir
513	slide	glisser
514	smile	sourire
515	spill	renverser
516	splash	éclabousser
517	stand	être debout
518	stay	rester
519	stop	arrêter
520	sweep	balayer
521	swim	nager
522	swing	balancer
523	take	prendre
524	talk	parler
525	taste	goûter
526		
527	think	penser
528	throw	jeter
529	tickle	chatouiller
530	touch	toucher
531	wait	attendre
532	wake	se réveiller
533	walk	marcher
534	wash	laver
535		
536	wipe	essuyer
537	wish	souhaiter
538	work	travailler
539	write	écrire
540	all gone	parti

541	asleep	endormi
542	awake	être réveillé
543	bad	pas bon
544	better	mieux
545	big	grand
546	black	noir
547	blue	bleu
548	broken	casse
549	brown	marron
550	careful	attention
551	clean	propre
552	cold	froid
553		coquin/e
554	cute	mignon/ne
555	dark	sombre
556	dirty	sale
557	dry	sec/che
558	empty	vide
559	fast	vite
560	fine	
561	first	premier/e
562	full	plein/ne
563	gentle	
564	good	bien
564		bon/ne
565	green	vert/e
566	happy	content/e
567	hard	dur
568	heavy	lourd/e
569	high	haut/e
570	hot	chaud/e
571	hungry	avoir faim
572	hurt	blesse
573	last	dernier/e
574	little	petit/e
575	long	long/ue
576	loud	fort/e
577	mad	fou/folle
578	naughty	vilain/e
578		méchant/e
579	new	neuf/ve
580	nice	gentil/le
581	noisy	bruyant/e
582	old	vieux/vielle
583	orange	orange
584	poor	
585	pretty	joli
585		beau/belle
586	quiet	tranquille
587	red	rouge
588	sad	triste
589	scared	avoir peur

590	sick	malade
591	sleepy	avoir sommeil
592	slow	lent/e
593	soft	doux/ce
594	sticky	collant/e
595	stuck	coince
596	thirsty	avoir soif
597	tiny	minuscule
598	tired	fatigue
599	wet	mouille
600	white	blanc/he
601	windy	venteux
602	yellow	jaune
603	yucky	dégoûtant/e
604	after	après
605	before	avant
606	day	jour
607	later	
608	morning	matin
609	night	nuit
610	now	maintenant
611	time	heure
612	today	aujourd'hui
613	tomorrow	demain
614	tonight	ce soir
615	yesterday	hier
616	he	il
617	her	lui
617		a elle/sa
617		elle
617		ses
618	hers	a elle/sa
618		ses
619	him	lui
620	his	a lui/son
620		ses
621	I	je
622	it	ça
622		il
622		elle
622		on
623	me	moi
624	mine	a moi
625	my	ma/mon/mes
626	myself	moi-même
627	our	notre
628	she	elle
629	that	ça
629	this	
630	their	leur
631	them	eux
632	these	ces

632	those	ceux
633	they	ils/elles
634	us	nous
635	we	nous
635		on
636	you	vous/tu
637	your	votre/vos/ta/ton
638	yourself	toi-même
639	how	comment
640	what	quoi
641	when	quand
642	where	où
643	which	le/la/les/quel(les)
644	who	qui
645	why	pourquoi
646	about	au sujet de
647	above	au dessus de
648	around	autour de
649	at	a
649	to	
650	away	loin
651	back	
652	behind	derrière
653	beside	a cote de
653	next to	près de
654	by	
655		chez
656	down	en bas
657	for	pour
658	here	ici
658		la
659	inside/in	a l'intérieur de
659	into	dans
660		
661		
662	of	de
663	off	
664	on	sur
665	on top of	au sommet de
666	out	au loin
667	over	par dessus
668	there	là-bas
669		vers
670	under	sous
671	up	en haut
672	with	avec
673	a	un/une
673	an	
674	all	tous/tout
674	every	
675	a lot	plein/beaucoup
676	another	un autre

677	any	
678	each	chaque
679	more	encore
680	much	un peu
681	not	pas
682	none	aucun/ne
683	other	autre
684	same	le/la même
685	some	du/de la/des
686	the	le/la/les
687	too	aussi
688	am	suis
689	are	sont
690	be	être
691	can	peux
692	could	pourrait
692	would	
693	did/did you	a fait
694	do	faire
695	does	fait
696	don't	ne pas
697	gonna/going to	aller
698		
699	gotta/got to	avoir a faire
699	hafta/have to	devoir faire
700	is	est
701	lemme/let me	laisse-moi
702	need/need to	avoir besoin de
703	try/try to	essayer de
704	want to	vouloir
705	was	était
706	were	
707	will	
708		
709	and	et
710	because	parce que
711	but	mais
712	if	si
713	so	donc
714	then	alors
715		entrée
716		garderie
717		crèche

APPENDIX IX

Lexical data (native parents only):

English **French** Total

16												
THO	**FLO**	**ANT**	**EMM**	**DUN**	**PEN**	**ELI**	**JAC**	**ANN**	**OLR**	**LIA**	**REB**	**OLB**
	28	**6**	**10**	**2**		**5**	**10**	**19**	**8**	**4**	**3**	**1**
	20	32	3	6		12	14	11	7	2	10	4
	48	38	13	8		17	24	30	15	6	13	5
17												
THO	**FLO**	**ANT**	**EMM**	**DUN**	**PEN**	**ELI**	**JAC**	**ANN**	**OLR**	**LIA**	**REB**	**OLB**
7	38	26		7	26	15	19	27	8		2	4
4	21	67		8	91	11	11	24	9		13	3
11	59	93		15	117	26	30	51	17		15	7
18												
THO	**FLO**	**ANT**	**EMM**	**DUN**	**PEN**	**ELI**	**JAC**	**ANN**	**OLR**	**LIA**	**REB**	**OLB**
	22	**22**	**4**	**5**	**41**	**17**	**20**	**50**	**13**			**5**
	34	**'3**	**16**	**9**	**116**	**23**	**21**	**28**	**15**			**5**
	56	115	20	14	157	40	41	78	28			10
19												
T)	**FLO**	**ANT**	**EMM**	**DUN**	**PEN**	**ELI**	**JAC**	**ANN**	**OLR**	**LIA**	**REB**	**OLB**
	30	**30**		**5**		**32**	**28**	**65**	**11**		**4**	**28**
	41	**1 4**		**11**		**62**	**25**	**36**	**16**		**11**	
	71	174		16		94	53	101	27		15	
20												
THO	**FLO**	**ANT**	**EMM**	**DUN**	**PEN**	**ELI**	**JAC**	**ANN**	**OLR**	**LIA**	**REB**	**OLB**
18	**45**	**31**	**10**	**6**		**57**	**28**	**107**	**21**	**6**	**21**	
14	**46**	**139**	**20**	**19**		**76**	**33**	**48**	**27**	**5**		**33**
32	91	170	30	25		133	61	155	48	11		
21												
THO	**FLO**	**ANT**	**EMM**	**DUN**	**PEN**	**ELI**	**JAC**	**ANN**	**OLR**	**LIA**	**REB**	**OLB**
26	**88**	**14**	**9**	**9**	**74**		**30**	**178**	**22**		**13**	**62**
24	**5)**	**130**	**21**	**19**	**203**	**126**	**40**	**90**	**43**		**23**	**57**
50	138	144	30	28	277		70	268	65		36	119
22												
THO	**FLO**	**ANT**	**EMM**	**DUN**	**PEN**	**ELI**	**JAC**	**ANN**	**OLR**	**LIA**	**REB**	**OLB**
21	**137**			**8**	**67**	**103**	**39**	**215**	**27**			
	171	**219**		**21**	**36**	**137**	**39**	**147**	**51**			
	308			29	429	240	78	362	78			
23												
THO	**FLO**	**ANT**	**EMM**	**DUN**	**PEN**	**ELI**	**JAC**	**ANN**	**OLR**	**LIA**	**REB**	**OLB**
35	**197**		**24**	**10**	**52**	**159**	**49**	**276**	**33**	**21**	**21**	**76**
19	**183**	**275**	**46**	**23**		**185**	**44**	**237**	**63**	**12**	**28**	**80**
54	380		70	33		344	93	513	96	33	49	156
24												
THO	**FLO**	**ANT**	**EMM**	**DUN**	**PEN**	**ELI**	**JAC**	**ANN**	**OLR**	**LIA**	**REB**	**OLB**
	214	**53**		**10**		**134**	**51**	**339**	**43**		**29**	**119**
26	**182**	**310**		**33**		**209**	**46**	**307**	**75**		**49**	**111**
	396	363		43		343	97	646	118		78	230

25												
THO	FLO	ANT	EMM	DUN	PEN	ELI	JAC	ANN	OLR	LIA	REB	OLB
52	244	137		10		218		364	47			175
53	319	388		43		235		410	82			166
105	563	525		53		453		774	129			341
26												
THO	FLO	ANT	EMM	DUN	PEN	ELI	JAC	ANN	OLR	LIA	REB	OLB
63	281			10	50		60	379			56	
88	303			51	422		59	429			54	
151	584			61	472		119	808			110	
27												
THO	FLO	ANT	EMM	DUN	PEN	ELI	JAC	ANN	OLR	LIA	REB	OLB
92	300			6			78	390	89	76		195
73	430			66			58	442	119	51		263
165	730			72			136	832	208	127		458
28												
THO	FLO	ANT	EMM	DUN	PEN	ELI	JAC	ANN	OLR	LIA	REB	OLB
116	347	241		18				417	134			243
1 3	4 6	5 8		61				480	186			340
219	793	789		79				897	320			583
29												
THO	FLO	ANT	FMM	DUN	PEN	ELI	JAC	ANN	OLR	LIA	REB	OLB
	403			23			133	439	226		98	338
	. 9			84			67		285			423
	912			107			200		511			761
30												
THO	FLO	ANT	EMM	DUN	PEN	ELI	JAC	ANN	OLR	LIA	REB	OLB
	487			23			202	436	218	129	153	366
	5			86			121	519	248	121	103	451
	993			109			323	955	466	250	256	817