Emergent academic language at home and at school

Laghzaoui, M.

Publication date:
2011

Emergent academic language at home and at school

Emergent academic language at home and at school

A longitudinal study of 3- to 6-year-old Moroccan Berber children in the Netherlands

P R O E F S C H R I F T

ter verkrijging van de graad van doctor
aan Tilburg University
op gezag van de rector magnificus,
prof. dr. Ph. Eijlander,
in het openbaar te verdedigen ten overstaan van een
door het college voor promoties aangewezen commissie
in de aula van de Universiteit

op vrijdag 9 september 2011 om 10.15 uur

door

Mohammadi Laghzaoui
geboren op 20 januari 1971 te Nador, Marokko

PROMOTIECOMMISSIE

Promotores	prof. dr. Sjaak Kroon prof. dr. Ton Vallen †
Copromotores	dr. Abderrahman El Aissati dr. Jeanne Kurvers
Overige leden	prof. dr. Fatima Agnaou prof. dr. Jan Blommaert prof. dr. Yamina El Kirat prof. dr. Mena Lafkioui prof. dr. Ulrich Mehlem prof. dr. Kris Van den Branden

ISBN	978-90-8891-316-7
NUR	616
Cover design	Proefschriftmaken.nl || Printyourthesis.com
Lay-out	Carine Zebedee, Tilburg
Printed by	Proefschriftmaken.nl || Printyourthesis.com
Published by	Uitgeverij BOXPress, Oisterwijk

To the memory of my beloved mother
To the memory of my supervisor Ton Vallen

Table of contents

Preface

A Moroccan Berber baby in the Netherlands hears its mother talking in Berber and starts to imitate her. Slowly, speech emerges. In time, one or more other languages will come into play - a rich inheritance that can ease learning and open doors but also a big load to carry. The period that may set the child up for success or failure is when that second language comes along, jostling with the first, pushing some language out, re-enforcing other parts, devaluing or revalidating past achievements.

This PhD dissertation is about this very period - the academic language development of Moroccan Berber children in the Netherlands before the age of six. It is part of a larger research project entitled 'Development of Academic Language in School and at Home' (DASH). The DASH project was funded by the Netherlands Organization for Scientific Research (NWO). The grant number of the Berber study is 411-03-063. I am grateful to NWO for the funding provided.

My scientific journey and the realization of this book could not have been finished without the help and support of many people. I am happy to acknowledge those who contributed to the completion of this work, although I know I cannot do them all justice.

First and foremost I owe my deepest gratitude to my late supervisor Ton Vallen for his guidance from the very beginning of this dissertation. We started this endeavor together but he passed away without witnessing the results of his inspiration and encouragement. I would also like to thank my supervisor Sjaak Kroon who, without any hesitation stepped in and took over the task of supervision. Sjaak, although you joined the DASH project at a later stage, I benefited a lot from your broad expertise and your critical comments. I also wish to thank my co-supervisor Jeanne Kurvers for her enthusiasm and dedication. Jeanne, exactly one year ago, you told me in the corridor: *'Dit is het jaar van de waarheid'* (this is the year of truth). Thank you for your supervision,

your methodological advice and your assistance with statistical analyses. Further, I want to express my thanks and gratitude to my co-supervisor Abderrahman El Aissati for his encouragements and useful comments. Abder, in the last year, it was a pleasure to share your office and enjoy the immense and interesting discussions which many times went beyond the area of my dissertation. '*Shukran aṭṭaṣ'* (thank you very much). I also want to express my thanks to Rian Aarts for her involvement and constructive comments. From the very start of the DASH project, you played an important role in coordinating the project at Tilburg University. In the last stage of my research, you were willing to read and comment on earlier drafts of the manuscript. Thank you very much.

Thanks are also due to my roommate Serpil Demir, with whom I shared the ups and downs of conducting research. It was a pleasure to work together with you and also to attend many conferences and courses together especially the LSA (Linguistic Society of America) institute at Stanford University in the USA. My thanks go also to Yahya E-rramdani, who worked as a post-doc within this project.

In conducting the fieldwork, I received the assistance of many individuals. My special thanks go to all children, parents and teachers for their loyal cooperation and persistence, staying involved in this project for more than three years. I gratefully acknowledge the group of research assistants who helped in collecting the data and transcribing parts of it. I particularly mention Senna Bouteba, Fatiha Elmassoudi, Véronique Verhagen and Naima Barkane.

I am also grateful for having been part of the DASH research group including colleagues with complementary interests and expertise. I would like to express my thanks to Paul Leseman, the project leader, for inspiring us with his ideas and his multidisciplinary interests. I also want to thank Aziza Mayo who did a good job coordinating all the projects especially when we started the data collection. My thanks also go to all the other members of the DASH project for their invaluable discussions especially during the first stage of the project.

I also would like to thank the manuscript committee, consisting of Fatima Agnaou, Jan Blommaert, Yamina El Kirat, Mena Lafkioui, Ulrich Mehlem and Kris Van den Branden for reading, assessing and commenting on my manuscript.

I feel privileged to have had the opportunity to study at Tilburg University and later on to work as a PhD researcher at the Faculty of Humanities and the

Babylon research center. It was and still is a distinct pleasure to be in a stimulating environment to conduct research. I would like to thank Guus Extra for his inspiration and his insightful lectures and discussions. Guus, I admire your boundless energy and your passion for the issues of multilingualism and multiculturalism. I also want to extend my thanks to Jan Jaap de Ruiter, who was my Master's thesis's supervisor, and with whom I still collaborate in publications and translations. I would also like to express my thanks to Ad Backus to whom I could always turn for advice and help. Ad, I appreciate your accessibility and the endless energy and time you give to research. I thank also Peter Broeder for his stimulation and the pleasant *pep talks*. In moments of frustration you often comforted me by saying: '*Het is maar een proefschrift!*' (it is just a dissertation).

I would also like to thank all the PhD students, those who finished and the ones who still yet have to complete their dissertations, for sharing their experiences with me and for the cheerful work environment. Among them are Juliëtte, Roel, Tamara, Seza, Max, Elma, Maria, Yonas, Geke, Kasper, Nadia, Dong Jie, Dorina, Danielle, Marianne, Jinling, Logan and Xuan.

I am grateful to have had the opportunity to present some of the results of this study at many international conferences. During my PhD period, I was also fortunate to attend many LOT (Netherlands Graduate School of Linguistics) winter and summer schools at different Dutch universities. I also appreciated the opportunity of attending the 2007 LSA institute at Stanford University – California. Many courses have been inspiring especially those taught by Michael Tomasello, William Labov, Eve Clark, William Snyder and Adele Goldberg.

Many thanks to Maeve Moynihan for her patience and for being ready to read all the chapters of this book. Her comments and suggestions were invaluable. I also thank Hans Verhulst for his comments.

I am also thankful to the department secretaries, Carine Zebedee and Karin Berkhout who were always there for help. Carine, thank you very much for the time you spent on the layout of this book and for the editorial work.

I wish to thank Pieter van Dongen, one of my non-linguist friends, for his pleasant talks. It was a pleasure to share your discussions about world cultures and many other subjects. To my youth friends Mimoun Kaddouri and Hassan Riani, I wish to extend my special thanks for their interest in all I do despite the distance.

I am glad to find my brothers Taieb and Miloud willing to be my paranymphs and stand by my side during the defense of my dissertation. Thank you so much for all your encouragements. I also owe a tremendous debt to my father for all he did for me. I likewise thank my family in Morocco, my sister Zahra and my sister-in-law Zahra for their support. A special thank you goes to my parents-in-law for their continuous stimulation.

The last phase of this dissertation was a turbulent period for me. With sad heart, I suddenly lost my dearest mother after a short sickbed. Dear Mom, *ymma inu i ɣizzen*, I will eternally be grateful for everything you have done for me. Shortly after, I lost my supervisor, Ton Vallen. Dear Ton, I cherish your discussions on academic but also so many other issues. You always were there to offer support. To you Mom, to you Ton, I dedicate this book.

Last and most importantly, a final word of thanks and sincere gratitude to my wife Esmah, my son Ilyes and my daughter Noor. I am incredibly lucky to have such a wonderful wife and partner who is always there for me in stressful moments. Esmah, writing a dissertation, as you yourself know, is a solo project but your being there has been fundamental for its completion. Thank you for your unconditional love and endless support. Ilyes, your talks about your MacBook and your rich presentations on skateboarding maneuvers were an enjoyable source of distraction. Noor, your arrival to this world may have delayed the publishing of this work but has indeed *lightened* our small family. I promise that I will read you more books and tell you more stories, after I have spent so many evenings analyzing the book reading sessions of other children for use in this book.

Mohammadi Laghzaoui
Tilburg, July 2011

CHAPTER 1

Introduction

Although language development in general has been widely investigated, only few studies have been carried out specifically focussing on the emergence and development of the academic language register. Not much is known in detail about the specific linguistic features and the determinants of this register, the impact of language input in the home and later in the school context, the role and impact of bilingual development in this respect, and the possible transfer from the first language (L1) to the second language (L2). It has been extensively documented that the language input to which children are exposed is related to their language development (Hart & Risley, 1995; Hoff, 2006; Huttenlocher *et al.*, 2002). However, still little is known about the specific language input needed to contribute to children's academic language development (Cox, Fang & Otto, 1997). We assume that children acquire academic language features as part of their language development in general. This assumption is based on the usage-based approach in studying children's language acquisition (Lieven, Behrens, Spears & Tomasello, 2003; Tomasello, 2000; Tomasello, 2003).

Cummins (1979) introduced the hypothesis of language interdependence, which suggests that in bilingual development linguistic skills can be transferred from L1 to L2 through a common mental mechanism underlying both languages. Scientific empirical evidence for this assumption, however, is mainly based on studies conducted on older bilingual middle-class children acquiring English and French, when both languages enjoy a high and equal social status; but when, for instance, national standard languages are paired with immigrant minority languages and dialects their status is not equal, though both are used in daily communication as well as in formal contexts (Bialystok, 2001). This is obviously different from the language situation of bilingual Moroccan Berber children growing up in the Netherlands. In the latter case, the status of L1, Berber, is much lower than L2, Dutch, which is acquired and used by these children for informal purposes as well as in the school context. In contrast to

Dutch, Berber is predominantly used orally in informal and non-academic contexts and hardly ever in education and the media. Furthermore, there are linguistic and typological differences between Berber and Dutch (see Chapter 2).

1.1 The DASH project

The study reported on in this book is part of the DASH (Development of Academic Language at School and at Home) project. This project is a large scale collaborative research program carried out at three universities in the Netherlands: Tilburg University, Amsterdam University and Utrecht University. This project is a longitudinal and interdisciplinary investigation on academic language development of 3-6-year-old Dutch, Moroccan Berber-Dutch and Turkish-Dutch children in communicative contexts at home and at school. The project consisted of one comprehensive study and three in-depth studies. The comprehensive study involved three ethno-linguistic groups, namely Dutch, Turkish and Moroccan Berber. The in-depth studies each focused on a subsample of these groups. This book reports on the in-depth study on the Moroccan Berber group.

The comprehensive study considered a large sample consisting of 58 Dutch monolingual children using Dutch as a home language, 56 bilingual Turkish children using Turkish as a home language, and 48 bilingual Moroccan Berber children using Berber, more specifically Tarifit Berber, as a home language. Dutch and Turkish are standardized languages with a long literary and a full fledged academic register; Tarifit Berber on the other hand is mainly an, only recently scripted and written, oral language without an anchored academic register. The official Berber script is not widely used for school purposes. In addition, children in the Netherlands do not have any acquaintance with it. As children from ethnic minority groups are expected to profit from their L1 in developing L2, this study aimed to examine how language status and access to the literate register of a language have an impact on academic language development through patterns of language use in the home setting (Scheele, 2010). By combining different disciplines, the comprehensive study aimed to understand, among other things, the impact of home language, literacy activities and socio-economic background on the language skills which are considered crucial for schooling. A number of instruments has been used to investigate

different aspects related to the development of academic language. In addition to general questionnaires designed for mothers and teachers, also linguistic and cognitive tests were developed to investigate children's language and cognitive development in L1 and L2. The study also aimed at examining possible effects of experiences with oral and literacy activities in the home language environment, the children's verbal short-term memory and possible positive cross-language transfers from L1 to L2 or a possible competition between L1 and L2.

The three separate longitudinal in-depth studies, conducted simultaneously with the general study, are of a linguistic nature. The primary goal of these studies is to gain an insight into features of academic language in the input of mothers and teachers as well as in the language output of the children. Each of these studies concentrated on one ethno-linguistic group (Dutch, Turkish and Moroccan Berber). To understand the specific academic language features in the input and their relationship to academic language output and development in each of these groups, detailed linguistic analyses covering lexical, morpho-syntactic and textual features were conducted. Of the 162 families that took part in the comprehensive study, sub-samples of each group were involved in the in-depth studies. The current study exclusively focuses on the Moroccan Berber group. The Turkish and Dutch groups are reported on in Demir (forthcoming), Aarts, Demir & Vallen (forthcoming) and Henrichs (2010).

1.2 Scientific relevance

It may not be surprising that Moroccan Berber children, not unlike other bilingual immigrant children seem to easily overcome problems with basic language vocabulary in L2 and other oral language skills needed in daily informal settings of interaction and in interpersonal communication in general. They, however, have considerable arrears at the level of specific vocabulary and other verbal skills required in formal conversations and instruction contexts. A similar pattern appears among bilingual immigrant children regarding the development of reading skills. As far as technical literacy skills are concerned, bilingual children seem to easily keep pace with their Dutch peers. At the level of reading comprehension, however, the differences between bilingual immigrant children and their Dutch peers seem to increase in the course of

primary education (De Jong & Leseman, 2001). This lag at higher linguistic levels creates a more complicated situation than simply a smooth transition from L1 at home to L2 at school.

The learning of the academic language register does not necessarily start with school attendance, but probably well before, at home. The degree of literacy of the parents on the one hand, and the literacy activities which parents carry out with their children on the other, are significant factors in this early process of school language development (Hoff & Naigles, 2002; Huttenlocher *et al.*, 2002). These factors influence the parents' type of language as well as the communicative styles they use with their children. In the case of bilingual children, it is assumed that the acquisition of school language skills in a second language is also affected by the literacy competences of the parents in the first language, i.e. that the use of the school register by both parents or by one of them in the first language stimulates the child to develop cognitive and linguistic competences in the first language, which can then be transferred to the second language when they go to school (Leseman, 2000; Kroll & De Groot, 2005; Bialystok, 2007).

In the last decades, many studies have concentrated on different aspects of school language development, such as narrative skills, technical vocabulary development, and emerging literacy (Leseman & Van den Boom, 1999). However, there are relatively few studies dealing with the acquisition and early development of specific linguistic structures, typical of school language, in the first language, let alone on their development in a language acquired and used as a second language. This concerns the acquisition of specific lexical, morphological and grammatical structures, which make complex language constructions possible, and accordingly facilitate the development of the academic language register.

Most studies on school language development of bilingual children in the Netherlands focus on the acquisition of Dutch as a second language in the higher grades of primary school and pay less attention to the pre-school and kindergarten period. Another limitation of the studies dealing with academic language development until now is that their focus is usually on tasks and activities that take place in the classroom without considering in-depth linguistic analyses of child-teacher or child-child interactions (Hajer & Meestringa, 1995). To date, there is hardly any longitudinal research dealing with bilingual children's school language development, wherein the children's

language output is studied in combination with the input of caretakers at home and at school (Weizman & Snow, 2001). A last aspect that did not receive sufficient attention in studies on language development of bilingual immigrant children in the school context, is the relationship between language acquisition and the home environment. Given the limitations of the studies presented above, this study attempts to build upon existing findings in order to investigate the social co-construction of the academic language register of Moroccan Berber children in the Netherlands. Furthermore, although the development of academic language received ample attention, most studies have been conducted outside the Netherlands in countries where English is the main national language, such as the United States, Australia, Canada and Great Britain. To the best of our knowledge, little research has been done on the specific characteristics of levels of academic language of Moroccan Berber children in the Netherlands where the national official language is the medium of instruction in education but, as a consequence of their immigration background, not the home language of all children.

1.3 Societal relevance

The main goal of this study is to investigate the development of the academic language register in Moroccan Berber children from the age of 3 to 6 years at home and at school. Most of these children grow up as bilinguals starting with Tarifit Berber at home and learning Dutch as their second language at a later stage. From the age of 4 onwards, a linguistic spurt takes place at the level of their Dutch language skills after entering primary school. The emerging academic language register, which is central in this study, is a socially constructed and dynamic phenomenon in which linguistic, cognitive and emotional factors are interwoven.

In the last decades, many researchers, educators and policy makers in the Netherlands expressed their concerns about the many and various problems of children with an immigration background in education. One of the persistent problems among Moroccan children, in general, is their substantial educational disadvantage. There are, however, hardly any studies on the determinants of this educational disadvantage. Different studies have shown that Moroccan and Turkish pupils are lagging far behind their native Dutch peers in school

achievements in Dutch primary education (Dagevos, Gijsberts & Van Praag, 2003; Tesser & Iedema, 2001). At the end of primary education, Moroccan and Turkish children on average leave school with a Dutch language delay of roughly two years as compared to native Dutch middle-class children (Dagevos & Gijsberts, 2007). This unfavourable educational position is usually associated with the fact that these children are exposed to another language than Dutch at home. We assume, however, that there are also other factors that might explain educational underperformances such as the (limited) acquisition and command of academic language skills. For many children, school language is different from the language used at home. Bilingual children especially face a double challenge: they have to learn not only a new language but also the school-based academic register of this new language, which may cause an extra obstacle for them. In doing research in this field, this study can possibly contribute to solving some educational and social problems for immigrant children in the Netherlands.

The current study focuses on the development of the academic language register of 3-6-year-old Moroccan Berber children of mainly low socio-economic background in their pre-school and kindergarten period. Using a multidisciplinary approach makes it possible to get better insights into both the home and the school environment and the transition between home and school, and helps in understanding the influence of these two different contexts on academic language development of bilingual children. This can contribute to new and suitable early education strategies for children with different social and cultural backgrounds and improve the position of immigrant minority children in Dutch education. In this in-depth study, specific lexical, morpho-syntactic and textual parameters will be investigated in both the input of caretakers and the output of children.

The family is considered as an 'informal context' in which the learning process takes place that prepares children to face the demands of formal schooling. Various studies on mono- and bilingual children have established that early informal linguistic input and communication patterns in the home environment are strongly related to children's development regarding intelligence, cognition, independent social functioning, communication skills and emergent literacy (Bus *et al.*, 1995; Gottfried *et al.*, 1998; Landry *et al.*, 2000; Weizman & Snow, 2001). Informal learning practices at home are related to the socio-economic status and ethnic-cultural background of the families. In

addition, the oral and written language use of the parents has an impact on the learning process in the family context.

1.4 Overview of the present dissertation

This book consists of eight chapters. The first three chapters provide general information on this study, a discussion of Berbers and their languages in Morocco and in the Netherlands and a theoretical discussion of the concept of academic language. Chapter 4 deals with the research design. The following three chapters (5, 6 and 7) present three empirical studies that deal with different aspects of academic language. Chapter 8, finally, provides general conclusions, a critical discussion and practical implications.

The present chapter has provided a general introduction of the current study and a brief summary of the DASH project on academic language development in the Netherlands. Additionally, the scientific and societal relevance of the project have been presented and discussed.

The objective of Chapter 2 is twofold. Apart from a brief introduction of the Berber language, this chapter deals with sociolinguistic aspects of Berber in Morocco and more specifically existing attitudes towards this and other co-existing languages in this North African country. Second, the position of the Berber language and the Berber community in the Netherlands are expounded. To this end, the status of Berber speakers in Dutch society and the status of Berber as a language in different sectors of society are described.

Chapter 3 presents the conceptual framework of this study and reviews different studies that dealt with academic language. In this chapter, the notion of 'academic language' is explained and the linguistic variables of academic language are described.

Chapter 4 deals with the methodology and design of this study. It presents the main research questions of the study. It also describes the procedure followed in the selection of the study's subjects. Subsequently, the research instruments and procedures used during data collection are outlined. Finally, data collection, data processing, construction of the database and the coding of the transcripts are explained.

Chapters 5, 6 and 7 present the results of three empirical studies. In Chapter 5, the lexical level of academic language is treated. Firstly, the lexical features

typical of school language and the measures used in their operationalization are highlighted. Then, the results regarding the use of lexical features in the input of caretakers at home and at school and in the output of children in both settings are presented. The chapter ends with a discussion of the outcomes.

Chapter 6 reports on the morpho-syntactic level of academic language. The aspect investigated here is the use of multi-clause sentences. The chapter opens with a theoretical discussion of clause combining and complex constructions. Then, the results of clause combining features in the input of caretakers at home and at school as well as in the output of children are presented. This chapter concludes with a summary and a discussion of the results.

Chapter 7 is devoted to the textual level of academic language. The aspect investigated here is decontextualization as a feature of academic language. First, the measures adopted in the operationalization of decontexualized language are introduced. After that, the results of decontextualized language features in mothers' and teachers' input, as well as children's output will be reported and commented on.

Chapter 8, finally, offers a summary of the findings of all empirical studies. On the basis of these results, general conclusions are drawn and possible implications with respect to educational practice are discussed. In addition, the findings and the limitations of the current study will be considered in a broader perspective to end up with suggestions and recommendations for future research on academic language development of bilingual children.

CHAPTER 2

Berbers and their languages in Morocco and the Netherlands

This chapter[1] gives an overview of the position of Berber and its speakers in Morocco and in the Netherlands. First, a description is given of the Berber language structure itself, with a focus on Tarifit. Then, the attitudes towards the languages spoken in Morocco are briefly described and some information is provided regarding language policy and language use in Morocco and regarding the implications of this policy for the position of Berber in education and Moroccan society as a whole. After that, the position of the Berber language and the Berber community in the Netherlands is described. Two dimensions will be distinguished in this respect: the status of the speakers of Berber in the Netherlands and the status of Berber as a language in the Dutch society. The first part will shed light on the demographic and immigration background of the Berber community and its members' position in the Dutch educational system and labour market. The second part will consider the status quo of the Berber language itself in the Netherlands. Aspects that will be examined are the position of Berber in social and cultural life of Moroccans, mass media, culture, religion and education. Finally, the chapter will deal with language acquisition, language proficiency, language choice, language dominance, and language preference within the Berber community in the Netherlands.

2.1 The Berber language

Berbers are the indigenous and the oldest known inhabitants of Morocco and many other parts of North Africa. Historians disagree as to the exact origin of the Berbers and the historical period in which they came to this part of Africa.

[1] This chapter is an adapted version of Laghzaoui (2007).

However, most references state that they date from about 3000 BC (Camps, 1980; 1987). The use of the term 'Berber' is arbitrary in the sense that it is of external origin and not a Berber word. The name was first used by the Greeks and Romans, and later by the Arabs. Among the native speakers of this language, instead of the term 'Berber', 'Tamazight' or 'Imazighen' are used which refer to the language and its speakers, respectively[2]. Due to the absence of documentation, it is extremely difficult to trace the exact history of Berber people. References to the history of Berber civilization are predominantly restricted to anthroponyms and toponyms. Despite many attempts of Berber linguists in promoting and developing the language, one of the big vulnerabilities of the Berber language is its dependence on an oral tradition and its lack of a written culture (Chafik, 2004).

Berber (or Tamazight) consists of a group of closely related languages, which are the mother tongues of diverse groups in northern Africa concentrated especially in Morocco and Algeria and of small numbers of native speakers scattered all over other North African countries. At a rough estimate, there are between 14 and 25 million speakers of Berber (Chafik, 2004). The Berber language belongs to the Afro-Asiatic language family, more specifically the Hamitic language family. It is the earliest language known in Morocco (Basset, 1952; Chafik, 1982; Chaker, 1989; Laroui, 2007).

In Morocco, Berber is, next to Moroccan Arabic, one of the major mother tongues used in the country. About 40% of the Moroccan population speaks Berber (Boukous, 1995; Ennaji, 1997). Traditionally and in terms of geographical distribution, Berber is spoken in three main dialectal varieties, which are associated with three regions of Morocco: Tarifit or Rif Berber is spoken in the northern Rif mountains and in the eastern part of the Middle Atlas. Tamazight is used in the Central Atlas mountains and the south-eastern parts of the country. The Tachelhit variety is used in the High Atlas and in the South west of Morocco, where the largest number of Berber speakers is concentrated. These three varieties share the same basic syntactic structure (Ennaji, 1997). The linguistic differences between these dialect groups are mainly at the level of the lexicon and phonology, which often makes them

[2] Instead of the term Berber, which bears negative connotations, people of Berber origin prefer using the term Tamazight to refer to the language, and Imazighen to refer to the people. In this study, however, the term Berber is used because it is more familiar among researchers in the western world.

mutually unintelligible. Tarifit Berber, as other Berber varieties, is characterized by a rich morphology and is a highly inflected language. The structure of Berber is mainly based on the verb and the noun (Sadiqi & Ennaji, 2004). Nouns in Berber are inflected for gender and number. In addition, nouns are distinguished in two states: free and construct states. The free state is the unmarked form of all nouns (Sadiqi & Ennaji, 2004) and the construct form is used after a proposition or when the subject is post-verbal. In Tarifit Berber, three verb classes are distinguished: finite, non-finite and auxiliary (Cadi, 1987; McClelland, 2000). With regard to word order, it is generally agreed that the canonical order of Tarifit Berber is Verb, Subject, Object (VSO) (Greenberg, 1966; Cadi, 1987). In spite of thousands of years of various forms of contact between Berber and Arabic, there are hardly any semantic or syntactic similarities between these two languages, except for some loanwords (Boukous, 2003).

In the last decades, there has been a strong movement among Berber intellectuals and researchers to codify and standardize the Berber languages as one language, viz. Amazigh. Berber has been a written language on and off for nearly 3000 years. Several invasions have disrupted the written tradition in the past. Berber was first written in Tifinagh script, which is also still used by the Touareg population in Niger. Subsequently, around the period between 1000 and 1500, Berber was written in the Arabic alphabet, particularly in the Tachelhit variety of the south of Morocco (Hammam, 2004). For practical reasons and to make this language widely accessible, at present Berber is often written in the Latin alphabet. In Morocco, after the creation of the Royal Institute of Amazigh Culture IRCAM (Institut Royal de la Culture Amazighe) in 2001, Berber was gradually introduced as a subject in primary education. In 2003, teaching Berber script was officially started in 317 public primary schools after its codification by using Tifinagh (El Aissati, Karsmakers & Kurvers, 2011).

The most important differences between Berber and Dutch are at the level of morphological typology, word order, the verbal system, and clause combining strategies. For instance, the canonical word order in Berber is VSO but this is more variable than the SVO word order in Dutch. In addition, Berber is a pro-drop language, making use of subject pronouns as affixes merely in limited pragmatic contexts, whereas these pronouns are compulsory in Dutch. Another difference is at the level of aspect and tense. In Berber,

aspect (i.e. the temporal flow and different perspectives in describing an event, in which usually two forms are distinguished: perfective or completive and imperfective or durative), is more relevant than tense (i.e. the fixing of events regarding a given time reference, such as the distinction between past and present tense). Moreover, in Berber both prepositions and postpositions are used as affixes while in Dutch only prepositions are available. In clause combining, Berber is characterized by parallel constructions and high frequency of coordination. Berber speakers, when learning Dutch as L2 may transfer this characteristic in contexts where subordination, for instance, would be more appropriate.

2.2 Language attitudes, policy and use in Morocco

One of the studies addressing language attitudes in Morocco is the well-known field study of Bentahila (1983), in which 200 Arabic-French bilinguals in Morocco participated. Most of the participants in this study considered Standard Arabic as the language of heritage and religion. So, this language has automatically acquired a superior status, no matter how difficult it is. Standard Arabic is also considered a 'pure' language and the most beautiful and richest one; pre-islamic poetry and the whole body of Arabic literature are mentioned by the participants in this regard. Moroccan Arabic, though, is seen as 'impure' and 'poor', because it depends on borrowing from other languages and, as a consequence, contains many loan words. Another argument which supports the high position of Standard Arabic, is that it conserves and strengthens the ties with the rest of the Arab world. Moroccan Arabic is, nevertheless, valued in other settings. Since the majority of the population uses it in daily interactions, it is obvious that it provides access to all domains and layers of the country. Bentahila (1983) asked his respondents what language they considered as the easiest and most practical in everyday life. Concerning 'easiness', 63% considered Moroccan Arabic as the easiest language, followed by French (24%), and Standard Arabic (10%); 3% was undecided. With regard to 'practicality', 72% chose Moroccan Arabic, 22% French, 3% Standard Arabic, and 3% was undecided. In addition to the practical role of Moroccan Arabic, it is usually seen as a simple and traditional language inherited from ancestors and protecting the distinct characteristics of Moroccan history. French is also sub-

ject to different attitudes. By some, it is called the language of the 'colonizer' but mostly it is considered the most modern and useful language for sciences, because most research reports and scientific literature in various disciplines are published in French (Abbassi, 1977).

Nortier (1989) reports that attitudes to different languages in Morocco are associated with several overlapping domains and values. Berber and Moroccan Arabic, for instance, share some values and domains such as ethnic solidarity, high affective value, oral language, home language, everyday use, in-group, intimacy and lingua franca. On the other hand, Standard Arabic and French are associated with modernism, formal education, middle and upper class, media, science and a successful career.

Concerning attitudes towards Berber in Morocco, some studies show that the restricted functions of this language in Moroccan society make the speakers form negative attitudes towards it (Boukous, 1995). Ennaji (1997) investigated students' attitudes towards Berber in Morocco. 152 Berber-speaking and Arabic-speaking students at the university of Fes participated in this study in which interviews and a questionnaire were used. On many aspects, different attitudes were found among Berber native and non-native speakers. Concerning the question as to whether Berber is a beautiful and rich language, only a small percentage of non-native speakers think that Berber is beautiful (19%) and rich (11%). Among Berber native speakers, however, 75% and 58% respectively believe that Berber is beautiful and rich. With regard to teaching Berber, 99% of native speakers and 81% of non-native speakers turned out to be in favour of teaching Berber at school. One of the general conclusions reported in this study is the positive shift in attitudes towards Berber among the youth and the intelligentsia in Morocco, in comparison with the 1970s and 1980s (Ennaji, 1997).

Another study which included attitudes towards Berber in Morocco was conducted by El Kirat (2001). In this study, the focus was on the status and future of the Berber language in the Beni Iznassen community as a consequence of the extensive contact of this variety with Moroccan Arabic. This Berber variety is spoken in the North-east of Morocco. El Kirat found that native speakers of Beni Iznassen developed a very negative attitude towards their mother language. Moreover, they consider their language a stigma and think that Moroccan Arabic is more useful than Berber (El Kirat, 2001).

In the years after the independence in 1956, there was hardly any awareness of the Berber identity in Morocco. The Berber language was considered no more than a medium of communication in the areas where it was used. Socially speaking, Standard Arabic and French were seen as the keys to an amelioration of a person's status. For this reason, many Berber parents encouraged their children to learn Arabic and French. In the last decades, however, the 'Berber world' has witnessed a cultural renaissance, especially in Morocco and Algeria. The Berber cultural movement in Morocco strives for the integration of the Berber language and culture in the educational system. It strives also to make it constitutionally a national language, reflecting its role as an important element of the Moroccan identity. As an example, the *Association Marocaine de Recherche et d'Echange Culturelle (*AMREC*)* has sent a letter to the late king of Morocco, Hassan II in 1999, asking him to give the Berber language a position in the national education system. However, before that, in 1994, king Hassan II had already showed his intention by announcing that the Berber language would be officially introduced into the Moroccan school system. Correspondingly, many associations in different parts of the country have engaged in promoting the position of Berber by introducing it into various institutions of the country. Efforts were also made to form a standard grammar and to encourage literary writing in this language. So over the long term, a positive change in the attitudes of native and non-native speakers towards Berber can be noted.

In 1999, a charter of reform of the Moroccan educational system was produced (*Charte Nationale d'Education et de Formation (*CNEF*), Royaume du Maroc*, 1999), including a range of proposed measures and recommendations to reform the educational system. As expressed in article 110, "the kingdom of Morocco adopts a clear, coherent and constant language policy in the field of education". This new policy is based on the reinforcement and improvement of the teaching and use of the Arabic language, a diversification of languages which are used in sciences and technologies and openness towards the Amazigh language. In article 115 which deals specifically with Berber, it is stated that:

> *Les autorités pédagogiques régionales pourront, dans le cadre de la proportion curriculaire laissée à leur initiative, choisir l'utilisation de la langue Amazigh ou tout dialecte local dans le but de faciliter l'apprentissage de la langue officielle au préscolaire et au premier cycle de l'école primaire.* (CNEF, article 115)

(The regional education authorities may, as part of the curriculum left to their own initiative, choose to use the Amazigh language or any local dialect in order to facilitate learning the official language during preschool and first cycle of primary school.)

Reference to the term 'Amazigh language' here is unique in the history of Moroccan language policy. Moreover, it implies that not all Moroccans speak Arabic and that Berber languages can facilitate the process of learning the official language for the native Berber speakers at school. Another important milestone in the Moroccan linguistic landscape is the decision of King Mohamed VI to create the Royal Institute for Amazigh Culture in 2001. The main tasks with which this institute was charged were the codification and standardization of Berber, the development of school materials, implementing the use of Berber in the educational system and determining which script should be used to formally learn and teach this language.

Tracing the historical and social meaning of Berber in Morocco, there has never been a cultural competition between Arabic and Berber (Khalafi, 2000; Laroui, 2007). For centuries, bilingualism and diglossia were commonplace in the Moroccan linguistic scene. Arabic was never considered a newcomer language opposed to the native language Berber. The two languages always had a harmonious and complementary distribution, Arabic being the sacred language and an instrument for practicing religious rituals and Berber fulfilling the role of a language for daily interactions. Traditionally, Berber as a home language was always used to facilitate the learning of religious sciences in both Islam and Judaism in Morocco (Zafrani, 1998). The transition which Berber children go through is, according to Grandguillaume (1995), different from, for instance, children who originate from Brittany as they enter French primary school, because these children totally have to discard their mother tongue. The Berber child was "involved in a *relation* with the sacred language which is obviously different from his daily conversational home language" (Grandguillaume, 1995: 85).

2.3 The Moroccan Berber community in the Netherlands

Multilingualism and multiculturalism are no recent phenomena in Dutch society. The Netherlands have witnessed immigration since the so-called

'Golden' seventeenth century. At that time, nearly 10% of the population was originally non-Dutch. From that time on, migration to the Netherlands for political and/or economic reasons continued (Laghzaoui, 2005). In view of the large demand for workmen in the sixties and seventies of the last century, the Netherlands as well as many other western European companies started recruiting cheap laborers in countries around the Mediterranean to perform mostly unskilled work. These so-called guest workers, among them the Moroccan Berber group, were brought in to do temporary jobs and were expected to return to their countries of origin after a limited period of time. Yet, and despite all predictions, their stay became more or less permanent. Van Amersfoort (1982) describes this process as follows:

> *The European host countries needed additional labour and assumed that they could reap the economic benefits from the employment of the foreign workers while at the same time keeping the social, cultural, and political consequences of the presence of these foreign populations within their borders to a minimum.* (Van Amersfoort, 1982: 187)

The following sections outline some characteristics of the Moroccan Berber community in the Netherlands in terms of both its demographic and educational perspectives and its position on the labor market.

2.3.1 Demographic and immigration background

Moroccan workers were recruited among others and came for work on a temporary basis to the Netherlands as 'guest workers'. However, in the seventies and eighties of the last century, these Moroccans started bringing their spouses and children to the Netherlands within the framework of family reunion regulations and in the course of time these families expanded. In the beginning of the 21st century, the Moroccan community has become one of the largest minority groups in Dutch society. As already shown in section 2.1., the sociolinguistic situation of the Moroccans in their home country is quite complex, and consequently, this complexity is reflected in the Moroccan community living in the Netherlands. Approximately 70% of Moroccan immigrants in the Netherlands originate from the northern part of Morocco and are native speakers of the Tarifit-Berber variety (Obdeijn *et al.*, 1999).

These Moroccans are called Rifians. Others stem from larger cities all over the country such as Tangier, Rabat and Casablanca. Most Moroccans have settled with concentrations in the large cities of the Netherlands. Since 1970, the number of Moroccans has grown rapidly. In 1971, their total number was 22,000; in 1987 the number reached 130,000 persons, mostly because of family reunion. According to the Dutch Central Bureau of Statistics, in 2010, the official number of the population of Moroccan origin in the Netherlands was 349,005 (CBS, 2011).

It should be noted here that it is difficult to get a clear-cut picture of the exact number of Moroccans in the Netherlands. Traditionally, 'country of birth' and 'nationality' have been considered as the most important criteria in collecting reliable information about the composition of minority groups in the Netherlands. The first criterion is rather objective and easy to establish. However, it holds solely for the first generation and does not work for second and third generations. There is intergenerational erosion through births in the host country (Extra & Maartens, 1998). The nationality criterion is also objective and relatively easy to establish. Yet, this criterion has a drawback because of naturalization processes. The declining significance of both the nationality and the birth country criterion led demographers to look for more valid criteria to establish reliable statistics on the number of people belonging to minority groups in the Netherlands. Complementary or alternative criteria, such as self-categorization (*To which ethnic group do you consider yourself to belong*?) and home language use (*Is any language other than Dutch used in your home?)* have been suggested and investigated (Broeder & Extra, 1999). These two criteria have been introduced notably in countries with a long and rich immigration history such as the USA, Canada and Australia. In the Netherlands, the official statistics of Moroccans, and immigrant minority groups in general, still rely on the first two criteria or a combination of both. The last two criteria have not yet been used on a national scale.

2.3.2 Position of Moroccan Berbers in Dutch education

The first generation of Moroccans who came to the Netherlands in the second half of the 20th century had received little education in Morocco. Most of them came from the Rif area, which is characterized by a high percentage of illiteracy (Obdeijn *et al.*, 1999). Children of Moroccan immigrants started to attend

Dutch elementary schools in great numbers, since education is compulsory in the Netherlands. Taking part in formal education was considered a crucial factor for their future chances on the labor market. Since the monolingual Dutch educational system was totally unprepared for this influx, different measures were taken to deal with it, such as offering Dutch as a second language and home language instruction (HLI) to Moroccan pupils as well as to children of other immigrant backgrounds. Different studies have shown that Moroccan children lag far behind their Dutch peers in school results (Dagevos, Gijsberts & Van Praag, 2003; Tesser & Iedema, 2001). They are, moreover, over-represented in lower vocational training, under-represented in general secondary education, and show high percentages of dropouts (Tesser & Iedema, 2001).

On the basis of the results of empirical research, different factors are highlighted to explain the low achievements of Moroccan pupils in Dutch education. Such factors include the characteristics of the pupils and their origins, duration of stay in the host country, command of Dutch, educational level of the parents and the family environment in addition to the quality of schools and teachers in general (Penninx *et al.*, 1998). But a positive change has been reported in more recent studies concerning the level of participation of Moroccans in Dutch education. According to national statistics, in 1996, 1,841 Moroccan students were registered in higher vocational education (HBO)[3] and 825 in university education (WO)[4]. In 2006, 6,922 students were registered in higher vocational education and 2,276 in universities – an increase of 73.4% and 63.7%, respectively, in one decade only. In the same period, their Dutch peers showed an increase of 13.1% in higher vocational education and 11.6% in university education, respectively (CBS, 2006). In 2010, the total number of Moroccan students in Dutch higher education was 12,004: 9,281 of them registered in higher vocational education and 2,723 in university education.

2.3.3 Position of Moroccan Berbers in the Dutch labor market

In its "Minority Memorandum" *(Minderhedennota)* published in 1983, the Dutch government proposed different ways to promote the social integration of ethnic minority immigrant groups in Dutch society. One of these ways

[3] HBO: Hoger Beroepsonderwijs (higher vocational education).
[4] WO: Wetenschappelijk Onderwijs (academic education).

concerned the fight against socio-economic disadvantages among minority groups and the attainment of better positions on the labor market.

The position of ethnic minorities, Moroccans in particular, on the Dutch labor market is unfavorable in comparison with the indigenous Dutch working population in the Netherlands. This is due to many factors, among which are the challenges at school previously discussed. In the 1970s, migrants in general, and Moroccans more specifically, witnessed almost no unemployment because the Dutch economy was in need of unskilled workers. Meanwhile, unemployment rates among Moroccans have become extremely high and alarming in the last decades. In 2010, the total number of Dutch people registered as unemployed was about 269,000 (approximately 4.2% of the active population on the job market). In the same year, the total number of Moroccans reported to be jobless reached a total of 17,000, which is 13.7% of the active population within this group (CBS, 2010).

2.4 Position of the Berber language in the Netherlands

As stated earlier, the language situation of Moroccans in the Netherlands is as complex as that in their home country. In Morocco, Arabic is institutionally the official language. However, for centuries, Berber-Arabic bilingualism has characterized the linguistic scene in Morocco. This multilingual profile in the home country is reflected in the Moroccan community in the Netherlands. As has already been indicated earlier, the majority of Moroccans living in the Netherlands originate from the Rif area where Tarifit Berber is spoken. A small minority of the Berberophone population in the Netherlands uses the southern Tashelhit variety.

The following sections give an account of the position and uses of the Berber language in different domains. Furthermore, results of studies on topics such as language acquisition, language attitudes, language maintenance and shift are presented. Finally, some aspects of language policy concerning the use of Berber in the Dutch educational system are set out.

2.4.1 Language acquisition, choice and use

As with any minority group, Moroccan-Berbers create their own community circles where family members, friends and acquaintances socialize. The first generation of Moroccan Berbers uses Berber as a medium of communication when all interaction participants have a command of the same variety of Berber. In other situations, since not all Berber varieties are mutually intelligible, Moroccan Arabic is used as a *lingua franca*. The use of Dutch by first generation immigrants is relatively small due to their limited proficiency in this language. Quite the reverse holds true in the case of the second generation. Their use of Berber is considerably low and Dutch is used as a *lingua franca* (De Ruiter, 1989; Broeder, & Extra, 1999). Generally speaking, Dutch is becoming more and more the dominant language of interaction between younger Moroccan Berbers.

Since the arrival of Moroccans in the Netherlands, many studies have been conducted into different linguistic aspects of the languages used by this immigrant group. These studies do not always treat the Berber community separately. They usually report on Moroccans in general. This section provides some results of studies focusing on the Berber community.

Concerning the first generation of Moroccan Berbers in the Netherlands, there are no accurate reports on language use. Most researchers focus only on children or youngsters. Generally speaking, first generation Moroccan Berbers use only their mother tongue because of their low proficiency in Dutch. In his research on language acquisition and language use among Moroccan youths, De Ruiter (1989) interviewed 80 Moroccan informants between 7 and 21 years old. Among other things, they were asked about the language they used in communication with their fathers and mothers. The general conclusion that appeared from the informants' answers was that the majority (80%) used their mother tongue, i.e., Moroccan Arabic or Berber. In the same study, it was reported that the mother tongue was used in only 44% of the contacts between brothers and sisters and in 34% of the contacts with Moroccan friends. Nearly all informants, except the 21-year-old group, were more proficient in Dutch than in the Moroccan languages. This pattern of language choice within the Moroccan community is confirmed by later studies, such as Broeder & Extra (1999).

Van der Avoird *et al.* (2002) conducted a large-scale survey among 6,000 Berber children in primary and secondary school in thirteen Dutch

municipalities. The majority of the children was born in the Netherlands. Their ages ranged from 4 to 17. With regard to language proficiency, almost all children (96%) said they did understand Berber and 92% of them said that they spoke it. One of the striking findings in this survey was that 28% of the children claimed to be able to read Berber and that 15% of them could write it. As to language choice, most children responded that they *always* or *often* spoke Berber with their mothers and fathers (respectively about 82% and 72%). With their younger as well as their older brothers and sisters, about 35% of the informants said to use Berber. Most children preferred using Dutch in interactions with their siblings since they master this language better than Berber. The same pattern was reported concerning their language choice with Berber speaking friends. With respect to language dominance, the study showed that 52% of the youngest pupils (4-5 years old) were dominant in Berber, whereas the older children (6-17 years old) were dominant in Dutch. By the end of primary education, however, children reported a growing balanced bilingualism. According to their language preference profile, young children in the lower grades preferred to speak Berber (45%), while children in higher grades preferred to speak Dutch (42%). It was also reported that a group of children (5-25%) had no preference for either language (Van der Avoird *et al.*, 2002).

Looking at these results, it can be concluded that there is an ongoing process of language shift especially among the younger generations in favour of Dutch, manifesting itself in all aspects of language: language proficiency, language choice, language dominance and language preference.

2.4.2 Berber in the mass media

Though its presence is so far only modest, Berber is used in a number of radio and television programs in the Netherlands. In 1969, the Dutch Network Corporation (NOS) started broadcasting its first radio program for Moroccans in Arabic and Berber (Kraal & Heelsum, 2002). This program mainly focused on news and actual events in Morocco because of the expected temporal stay and the return to the home country. After 1983, more time was devoted to radio programs for minority groups and to topics related to the long-lasting stay of Moroccans and other relevant socio-economic issues for this group. Since 1995, the Dutch Program Corporation (NPS) has been broadcasting a follow-

up of this program with the aim of supporting and promoting the participation and integration of Moroccan migrants in Dutch society. The languages used in this program are mainly Moroccan Arabic and Tarifit Berber. In recent years, however, the broadcasting time of this program has been reduced as a result of political discussions about the supposed failure of the multicultural approach to society.

Another area where Berber is frequently used is the Dutch Muslim Broadcasting Corporation (NMO). NMO is a religious and cultural broadcasting organization producing radio and television programs to cater for the needs of the entire Muslim community in the Netherlands. The language mainly used in the programs is Dutch. Berber is used, for example, in interviews with first generation Moroccans or in informative programs addressing Moroccans. Nowadays with satellite dishes, it is possible for the Dutch Berber community to receive radio and television programs broadcasted by the Moroccan Radio and Television (RTM). These programs are, according to journalists of the Berber editorial staff, very popular among Berbers in the whole of Western Europe (El Aissati & De Ruiter, 2004). Another possibility for audiences to follow Berber programs is through the Radio Télévision Berbère (BRTV). This channel launched its first programs in 2000 from France and can be received all over Europe and North Africa via satellite with a membership. It is a bilingual station using both Berber and French. Its aim is to promote Berber culture in general. However, this channel broadcasts mostly in Kabyle, the main Berber variety spoken in Algeria. A more recent initiative in the Netherlands in this respect was the start of the Amazigh TV channel in 2006. This channel broadcasts from the city of Utrecht. Its main goal is to produce programs for the Berber and Moroccan community in Europe, North Africa, and especially the Netherlands. The programs are in Tarifit Berber and Dutch.

2.4.3 Berber in culture and religion

The permanent stay of Moroccans in the Netherlands has led to the emergence of all kinds of social and cultural organizations in the beginning of the 1990's. Berber activists have created cultural associations to support and preserve Berber language and culture, in addition to issuing newsletters and news bulletins in Dutch and Berber. Among these associations are *Adrar*, *Syphax*, *AbiN*, *Tamaynut*, *Nokour* and *Tawiza*. These organizations have published

newsletters, organized lessons in Berber and have spread information via websites. These newsletters and websites, however, are rarely in Berber. As far as literature is concerned, Berber is used in different genres such as novels, short stories and poetry. The latter is the most known type among the Berber community in the Netherlands, because of the oral nature of the Berber language. In recent years, many poets have taken part in literary evenings and have published their work especially in Latin alphabet so as to reach as many people as possible. Famous Berber poets in the Netherlands are Ahmed Essadki (1997), Mohamed Chacha (1995, 1997, 2000), Mimoun El Walid (1994, 1996) and Ahmed Ziani (1993, 1997). Another new domain where Berber is widely used is the Internet. The appearance of this medium has created a new space for the Amazigh language. Different organizations and associations create their own sites with information in Dutch and less frequently in Berber, mostly in Latin script. Well-known Berber websites are www.tawiza.nl and www.amazigh.nl.

All over the Muslim world, Arabic is considered the language of religion *par excellence*. As a consequence, nearly all liturgical occasions such as praying, Friday sermons or reading the Qur'an take place in Arabic, the sacred language of Islam. Accordingly, Berber plays a limited role in religious matters in the Netherlands. It is used sporadically during religious lectures and informal discussions about religion. It should be noted here that Islam as a religion does not exclude languages other than Arabic. In many places in the Quran, multilingualism and multiculturalism are praised as shown in the following verse:

> *And among His Signs is the creation of the heavens and the earth, and the variations in your languages and your colours: verily in that are Signs for those who know.* (Quran: The Romans, verse 22) (Yusuf Ali, 1998)

2.4.4 Berber in education

At first, Dutch policy towards Mediterranean migrants was characterized by the belief that these groups were staying only temporarily in the Netherlands. At the same time, Dutch authorities had to guarantee equal treatment at the level of education, job market and social facilities, regardless of the countries of origin.

In 1979, the Scientific Council for Government Policy (Wetenschappelijke Raad voor het Regeringsbeleid) published the "Ethnic Minorities Report" in which the government was advised to take appropriate measures and to develop a clear policy towards immigrant ethnic minorities in the Netherlands (WRR, 1979). In this context, the Ministry of Education introduced the so-called *Onderwijs in Eigen Taal en Cultuur* (OETC), Home Language and Culture Instruction (henceforth, HLCI). The main objective of HLCI in the first years was to prepare the children of immigrants to maintain language and culture in order to be able to fit in the educational system of the country of origin after return. However, after the mid-eighties, it became clear that most migrants were choosing to settle in the Netherlands for a long time or permanently. Consequently, HLCI was transformed into a means for fighting against low educational achievement among children of immigrants as well as for bridging the gap between the home and the school environment in order to prepare these children for full participation in Dutch society. Moreover, learning their mother tongue was expected to help the children to know more about their cultural and linguistic heritage and would contribute to the development of their cognitive abilities. For Moroccan children, it was obvious that Arabic would be provided in the framework of HLCI, since Arabic is the official language of Morocco. Thus, the Berber community had no other choice than standard Arabic. A number of language surveys investigated the level of interest in attending these lessons. The largest survey was conducted in the municipality of The Hague (Van der Avoird *et al.*, 2002). Concerning primary education, it showed that among the 1,830 parents who were interviewed, about 12% (223) would like their children to follow lessons in Berber. At the level of secondary school, 10% of 967 Berber children indicated that they were interested in attending Berber lessons. Concerning methods and teaching materials, some Berber programs have been developed such as *Iles Inu* (E-rramdani *et al.*, 1998). Also other teaching materials have been translated from Dutch into Berber such as parts of *Schatkist* (Mommers & Verhoeven, 1989) and the home-based preschool program *Opstap-Opnieuw* (Opstap-Opnieuw, 1992).

CHAPTER 3

Academic language

When entering school, children are faced with many challenges that are crucial for their educational career. In the Netherlands, as children reach the age of four, they make a transition from the familiar, informal home environment to the relatively unknown, formal context of school[1]. The challenges children face in school are of a cultural, socio-emotional and cognitive nature (Dunlop & Fabian, 2007). Children at this age are expected to also deal with another challenge, namely a new language register and its use in the context of schooling, where they are expected to convey information in a structured and conventional way. This new language register is usually labeled academic language or school language. It differs in many respects from the language used at home. Variations between children in the acquisition of and proficiency in this specific register have been associated with later school achievements (Heath, 1983). Also in the Dutch situation, a good command of the Dutch school language is crucial for school success. If children's skills in this domain are not sufficient, it may prevent them from profiting sufficiently from the content offered in the curriculum. For many children, the academic register is not only different from the language used at home. Young bilingual pre-school children in particular, learning Dutch as a second language, face a dual challenge. They not only have to learn a specific language register used in cognitively demanding school contexts, but they also have to learn this new register in a relatively unknown language that is not their L1.

Communication between young children and their parents and/or caregivers at the pre-school age is crucial with respect to children's language development along with the social, emotional and cognitive support provided. Previous research reveals that the language development of children is related to the

[1] Many children start even earlier with different kinds of formal schooling when enrolled in preparatory preschool intervention programs in day-care centres.

language input to which children are exposed in early childhood in the home environment (Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991; Weizman & Snow, 2001; Tabors, Snow & Dickinson, 2001). These findings lend support to the theoretical framework of a usage-based approach (Tomasello, 2000), which argues that young children mainly learn language from imitating the input in their language environment. Many factors such as the socio-economic status of the parents, their literacy level and school-like activities conducted with their children at home may affect this process of language learning.

In section 3.1, a theoretical account of the role of parent-child interaction and input is given. After that, in section 3.2 we will go into the theoretical notion of academic language, followed by a description of the academic register in section 3.3. In the following sections, the specific linguistic variables which constitute the academic language register on different levels will be explained, starting with the lexical level (3.4), then the syntactic level (3.5) and the textual level (3.6). Finally, section 3.7 discusses the academic register in Tarifit Berber .

3.1 Impact of interaction and input

Children learn language in a social context in which they interact with parents. The content of these interactions and more specifically the input parents provide to their children, is crucial in the language learning process that children go through. Children whose social experiences provide more communicative opportunities and richer input build their vocabulary at a faster rate than children with less communicative experiences and poorer input (Hoff, 2006). Language input at home is stimulated when children are involved in verbal interaction and in literacy activities with their parents, such as shared picture book reading, dinner table conversations, watching educational television programs and singing songs. All these conversations are characterised by the use of a rich lexicon, complex sentences, and the use of specific (non-deictic) reference to time and space, the kind of language which is also expected and needed at school.

The way in which parents interact with their children and the language input they provide, differs between families. One of the variables that may cause these differences is the socioeconomic status (SES) of the parents. Many studies provide evidence that lower educated and economically disadvantaged

parents talk much less to their children than well-educated and economically advantaged parents do (Heath, 1983; Ward, 1971; Schachter, 1979; Ninio, 1980). Earlier research concerned with the influence of SES shows that the qualitative and quantitative variations in parental input related to SES explain most of the differences in children's language proficiency (Hoff, 2006; Raviv, Kessenich & Morrison, 2004). These studies suggest that children from higher SES levels, compared to those from lower SES families, have more and easier access to the language input that supports successful and fast linguistic development. For example, children from high SES families more frequently take part in home verbal activities such as shared book reading and other activities where stimulating parent-child interaction takes place. These types of conversations are characterized by the use of rich vocabulary, complex and dense sentences, and by semantically interconnected discourse (Hoff & Naigles 2002; Leseman, Scheele, Mayo & Messer, 2007). These are characteristics that can be considered as features of the academic register, as will be discussed in the following sections.

3.2 Academic language: theoretical considerations

Academic language can be defined as the use of language for specific, decontextualized and cognitively demanding communication in school-like tasks (Schleppegrell, 2004). For many children, this register is very different from the language used at home. It involves communicating about subjects and things that do not belong to here-and-now reality, mostly without the support of concrete contextual or immediate elements. Many studies have shown a positive relationship between an early exposure of children to this specific language register and school success (Cummins, 1991; Hoff & Naigles, 2002). Once at school, children are expected to master the academic language register and to be able to receive and provide information that is structured according to conventions different from the home language. Academic language encompasses lexical, morpho-syntactic and textual characteristics, which enable a structurally more elaborated and more abstract communication, whereas informal interactional language, such as at home, is more context-dependent.

The differences between the school register and the informal register have been widely and extensively studied from different perspectives. In his seminal

work, Ferguson (1959) classifies situations that make use of two different codes in a complementary way: a High Variety that belongs to the top of the social scale and a Low Variety which is at the bottom. Ferguson (1959: 435) originally described *diglossia* as:

> *a relatively stable language situation in which, in addition to the primary dialects of the language (which may include a standard or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety, the vehicle of a large and respected body of written literature, either of an earlier period or in another speech community, which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any section of the community for ordinary conversation.*

The notion of diglossia in Ferguson's frequently-quoted article has become a basic framework for many studies that deal with language diversity in relation to education and literacy as a whole.

Bernstein (1971) introduced a 'code theory' about language and social class, in which he distinguishes two types of language codes: the elaborated code versus the restricted code. The elaborated variety is associated with middle and upper-middle class families and is said to be used more in formal settings and for school-like tasks, also called the 'educational code'. It is characterised by the use of more subordinate clauses, logical connectives and explicit references. The restricted code, on the other hand, is associated with working class families and is used more in informal situations and daily conversations where there is a lot of shared and taken-for-granted knowledge among the speakers. It is distinguished by the use of pronouns, unfinished sentences, tag questions and use of gestures. As emerges from above, the differences between the two codes are not absolute and clear-cut but gradual. Bernstein suggested this theory as an explanation for the poor school performance of working-class children especially on language-oriented subjects. His assumption is that children who are not sufficiently exposed to the elaborated code, which is common in school contexts, would be disadvantaged in their school career. Despite the criticism this theory has received, it has enjoyed large influence among linguists, educationalists and policymakers.

Chafe (1982) has proposed two fundamental differences between typical speaking and writing styles. Speaking goes faster than writing, and speakers

interact with their audiences to a greater extent than do writers. Chafe suggests an underlying dimension associated with each of these situational differences: integration/fragmentation and detachment/involvement. At the level of the integration/fragmentation dimension, features such as nominalizations, participles and series of prepositional phrases mark integration, while fragmentation is marked by successive clauses without connectives or coordinating conjunctions. As to the detachment/involvement dimension, detachment is marked by passives and nominalizations, while involvement shows more first person pronouns and emphatic particles. In doing so, Chafe (1982) describes conversational texts as fragmented and involved, academic texts as integrated and detached. In another study, Tannen (1982) reported that written and formal language has more subordinate clauses than spoken language. Moreover, decontextualized academic language tends to be more hypotactic by using subordination, relative clauses and adverbial clauses, while contextualised interactional discourse is more additive and paratactic by using for example ellipsis.

Dealing with similar differences between formal and informal language, Cummins (1991) postulates a distinction between two types of language proficiency: *Basic Interpersonal Communicative Skills* (BICS) and *Cognitive Academic Language Proficiency* (CALP). BICS refers to the conversational language that is needed to interact in everyday communication. It usually takes place in socially shared contexts (context embedded) and is cognitively undemanding. CALP refers to a language proficiency that is acquired and needed by children to cope with the formal academic requirements in school settings. It is more complex and has less contextual support (context reduced).

In a more recent approach, Schleppegrell (2004), inspired by systemic functional linguistics (Halliday, 1994), introduced a framework in which she deals with the notion of the academic language register. This framework, which has been applied in our study, will be dealt with in more detail in the following section.

3.3 The academic language register

Functional linguistics is a linguistic theory that perceives language as a social process that contributes to the realization of different social contexts (Schleppegrell, 2004). In other words, the functional approach explains how

linguistic entities realize social meanings and how these meanings account for a variety of contexts. An important element of functional analysis is the notion of register which explains the relation between language and context. According to Halliday & Hasan (1989), a register is a constellation of lexical and grammatical features that realizes a particular situational context. Depending on the context, register can be described in terms of *field* (the subject matter being talked about), *tenor* (how speaker/hearer or writer/reader are related to each other), and *mode* (how textual choices are organized). In every phrase, sentence or text, and by making lexical and grammatical choices from the three areas mentioned above, three kinds of meanings can be construed: the *field* of discourse realized in *ideational* resources, the *tenor* of discourse realized in *interpersonal* resources, and the *mode* of discourse realized in *textual* resources (Schleppegrell, 2004).

Schleppegrell (2004) made an analysis of school texts, based on functional linguistics (Halliday, 1994), and derived from that the linguistic features of academic language. Table 3.1. illustrates how different variables of academic language are realized in linguistic features.

Table 3.1: Academic Register elements and their realizations (source: Schleppegrell, 2004)

Contextual Variable	Linguistic Realization
Field (Presenting ideas)	Ideational Choices – Noun phrases/nominal groups (participants) – Verbs (process types) – Prepositional phrases, adverbial adjuncts, and other resources for information about time, place, manner, etc. (circumstances) – Resources for making logical relationships
Tenor (Taking a stance)	Interpersonal Choices – Mood (statements, questions, demands) – Modality (modal verbs and adverbs) – Intonation – Other resources for evaluative and attitudinal meaning (e.g., resources for appraisal)
Mode (Structuring a text)	Textual Choices – Cohesive devices, including conjunctions and connectors – Clause-combining strategies – Thematic organization

Table 3.1 shows that ideational choices (field dimension) are achieved mainly by using nouns, verbs and other content words. Academic language is usually characterized by the use of technical, complex, specific and abstract vocabulary. Hence, relying on nominal elements helps to condense the information being conveyed. For instance, the use of nominalization serves to combine two or more nominal entities in a relational process and presents the argument in a single clause instead of two or more. Children at school are supposed to deal with language in their school tasks (oral/written) without depending on any contextual support or shared knowledge where generally deictic words are used. This leads to the use of a dense and diverse lexicon, a relatively high level of explicitness through the use of more lexical words and less functional words to present ideas, describe situations, and refer to time and space. It should be mentioned in this respect that many features of academic language being discussed here are observed at a later stage in schooling. Young children are not expected to use that many features of academic language. Yet, it is important to trace these features at an early age and try to understand how they develop from the outset.

According to Schleppegrell (2001; 2004), features of academic language concerning the field dimension can be summed up as follows:

– the use of a varied, rich vocabulary and the use of relatively infrequent and specific or technical words and concepts;
– the use of expanded nominal groups that condense information;
– specific and explicit reference to objects, situations, persons by using as many content words as possible (lexical density in clauses);
– explicit indication of time and place to elucidate the context of reference, by means of specific reference words, varied use of verb tense and aspect, and by using content words to describe events and circumstances explicitly;
– decontextualized use of language where shared knowledge and contextual support do not, or cannot play a role. The speaker/writer who uses language for academic purposes has only linguistic context at his disposal and refers to persons and events outside the concrete here-and-now context.

The interpersonal choices that realize the *tenor* of discourse account for the relationship between speaker/hearer or writer/reader and the stances and judgements of the speaker or writer. In the context of schooling, children are expected to present and receive information in a more reasoned, distanced and

authoritative way which is unlike informal language. Grammatical tools that realize this authoritativeness are declarative mood and third person to enable an impersonal presentation (Schleppegrell, 2004). This kind of strategy leads to the realization of an assertive speaker/writer who exposes himself as an authoritative person dealing with objective information. The main traits of academic language in the domain of tenor can be summarized as follows (based on Schleppegrell, 2001; 2004):

– adopting impersonal points of view by relying on declarative mood to express statements, requests and demands;
– use of resources for presenting stance, including control of explicit and implicit objective options for attributing commitment to a proposition;
– use of modal auxiliary verbs and adverbs;
– explicit statement of the communicative goal towards the listener by means of varied use of verb mood and tense (for instance, reporting, asserting, convincing, requesting) and the use of adverbs and auxiliary verbs in order to express epistemological attitudes (e.g. possibility, probability);
– intonation (in speech) and other resources for attitudinal meaning.

Textual choices that bring about *mode* refer to the way in which a text or a conversation is organized. Unlike informal conversational settings, academic contexts require highly structured components of grammar to assure the adequate transmission of information. In school-based activities, children need to be able to express their ideas in the nominalised, condensed structures of academic writing, rather than relying on the clause-chaining that is typical for informal interaction, where simple clauses with finite verbs tend to be linked with conjunctions and elaborated with adverbial clauses (Schleppegrell, 2004). A highly structured text for academic ends is realized through clause-combining strategies and cohesive devices. Such strategies lead to a high lexical density, logical connections and cohesive relationships. In informal contexts, meaning is realized by mutual co-construction of the participants. Conversely, in academic contexts, meaning is reached by controlling the flow of information without co-construction. Some characteristics of academic language in the scope of mode, based on (Schleppegrell, 2001; 2004) can be described as:

– use of clause-combining strategies which result in condensation of information;

– explicit coherence between separate utterances and the narrative, hierarchical or logical construction of narration or discourse, through co-reference and nominalizations, synonyms and conjunctions in order to create semantic and logical relations between utterances;
– shifts from general and abstract to specific and concrete to present arguments that are more conventionalized.

Drawing on the above-mentioned categories, academic language can now be closer defined as a particular register that substantially differs from conversational informal language. Up to this point the features of this register have been explored by looking at the contextual variables of field, tenor and mode. Now we will go in more detail into the specific linguistic choices speakers make in order to realise context. In the next sections, the linguistic variables that are characteristic of academic language will be described. These variables are situated on different language levels: there are variables on the word level (lexical level, section 3.4), variables on the level of the sentence (syntactic level, section 3.5) and variables on the level of the text (textual level, section 3.6). In each section, a description of the features on this level is provided, followed by an overview of the available research on the acquisition of these linguistic features in children's language development.

3.4 Lexical features of academic language

3.4.1 Lexical richness

The lexical level is closely linked with the field dimension in functional grammar, because the lexical elements in a sentence realize the ideational part of the discourse. As commented on before, a characteristic of academic language is the use of a varied, rich vocabulary and the use of relatively infrequent and specific or technical words and concepts. Besides that, references to objects, situations and persons are made explicit by using as many content words as possible (Schleppegrell, 2004). These features of academic language at the lexical level can be captured in the notion of lexical richness, consisting of lexical diversity, lexical density and lexical specificity. Lexical diversity generally deals with the number of different words as a proportion of the total number of words used in a text or a conversation. The common

instrument that traditionally has been adopted for analysis is the type-token ratio (TTR), which computes lexical diversity by dividing the number of different words (Types) by the total number of words (Tokens). However, this traditional measure suffers from some limitations since it is sensitive to differences in language sample sizes (Van Hout & Vermeer, 2007). It appears from different analyses that larger samples give lower values of TTR. Recent studies offer alternatives to overcome this problem (Malvern & Richards, 2002; Durán, Malvern, Richards, & Chipere, 2004). A newly developed measure is called the index *D*, which uses repeated calculations based on sub-samples to establish how TTR changes as the sample size increases.

The second measure is lexical density. Academic texts are characterized by more lexical density than texts occurring in informal and non-academic settings (Schleppegrell, 2001). Academic language is more dense than informal language and contains a larger amount of nouns, verbs and adjectives. Different methods have been used to measure lexical density. There are approaches which focus on the content/function word dichotomy (Laufer & Nation, 1995; O'Loughlin, 1995; Read, 2000), while others consider the proportion of content words in relation to all words used in a transcript or text (Egging, 1994; Fang, 1997).

The third measure at the lexical level is lexical specificity. This is realized by the use of infrequent, technical and rare vocabulary items. In some studies, it is indicated as sophisticated vocabulary (Weizman & Snow, 2001). In English, for instance, the term sophisticated vocabulary is used to refer to words in general use by the language community that fall outside the 3,000 most common words and their various inflected forms. In the school context, children are expected to use specific lexical items to be more explicit and precise in dealing with information. Determining a valid and reliable measurement for lexical specificity is a great challenge, especially when dealing with languages, such as Berber, that do not have an existing list of sophisticated low-frequency vocabulary. For the latter reason mainly, we decided to use only lexical diversity and lexical density in this dissertation.

3.4.2 Research on lexical richness

A large amount of research on language acquisition and vocabulary development within monolingual and bilingual families has shown that vocabulary size and the early language skills of children are strongly related to

the language input children are exposed to in early childhood in the home environment (Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991; Weizman & Snow 2001; Tabors, Snow & Dickinson, 2001). Different studies provide evidence for the influence of the amount of parental speech on children's vocabulary growth. First of all, Smolak and Weinraub (1983) examined mothers' speech of two groups of children with large and small vocabularies. The results showed that mothers of children with a large vocabulary produced significantly more speech than mothers of children with less vocabulary at their disposal. Tomasello *et al.* (1986) studied mother-child interactions in experimental laboratory play sessions and found that the number of mothers' utterances significantly correlates with the number of word types produced by their toddlers. Similarly, the studies of Ninio (1993) and Brent and Siskind (2000) show that about 60% of children's early vocabulary are the same words which their parents often use, especially single word utterances such as proper nouns, action verbs, objects and interjections. In their well-known longitudinal research, Hart and Risley (1995) studied parents' interaction with their children from the age of about 10 months up to 3 years in families with different SES backgrounds. This study revealed clear differences at the level of vocabulary richness of the input received by children from low-SES families compared to children who stem from high-SES families. This study included three categories of families from different socioeconomic strata: professional families (middle and high SES), working class families and families on welfare. It turned out that the amount of parents' talk correlates with families' SES, with professional parents as talkative and welfare parents as taciturn. In the beginning, children of all the groups started to talk around the same time. However, children from higher SES families heard more words per hour, resulting in a large vocabulary. For instance, children from professional families heard on average 2,153 words per hour, children from working class families heard on average 1,251 words per hour and children from welfare families heard on average 616 words per hour. Hart and Risley (1995) stated that children need to be exposed to rich and varied speech in order to understand what objects are being talked about and how words are used in different contexts. They concluded that the amount of language in parent-child interactions in the early years of life is the major source of children's vocabulary growth and later language proficiency in general.

All these studies provide evidence that the exposure to a relatively large amount and variety of language facilitates children's vocabulary development.

3.5 Syntactic features of academic language

3.5.1 Syntactic complexity

In functional grammar, academic language can be manifested through the choice of a particular clause subject or clause object (field). It is usually expected that more pronominal subjects and objects will be used in informal language than in school language. At the same time, in academic texts, subjects and objects are often lexicalized, nominalized or even realized by using relatively long noun phrases.

Furthermore, the academic register is characterised by specific choices concerning the aspect and mood of the verb used in a text or conversation. Schleppegrell (2004) claims that the declarative mood is chiefly used in formal school language to convey information, while other types of mood, such as imperatives and interrogatives are more typical for daily conversational language.

Another element of academic language which can be situated at the morpho-syntactic level is reference to time and space. In academic language, references to time and place are often made explicit by using content words, whereas in informal conversation there are more non-specific and deictic references. Hence, in conversational interactions, deictic reference words like 'this' and 'that' are used to refer to referents being present in the immediate context.

Finally, syntactic complexity is achieved in academic discourse by using clause combining strategies. Schleppegrell (2001; 2004) argues that the high degree of syntactic complexity expected in school-based tasks is realized by condensing information through clause combining strategies. Using these strategies, speakers can pack more information in one sentence or utterance in order to achieve information density. In school contexts where children are exposed to the academic register, children are expected to pack information effectively in their utterances. Different clause combining strategies can be used to achieve different goals. These goals may range from conveying attitudes, as in complement clauses, to defining referents in an elaborate way, as in relative clauses, to pointing out the spatial-temporal-causal relations among events, as in adverbial clauses (Tomasello, 2004). In addition, speakers or writers can use clause combining to talk about specific events and relations to produce a coherent and well-structured discourse.

Two main types of complex sentences can be distinguished: coordinate and subordinate ones (Bowerman, 1979; Bloom *et al.*, 1980). Coordination is where two equivalent independent clauses are linked, whereas subordination is where two (or more) unequal and asymmetrical clauses are involved and in which one dependent clause is used to modify or complement another matrix (or main) clause. Within the subordinate category, three types can be distinguished: complement clauses, relative clauses and adverbial clauses. A complement clause is defined as an argument of the predicate in the matrix clause. A relative clause can be described as an attribute of a noun or noun phrase. An adverbial clause is defined as some kind of modifier of the associated matrix clause or verb phrase. Complement and relative clauses often function within another clause. They usually operate as a nominal group or part of a nominal group and are called embedded clauses. On the other hand, adverbial clauses can participate independently in discourse structuring and are then called hypotactic clauses (Schleppegrell, 2004).

3.5.2 Research on syntactic complexity

Existing research on language input and syntax indicates that the use of clause combining strategies in parents' and teachers' speech significantly influences the later syntactic development of children (Huttenlocher, *et al.*, 2002).

At an early age children begin to combine nominal constructions and other separate clauses. Many studies report that children start using multi-clausal utterances during their third year of life (Brown, 1973; Bowerman, 1979; Bloom *et al.*, 1989; Diessel & Tomasello, 2001; Diessel, 2004). Concerning the order of acquisition, many studies carried out, especially into English, have shown that complement clauses are the first constructions which emerge at the age of two, followed by coordinate and adverbial clauses during the third year, and finally by relative clauses after the age of three (Bowerman, 1979; Bloom *et al.*, 1980; Dromi & Berman, 1986).

Complement clauses can be divided into *infinitival* and *sentential* complement constructions. The first multiple clauses children produce are sentences which include a finite verb and an infinitive such as in the sentence *I wanna eat*. At this stage, these constructions take place with a limited number of *light* and complement-taking verbs such as *want*, *go*, *have* and *like*. Diessel (2004) notes that although these verbs behave grammatically like a matrix, they semantically

function like modals used for purpose/intention and compulsion. Following the stage of infinitival complement clauses, children start producing utterances containing sentential complement constructions (Bloom *et al.*, 1989). In the latter study, it is also shown that these constructions mainly contain epistemic verbs indicating knowledge, perception and certainty like *think*, *know* and *see*.

Relative clauses are in the first stage of production realized without relative pronouns by using empty nouns like *ones*, *kind* or *thing*. After this stage, children start producing clauses using a combination of full nouns (such as *book*, *pen*) and relative pronouns (*who*, *that* and *which*).

Concerning the acquisition of adverbial clauses, children start talking about sequences of events well before age 2 (Clark, 1970) by describing successive events and actions relying on the order-of-mention strategy with no explicit linking tools between them. From the age of 3, children start to produce a variety of connectives to mark different sequential relationships. The most frequent conjunctions found by many researchers in spontaneous child speech are *but, when, so, because, if, while, after, before* and *since* (Diessel, 2004).

By the age of 4, children are using most of the complex sentence types in their spontaneous speech (Bowerman, 1979). After the age of 5, children's grammatical complexity continues to increase as sentences become longer and structurally more complex in addition to more expanded noun phrases (Hoff, 2009).

3.6 Textual features of academic language

3.6.1 Decontextualization

On the textual level, the academic register implies a certain organization of texts. Texts need to be coherent, a characteristic that can be achieved by using certain cohesive devices. Coherence implies the way a text or a conversation is temporally and causally organized into a sequence that is meaningful by drawing on shared knowledge (Shapiro & Hudson, 1991). Accordingly, coherence refers semantically to the overall structure of a text and how the events and different parts of a narrative are interrelated and organized in a meaningful way. Cohesion, on the other hand, refers to the explicit way in which coherence is realized by using linguistic devices. These devices link utterances and sentences together at a local level by indicating the semantic

relationships between the events. Speakers may use different cohesive devices like reference, ellipsis, substitution, repetition, summation, inference and apposition. In order to develop good narrative techniques, children have to acquire the skills of using cohesive devices in establishing coherence.

In addition to cohesion in interactions, the degree of abstraction is relevant within the textual level. Abstraction level, also called decontextualization (Van Kleeck *et al.*, 1997), refers to the use of language without relying on shared knowledge and immediate spatial or temporal context. Decontextualized language can be defined as an 'expression of thinking that is increasingly disembedded from the supportive perceptual context' (Van Kleeck, Gillam, Hamilton & McGrath, 1997: 1262). Different terms have been used to indicate this type of language use, such as non-immediate speech (DeTemple & Beals, 1991; DeTemple, 2001), non-present talk (Dickinson, 2001), and decontextualized language (Snow, 1983; Curenton, Graig & Flanigan, 2008). In this study, the term decontextualized language will be used. Throughout this chapter, decontextualized language is used to refer to abstract language as will be explained later on.

Language can be viewed as a spanning continuum that ranges from contextualized to decontextualized discourse (Curenton, Craig & Flanigan, 2008). Contextualized discourse refers to conversations between participants that share the same knowledge and environment where messages are conveyed also via gestures and facial expressions, whereas decontextualized discourse refers to talks about abstract objects and events that are not part of the present environment. In contextualized language, meaning is represented by using gestures and cues that are present in the context of the conversation, in addition to the shared experiences among speakers (Pellegrini, 1985). In decontextualized discourse, on the other hand, meaning is conveyed through specific and explicit vocabulary, grammatical devices and explicit comparisons which demand higher levels of reasoning (Curenton, Graig & Flanigan, 2008).

Given the differences between (de)contextualized discourse, it is of importance to highlight the notion of decontextualized language and its relation to the context of literacy and schooling. According to Denny (1991), modern Western thought is characterized by decontextualization as a distinctive feature, distinguishing it from thought patterns in both agricultural and hunter-gatherer societies. Decontextualization meant here 'is the handling of information in a way that either disconnects other information or backgrounds it' (Denny, 1991:

66). This and many other studies found that agricultural communities usually show high level of contextualization, compared to the industrial Western communities (Denny, 1978; McIntyre's, 1976; Reuning & Wortley, 1973; Berry *et al*., 1986; Greenfield, 1972; Scribner, 1977; Hutchins, 1980; Perkins, 1980). According to these studies, decontextualization is not only caused by literacy but is mainly attributed to the change of human communities from small-sized groups where members share the same knowledge and background and usually communicate through face-to-face interaction, to larger and more complex societies which include thousands and millions of people who seldom share any information or context. This claim and dichotomous framework, however, is too simplistic and relative. For instance, considering the Berber community which is an example of an agricultural society, the oral tradition which characterizes the Berber language relies also on decontextualized language in fairy tales, folk stories and traditional accounts.

Decontextualized language was also subject to investigation in the well-known study of Bernstein (1960). Bernstein argues that even within one industrial society, in this case the UK, groups differ in language use. Marginal groups and working-class parents use contextualized language with their children which is termed 'restricted code', being shaped by shared assumptions and immediate context. On the other hand, middle-class parents tend to use a more decontextualized language which is termed 'elaborated code'. Yet, it should be emphasized that in his later work, Bernstein modified his strict distinction between the restricted and elaborated codes by admitting the degrees found with both codes and asserting that even children with working-class backgrounds might use the elaborated code in certain contexts and vice versa.

The language that is required in the classroom forms a challenge and even causes learning problems for many children. According to Donaldson (1978), the decontextualized and abstract nature of school language is characterized by a disembedded quality, as described in the following quotation:

> *It is when we are dealing with people and things in the context of fairly immediate goals and intentions and familiar patterns of events that we feel most at home. […]. Thus even pre-school children can frequently reason well about the events in the stories they hear. However, when we move beyond the bounds of human sense there is dramatic difference. Thinking which does move beyond these bounds, so that it no*

> *longer operates within the supportive context of meaningful events is often called 'formal' or 'abstract'[…] to reduce the risk of confusion […] I shall speak rather of 'disembedded thinking' […]. The better you are at tackling problems without having to be sustained by human sense the more likely you are to succeed in our educational system.* (Donaldson, 1978: 76-78)

A framework for abstraction levels was developed by Blank, Rose & Berlin (1978). This theory has been used in many studies (Sorsby & Martlew, 1991; Van Kleeck *et al.*, 1997; Massey, 2004; Groenendaal, Goudena & Olthof, 1992; Nap-Kolhoff & Van Steensel, 2005) and consists of four levels of abstraction (see Chapter 4 and 7). In these levels of abstraction the extent to which the content of the interactions abstracts from the directly visible here-and-now differs. Figure 3.1 represents these abstraction classifications. As can be seen, there is an increasing separation between the perceptual world and the language used in conversations.

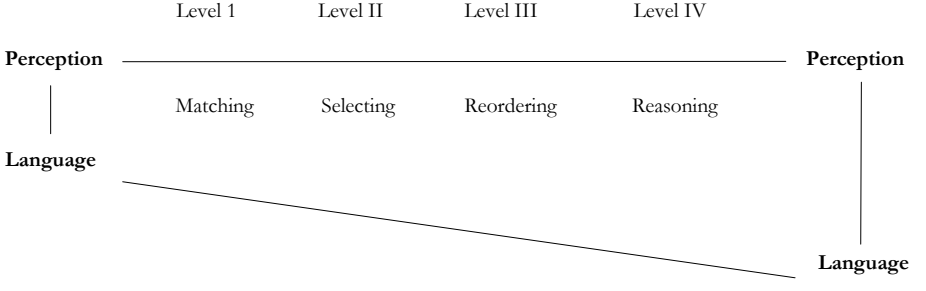

Figure 3.1: The concept of perceptual-language distance underlying the levels of abstraction (adapted from: Blank, Rose & Berlin, 1978)

3.6.2 Research on decontextualization

From the age of 2, children get involved in conversations that concern not only the here-and-now but also the there-and-then (Sachs, 1983). Until this point of development, children rely on the immediate shared world, in time and space, to shape their communicative skills. Soon after this stage, children start dealing with events and topics that happened in the past or that are not present in the immediate world. With this new situation, children learn to use new devices and strategies to refer to different events and objects not belonging to the shared

physical context. They are going to use decontextualized language more and more, and when entering school the use of decontextualized language is required from them.

The mastery of decontextualized language has often been associated with school achievement (Feagan, 1982; Snow & Dickinson, 1990; Snow, Burns, & Griffin, 1998; Dickinson & Tabors, 1991; Reese, 1995), since it requires children to rely on higher levels of reasoning, to make inferences and use abstract concepts that are disembedded from the supportive and immediate environment. According to Snow *et al.* (1989), the use of decontextualized language is crucial to success in classroom discourse and successful acquisition and development of literacy. The extent to which children are exposed to decontextualized language in the home environment has been demonstrated by many studies to be related to the children's school achievement.

In her pioneering ethnographic studies Heath (1982; 1983) argues that the amount and the quality of interactions during shared parent-child book reading in the pre-school period is an important predictor for academic success of the children. In these studies, both middle-class and working-class communities were involved. Heath found that children of middle-class parents were engaged in literacy practices from an early age, compared to children with working-class background. These studies also revealed that middle-class parents not only read picture books to their preschool children, but go beyond the visible context to talk about related events and ask questions about abstract concepts. Children were also asked questions about specific events in the book and how they related to the children's world and experiences.

In another study, Snow, Cancini, Gonzalez and Shriberg (1989) found that decontextualized language is crucial for academic success and that there is a strong correlation between decontextualized skills and school achievement. In addition, a wide variation has been found in both home and school environments in terms of providing and supporting these skills. Snow's (1983) study has demonstrated that parents' use of decontextualized strategies during conversations with their children can facilitate the process of early reading and writing. These outcomes are very much in tune with Snow's (1991) conclusions that mothers' use of decontextualized language to their preschool children during book reading is related to children's vocabulary development, reading comprehension abilities and literacy. Furthermore, other research suggests that children with low-SES backgrounds have difficulties using abstract language

that deals with non-immediate situations, because they are not exposed to this kind of speech in their home environment (Hart & Risley, 1992; Heath, 1994; Aarts, Demir, Kurvers, Laghzaoui & Henrichs, 2006). Children from low-SES families also consistently show weakness in tasks and activities such as explaining and defining things, retelling events and stories, and talking about objects that are not in the immediate setting.

In the school environment as well, the extent to which children are exposed to decontextualized language has been shown to be related to the development of different skills that shape their school career. The exploration of teacher-child conversations has demonstrated that the frequency and the type of talk to children are important determinants in children's oral language skills and literacy development (Massey, 2004). Smith and Dickinson (1994) have also demonstrated that cognitively challenging conversations have beneficial effects on children's literacy development and problem-solving skills.

Home-school comparisons in a study of 4-year-old children on book reading with preschoolers (Dickinson, De Temple, Hirschler & Smith, 1992) showed that mother-child conversations at home during book reading sessions were dominated by immediate talk, while in school teacher-child conversations were characterized by non-immediate utterances. Immediate talk consists of utterances where the speaker labels, counts, names colours, and comments on information immediately available in the book. This category contains many 'what' and 'where' questions. Non-immediate talk, on the other hand, refers to utterances that move away from what can be seen in the book and includes analysis, recall, discussion, or connection between the story and the child's own world. This category contains many 'why' questions (Dickinson, De Temple, Hirschler & Smith, 1992). The comparison of home and school contexts showed that mothers provide a basic introduction to book reading that teachers expand by engaging children in talks of a cognitively challenging type.

3.7 Academic language in Berber

As was pointed out in Chapter 2, the Berber language was introduced into public education in Morocco only in 2003. Since then, school textbooks and different educational materials have been regularly published, mostly for use inside the classroom. These texts are written in the official Tifinagh script, an

adaptation of the original north-African script found in many scriptures (see Chapter 2, and El Aissati, Karsmakers & Kurvers, 2011). No analysis has yet been made regarding the properties of these texts as part of academic language use. However, one can safely assume that they contain certain characteristics of academic language, since they are primarily meant for use in a school context.

The first surviving text in Berber, although more in the form of word list, dates back to the 12th century, when Ibn Tunart composed the first known Arabic-Berber dictionary (Van den Boogert, 1998). According to Van den Boogert, the Berber variety in which the dictionary was written is very close to the southern variety (Tashelhit), although the author himself is originally from the north of Algeria (Qal'at beni Hammad). The *Fonds Roux*[2], of the University of Aix-en-Provence has a total of 150 texts in Berber, from translations of works on Islam to dictionaries of medicinal plants and religious poetry (Van den Boogert, 1995).

For our purposes, it is noteworthy that there are no known manuscripts prior to the 19th century in Tarifit Berber, the variety which is the focus of the present study (see for example, Albino, 1859; Sarrionandia, 1905). During the protectorate period in Morocco (1912-1956) a number of similar works were published in Tarifit Berber, with annotated translations in French or Spanish (e.g. Biarnay, 1917, Basset, 1920). These were all written in an adapted Latin script.

Present day academic Berber, as mentioned above, is found mostly in publications of the Royal Institute for the Amazigh Culture (IRCAM), and in a number of publications of different authors, mostly poetry, plays and novels. The latter are found in all three major varieties of Berber, including Tarifit. Just how 'academic' is the language in which they are written is an empirical question that could be answered by studying them. Since this has not yet been done, a few examples from some of these works will be given below to illustrate the type of language in which they are written.

A few examples from El Aissati (2001) can illustrate the differences between the discourse of everyday usage and academic Tarifit Berber. Example 3.1 shows the type of sentences that could be used in everyday conversations.

---

[2] The Fonds Roux is a foundation at the University of Aix-en-Provence named after the French Arabist and Berberologist Arsène Roux (1893-1971). This foundation contains more than two hundred Berber and Arabic texts and manuscripts.

3.1 Aqzin nni y-arwel; qeloent idd x-as tzizwa
Dog that 3m.run away; chase-3.f.perf directional-particle on-him bees
(the dog ran away [because] bees had chased him [away from the tree])

In typical academic language, the conjunction 'minzi' (because) would be used to join the two clauses (see example 3.2):

3.2 Aqzin nni y-arwel, minzi qeloen-t idd x-as tzizwa
Dog that 3m.run away; because chase-3.f.perf directional-particle on-him bees
(the dog ran away because bees had chased him [away form the tree])

Example 3.3 shows the absence of the use of the conjunction 'u ca' (so, then):

3.3 Aqzin nni i-ssenhezz tacejjart nni; t-ewd2a-dd teghrast nni
Dog that 3.m-shake tree that; 3.f-fall down-here beehive that
(the dog shook the tree; the beehive fell down)

The two clauses would be more typical of academic discourse if they were connected by the conjunction 'u ca' (then; therefore; so) (see example 3.4):

3.4 Aqzin nni i-ssenhezz tacejjart nni, u ca t-ewd2a-dd teghrast nni.
Dog that 3.m-shake tree that; then 3.f-fall down-here beehive that
(the dog shook the tree; then the beehive fell down)

These examples are taken from narratives told by youngsters aged 12 to 14, based on a picture book used in the El Aissati study. Another set of examples given below is taken from the series 'Iles Inu' (my language) which was developed as supporting material for learners of Tarifit (E-rramdani, Raha & De Santos-Tirado, 1998):

3.5 […] tugha ij n ugharda i-zeddegh di taddart tes2beh2 di temdint di weh2fur nnes ameonan, i y-bna sadu lmaryu n kuzina.
Was one of mouse-gen 3.m-live in house nice in city in hole his small, which 3.m-build under wardrobe of kitchen
(there was a mouse which lived in a beautiful house in the city, in a small hole which he had built/dug under a kitchen cupboard)

Sentence 3.5 illustrates an important characteristic of academic discourse, namely complex sentence structure exemplified by a relative clause.

Chapters 5 through 7 are concerned with the emergence of academic discourse among children aged 3 to 6 and the academic discourse properties exhibited by their mothers in their speech, at the lexical level, the morpho-syntactic level and the textual level.

CHAPTER 4

Research design and methodology

The main purpose of this study is to gain more insight in the characteristics of the emergent and early development of the academic language register of bilingual Moroccan Berber children growing up in the Netherlands. This in-depth study is based on the analysis of factors related to the linguistic input at home and at school. It is aimed to provide a detailed analysis of *which* features of the academic language register in both L1 and L2 are provided in the input of parents and teachers, *which* features of academic language are detectable in the language use of the children between the ages of three and six, and *what* the relationship is of all these with input features and background variables such as SES and literacy. This chapter starts by presenting the main research questions guiding the whole study (section 4.1). Section 4.2 provides information about the recruitment procedure of the families and schools involved in the current study. In section 4.3, an overview of the instruments used in this study is presented. Section 4.4 gives an overview of the informants in this study: the children, their families and their teachers. Finally, section 4.5 contains a description of data collection, transcription, coding and analysis.

4.1 Research questions

The objective of the current study is to examine the co-construction of academic language of Moroccan Berber children in communicative contexts at home and at school. In order to do so, the input of caretakers and the output of children will be analyzed on lexical, morpho-syntactic and textual levels. We will do this by trying to find answers to the following main research questions:

1 To what extent do Moroccan Berber mothers use features of academic language in their input to their preschool children between the age of 3 and 6?

2 Is the use of academic language features by the mothers related to SES and home literacy practices?
3 To what extent do teachers use features of academic language in their input to Moroccan Berber school children?
4 To what extent do children use features of academic language between the ages of 3 and 6 in the home and school settings?
5 Is there a difference between mothers and teachers regarding the use of academic language features?
6 Is there a relationship between the parents' and teachers' input and the children's use of academic language features?

On the basis of these comprehensive questions, more specific underlying research questions will be formulated in each of the empirical studies in the following three chapters.

4.2 Recruitment procedures

4.2.1 Recruitment of families

The in-depth study reported on here is part of a larger longitudinal study on academic language development at home and at school. 12 Moroccan Berber children (four boys and eight girls) and their caretakers living in the Netherlands in the cities of Utrecht, Ede, Tilburg and Oosterhout were followed during a period of 3 years, starting at the age of 3 until the age of 6. The families were recruited by asking the municipalities to provide us with the addresses of Moroccan children born between October 2001 and May 2002. After receiving the addresses, parents were contacted with a letter in Dutch and Arabic explaining the purpose of the research project. Subsequently, the families were visited by trained female research assistants belonging to the same ethnic group and speaking Tarifit Berber as well as Dutch. These assistants were also trained to use the scripted form of Tarifit Berber to be used in administering the questionnaires and assessments.

Among the criteria considered for the selection of children is that none of them should belong to a single-parent family or has developmental delays or medical problems related to speech or hearing. In addition, children attending playgrounds or day care centers for more than two days per week were

excluded. Another selection condition was that parents used Tarifit Berber for at least 70% of the time in their daily interactions at home. The positive response rate among the contacted Moroccan families was 45%. First of all, part of the initially contacted families was not speaking Berber, but Moroccan-Arabic. Other reasons for not participating for the Berber-speaking families were mainly due to the presence of the video camera and other unexplained or unexpected privacy matters. Further reasons for refusal of cooperation were lack of time or interest. Table 4.1 gives an overview of the number of informants during the whole process of data collection.

Table 4.1: Number and mean age of children during the three measurement times

	Time 1	Time 2	Time 3
Mean age children (years)	3;2	4;2	5;10
Number of children	17	14	12

As can be seen from Table 4.1, at the beginning of the study, a total of 17 Berber families were involved in this project. Over the course of data collection, several children dropped out. After measurement time 1, two families informed us that they could not continue because they found data collection too time-consuming. In measurement time 2, we received a withdrawal from one teacher. The reason given for this was that the teacher had many other tasks next to teaching. Just before measurement time 3, two other families withdrew from participating in the last round of data collection. One of these families did not want to participate anymore because of the presence of the video camera and the other family did not give a clear motivation. These five children who missed one or more measurement rounds of data collection were omitted from the analysis and only the twelve families that participated in all measurement times were kept.

4.2.2 Recruitment of schools

Just before the participating children reached school age (4 years old), and after the first home data collection was finished, parents were asked to sign a written permission form which allowed us to visit their children in school and videotape some school lessons they would attend. After the parents' agreement, schools were contacted by phone and the objective of the research was

explained to the headmasters. This was followed by a written letter for the class teacher which explained the aim of our research and how the data collection would proceed. Initially, all teachers agreed to participate. However, after measurement time 2, one teacher did not want to take part in the last round of the data collection. The core group of children which participated in all school measurement times consisted of 12 children.

4.3 Instruments

The following instruments were used to obtain the data used in this project: (1) interaction tasks designed for mother-child and teacher-child conversations in naturalistic settings to get in-depth linguistic data; (2) a detailed questionnaire dealing with family background, language use and literacy activities in the home context and a short questionnaire for the teachers; (3) vocabulary tests in the first (Tarifit Berber) and second language (Dutch). All items of the original Dutch interviews, which were used during the home data collection, were translated into Tarifit Berber

4.3.1 Home interaction tasks

At home, mother-child dyads were observed while involved in two interaction tasks of different genres during each of the three visits: a book task and a picture task. The age of children at the three measurement times was respectively 3;2, 4;2 and 5;10 years. The mothers were instructed to act as they would normally do. All the sessions were videotaped. It should be mentioned that initially, next to the book task and the picture task, two other tasks were designed to be used at home: block construction and mealtime conversation. However, the mealtime conversation task was used only at measurement time 1, because most parents felt very uneasy about their meal time being videotaped. And although the block construction task was used during the whole data collection, we decided to exclude this task as well and concentrate on the book and picture task as it appeared that transcribing and coding all the data was very labor-intensive (see section 4.5).

Book task

In the book sessions, mothers were asked to 'read' to their children (or tell the story of) a book that the research assistant brought with her. Since Berber was not known as a written language and mothers therefore were not used to read Berber, the book session was first prepared. Together with the mother, the research assistant went through the book while the child was not present. Two procedures were adopted: (1) for mothers who could read: they read the book in Dutch and the assistant provided them with the Berber translation. (2) For mothers who could not read: the research assistant read the book twice in Berber to the mother, while the mother looked at the pictures in the book and, if necessary, asked questions.

In the session with the child, the mothers were asked to first read the book to their child or tell the child the story in Berber. After that, they were asked to go through the book once again and engage their child in a talk about the content of the book. This interaction with the children was the data-source for transcription and analysis.

During the first measurement time, the book used was "*Tijn op de fiets*" (Tijn on the bicycle) (Sluyzer & Oud, 2003). *Tijn op de fiets* is a children's book about traffic and transport, and what you may encounter on the road. Tijn and his mother go together on a bike to visit his grandparents. Tijn sits on the front seat together with his bear. On their way they encounter a truck, a train, a boat and a plain. Furthermore, they buy fruit from the market. From his grandparents, Tijn gets his own means of transport: a learner bike!

In the second measurement time, the book "*Tim op de tegels*" (Tim on the paving stones) (Veldkamp & De Boer, 2004) was used. *Tim op de tegels* is a humorous picture book where the toddler Tim takes his father's command to stay on the pavement very literally and as a consequence gets involved in dangerous situations. Tim wants to play outside but as his father has no time to join him, he commands him to stay on the sidewalk. As promised, Tim stays on the paving stones, even when they are lifted by pavers, by a truck driver, dock workers and a crane. When father looks out of the window, he sees Tim being carried away. He leaves everything and rushes after his son. This leads to hilarious and exciting chase scenes.

The last book "*Emma in het spookhuis*" (Emma in the ghost train) (Waechter, 2007) was used in measurement time 3. *Emma in het spookhuis* is about Emma, a rabbit, who dreams at night of scary monsters. This frightened rabbit is smart

enough to get help from a doctor who specializes in bad dreams and nightmares. The dream expert advises her to read a self-help book about monsters. After reading this book and doing the exercises, Emma wants to put her freshly acquired knowledge into practice. She decides to go into the haunted house. Inside it is "just as scary as she had imagined". What follows is a hilarious fight in which Emma beats the monsters. However, as a result she is denied access to the haunted house for one year. But she has lost her fear and nightmares for ever!

Picture task

For the picture interaction tasks, three different pictures, taken from Leiber (1999), were used in the three home visits. These were semi-structured tasks on the basis of which mothers were instructed to talk with their children. The mothers were asked to interact with their children as they would usually do. The duration of the picture task in the three measurement times took between 5 to 10 minutes of time. These tasks elicited both personal and impersonal genres since children could make associations between what they see and their own experiences. In the first measurement round, the picture "*Op straat*" (In the street) was used. It is a general view of a busy street where a range of activities takes place, such as parents accompanying their children to school, people walking or driving, and work activities. The picture used for the second measurement round is "*Circus Pepo*". It is an arena with various scenes, performers and visitors. It contains acrobats, clowns and animals like lions, giraffes, horses and the like. Finally, the picture "*Het klaslokaal*" (The classroom) was used in the third measurement round. This picture depicts children in different settings in the classroom. In the presence of their teacher, children form small groups and are involved in different activities: drawing, playing music, constructing with blocks, using the library, etc.

4.3.2 School interaction tasks

Children were visited at school during two measurement times at the age of 4;2 years and 5;10 years respectively. At the beginning of the data collection at the schools, we started recording three settings: book reading, circle time (group discussion), and a crafts session. However, due to time and labor constraints, we report in this study only on two school interaction tasks: book task and

circle time. These tasks elicit both personal (circle time) and impersonal (book task) narratives. The teachers were invited to follow as much as possible their daily routines, for example in working with smaller groups or the whole class, or in the ordering of activities. Since the aim of the study is to report on language use common in classroom practices, teachers were asked to act as they normally would do.

Book task

For this task, two books were selected which the children did not know before. Teachers were asked to read the book to the children (or tell the story) in the way they usually do. This could be in reading to the whole class or to a smaller group of children (including the target child). After that the teacher was asked to engage the children in a talk about the content of the book. The duration of the book task in both visits took between 10 to 15 minutes.

In the second measurement time, the book used was "*De Verrassing*" (The surprise) (Van Ommen, 2003). A woolly sheep weighs herself and measures how thick her wool is. Then she goes on her scooter to buy a bright red color paint. Back home she colors and shaves her wool. The sheep has a plan for her beautiful red wool, but that remains a surprise until the end.... Her friend giraffe gets a gift, a beautiful woolen sweater made entirely by the sheep.

In the third measurement time, the book "*Wie is er hier het dapperst?*" (Who is the bravest?) (Pauli & Schärer, 2006) was presented. The book is about mouse, snail, frog and sparrow who are sitting near the lake wondering what to do. Frog suggests having a contest to find out which one of them is the bravest. So, the mouse swims underwater back and forth, the frog eats a lily pad and the snail comes out of its shell to crawl. And then it is the turn of the sparrow... After a long reflection, she lets the other animals know that she will not join the contest! Then the animals decide that sparrow is the bravest, because daring to say "no" is very brave.

Circle time

In the circle time, the teachers were asked to form a circle and talk with children in a group discussion in the way they do on a daily or weekly basis. The subject matter of the discussion could be about what the children did in the previous day or weekend or about a certain theme which the teacher could freely choose. The teacher was asked to sit next to the target child so that we

could videotape them together. The teacher was also asked to involve the target child in the discussion by asking questions and reacting to what she or he said. The task took between 10 to 15 minutes and was recorded on video tape.

4.3.3 Questionnaires

Home questionnaire

The questionnaire used for the whole DASH study contained questions about different demographic characteristics of the families (ethnicity, education, work and literacy), about daily informal education (family outings, media, songs and stories, reading behavior within the family, games, language use with siblings) and about their pedagogical background (parenting and daily hassles, the presence of social support from family and acquaintances concerning parenting, parental stress, child rearing practices, parental attitudes towards children's upbringing and education, children's behavior and children's motor development).

Moroccan Berber mothers, as the main caregivers, were interviewed by trained female research assistants originating from the same ethnic group. As already mentioned before, these assistants had a command of Tarifit Berber and Dutch. As some mothers were illiterate, the questionnaire was administered in the form of a personal interview in the language the mothers preferred, Dutch or Tarifit Berber. The background variables used in this study, socioeconomic status of the family and reading and writing behavior (literacy), will be described here.

Socioeconomic status

Socioeconomic status (SES) is a compound variable, usually comprising educational level, occupational prestige and income (Hoff, 2006). In dealing with the cognitive and linguistic development of children, the educational level of the mother is often adopted as the most important index to define SES (Ensminger & Fothergill, 2003). In the current study, family SES was calculated by relying on two constituents: the highest received educational training of both parents and the status of their current occupation. Concerning education, the mothers were asked about their, as well as their husbands', educational career in both the Netherlands and Morocco. Equivalence lists were used to judge the education in Morocco. The educational level of the parents was rated on a scale

that ranged from 1 (no education) to 7 (college or university degree). As to employment, parents were asked about their current or last job, which were given scores ranging from 1 (unemployed) to 6 (academic level job). Professions were categorized according to the Dutch standard job index (Standaard Beroepenclassificatie, CBS, 2001). Family SES was, thus, computed as the mean of parents' highest received education and the current occupation.

Home language and literacy practices

The children's exposure to language input through literacy activities in the family setting was assessed by a questionnaire based on the Early Childhood version of the HOME observation scheme developed by Bradley and Caldwell (1984). The list of the questions addressed to the mothers covers the activities which are thought to be a stimulant for vocabulary growth and language development of children, such as shared book reading, name and letter writing, dealing with printed texts, counting, oral story telling, sharing talks during dinner time, talking about general topics that might happen at school or when playing with children, singing songs, reciting rhymes, watching television (for both entertainment and educational purposes) and using the computer. The mothers were asked to estimate how often they are engaged with their children in these kind of activities on a five point scale ranging from 1 (no, never) to 5 (yes, at least once a day). Because book reading in Berber is practically unknown to these parents, the mothers were asked to report on these practices in Dutch, and whether they used Berber at all, for example to explain new concepts. The mothers were also asked about the frequency of reading and writing of both parents, rated on a scale from 1 to 7.

School and teacher questionnaire

A questionnaire was also presented to the teachers. Teachers were asked about their educational training and their experience in teaching in primary education. In addition, teachers were asked whether they learned anything related to first and second language acquisition and multilingualism during their training. Teachers were also asked if they were fluent in any of the minority languages spoken by one of their pupils. Finally, teachers were asked questions about the composition of their classes, such as class size, number of second language learners and number of Moroccan-Berber children and whether assistance was available.

4.3.4 Vocabulary tests

The children's Tarifit Berber (L1) and Dutch (L2) vocabulary was assessed by using the *Diagnostic Test for Bilingual Development*, developed by the Central Institute for Test Development (CITO) (Verhoeven *et al.*, 1995). This test, later computerized by the DASH research group, is used to measure, among other things, children's receptive vocabulary and has been designed for researchers and educators dealing with bilingual children in Dutch education. The test contains 30 items (Cronbach α=.83) and the words originate from a database of vocabulary that is considered by teachers to be important for young children to be able to deal with school tasks and participate successfully in Dutch elementary education. On a computer screen, four pictures are shown to the child. The research assistant mentions the target word and the child is asked to match it with one of the four pictures. The test was administered during each of the three home visits by a trained research assistant who belongs to the Moroccan Berber community. Testing went on until the child failed five consecutive items or otherwise finished all 30 items of the test. The test score is the number of items correctly identified by the child.

4.4 Participants

Families

The twelve families that took part in all the measurement moments of this study were all families where the father and mother were living together with their children. All the mothers said they were raising their children together with their husbands. All mothers were born in Morocco and their mean age was 31.3, ranging from 23 to 41 years old at the beginning of the data collection in 2005. The mothers had lived in the Netherlands for an average of 16.6 years, varying from 5 to 26 years. The educational level of the mothers ranged from none to higher vocational education. More specifically, three mothers indicated that they had never attended school. Two mothers finished primary education in Morocco. Two mothers followed secondary education. Four mothers reported that they finished middle-level secondary vocational level (MBO). Finally, one mother attended higher vocational education. With regard to the profession of the mothers, it turned out that five mothers were housewives, three mothers were shop assistants, three mothers were assistant carers with the

elderly or geriatric nurses and one mother was a pharmacy assistant. It should be mentioned that four of the seven mothers with professions mentioned here were jobless or disabled at the time of data collection.

The fathers of the participating children were also all born in Morocco. Their mean age was 38.9 years, ranging from 30 to 57 years. The number of years the fathers had lived in the Netherlands was on average 14.8 years, ranging from 4 to 30 years. Concerning the educational levels of the fathers, six of them indicated that they finished primary education and one attended Quranic school. two fathers finished secondary education, two attended higher vocational education (HBO) and one attended university. As to professions of the fathers, seven of them were factory workers or cleaners, one worked as a truck driver, one as a flower picker and one as a baker. Finally, one father was a claim handler and one an assistant manager. Two of the fathers were disabled at the moment of data collection.

Schools and teachers

The total number of teachers in the two measurement times conducted at schools was 17. Two children were taught by two part-time teachers, and three children had different teachers during measurement time 2 and 3. All teachers were females. As to the education of the teachers, eight of them obtained a bachelor of education at a teacher training institute for primary teachers (*Pedagogische Academie voor het Basisonderwijs*), three teachers had a degree of a preschool teacher at a training institute (*Kleuterleidster Opleiding School*), and one teacher had an academic degree. Their teaching experience ranged from 1.5 to 30 years (mean=13.6). Eight teachers indicated that they had followed some courses related to language acquisition. The same number of teachers said that they had followed courses on teaching Dutch as a second language. Furthermore, five teachers reported that they followed classes on multilingualism and second language acquisition during their training. Only one teacher indicated that she speaks a minority language, Tarifit Berber, as well as Dutch. The total number of the children in the classrooms of the target children ranged from 12 to 27, with a mean of 20.42. The total number of children with a non-western background ranged from 3 to 22 pupils (mean=12.42). Finally, the number of Berber children in the classrooms ranged from 1 to 9 children (mean=4).

4.5 Data collection

As already mentioned, the data collection consisted of three measurement times at home and two at school. Concerning the home visits, the first measurement time took place when the children were 3;2 years old, the second measurement time at the age of 4;2, and the third measurement time at the age of 5;10. During each data collection round, the families were visited twice by the assistants. The first visit, which took about 2.5 hours, was dedicated to administering the questionnaire including general family demographics, language profile, language choice and language use at home. The second visit was devoted to the mother-child interaction tasks and to administering the standardized vocabulary tests. Fathers were not involved in the interviews. Since mothers spend most time with children and are considered as primary caregivers within the Moroccan community and subsequently as providing the most input for children, we decided not to involve the fathers. After every data collection round the families received a gift voucher and a children's book as a reward for their participation.

As for the data collection at school, children were visited and videotaped at school after we received the parents' written permission. The first school visit took place when the children were 4;2 years old and the second one at the age of 5;10. At the schools, the data were collected by the research assistants and the investigator/author of this book. All school visits took place in the mornings and data collection sessions took on average 2 hours.

A total of 60 videotaped sessions of mother-child and teacher-child conversations resulted in a language corpus which has been transcribed for the current study. Table 4.2 provides an overview of the total number of utterances (see 4.6 for explanation) in the transcripts, resulting from the interaction tasks in all measurement times at home and at school.

Table 4.2: Total number of utterances in home and school transcripts during 3 measurement times

		Time 1	Time 2	Time 3	Total
Mothers (N=12)	Book task	1,265	952	749	2,966
	Picture task	1,057	1,295	908	3,260
Teachers (N=12)	Book task	–	1,282	1,540	2,822
	Circle time	–	1,187	1,359	2,546
Children (home) (N=12)	Book task	469	391	301	1,161
	Picture task	624	689	702	2,015
Children (school) (N=12)	Book task	–	170	201	371
	Circle time	–	150	218	368
Total		3,415	6,116	5,978	15,509

The whole corpus consists of 120 transcripts and 15,509 utterances. The home mother-child interactions delivered 72 transcripts and 9,402 utterances (mothers=6,226; children=3,176). School tasks resulted in 48 transcripts and 6,107 utterances (teachers=5,368; children=739).

All the home and school data were transcribed verbatim. The interactions have been transcribed using the *Child Language Data Exchange System* programme (CHILDES), following the conventions of *Codes for the Human Analysis Exchange System* (CHAT; MacWhinney, 2000). This database has become an important source for the study of child language. The choice of CHILDES for transcribing our data was due to many reasons. The program consists of several tools which enable automatic treatment of the data. Another reason is that computer-based files are easy to exchange and different data can be easily compared. In addition, the uniformity of the rules makes transcripts reliable.

4.6 Transcription

Transcription of all interactions started when the assistant or researcher had finished the instructions, and the mother or teacher got ready to start the task, for instance by asking the child to sit down, pointing to the book or picture. Sometimes the task began immediately after the camera was started. In this case transcription starts directly at beginning of the recording. The transcription

stopped when the task was finished. However, to make sure the transcripts did not over-run, a time limit was set for every task, as will be explained in section 4.6.1.

Every transcript is made of utterances. An utterance is one single discourse act, expressing one complete, sometimes complex, meaning proposition and/or intention. An utterance is mainly recognized by the intonation contour (what sounds as belonging together) and by the fact that an utterance is mostly a complete statement or a complete piece of information, by a clear pause or a turn-taking by another speech partner. Every utterance is transcribed on a separate line. Transcription was done by using the standard orthography, not phonetic symbols. We transcribed as literally as possible what the participants were saying, including mistakes, hesitations, false starts, corrections and repetitions.

4.6.1 Transcription of home data

In order to prevent the transcriptions from getting too long and make sure they have comparable length, a transcription time limit was set for every task. For the book task, the maximum time to be transcribed was 9 minutes. For the picture task, the maximum number of minutes was 8. Transcription of Berber conversations turned out to be not an easy task during this project. Although the author as well as the research assistants are native speakers of Tarifit Berber, they did not have previous experience of using this language in a written form. Previously, they had used Tarifit Berber for conversational rather than academic purposes. Given this, a number of decisions had been taken to ensure uniform and consistent transcripts.

The default language is given in a header tier: @Languages: (be, nl). For the home setting, Berber is the language in which the mother was asked to interact with her child. If a word was said in another language, for instance in Dutch, this was coded with the symbol "@nl" directly following the word. For example:

*MOT: waniti d bus@nl t-zri-d.
(this is a bus, you see)

If a complete utterance was in Dutch, we added the symbol [$nl] at the end of the utterance, but the default language remained Berber. In school, the default language for all children was Dutch.

While transcribing, and due to Berber's rich morphology, consistent word segmentation was developed and adopted by distinguishing the stem from prefixes and suffixes. Word stems were consistently distinguished from affixes by using a hyphen. Another challenge we faced during the transcription was that the informants spoke local Berber varieties that phonologically slightly differed from each other. The variety used for transcription in this study is that of the central Rif area as it is the mother tongue of most Berbers living in the Netherlands. In Appendix 1, by way of example, a transcript of a mother-child interaction in Tarifit Berber is presented. Since Berber has several phonemes that do not occur in English or Dutch and the CHILDES system uses ASCII characters, it was decided to use only these characters instead of phonetic symbols. Table 4.3 presents an overview of symbols used to transcribe Berber consonants which do not have equivalents in ASCII system.

Table 4.3: Unconventional Tarifit Berber consonants as used in the CHAT program

Berber consonant (IPA)	The form of transcription in CHAT program	Example
Voiced pharyngeal fricative [ʕ]	/o/	oan (to push)
Voiceless pharyngeal fricative [ħ]	/h2/	ah2enjir (boy)
Voiceless velar fricative [X]	/x/	tixsi (sheep)
Voiced velar fricative [ɣ]	/gh/	aghrum (bread)
Voiceless alveo-palatal fricative [ʃ]	/c/	ancuc (lip)
Voiceless uvular stop [q]	/q/	aqrab (bag)
Voiced emphatic dental stop [dˤ]	/d2/	d2ad2 (finger)
Voiceless emphatic dental stop [tˤ]	/t2/	tghat2 (goat)
Voiceless emphatic dental fricative [sˤ]	/s2/	s2abet (harvest)
Voiced alveolar fricative [z]	/z2/	Anz2ar (rain)

4.6.2 Transcription of school data

Like in the home interaction tasks, also the school data were transcribed by applying the conventions of CHAT and CHILDES. Since school interactions were part of the lessons, they usually were longer than mother-child inter-

actions. Because of time constraints, we decided to transcribe only parts of the recorded conversations. From each school task, 8 minutes of time were transcribed. The total time of the transcribed sample comprises three episodes: three minutes at the beginning of the task, three minutes in the middle of the task and two minutes at the end of the task. As to the middle fragment, three minutes were selected beginning where the target child talks or is invited to speak. In Appendix 2, an example of a transcript of teacher-child interaction is presented.

4.7 Coding the transcripts

To be able to allocate the distinctive features of academic language in the recorded interactions, a coding scheme has been developed by the DASH research group (DASH, 2006). This coding scheme is applicable to the three languages used by the subjects of the three in-depth studies: Berber, Turkish and Dutch. In total, four linguistic levels are distinguished within the coding scheme. In the following sections, this coding scheme encompassing the three essential dimensions of functional grammar (field, tenor and mode) will be discussed. The whole coding scheme is discussed here with a detailed focus on the aspects which are used in the current study. The DASH coding scheme (DASH, 2006) contains four main levels: the lexical level, the morpho-syntactic level, the textual level and the socio-pragmatic level. In each of these four levels, elements can be retrieved of the functional grammar division of field, tenor and mode, as covered in Chapter 3. However, in the present study, only three levels are selected for analysis, namely lexical, morpho-syntactic and textual levels.

Lexical level

Measures we will use on the lexical level are lexical diversity and lexical density (see Chapter 5). In order to calculate lexical density, we need to distinguish content words from function words and from other categories such as question words and formulaic openings and endings. The term 'content word' here signifies all parts of speech that bear a meaning itself or refer to objects, actions and characteristics. 'Function words' have hardly any lexical meaning or are ambiguous. Instead, they serve primarily to organise and form the meaning of a

whole sentence or to express grammatical relationships with other words within a sentence. They can also specify the attitude or mood of the speaker. The difference between content and function categories (also called lexical/ grammatical, major/minor parts of speech) corresponds to the distinction drawn between open and closed classes (Deuchar, 1999).

Within the category of content words, four word classes are distinguished: nouns, verbs, adjectives and adverbs. The categories we coded as function words are articles (*the, a...*), pronouns (*I, you, my, your...*), prepositions (*on, under, in...*), conjunctions (*because, when, thus, except...)* copulae (is, are...), in addition to other words, particles and adverbs with little lexical meaning. Counting (1, 2, 3 etc.) is considered as a function word; but in sentences such as: *I have two books*, *two* is a content word. Onomatopoeic phrases are considered as content words where they replace a proper content word or noun. For example: *Do you see a wafwaf*? Proper names (*Peter*), and also family names (*grandma, daddy*) are considered as content words. Modal verbs, like *kunnen (can), mogen (may), moeten (must), hoeven (have), willen (will)* are coded as content words. Compound words in Dutch that are written as one word such as *appelboom* (*apple tree*) are coded as one content word. In Berber, there are compounds in which the words are linked by grammatical affixes or propositions such as *tazribt n tghatten (stable of the goats)*. In this case, we code two content words. Compound verbs, which are frequent in Dutch, like *boodschappen doen (to shop)* are one concept and we code it as one content word (verb). Criteria for taking compound verbs as one concept apply wherever it is a fixed expression, and the verb does not have an independent meaning on its own. Examples are: *indruk maken (to impress), pijn doen (to hurt), blijven zitten (to repeat a class)*. Another frequent verb category in Dutch is the separable verbs (*scheidbare werkwoorden*) where the verb is composed of a verb stem and a separable affix such as the verb *aftreden (to resign)*. These verbs are in fact one concept but can get separated in the sentence, as in this example: *Hij treedt morgen af. (He will resign tomorrow*). These verbs are coded as one word. Every word is coded in order of appearance in the sentence. In the following example, coding at the lexical level is illustrated:

*MOT: ze zijn de riolering aan het maken. (*they are repairing the sewage*)
%lex: $F $F $F $C:noun $F $F $C:verb

Morpho-syntactic level

In our analysis a distinction is made between lexicalized, nominal and pronominal categories for both subjects and objects. On the morpho-syntactic level, we concentrated on coding clause combining, distinguishing between four types of clause-combining constructions: coordinate clauses (coor), complement clauses (compl), relative clauses (rela) and adverbial clauses (adve). A coordinate clause is the combination of two clauses that can stand on their own, independently. For instance: *He tried hard, but he failed.* An example of coding of coordinate clause:

MOT:* hij stond op en ging [:weggaan] weg. *(he stood up and left)*
%ms2: $CC:coor

Complement clause is defined as an argument of the predicate in the matrix clause. For instance: *Peter promised that he would come.* An example of coding a complement clause is:

*MOT: vertel me eens wat je ziet. *(tell me what you see)*
%ms2: CC compl

Relative clause is defined as an attribute of a noun or noun phrase. For instance: *Anne bought the bike that was on sale.* An example of coding the relative clause:

*MOT: de man die ging fietsen is terug. *(the man who went cycling is back)*
%ms2: $CC:rela

Adverbial clause is defined as some sort of modifier of the associated matrix clause or verb phrase. For instance: *He arrived when Mary was just about to leave.* An example of coding the adverbial clause:

*MOT: ik las de mail toen ik thuis was. *(I read the e-mail when I was at home)*
%ms2: $CC:adv

As a general rule, if the mother, teacher or child literally repeated what the other one said, we coded it on the morpho-syntactic level. Self-repetitions were coded only once on the morpho-syntactic level. Formulaic speech or phrases

that have become idiomatic such as *alsjeblieft (please)* or *goed zo (well done)* are not coded on the morpho-syntactic level. Also interjections like *hmm* are not coded on morpho-syntactic level. Insertions of another language or the use of a word from another language in an utterance is coded on the morpho-syntactic level.

Textual level

On the textual level, the focus is on the abstraction level, more specifically in reference to contextualized vs. decontextualized context. The level of abstraction is defined as the extent to which the content of the interactions abstracts from the directly visible here-and-now. We use a four-level classification of abstraction developed by Blank, Rose and Berlin (1978).

The first level is *matching* perception (coded as $AL:mat), which is the lowest level of abstraction. It means focusing on only one item of perceptually available information. It is used for labeling, naming objects, saying or asking about the name of an object or person. For example:

MOT: wat zie je? *(what do you see?)*
%tex: $AL: mat

The second level is *selective analysis* of perception (coded as $AL: sel). At this level relationships between perceptually available items are being drawn or integrated. The utterance does not go beyond what is perceptually available. This includes questions or comments that require the child to focus on specific aspects of objects and events and to integrate different components. This level can be used for describing characteristics such as number, size and color or describing scenes and actions. It also contains recalled information previously mentioned in a text or a conversation. For instance:

*MOT: pak het groene blokje. *(take the green block)*
%tex: $AL: sel

The third level is that of *reordering* about perception (coded as $AL: reo). At this level the utterance goes beyond immediate perception, making deductions about what happened, identification of oneself with someone else, recalling one's experiences and the like. Utterances on this level include defining meanings, summarizing stories or events, providing a point of view, identifying

similarities and differences, recalling one's own experiences, expressing a desire to do something or to have something and evaluating and judging something. For instance:

*MOT: dat was een mooi book. *(that a was a nice book)*
%tex: $AL: reo

The fourth level is *reasoning* about perception (coded as $AL: rea). It is used when relating the actual situation to what might, would or could happen (prediction) to the perceived objects or people. At this level, which is the most abstract one, language can be used to make predictions about what will happen next in a story or hypothesize about the outcome of an event; to consider the causes of events in order to formulate solutions to problems; to explain story concepts and actions and reflect on factual world knowledge. In addition, utterances can provide general information that helps the child understand the story. For example:

*MOT: wat eten konijnen? *(what do rabbits eat?)*
*tex: @AL: rea

For all levels, only task-related utterances are coded. Procedural utterances (for instance: *let's get started*) are not coded for this level. If the utterance is an exact repetition of the foregoing utterance of the other speaker, it is not coded for abstraction level. Repetitions of one's own utterances are coded only once.

4.8 Data analysis

After transcribing and coding all data by using the coding scheme, the frequencies of all coded levels were computed by using the language analysis program CLAN (MacWhinney, 2000). For calculating vocabulary diversity, the index D was generated by using the software program *vocd* which belongs to CLAN. The software program and the CLAN manual are freely available and downloadable (http://childes.psy.cmu.edu/). The CLAN manual describes the *vocd* program and its computational attributes.

The quantified data were manually exported to the SPSS program (version 15.0 and 16.0) and analyzed by means of different statistical tools. To be able to

answer the research questions formulated in each empirical chapter on different aspects of academic language features, we use different statistical procedures. After presenting the descriptives of the mothers' input, the time effect and the task effect (in Chapter 5) are analyzed using ANOVAs with repeated measures. After that, a Kendall's Tau correlation analysis of SES and home literacy practices with the linguistic variables is conducted (Field, 2009). At the level of the teachers' input, ANOVAs with repeated measures are also used to measure time and task effects. In addition, the input of mothers and teachers is compared by using the non-parametric Mann-Whitney U test. The same procedure is followed with respect to children in the home and school settings. After presenting the descriptive statistics of the academic language measures, the effect of time and tasks are analyzed by using ANOVAs with repeated measures. A Kendall's Tau correlation analysis is also conducted on SES and literacy background with academic language variables. To compare children's features of academic language at home and at school, Wilcoxon Signed-rank tests are conducted. Task effects are not treated in Chapters 6 and 7, as we use only one task in those chapters, namely the book task. To relate features of academic language in the input of mothers and teachers to features in the language of the child, Kendall's Tau correlation analysis is conducted.

CHAPTER 5

The lexical level of academic language: density and diversity

The main goal of this chapter is to study the lexical component of academic language. We examine the extent to which Moroccan Berber children are exposed to academic language at home and at school. In the current study, we hypothesize that children who are exposed to dense and diverse vocabulary, which are features of academic language, will more easily develop a dense and diverse vocabulary, which is required in the school context.

The organisation of this chapter is as follows. First, the research questions are formulated. After that, the operationalization of the lexical features of academic language is presented in section 5.1. In section 5.2, the findings concerning the lexical richness in mothers' input is examined. In section 5.3, the results of the lexical richness in the teachers' input are reported. In section 5.4, the input in the home and the school settings will be compared. Section 5.5 deals with the lexical richness in the children's language output. In section 5.6, the academic language input-output relationship will be explored. Finally, in section 5.7, some general conclusions are drawn on the basis of all these results.

This study covers lexical characteristics of the language used by both children as well as their parents and teachers in home and school environments. Our first assumption is that the use of a high quantity of speech, both at home and at school, towards children on different subjects provides and stimulates the kind of language expected at school. In addition, we hypothesize that family SES and literacy activities at home are related to the input of the mothers as well as to the children's output. Against this background, the aim of this chapter is to investigate the quantity and quality of the lexical input of the caretakers (mothers and teachers), of children, and to relate features of the input to output features of the child's language. This will be done by answering the following research questions:

1 To what extent do Moroccan Berber mothers use lexically rich language in their input to their preschool children?
2 Is the use of lexically rich language by the mothers related to their SES and home literacy practices?
3 To what extent do teachers use lexically rich language in their input to Moroccan Berber preschool children?
4 Is there a difference between mothers and teachers regarding the use of lexically rich language?
5 To what extent do these children use lexically rich language at home and at school?
6 Is there a relationship between the lexical richness of parents' and teachers' input and the children's use of lexically rich language (output)?

5.1 The lexical features of academic language operationalized

As explained in Chapter 3, different distinctive features at the lexical level are considered to be typical of academic language. Before proceeding, it is worth mentioning that there is a controversy in the literature concerning vocabulary and lexical assessment. In some studies, 'the terms lexical diversity and vocabulary richness appear to be synonymous' (Durán *et al.*, 2004: 3) and diversity is determined by counting the total number of different words in a sample. In other studies, the degree of difficulty or rarity of words is considered (Vermeer, 2000; Daller, Van Hout & Treffers-Daller, 2003). Still other studies state that the components mentioned before are complementary to each other (Malvern & Richards, 2000; Read, 2000). In the current study we consider two lexical variables in order to measure lexical richness: lexical density and lexical diversity. Lexical richness is used here as an umbrella term for lexical density, lexical diversity and lexical rarity. However, in this chapter, we only use lexical density and lexical diversity.

Lexical density

Lexical density refers to the ratio of lexical words (or content words) used in a text or conversation to convey a message. According to Ure (1971), lexical density is the proportion of tokens that are content words as opposed to function words. This notion of lexical density was used to measure 'literacy' in

written and spoken language. Academic texts are characterized by more lexical density than texts occurring in informal and non-academic settings (Schleppegrell, 2001). Academic language is more dense than informal language and contains a larger proportion of nouns, verbs, adjectives and adverbs. Due to their lexical density, academic texts efficiently pack more information into each sentence, while interactional texts are less characterized by lexical *thickness*. School-oriented texts are typically more dense than texts intended for informal communication. The latter contain more short utterances with a low proportion of nouns and less subordinate clauses and nominalizations.

To calculate lexical density, different methods have been applied. According to Halliday (1994), lexical density can be described as the number of content words per non-embedded clauses in a text. An alternative is the number of content words as a proportion of the total number of words in a text (Egging, 1994; Fang, 1997). Another way to calculate lexical density is to divide the total number of content words by the number of function words (Laufer & Nation, 1995; O´Loughlin, 1995; Read, 2000). Since we are dealing in this study with spoken language which is segmented into single utterances, lexical density is measured by considering the total number of content words in relation to the number of utterances (cf. Halliday, 1994; Schleppegrell, 2001; 2004). In general, lexical density says something about the way in which information is conveyed by the use of content words. In this study, lexical density is established as the total number of content words divided by the total number of utterances. This choice is made because the utterance is the central point of our analysis. Since Berber and Dutch interactions at home and school were transcribed in fragmented utterances, it is of importance to consider the number of content words in relation to single utterances and not to all words or to the whole text. In addition, the exclusion of morphemes and function words in calculating lexical density is necessary because of the differences between Berber and Dutch in terms of morphological richness and the degree of verb inflections.

Consider the following two examples, excerpt 5.1 and 5.2, taken from our data during a book reading session at measurement time 2 at home and at school respectively. Only the utterances of the mother and teacher are shown here and in both examples we consider the same number of utterances (7). With the excerpts and examples used throughout this study, a translation into English is provided for Dutch as well as Berber sentences. An additional gloss is provided for Berber sentences because the structure and typology of this

language, compared to Dutch, is different from English (a list of abbreviations used in the glosses is presented in Appendix 3).

Excerpt 5.1: Mother of Ilyas (Time 2, book reading session)

MOT: u-**si**-n-d yini i-**oeddr**-en **sira** nni **nna**-n-s min da t-gg-ed ?
Come.PERF- 3P.M directional particle those repair.IMPERF-3P.M- pavement that
tell.PERF- 3P.M what here 2S.M-do.IMPERF
(those who repair the pavement came and asked him: what are you doing here?)

MOT: y-**nna**- s- en qa **qqim**-egh x **lusa**.
3S.M-tell.PERF CL.him them T.A sit-AOR-1S.M on the paving stones
(he told them I am sitting on the paving stones)

MOT: y-**nna**-yi **baba qim** da x **lusa**.
3S.M-tell.PERF-CL-me my father sit-AOR here on paving stones
(my father told me to sit on the paving stones)

MOT: haqqa **ksi**-n-t di **kamyun**.
P.A take.PERF-3P.M-1S.M in truck
(there they took him in the truck)

MOT: haqqa **baba**-s i-**xezzar** day-es.
P.A father-his 3S.M-look.IMPERF at-him
(there his father is looking at him)

MOT: haqqa **ksi**-n-t.
P.A take.PERF-3P.M-1S.M.
(there they took him)

MOT: haqa-t.
P.A -he
(there is he)

Lexical density: 16/7 = 2.28

Excerpt 5.2: Teacher of Ilyas (Time 2, book reading session)

*TEA: ik **heb** een **boekje** voor jullie.
(I have a book for you(PL))
*TEA: ik **wacht even** tot jullie allemaal **stil** zijn.
(I wait for a while until you all are quiet)
*TEA: **anders kan** ik **natuurlijk** niet het **verhaaltje** gaan **vertellen**.
(otherwise I will not be able to tell the story)
*TEA: Ashraf, wat is dat voor een ***dier***?
(Ashraf, what sort of animal is that?)
*TEA: is dat een **moeder**?
(is that a mother?)
*TEA: een **schaap**.
(a sheep)
*TEA: en waar **zit** dat **schaap** op?
(and what is that sheep sitting on?)

Lexical density: 15/7 = 2.14

In excerpt 5.1. lexical density is 2.28: 16 content words, which are in bold, are divided by 7, the total number of utterances. In excerpt 5.2. lexical density is 2.14: 15 content words are divided by the total number of utterances (7).

Lexical diversity

The other feature examined in this chapter is lexical diversity. Mother-child interactions during daily contacts are characterised by the use of frequent vocabulary, while at school, children are expected to use low-frequency vocabulary (Schleppegrell, 2004). In the school environment children are expected to use, and are exposed to, a more sophisticated, technical and less frequent vocabulary. For school tasks, children are encouraged to use different and varied words to talk about objects and events. This helps make their vocabulary more diverse. Lexical diversity indicates the variation of the vocabulary in a written or spoken text. It has been traditionally measured by using the type-token ratio (TTR), that is the number of different words in a sample of spoken or written text divided by the total number of tokens. However, due to TTR's sensitivity to sample size, new methods have been suggested such as Guiraud's index, also known as RTTR. This index is calculated by the number of types divided by the square root of the number of

tokens (Guiraud, 1960) and is independent of text length. Recently, an even more advanced method has been introduced that takes into account the limitations of the older measures. This new algebraic method, which is used in the current study, is called the index D. It has been developed by Malvern, Richards and McKee (Malvern & Richards, 2002; McKee *et al.*, 2000) and is calculated by using the computer program *vocd*, which is part of the CLAN system (MacWhinney, 2000).

Vocd takes 100 random samples of 35 tokens and calculates the mean type-token ratio (TTR) for these samples. This calculation is repeated with 16 different samples, ranging from samples of 35 to samples of 50 tokens. Subsequently, the program plots the mean TTR scores for each sample size. The program, then, uses a curve-fitting formula to reach the best fit between the theoretical curve and the observed data of the random-sampling TTR curve (Malvern *et al.*, 2004). As a result, the D value is the parameter used for lexical diversity. According to Malvern and Richards (2002), the values of D range from D=5 for a five-year-old child with a language impairment, to D=120 for a sample of academic writing.

5.2 Lexical richness in mothers' input

Maternal lexical input is analyzed in the current chapter by considering two dimensions. The quantity (number of utterances, tokens, content words and function words) and the quality (density and diversity). Table 5.1 presents the means and standard deviations of different input variables of the mothers during the three measurement times. As shown in previous chapters of this study, the measurement time 1 was when the children were 3;2 years old, measurement time 2 when the children were 4;2 years and measurement time 3 when the children were 5;10 years old. These descriptive statistics show maternal lexical input during two tasks at home.

Table 5.1: Mothers' lexical input in two interaction tasks during three measurement times

N=12		Time 1 (Age 3;2 years)[1]		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
		M	SD	M	SD	M	SD
Utterances	Picture	88.1	25.5	107.9	33.5	75.7	22.6
	Book	105.4	39.8	79.3	19.5	62.4	23.2
Tokens	Picture	669.3	292.0	838.2	329.4	458.8	272.5
	Book	1180.0	511.9	950.9	321.9	716.0	342.0
Word types	Picture	335.6	153.4	361.8	143.3	277.7	109.4
	Book	521.2	218.0	443.2	128.6	341.2	165.9
Content words	Picture	116.9	51.2	162.7	65.7	96.5	51.3
	Book	275.4	130.0	232.1	84.6	175.8	88.3
Function words	Picture	136.4	66.7	227.0	104.5	185.7	61.2
	Book task	288.7	129.2	298.7	124.4	184.6	89.0

The mean number of utterances ranges from 62 (book task, Time 3) to 107 (picture task, Time 2); the number of tokens from 458 (picture task, Time 3) to 1,180 (book task, Time 1); the mean number of types from 277 (picture task, Time 3) to 521 (book task, Time 1); the mean number of content words from 96 (picture task, Time 3) to 275 (book task, Time 1) and the mean number of function words from 136 (picture task, Time 1) to 299 (book task, Time 2).

At the level of utterances, there is a slight variation over the two tasks. At measurement time 1, mothers produced more utterances during the book task than during the picture task. At Time 2 and 3, however, the total number of mothers' utterances during the picture task was higher than in performing the book task. While the number of utterances in the book task decreased over time, for the picture task it first increased and then decreased.

The number of word tokens decreased from Time 1 to Time 3 both for the book task and for the picture task. The total number of word tokens was higher in the book task than in the picture task during all three measurement times. It is of note that in the last measurement time (children's age is 5;10) mothers produced the lowest number of utterances and word tokens compared to Time 1 and 2. The same pattern can be seen at the level of word types. This might be

[1] Henceforth, when we report the results of children, mothers and teachers, the age of children is mentioned next to each measurement time.

due to the fact that children start talking more at this age and consequently take more time and turns than their mothers. Mothers used more word types during the book task conversations than during the picture task. Again at Time 3, and in both tasks, mothers produced less word types than in Time 1 and 2. Regarding content words, over the three measurement times, mothers used more lexical words during the book reading task than during the picture description task. In the last measurement time, mothers used the lowest number of content words compared with the Times 1 and 2 when the children were younger than 5 years old. The absolute numbers of all tokens, lexical density and lexical diversity for each individual mother, teacher and child during the book task and the picture task are presented in Appendix 4.

After considering these quantity measures, we now turn to the quality variables density and diversity. These two measures are adopted as academic language components. Table 5.2 shows the mothers' mean values and standard deviations for both density and diversity during two different tasks at three measurement times.

Table 5.2: Means and standard deviations of lexical richness measures of mothers at three measurement times

N=12		Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
		M	SD	M	SD	M	SD
Density	Book	2.54	0.48	2.97	0.99	2.74	0.74
	Picture	1.29	0.35	1.49	0.40	1.24	0.45
Diversity	Book	85.92	23.09	78.50	27.46	79.50	18.23
	Picture	59.45	23.61	61.28	30.48	57.49	29.53

The mean density of the mothers ranges from 1.24 (picture task, Time 3) to 2.97 (book task, Time 2). The latter might be flattened because one of the mothers was an extreme outlier with a density code of 5.71 (meaning on average nearly six content words per utterance). In all measurement times, mothers' lexical density during the book reading was, on average, twice as high as during the picture task.

Longitudinally, in both tasks, mothers do not show much variation across the three measurement times. Note that in both tasks, mothers' lexical density increased from Time 1 to Time 2 but decreased from Time 2 to Time 3. Both

home tasks in measurement time 2 yielded the highest values of lexical density. ANOVAs with repeated measures with time and task as within-subject factors and density as dependent variable showed no significant effect of time ($F(2,22)=1.48$, $p=.25$, $\eta_p^2=.12$), but did show a significant main effect of task ($F(1,11)=101.58$, $p=.000$, $\eta_p^2=.90$). The interaction between time and task was not significant ($F(2,22)=.39$, $p=.69$, $\eta_p^2=.03$). This means that the density of the mothers' utterances was more or less stable over time, but that the language used during the book task was on average much more dense than the language used during the picture task.

As for lexical diversity, it is not divergent from the general pattern so far. The mothers' language was more diverse when involved in book reading than in picture description and there is no clear increase over time. ANOVAs with repeated measures with time and task as within-subject factors and diversity as dependent variable showed no significant effect of time ($F(2,22)=.23$, $p=.79$, $\eta_p^2=.02$), but did show a significant main effect of task ($F(1,11)=13.62$, $p=.004$, $\eta_p^2=.55$). The interaction between time and task was not significant ($F(2,22)=.35$, $p=.71$, $\eta_p^2=.03$). This means that the diversity of the mothers was more or less stable over time, but that the language used in talking about a book was on average much more diverse than the language used during the picture task.

The lexical differences between the book task and the picture task are illustrated below by using excerpts taken from two different transcripts of the same mother in two different tasks. Excerpt 5.3 is taken from a transcript of Selma and her mother during the book task and excerpt 5.4 is part of a transcript of a conversation between the same mother and her daughter during the picture task.

Excerpt 5.3: (Mother of Selma: book task; Time 2)

*MOT: war ghar-i bu rweqt y-nna baba-s.
NEG to have-1.m/f NEG time 3S.M-tell.PERF father-his
(his father told him he has no time)

[…]

*MOT: nni-gh- ac qim x zzellij!
tell.PERF-1S to you sit-AOR on paving stones
(I told you to stay on the paving stones)

*MOT: xatar [x2] ad t-rah2-ed x zzellij y-nna baba-s.
dangerous to 2S-walk.IMPERF-2M on paving stones 3S.M-tell.PERF father-his
('it is dangerous to walk on the paving stones', told his father)
*MOT: waxxa amenni i-ruh2 Tim ghar barra.
although that 3S.M-go.PERF Tim to outside
(nevertheless, Tim went outside)
*MOT: ta xzar ghar barra dini yinni i-oeddr-en zzellij.
Just look-AOR directional particle outside there those 3P.M-make-IMPERF pavement
(look outside, there you see those who make the pavement)
*MOT: y-nna-s i Tim qim danita.
3S.M-tell.PERF-3P to Tim sit here
(he told Tim to sit here)
*MOT: haqa oawed u-si-n-d ghar-s nna-n-as min da t-gg-ed?
Here again come.PERF-3PL directional particle-to him tell.PERF-3P.M what here 2S-do.IMPERF-2S
(look, they came again and told him what are you doing here?)
*MOT: u-si-n-d ghar-s yini i-xedm-en
come.PERF-3PL directional particle-to him those 3P-work-IMPERF-3PL
di ccario nna-n-as min da tt-gg-ed cek?
in street tell.PERF-3PL him what here 2S-do.IMPERF-2S you
(street workers came and told him: what are you doing here?)

Excerpt 5.4: (Mother of Selma: picture task; Time 2)

*MOT: min tt-wari-d dini?
what 2S-see.IMPERF-2S there
(what do you see there?)
*MOT: min tt-wari-d naghni?
What 2S-see.IMPERF-3S else
(what else you see?)
*MOT: nec tt-wari-gh ijen xxx.
I 1S-see.IMPERF-1S one (xxx=unintelligible)
(I see one […])
*MOT: mani i-lla &uh mani i-lla giraffe@nl?
where 3S.M-to be Where 3S.M-to be giraffe.
(where is the/a giraffe?)

*MOT: mani i-lla?
Where 3S.M-to be
(where is it?)
*MOT: wah.
(yes)
*MOT: tt-wari-gh grote@nl man@nl heel@nl.
1S-see.IMPERF-1S big man very
(I see a big man, very)

These two transcripts show the lexical differences in terms of the number of content words, word types, density and diversity. As can be seen in Table 5.3, mother's input during the book task contained more content words and types and was lexically more dense and diverse than in the picture task.

Table 5.3: Lexical input of the mother of Selma during two different tasks in Time 2

	Content words	Word types	Lexical density	Lexical diversity
Book task	175	369	2.22	66.60
Picture task	148	246	1.85	54.04

5.2.1 Relationship of mothers' input with SES and literacy

To address the possible relationship between the mothers' use of academic language features and their family background, correlations were calculated between the input measures and SES and literacy practices at home respectively. The results are presented in Table 5.4.

Table 5.4: Kendall Tau correlations of SES and literacy with tokens, density and diversity of the mothers in three measurement times

N=12		Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
		SES	Literacy	SES	Literacy	SES	Literacy
Tokens	Book task	.13	.14	.35	.37	.51*	.46*
	Picture task	.00	.02	.38	.30	-.14	-.10
Density	Book task	.22	.40τ	.31	.27	.31	.40τ
	Picture task	.00	.08	.22	.18	-.06	-.02
Diversity	Book task	.51*	.40τ	.47*	.53*	.32	.33
	Picture task	.38	.59*	.38	.30	.38	.50*

* p<.05; τ tendency

At the quantity level, in Time 1 the number of tokens showed a low correlation with the SES and literacy activities during both interaction tasks (book reading and picture task). In Time 2, the number of tokens positively correlated with the SES and literacy of the family during both tasks. In Time 3, the number of tokens significantly correlated with the SES and literacy during the book task. However, the tokens during the picture task did not correlate with the two background variables.

Lexical density, showed overall positive correlations with both SES and literacy for the book reading task (although not significant) but hardly any correlation for the picture task. In Time 1, a positive relationship was established with SES and literacy during the book task. During the picture task no correlation was found. In Time 2 and in both interaction tasks, density positively but not significantly correlated with SES and literacy variables. In Time 3, a positive correlation was found with density in the book task and no correlation with the picture task. For diversity the overall relationship with SES and literacy is clearer: all correlations are high (range .30 - .59) and some are significant for the book task and the picture task.

The influence of SES and literacy activities is illustrated by discussing the following utterances taken from two different transcripts of two mothers with different SES and literacy backgrounds. The transcripts originate from the book reading task in measurement Time 3. Excerpt 5.5 is of the mother of Ibtisam, who is of relatively high SES and literacy (both score 4). Excerpt 5.6 is of the

mother of Ilyas, who is of low SES and literacy (both score 1). From each dyad's conversation, only 9 utterances of the mothers are shown here.

Excerpt 5.5: (Mother of Ibtisam: book task; Time 3)

*MOT: emma.
(Emma)
*MOT: t-xezzar.
3S.F-look.IMPERF
(she is looking)
*MOT: di ijjen aqemum heel@nl groot@nl.
in one mouth very big
(in a very big mouth)
*MOT: nee@nl cem ghir sh2es waha waxxa!
No 2S.F-you just 3S.F-listen.AOR only ok
(no, you just listen ok!)
*MOT: ghar-es tighmas t-imeqran-in qedoa-n-t.
have.IMPER-3S teeth F-big-PL sharp-PL-F
(it has big sharp teeth)
*MOT: safi t-nna-s snaoat t-aneggarut t- as-ed.
ok 3S.F-say.PERF-3S 2S.M-show.AOR F-last-F then come.AOR-directional
(she said ok, just show the last one and come)
*MOT: safi ucha t-faqq-ed zig ides.
ok then 3S.F-wake.AOR from sleep
(and then she woke up)
*MOT: kurchi t-uf s tidi xxx.
all 3S.F-wet with sweat
(she was completely wet from swea.)
*MOT: safi war zemmar-gh ucha t-jjiwen.
enough NEG can.AOR-1S then 3S.F-satiate.PERF
(enough, I can't and her hunger was satiated)

Excerpt 5.6: (Mother of Ilyas: book task; Time 3)

*MOT: taddart n amz2iwen.
house of ogres
(the house of ogres)

*MOT: uca.
then
(then)
*MOT: di taddat n amz2iwen.
in house of ogres
(in the house of ogres)
*MOT: ha- yqa- t.
P.A. 3F.S
(there is she)
*MOT: wah mara amziwen eni.
yes if ogres those
(yes, if those monsters)
*MOT: raja.
wait.AOR
(wait)
*MOT: ha- qa- t.
P.A. 3F.S
(there is she)
*MOT: t-wari-d t?
2S-see.IMPERF-2S 3S.F-her
(do you see her?)
*MOT: ha- qa- t rexxu t-qqim ghar rbuki nnes a-s t-kemmer.
there to be 3S.F-her now 3S.F-sit.IMPERF to book her Future particle it to 3S.F-finish.AOR
(now she's sitting down to finish her book)

In the whole conversation from which excerpt 5.5 is taken, the mother of Ibtisam produced 58 utterances and used a total number of 361 word types. Lexical density of the mother is 3.81 and lexical diversity is 91.10. In the complete transcript from which excerpt 5.6 is taken, however, the mother of Ilyas produced 32 utterances, and used a total number of 129 word types. The lexical density of the mother is 1.94 and lexical diversity is 47.55. It can thus be concluded that the input of this less educated mother is lower in quantity, lexically less dense and less diverse than the input of the more educated mother.

5.3 Lexical richness in teachers' input

Lexical richness in the school setting was studied by using two activities: a more structured activity, book reading, and a less structured one, circle time. The book task is used in both the home and school setting, which makes comparison possible. In this section, the lexical input of teachers is addressed. Data analyses concern both lexical quantity and lexical quality. In Table 5.5 means and standard deviations of the different measures are presented. As explained above, children were observed twice at school: at the age of 4;2 and 5;10.

Table 5.5: Teachers' lexical input in two interactions tasks during two measurement times

N=12		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
		M	SD	M	SD
Utterances	Circle	98.9	16.0	113.3	17.8
	Book	106.8	32.2	128.3	27.2
Tokens	Circle	1069.9	243.4	1183.5	229.2
	Book	1375.7	321.8	1416.8	238.0
Word types	Circle	548.3	103.6	632.5	101.7
	Book	656.3	161.6	397.2	232.4
Content words	Circle	184.8	46.6	216.3	50.2
	Book	277.0	85.9	327.1	71.9
Function words	Circle	295.1	115.2	331.1	63.0
	Book	334.6	104.5	393.3	110.5

The average number of utterances of the teachers ranged from 99 (circle time, Time 1) to 128 (book task, Time 3), the number of tokens from 1,070 to 1,416, the number of word types from 397 (book task, Time 2) to 656 (book task, Time 1), and the number of content words from 185 (circle, Time 1) to 327 (book task, Time 2). In both measurement times teachers produced more utterances, tokens, word types and content words during the book reading task than during the circle time.

From a longitudinal perspective, the teachers' input at the level of utterances, tokens and content words increased from Time 2, when the children had just entered primary school, to Time 3. During the book reading task, teachers produced considerably more talk in Time 3 than in Time 2.

Turning from quantitative data to qualitative, an overview of the quality of the teachers' lexicon in terms of lexical density and diversity is provided in Table 5.6.

Table 5.6: Lexical richness of the teachers in two measurement times

N=12		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
		M	SD	M	SD
Density	Circle	1.86	0.34	1.95	0.46
	Book	2.72	0.91	2.58	0.40
Diversity	Circle	76.16	12.97	88.21	9.58
	Book	71.55	11.76	85.34	7.22

Concerning lexical density, differences have been found between the book task and the circle time. In both measurement times, teachers' mean values for lexical density in the book reading task was higher than in the circle time. In other words, teachers use more dense vocabulary when talking about a book than during circle time. Longitudinally, at the level of book reading, the teachers' lexical density in measurement time 2 was higher than in measurement time 3. ANOVAs with repeated measures with time and task as within-subject factors and density as dependent variable showed no significant effect of time ($F(1,11)=.03$, $p=.87$, $\eta_p^2=.003$), but did show a significant main effect of task ($F(1,11)=16.96$, $p=.002$, $\eta_p^2=.60$). The interaction between time and task was not significant ($F(1,11)=.59$, $p=.46$, $\eta_p^2=.05$). On average, the teachers' density was more or less stable over time, but the language used in talking about the book was on average much more dense than the language used during the circle time.

A somewhat different pattern emerged with regard to lexical diversity. The teachers mean values for lexical diversity during book reading was somewhat lower than during the circle time in both measurement times, which means that the circle time elicited a more diverse vocabulary than book reading. From a longitudinal perspective, during the book reading task, the lexical diversity of the teachers increased from Time 2 to Time 3. ANOVAs with repeated measures, with time and task as within-subject factors and diversity as a dependent variable did show a significant main effect of time ($F(1,11)=23.47$, $p=.001$, $\eta_p^2=.68$), but did not show a significant main effect of task

($F(1,11)=1.28$, $p=.28$, $\eta_p^2=.10$.) The interaction between time and task was not significant ($F(1,11)=.64$, $p=.44$, $\eta_p^2=.06$). On average, the teachers used a more diverse vocabulary when the children were 5 years old than when they were 4, and this pattern was visible in both the book task and in circle time.

The differences at the level of teachers' vocabulary according to the age of the children are illustrated by excerpts 5.7 and 5.8. These fragments are from transcripts of the same teacher and the same child, Abdelilah, in two different book reading sessions. Excerpt 5.7 is taken from a long conversation between the teacher and the pupils around the book '*De verrassing*' (The surprise). The age of Abdelilah at this measurement time is 4;2. Excerpt 5.8 is taken from another conversation between the teacher and the pupils about the book '*Wie is er hier het dapperst*' (Who is the bravest here?). The age of Abdelilah at this measurement time is 5;10.

Excerpt 5.7: (Teacher of Abdelilah: book task; Time 2)

*TEA: jij denkt over een schaap.
(you think it's about a sheep)

[…]

*TEA: wat denk jij Kawtar?
(what do you think Kawtar?)
*TEA: een verrassing, jij denkt dat het over een verrassing gaat.
(a surprise, you think it's about a surprise)
*TEA: hoe weet je dat dan?
(how do you know that, then?)
*TEA: ken jij het boek?
(do you know the book?)
*TEA: jij denkt dat er een verrassing in zit.
(you think there is a surprise in it)
*TEA: en jij Ikrame?
(and you Ikrame?)
*TEA: truien maken?
(to make sweaters?)
*TEA: kleren maken.
(to make clothes)

Excerpt 5.8: (Teacher of Abdelilah: book task; Time 3)

*TEA: mus is hier het dapperst.
(the sparrow is the bravest here)
[…]
*TEA: &oh waarom is ie nou dapper?
(oh, why is he brave?)
*TEA: dat hij, dat hij een beetje durft zielig te zijn.
(that he, that he dares to be pathetic)
*TEA: &nou dat vind ik echt knap.
(well, I think that's really smart)
*TEA: <hij durft te zeggen>//<hij durft zielig> te zijn.
(he dares to say, he dares to be pathetic)
*TEA: hij durft als enige te zeggen ik doe niet mee.
(he is the only one who dares to say I am not participating)
*TEA: als je niet mee wilt doen dan doe je niet mee.
(if you don't want to participate then you don't participate)
*TEA: en dan ben je eigenlijk ook wel een beetje dapper &he.
(and then you actually are a little bit brave, uh)
*TEA: kijk en ze zingen en ze springen en ze vinden de vogel het dapperst.
(watch, and they sing and they jump and they think that the bird is the bravest)

It is clear that during measurement time 2, when the child's age is 4;2, the teacher used shorter utterances and less diverse vocabulary than in measurement time 3. In the complete conversation of Time 2, from which excerpt 5.7 was taken, the teacher produced 86 utterances, 1,674 tokens, 814 word types and 371 content words. The lexical diversity was 67.75. During the conversation of measurement time 3, to which excerpt 5.8 belongs, the teacher produced 152 utterances, 1,889 tokens, 907 word types and 439 content words. The lexical diversity was 91.91. So, this teacher used more diverse language when the target child and the children in the group were older.

5.4 Input in two settings: mothers and teachers compared

To gain insight into the input that children receive in the home and school environment, a Mann-Whitney U-test was carried out to compare the mothers' and teachers' input variables at two measurement times. This was done by using

one sort of interaction task, book reading, which had been used in both settings. The descriptive statistics and significance levels are displayed in Table 5.7.

Table 5.7: Overview of the Mann-Whitney U-test results on different variables of the mothers' and teachers' lexical richness in two measurement times

N=24 (mothers = 12 teachers = 12)		Time 2 (Age 4;2 years)		Significance level	Time 3 (Age 5;10 years)		Significance level
		M	SD		M	SD	
Utterances	Mothers	79.3	19.5	U=27.0* (p=.008)	62.42	23.16	U=3.50** (p=.000)
	Teachers	106.8	32.2		128.33	27.19	
Tokens	Mothers	945.3	321.9	U=23.0* (p=.004)	716	341.97	U=6.0** (p=.000)
	Teachers	1375.7	321.8		1497.83	225.47	
Density	Mothers	3.0	1.0	U=64.0 (p=.66)	2.74	.74	U=64.0 (p=.67)
	Teachers	2.7	.9		2.58	.40	
Diversity	Mothers	78.5	27.5	U=65.0 (p=.71)	79.50	18.22	U=50.0 (p=.22)
	Teachers	71.6	11.8		85.34	7.22	

* p<.05; ** p<.01

Mann-Whitney U results show that when the children were four years of age, the teachers produced significantly more utterances (p=.008) and word tokens (p=.004) than did the mothers. Lexical density and lexical diversity at that time were higher for the mothers, but not significantly (p=.66 and p=.71 respectively). At Time 3, concerning the total number of utterances, a significant difference between mothers and teachers was found. It can be seen that teachers generated twice as many utterances as the mothers (p=.000). A somewhat similar pattern emerged with regard to the total number of word tokens (p=.000). Concerning lexical density, mothers used lexically a somewhat more dense language than the teachers, though the difference is not significant (p=.67). Finally, a slight but not significant difference can be seen at the level of lexical diversity. Teachers used a little more diverse speech than mothers in measurement time 3, but not significantly (p=.22).

5.5 Lexical richness in children's output

5.5.1 Home data

Having dealt with the mothers' and teachers' input, we now turn to deal with the children's language use and in particular their lexical output. It should be noted that the language used by children during interactions with their mothers was a mixture of Berber and Dutch. As this is their means of communication in their home environment, we decided to consider the whole transcripts instead of selecting only Berber words (see Chapter 4). The means and standard deviations are set out in Table 5.8.

Table 5.8: Descriptive statistics for children's lexical measures in two interaction tasks during three measurement times at home

N=12		Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
		M	SD	M	SD	M	SD
Utterances	Book	39.1	20.4	32.6	14.2	25.1	16.7
	Picture	52	22.8	57.4	21.0	58.5	14.7
Tokens	Book	171.8	103.7	224.3	122.7	189.5	146.7
	Picture	228.9	130.2	317.8	174.7	331.0	216.0
Word types	Book	90.3	51.7	110.4	59.2	86.4	65.9
	Picture	102.8	59.2	137.7	75.8	214.2	100.1
Content words	Book	46.8	25.7	56.3	28.1	42.3	38.2
	Picture	47.4	26.4	66.3	36.6	86.8	40.7
Function words	Book	33.1	27.9	68.2	43.2	38.5	33.3
	Picture	37.4	26.8	67.8	40.5	116.8	61.3

On the level of utterance, an important finding was that children show a decrease across the three measurement times during the book reading task. On average, children produced the highest number of utterances in Time 1, followed by Time 2 and Time 3. At the level of word tokens, however, children used most tokens in Time 2, followed by Time 3 and Time 1. Concerning word types and content words, measurement time 2 again witnessed the highest scores, followed by Time 1 and Time 3. Regarding the picture description task, all variables showed an increase corresponding with the increase of age. The

total number of utterances increased slightly over the three measurement times. However, the scores of the other variables (word tokens, word types, content words and function words) became considerably higher as the children got older. Finally, it should be noted that, in contrast to the mothers, children's scores on different variables were higher during the picture task than during the book task. This can be attributed to the nature of both tasks. In the book task, it was usually the mothers who took the initiative in discussing the story of the book, while the picture task invited the children to produce more talk.

In the following examples (Excerpt 5.9 and Excerpt 5.10) taken from two conversations with the same child, Yasmine, some differences are illustrated concerning the lexical features in two different home tasks. The age of Yasmine in this measurement time was 5;9.

Excerpt 5.9: (Yasmine: book task; Time 3)

*CHI: ijen taqenayt.
one rabbit
(a rabbit)
[…]
*CHI: yih.
(yes)
*CHI: d-t-ameqran-t.
COP-F-big-F
(it is big)
*CHI: yih.
(yes)
*CHI: ghar lkutubat.
(to the books)
*CHI: ghar kassa@nl!
(to the cash desk!)
*CHI: circus@nl!
(circus!)

Excerpt 5.10: (Yasmine: picture task; Time 3)

*CHI: tin t-xezzar di lkitab.
that one-S.F 3S.F-IMPERF-look in the book
(that one is looking in the book)

*CHI: tin t-oyar.
that one-S.F 3S.F-IMPERF-play
(she, there, is playing)
*CHI: tin t-oyar ig tumubinat.
that one-S.F 3S.F-IMPERF-play with cars
(she is playing with cars)
*CHI: i yina min t-egg-en?
and these what do-IMPERF-3P.M
(and these, what are they doing?)
*CHI: da ijen baby@nl.
here one baby
(here is a baby)
*CHI: hun@nl spelen@nl piraatje@nl.
(they play little pirate)
*CHI: lijkt@nl het@nl meer@nl op@nl een@nl dagverblijf@nl.
(it looks more like a day care)

An overall factor when considering these two fragments is that the child uses both Berber and Dutch. As can be seen in excerpt 5.9, taken from a book task, the child uses short utterances and less content words than in excerpt 5.10, which is taken from a picture task. In the complete conversation of the book task, Yasmine produced 13 utterances. Concerning types and tokens, Yasmine used 22 different word types and 56 word tokens. The total number of content words was 10. During the picture task, Yasmine produced 46 utterances. As to the number of types and tokens, Yasmine used 126 different word types and 268 word tokens. The total number of content words in this conversation was 87.

In order to assess the children's lexical richness in the home setting, quality variables will be considered. In Table 5.9, the mean scores and standard deviations for lexical density and lexical diversity are presented.

Table 5.9: Lexical richness of children at home during three measurement times

N=12		Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
		M	SD	M	SD	M	SD
Density	Book	1.19	.40	1.70	.39	1.52	.56
	Picture	.89	.23	1.09	.35	1.51	.63
Diversity	Book	51.39	21.55	60.28	26.15	77.27	35.46
	Picture	59.22	18.06	48.02	16.54	56.19	38.95

With regard to lexical density, the children's language in Time 1 and Time 2 during the book reading task was more dense than in the picture task. In Time 3, the lexical density values during both tasks were similar. In general, the children's talk during the book task had a more dense vocabulary than during the picture tasks. With the picture task, lexical density increased over time, but with the book task it increased only from Time 1 to Time 2. ANOVAs with repeated measures with time and task as within-subject factors and density as dependent variable did show a significant main effect of time ($F(2,22)=7.08$, $p=.004$, $\eta_p^2=.39$), and also a significant main effect of task $F(1,11)=14.01$, $p=.004$, $\eta_p^2=.56$). The interaction between time and task was not significant ($F(2,22)=3.15$, $p=.06$, $\eta_p^2=.06$), but showed a trend. This means that, on average, the children used more dense language as they grew older (especially in the picture task).

Furthermore, children used more dense language during the book task compared to the picture task especially, when they reached 3 and 4 years of age. Graphic presentations of the results in two home tasks are shown in Figure 5.1.

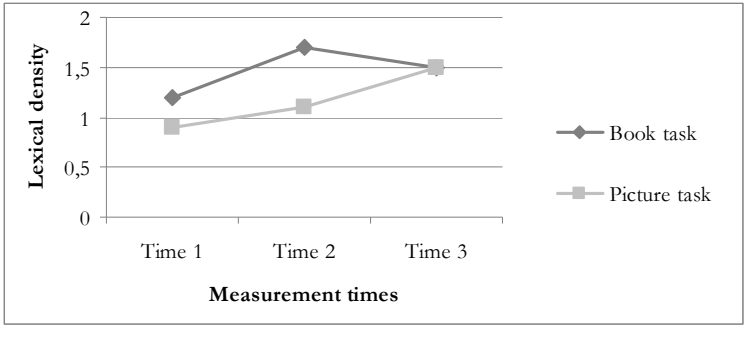

Figure 5.1: Children's lexical density in three measurement times in the home setting

As far as lexical diversity is concerned, the average lexical diversity scores increased over time, but for the picture task, diversity was higher at Time 1 and Time 3 compared to Time 2. ANOVAs with repeated measures with time and task as within-subject factors and diversity as dependent variable did not show any main effect of time ($F(2,22)<1$) or task ($F(2,22)<1$). It must be noted that the lexical diversity of many of the children in different measurement times could not be calculated because of the low number of utterances they produced. The children's average scores at the level of lexical diversity over three measurement times are plotted in Figure 5.2.

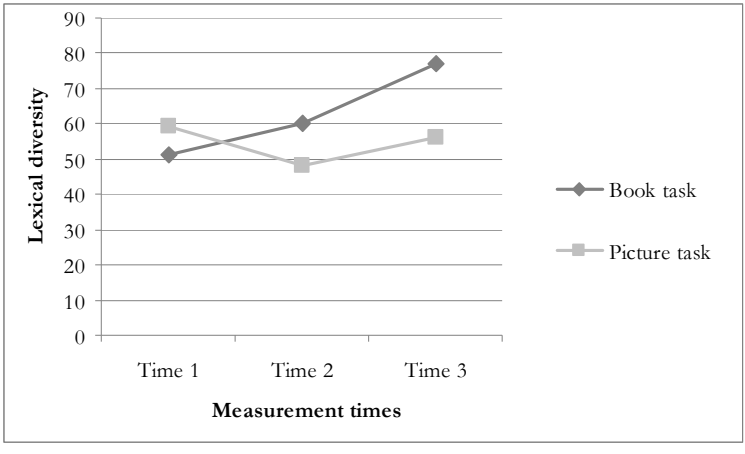

Figure 5.2: Children's lexical diversity in three measurement times in the home setting

When we consider the scores of children's lexical diversity during the book task, we can see a considerable progress across the three measurement times. The score values of lexical diversity were 51.4, 60.3 and 77.3 respectively in Time 1, Time 2 and Time 3. Lexical diversity during the picture task shows a different trend. In Time 1, children used the most diverse language (mean=59.2), followed by a decrease in Time 2 (mean=48.0) and an increase in Time 3.

5.5.2 Relationship of children's output with family SES and literacy

This section discusses whether lexical features of academic language in children's language can be associated with family SES and home literacy

activities. Table 5.10 gives an overview of the Kendall Tau correlations between lexical variables and both SES and home literacy background variables.

Table 5.10: Kendall tau correlations of SES and literacy with academic language variables in the children's' vocabulary at home

N=12		Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
		SES	literacy	SES	literacy	SES	literacy
Tokens	Book	.35τ	.11	.22	.14	-.19	-.14
	Picture	-.25	-.49*	.06	.17	.11	.16
Density	Book	.27	.22	-.09	-.17	-.33	-.29
	Picture	.16	.24	-.25	-.11	.09	-.01
Diversity	Book	.36τ	.17	.08	.14	.54*	.25
	Picture	-.18	.06	.13	.07	.38τ	.33

* p<.05; τ p<.10

In Time 1, during the book task, all the variables correlated positively with SES and literacy background of the children. During the picture task, however, somewhat different outcomes were found. The number of tokens correlated negatively with SES as well as with home literacy activities. Diversity also correlated negatively with SES. In Time 2, in both interaction tasks, low correlations were established between different academic language variables and SES and literacy. In Time 3, the academic language variables show different correlations with SES and literacy. The number of tokens during both tasks shows a negligible relation with SES and home literacy activities. Regarding density, a negative correlation was established with SES and literacy during the book task but almost no correlation was found during the picture task. Diversity, on the other hand, shows a strong and positive correlation with the two family background variables.

The data of two of the children illustrates the fact that there was no straightforward relationship between family SES and literacy and academic language features in the children's output. Consider two children with different socio-economic backgrounds and literacy practices at home. Selma (SES = 1; literacy = 1) and Ibtisam (SES = 4; literacy = 4) were involved with their mothers, in dyadic conversations during the picture task. The mean age of the children at this point of the data collection was 3;2. Confirming the results

shown in Table 5.10 and despite the lower SES level and literacy of Selma's family, she scored on several lexical components higher than did Ibtisam, who lived in a family with a higher SES level and literacy. In the conversation with her mother, Selma produced 75 utterances, 182 word types and 411 word tokens. The total number of content words she used was 99. Her lexical density and diversity were respectively 1.32 and 85.48. Ibtisam, on the other hand, produced 44 utterances, 72 word types and 182 word tokens. The total number of content words she used was 44. Finally, Ibtisam's lexical density and diversity were 1.00 and 57.48, respectively.

5.5.3 School data

In this section the children's linguistic performances in school interactions will be presented and discussed. Table 5.11 gives an overview of the means and standard deviations of the lexical variables in the two school measurement times.

Table 5.11: Descriptive statistics of the children's lexical measures in two interaction tasks during two measurement times at school

N=12		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
		M	SD	M	SD
Utterances	Book	14.17	11.70	16.75	12.07
	Circle	12.00	6.79	18.17	9.87
Tokens	Book	63.92	56.32	112.17	104.09
	Circle	84.33	54.17	118.33	92.30
Types	Book	34.00	31.55	56.50	53.69
	Circle	37.75	26.29	73.08	50.15
Content words	Book	17.17	15.16	27.33	24.43
	Circle	16.08	9.64	27.83	17.63
Function words	Book	15.42	17.07	35.33	28.76
	Circle	21.42	16.48	32.25	29.51

As can be seen from Table 5.11, all quantity measures (utterances, tokens, types and content words) increased from Time 2 to Time 3. With regard to task differences, on average the children produced more utterances during the book

task than during the circle time in Time 2. In contrast, in Time 3, more utterances were produced in the circle time than in the book task. At the level of tokens and word types, and in both measurement times, children produced more during the circle time than during the book task. As to content words, an increase can be seen between Time 2 and 3, but almost no differences are found between the two tasks. A general but expected trend in children's language use at school compared to the home setting is of interest. As can be seen in Table 5.7 and 5.11, children generally produced more talk (utterances, words etc.) during home tasks than during school task. In home interactions, children are involved alone in dyadic conversations with their mothers while at school they are, with other children, part of a group. The results regarding the lexical richness in the school context are presented in Table 5.12.

Table 5.12: Lexical richness of children at school during two measurement times

N=12		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
		M	SD	M	SD
Density	Book	1.20	.34	1.52	.39
	Circle	1.34	.53	1.63	.86
Diversity	Book	43.48	14.94	44.98	18.79
	Circle	50.84	9.15	53.12	23.08

To begin with, in both school measurement times, children used on average more dense language in the circle time than in the book task. And for both tasks, density scores increased from Time 2 to Time 3. ANOVAs with repeated measures with time and task as within-subject factors and density as dependent variable did show a significant main effect of time ($F(1,11)=5.93$, $p=.03$, $\eta_p^2=.35$), and no main effect of task $F(1,11)=.35$, $p=.56$, $\eta_p^2=.03$). There was no interaction between time and task ($F(1,11)=.02$, $p=.90$, $\eta_p^2=.00$). This means that, on average, the children used more dense language in both tasks when they grew older. The results in two school tasks are shown in Figure 5.3.

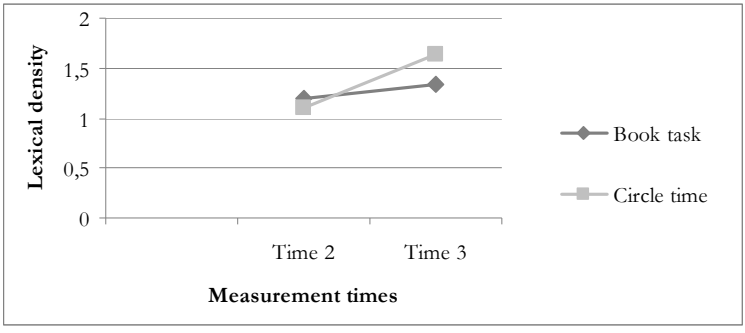

Figure 5.3: Children's lexical density in the school setting

With regard to lexical diversity, children used lexically more diverse language during the circle time than during the book task. Lexical diversity only increased to a small amount over time. ANOVAs with repeated measures with time and task as within-subject factors and diversity as dependent variable did not show any main effects for time ($F1,11<1$) and task ($F1,11<1$). This is mainly because the lexical diversity of many of the children could not be calculated as they produced too small number of utterances. The children's scores for lexical diversity during the two school tasks are graphically illustrated in Figure 5.4.

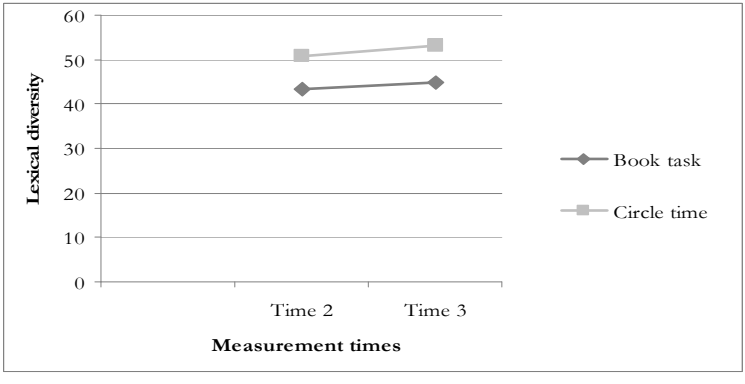

Figure 5.4: Children's lexical diversity in the school setting

Children show roughly the same pattern of lexical diversity as with lexical density. In both tasks, we can see a slight increase from Time 2 to Time 3.

Excerpts 5.11 and 5.12 illustrate the lexical differences in two measurement times at school. In these two examples, 10 utterances of the same child (Anissa)

are taken from two complete transcripts during the circle time in Time 2 and Time 3.

Excerpt 5.11: (Anissa: Time 2, circle time)

*CHI: allemaal *zelfde bij*.
(it is all the same bee)
*CHI: <ik *deed*> [/] ik *deed* het voor haar zo.
(I did that for her in this way)
*CHI: <ik *deed*> [/] ik *deed* zo *deed* ik ook.
(I did that in this way too)
*CHI: is die voor jou?
(is that for you)
*CHI: ik *wil*.
(I would like to)
*CHI: mijn *kaartje*.
(my picture)
*CHI: *bij*.
(bee)
*CHI: &xxx *lieveheersbeestjes*.
(ladybirds)
*CHI: en *Asan*?
(and Asan?)
*CHI: een van *beestjes*.
(one of beasts)

Excerpt 5.12: (Anissa: Time 3, circle time)

*CHI: *zo lijkt* ie *boos*.
(he seems angry in this way)
*CHI: ik *weet* het.
(I know)
*CHI: maar deze is *heel moeilijk*.
(but this is very difficult)
*CHI: *afwassen*.
(to wash up)
*CHI: *spullen* van de *tafel* naar de *keuken brengen*.
(to take stuff from the table to the kitchen)
*CHI: de *spullen* van als ik *ga eten* de *spullen* van de *tafel* naar de *keuken brengen*.
(taking the stuff from the table to the kitchen if I eat)

*CHI: en de *flesje* van mijn *baby* naar de *keuken brengen*.
(and take my baby's bottle to the kitchen)
*CHI: en die *glazen* die *kunnen gewoon* in die *wasmachine*.
(and those glasses can simply be put in the washer)
*CHI: *soms mag* ik ook *meehelpen* met de *deuren*.
(sometimes I can also help with the doors)
*CHI: met de *deuren zo wassen* met de *doek*.
(to wash the doors with the cloth in this way)

Excerpt 5.11 was produced by the child at age 4;2 and excerpt 5.12 at age 5;10. In both transcripts the content words are italicized for emphasis. In the transcript from which excerpt 5.11 was taken, the child used in total 15 content words in 11 utterances. Thus, lexical density is 15/11=1.36. Concerning the transcript from which excerpt 5.12 was taken, a total number of 38 content words were produced in 21 utterances. The lexical density is 38/21=1.80.

5.5.4 Children's home and school data compared

To gain more insight into the language behaviour of the children in the two different contexts, namely home and school, Wilcoxon Signed-rank tests were conducted to compare lexical density and diversity, which typify the academic register during a variety of tasks, both at home and at school. The most important task for consideration is the book task that was done in both settings since more academic language use is expected during this semi-structured task. The less structured tasks are the picture task at home and the circle time at school. Although not similar, they both were intended to elicit (im)personal narratives of children engaged in spontaneous speech. The results of the book tasks in both settings are presented in Table 5.13.

Table 5.13: Overview of Wilcoxon Signed-rank test results of children's lexical density and diversity during the home and the school book tasks

N=12		Time 2 (Age 4;2 years)		Significance level	Time 3 (Age 5;10 years)		Significance level
		M	SD		M	SD	
Density	Home	1.69	.39	Z=-2.59* (p=.01)	1.52	.56	Z=-.08 (p=.94)
	School	1.20	.34		1.52	.39	
Diversity	Home	60.28	26.15	Z=-.73 (p=.47)	77.27	35.46	Z=-1.60 (p=.11)
	School	43.48	14.94		44.98	18.79	

* p<.01

Note that the mean score can differ (especially for diversity) from those presented in the tables before, due to a loss of a number of children in the calculation. In measurement time 2 (age 4;2) children used a significantly more dense lexicon at home than at school (p=.01). In measurement time 3, however, lexical density values were identical at home and at school (1.52; p=.94). As for lexical diversity, in both measurement times, the children's means at home were higher than those at school, though statistically not significant. Next to the quantitative data shown in Table 5.13, we provide illustrative examples of utterances of the child Anissa in the home and school settings. Both transcripts resulted from book tasks at the age of 4;2.

Excerpt 5.13: (Anissa: home task; book reading, Time 2)

*CHI: *mama* maymmi?
Mom why
(mom, why?)
*CHI: *mama mag*@nl ik@nl xxx am *nni*-gh qao?
Mom may I you tell.PERF-1S.M/F all
(mom, may I xxx I all told you?)
*CHI: *ga*@nl je@nl mee@nl naar@nl *buiten*@nl.
(are you joining me outside?)
*CHI: nee@nl ik@nl heb@nl geen@nl *tijd*@nl.
(no, I have no time)
*CHI: *tim* y-*isi*-d *step*@nl i-*ruh2* zzay-s.
tim 3S.M-take.PERF-directioanl particle scooter 3S.M-go.PERF with-CL.it
(Tim took a scooter and went with it)

*CHI: y-*nna* s *papa*-s xxx.
3S.M-tell.PERF to him dad-CL.his
(his dad told him xxx)
*CHI: y-*nna* s *baba*-s *qim* din.
3S.M-tell.PERF to him father-CL-his 2S.M/F-sit-AOR there
(his dad told him you should sit there)
*CHI: *qim* din.
2S.M/F-sit.AOR there
(sit there)
*CHI: y-*nna* s *baba*-s.
3S.M-tell.PERF to him father-CL-his
(his father told him)
*CHI: y-*ksi* *step*@nl y-*azzer* ghar.
3S.M-take.PERF scooter 3S.M-run.AOR to
(he took the scooter and ran to)

Excerpt 5.14: (Anissa: school task; book reading, Time 2)

*CHI: *konijn*.
(rabbit)
*CHI: *schaap*.
(sheep)
*CHI: met de *motor*.
(with a motor)
*CHI: een *cadeautje*.
(a gift)
*CHI: daar bij *achter*.
(there behind)
*CHI: *hier*.
(here)
*CHI: &xxx *rijden*.
(&xxx ride)
*CHI: s*chaap gaat* op de *wegen*.
(the sheep goes on the scales)
*CHI: hoe *dik* is hij?
(how fat is it?)
*CHI: &oh maar *optellen*.
(&oh! but count)

Excerpts 5.13 and 5.14 contain 10 utterances taken from dyadic conversations of Anissa with her mother and her teacher, respectively. In both transcripts, content words are put in italics. In excerpt 5.13, the child uses 22 content words during the home book task, while in excerpt 5.14 belonging to the school book task, the child produces only 12 content words. In the complete transcripts from which these two excerpts are taken, the lexical density of the child was 1.80 and 1.31 respectively during the home and school interactions.

Lexical density and lexical diversity were also measured during the picture task and the circle time. A comparison between the home and school results is illustrated in Table 5.14.

Table 5.14: Overview of the Wilcoxon Signed-rank test results of children's lexical density and diversity during the home picture task and the school circle time

N=12		Time 2 (Age 4;2 years)		Significance level	Time 3 Age 5;10 years		Significance level
		M	SD		M	SD	
Density	Home	1.09	.35	Z=-1.41	1.51	.63	Z=-.16
	School	1.34	.53	(p=.16)	1.63	.86	(p=.88)
Diversity	Home	48.02	16.54	Z=.00	56.19	38.95	Z=-.42
	School	50.84	9.15	(p=1.0)	53.12	23.08	(p=.67)

Concerning lexical density, children used more dense vocabulary during the school circle time than during the home picture task but it should be noted that the differences in both measurement times were not significant (p=.16 and .88). At the level of lexical diversity, the mean score in measurement time 2 was slightly higher at school (50.84) than at home (48.02). The reverse can be seen in measurement time 3, where the children's language in the home task was more diverse (56.19) than in the school task (53.12). However, these differences were statistically not significant (p= 1.0 and .67 respectively).

5.6 Input-output relationships

In this chapter, we were also interested in the parents' and teachers' use of lexical academic language features and its relation to the children's use of these features. For this, correlation analyses were conducted to see whether a

relationship could be found between the number of tokens (quantity measure), lexical density and lexical diversity (quality measures) in the input of caregivers at home and at school and in the language used by the children. In Table 5.15 the results of the correlation analysis between the mothers' and children's tokens, lexical density and lexical diversity during the book task are presented. Only the correlations between input at all the measurement times and output at the same measurement time or later are presented.

Table 5.15: Correlations (Kendalls tau) among mothers' and children's lexical features during the book task (N=12)

		Children								
		Time 1			Time 2			Time 3		
		Tokens	Density	Diversity	Tokens	Density	Diversity	Tokens	Density	Diversity
Mothers										
Time 1	Tokens	.18			.00			-.36		
	Density		.26			-.03			-.05	
	Diversity			.56*			.17			.57*
Time 2	Tokens				.21			-.46*		
	Density					-.33			-.05	
	Diversity						.56*			.00
Time 3	Tokens							.00		
	Density								-.05	
	Diversity									-.29

* p<.05

As can be seen from Table 5.15, the number of tokens used by the mothers was not related to the number of tokens used by the children in all the measurement times. Only one significant negative correlation was found between the number of tokens used by mothers in Time 2 and that of children in Time 3. With regard to lexical density, no relationship could be found between the mothers' and children's use of dense utterances in all the measurement times. But finally, we found significant correlations in Time 1 and Time 2 between lexical diversity in parental input and in children's language. In Time 3, however, no

relationship could be established between the mothers' and children's lexical diversity.

We also studied the relationship between parental input and children's output during the picture task. Correlations are displayed in Table 5.16.

Table 5.16: Correlations (Kendalls tau) among mothers' and children's lexical features during the picture task (N=12)

		Children								
		Time 1			Time 2			Time 3		
		Tokens	Density	Diversity	Tokens	Density	Diversity	Tokens	Density	Diversity
Mothers										
Time 1	Tokens	.06			.24			-.01		
	Density		.15			.09			.00	
	Diversity			.11			.16			.27
Time 2	Tokens				.00			.08		
	Density					.09			.18	
	Diversity						.29			.21
Time 3	Tokens							.46*		
	Density								.24	
	Diversity									.18

* p<.05

During the picture task, only one significant association was found; the number of tokens used by mothers in Time 3 was positively associated with the number of tokens used by the children. Concerning other variables, positive but non-significant relationships were found between mothers and children.

In the school context, we explored the relationship between the teachers' input and the children's language use during the book task and the circle time. Table 5.17 provides an overview of correlation analyses between teachers' and children's total number of tokens, lexical density and lexical diversity during the book task.

Table 5.17: Correlations (Kendalls tau) among teachers' and children's lexical features during the book task (N=12)

		Children					
		Time 2			Time 3		
		Tokens	Density	Diversity	Tokens	Density	Diversity
Teachers							
Time 2	Tokens	.03			-.07		
	Density		.23			.17	
	Diversity			.20			-.07
Time 3	Tokens				.20		
	Density					-.03	
	Diversity						-.07

It can be seen from Table 5.17 that no significant relationship was found between teachers' and children's scores at the level of tokens, lexical density and lexical diversity. During the circle time, only one significant relationship was found. This was between teachers' and children's tokens in measurement time 3 (see Table 5.18).

Table 5.18: Correlations (Kendalls tau) among teachers' and children's lexical features during the circle time (N=12)

		Children					
		Time 2			Time 3		
		Tokens	Density	Diversity	Tokens	Density	Diversity
Teachers							
Time 2	Tokens	.18			-.30		
	Density		-.18			.06	
	Diversity			.40			.14
Time 3	Tokens				.46*		
	Density					-.03	
	Diversity						-.14

As already mentioned, the small number of participants and the variation within mothers, teachers and children, could possibly not lead to high and stable correlations.

5.7 Discussion

In this chapter, the lexical features of academic language have been investigated. Next to quantity measures, quality variables (density and diversity) in the language input of mothers and teachers as well as of the children were studied during two different interaction tasks at different times at home and at school. In this section, the outcomes presented above will be discussed in the light of research questions formulated earlier.

The first question addressed the extent to which Moroccan Berber mothers use lexically rich language in their input to their preschool children. With regard to the quantitative aspects, the data resulting from the two interaction tasks, picture description and book reading, showed differences across the three measurement times. The mean number of tokens, word types and content words produced by the mothers during the book task was higher than during the picture task in all measurement times. This meant that mothers produced more talk during the structured task, book reading, than during the picture task. In considering the effect of time for both interaction tasks, mothers' mean number of tokens, word types and content words decreased as the children got older. Especially in Time 3, when children were nearly 6, mothers seemed to talk less in both tasks. This might be attributed to the increasing participation of children in the mother-child conversations and thus the decrease of the mothers' number of utterances in general. In a similar way, the findings at the level of the quality variables which we distinguished as features of academic language showed substantial differences between the book task and picture task. In the three measurement times, mothers showed a higher lexical density and lexical diversity during the book task than when involved in the picture task. Henrichs (2010) in contrast found that all lexical academic register features occurred more often during the activity of mealtime conversation – less structured and more spontaneous. What might account for this finding is that the book tasks contained a more directive structure with a clear story and characters which elicited rich and varied vocabulary. On the other hand, the picture tasks consisted of isolated pictures and less varied themes, and thus

encouraged less dense and diverse talk. It must be noted that as with quantity variables, mothers' lexical density and diversity were not subject to increase over measurement times. Lexical density and lexical diversity in mothers' language increased until children reached the age of 4;2. As children reached age 5;10, mothers' input became less dense and diverse.

The second research question concerned the relationship between the academic language features in mothers' input and their family background (SES and literacy). At the level of quantity, positive and significant correlations were found between the total number of tokens in mothers' speech and SES and literacy. Especially during the book task and when children reached the age of 5;10, mothers with high SES and literacy tended to use more tokens. With regard to lexical density, and especially in the book task, mothers' SES and literacy level were positively related to the use of dense vocabulary. The other lexical academic language feature, diversity, showed a significant correlation with SES and literacy in both tasks in almost all measurement times. This outcome suggests that, probably, reading activities among Berber mothers with higher SES and literacy levels is a common practice, specifically in Dutch, as there are very few children's books in Berber. This habit is, for these mothers, a helpful instrument in providing a rich and varied vocabulary.

As regards the input of the teachers (research question 3), it was demonstrated that the quantity and quality variables in both interaction tasks showed an increase from Time 2 to Time 3. School data also demonstrated that teachers talk more and use a richer vocabulary as children grow older. At the level of quantity variables clear task effects were found. Teachers tended to talk more when they were involved in book reading than when involved in the circle time. However, this effect becomes less straightforward when we consider lexical density and lexical diversity. While teacher's lexical density was higher during the book task than during the circle time, the opposite can be seen with lexical diversity; this was higher during circle time than during the book task. This might be an effect of the more diverse topics during the circle time.

As far as the children are concerned, the home data showed that the picture task quantitatively generated more talk than the book task. In all measurement times, children produced more utterances, tokens and types while involved in the picture task than when talking about the book. Another interesting difference between the two home tasks is that all quantity variables decreased during the picture task as children grew older; the book task in contrast showed

drops and rises. However, when we consider the lexical academic register features, density and diversity, we see a different pattern in the two home activities. On the whole, children seem to use more dense and more diverse language during the book task than during the picture task. As expected, the book tasks, usually associated with schooling, facilitate the use of a richer vocabulary. Looking at the home interaction tasks, we were also interested in the effect of SES and literacy backgrounds of the family on the use of lexical features of academic language. A positive and sometimes significant impact of both family background factors were found during the book task. Lexical diversity in particular was strongly correlated with the socio-economic status of the family in all measurement times.

In the school setting, children produced less utterances than in the home setting. This was expected, as children in school interactions were involved, together with other children, in tasks with their teacher and they had to talk Dutch only. With regard to task effect, no striking differences were found at the level of the language production between the book task and the circle time. It is worth noting that quantity variables showed a considerable increase from Time 2 to Time 3. When we look at the lexical features of the academic language, density and diversity, we see a task effect. Unlike in the home interaction tasks, the circle time elicited more dense and more diverse language than the book task. Children's lexical density and diversity also witnessed an increase from Time 2 to Time 3.

On the basis of the findings in the present chapter, it can be concluded that the language input of the caretakers at home and at school contained many aspects of the lexical features of academic language. The comparison of mothers' and teachers' language use revealed that the main differences were found at the level of quantity variables while no significant differences were found at the level of density and diversity. Furthermore, children showed a rapid increase on many lexical aspects in the home as well as in the school context particularly between the ages of 4;2 and 5;10. It is clear, thus, that these academic features emerge and develop as children start their school career.

As for the input-output relationships, the data of this (small) group of participants could not convincingly point to a strong relationship between academic features in the input of caretakers, and the same features in the output of children.

CHAPTER 6

The morpho-syntactic level of academic language: clause combining

To study morpho-syntax and its relation to the development of academic language of Moroccan Berber children, we focus in this chapter on clause combining[1] strategies in the input of the mothers and teachers as well as in the language production of the children. Section 6.1 deals with the operationalization of clause combining. In section 6.2, the results concerning the use of clause combining in the input of the mothers are presented. Clause combining in the input of the teachers is addressed in section 6.3. In section 6.4 the results of a comparison between mothers' and teachers' input are provided. The results of children's output in the home and school settings follow in section 6.5. Section 6.6 deals with the input-output relationships. Finally, section 6.7 contains a discussion of the results.

The objective of this chapter is first of all to investigate how Moroccan Berber children are exposed to complex structures during conversations with their parents and teachers and how this influences their language use. Evidence for positive effects of the use of clause combining by parents and teachers on the children's syntactic growth has been found by Huttenlocher *et al.* (2002). The focus in this chapter is on the use of simple and complex sentences by the mothers and teachers while talking with the children in the home and school environments, respectively. Of course, the results from the children's output will also be presented. We assume that the frequency of complex sentences used in caretakers' speech at home and at school provides a basis for the development of academic language required and expected at school. In addition, it is expected that the socio-economic background and the literacy practices of the families play an important role in shaping the children's

[1] In this study, terms referring to clause combining such as complex constructions and syntactic complexity are used interchangeably.

academic language, or what Christie (1985) called 'the hidden curriculum' of schooling. Therefore, these two parameters are included. Bearing these conditions in mind, the following research questions are formulated:

1 To what extent do Moroccan Berber mothers use complex structures in their input to their (pre)-school children during book reading activity?
2 Is the use of complex structures by the mothers related to SES and home literacy practices during book reading activity?
3 To what extent do teachers use complex structures in their input to Moroccan Berber school children while reading books?
4 To what extent do these children use complex structures at home and at school during book reading tasks?
5 Is there a difference between mothers and teachers regarding the use of syntactically complex features during book reading?
6 Is there a relationship between the parents' and teachers' input and the children's use of complex structures (*output*) during book reading activities?

To answer these questions, language data were gathered longitudinally at home and at school. By means of a book reading task, three data collection rounds were carried out at home and two at school. A detailed description is given in Chapter 4.

6.1 The morpho-syntactic features of academic language operationalized

The use of clause-combining strategies in spoken or written language generates complex sentences. The latter is considered to be one of the features of academic language (Schleppegrell, 2004). Huttenlocher *et al.* (2002) report that the use of clause combining in parents' and teachers' speech to children positively influences syntactic growth of children. Learning different constructions enables young children to use them to communicate effectively in different contexts. Clause-combining strategies are needed to talk about events or describe objects in single utterances. The way speakers pack information affects the way their addressees interpret what is said (Clark, 2003). Speakers of all natural languages possess devices and strategies to talk about various related

events and situations. For different discourse functions, speakers can use various strategies by linking clauses in different ways. Such clauses can function as an argument of a predicate or a matrix verb (complement clauses), or as an attribute that modifies a noun or a noun phrase (relative clause), or to specify temporal, spatial, causal, and conditional relations between clauses (adverbial clauses).

Early forms of subordination are usually hard to detect due to the absence of complementizers (e.g. that, to, for), conjunctions (e.g. because, when, before) or relative pronouns (e.g. which, who, that). Such clauses without complementizers are also called 'pre-conjunctional subordinate clauses' (Penner & Mueller, 1992). Consider the following example[2]:

6.1 y-qen tisira, h2enna (Ilyas, 3;2)
3S.M-put on.PERF shoes grandma
(he puts on shoes, grandma)

This Tarifit Berber utterance was produced by a Moroccan Berber child at age (3;1,4) while talking with his mother during a picture book reading task. The child was actually describing a scene where a boy was putting on his shoes to go and visit his grandmother. In the adult-like use of complementation and subordination, this sentence would consist of a main clause and an embedded complement clause linked with the complementizer *h2uma/bac* (meaning: in order to). The complete sentence would look like:

6.2 qa y-t-qen tisira **bac** ad i-rah2 ghar h2enna-s
IMPER 3S.M-put on.PERF shoes to FUT 3S.M-go.AOR to grandma-his
*(he is putting on his shoes **to** visit his grandmother)*

With the absence of markers, only the context assists to recognise this initial phase of complementation and subordination produced by young children. Complementizers such as *that* do not appear in the early stages of language development.

Another challenge we face in dealing with clause combining while studying a bilingual corpus of Berber and Dutch is that the notion of subordination is

[2] Note that not all examples used in this section are from the samples used in the current study. Utterances which belong to our informants will be explicitly mentioned.

mainly developed on the basis of Indo-European languages (Cristofaro, 2003). In Berber, it is sometimes hard to detect or to consider a sentence as a complex one because of the absence of connectives.

Complex sentences are grammatical entities consisting of multiple clauses. These grammatical entities describe a certain relationship between two (or more) situations/events in two (or more) clauses. The four clause-combining types will be discussed in the following sections.

Coordinate clauses

Coordination is the combination of two or more elements known as conjuncts or conjoins, that can stand independently as full propositions. According to Dik (1968: 25), coordination can be defined as 'a construction consisting of two or more members which are equivalent as to grammatical function, and bound together at the same level of structural hierarchy by means of a linking device'. In this process, a string of two or more coordinated constituents (clauses, phrases, or words) are linked together by co-ordinating conjunctions (e.g. or, and, but) on the same functional level. Examples of connectives used in Dutch are *en*, *maar*, *of* (and, but, or). In Berber, the following linking devices can be used: *d/u, maca*, *nigh* (and, but, or). However, in Berber as in many other languages, connectives are not always used explicitly to express coordination (Ennaji, 1985). Instead, coordinated items are temporally juxtaposed next to each other with empty conjunction (ø) as in the following example:

6.3 Hadda t-isi-d stilu ø ymma-s t-arzu-d tawriqt
Hadda 3S.F-bring.PERF-directional particle pen mother-his 3S.F-look for.IMPERF-directional particle paper
(Hadda brought a pen [and] her mother looked for a paper)

Coordinate clauses are known to develop as one of the first structures in the language acquisition process (Bowerman, 1979). In addition, they provide the basis for the development of other complex constructions such as adverbial and relative clauses.

Complement clauses

The first complex constructions which appear in children's speech are complement sentences which mark the transition from the use of simple

isolated clauses to the more complex aspects of grammar. Generally speaking, a complement clause can be defined as a clause which is the complement of another constituent. Therefore, a clause can be a complement of a noun, verb or adjective. Complementation is the special instance of complex sentences in which one proposition serves as an argument within another proposition (Bloom *et al.*, 1989). Consider the following examples:

6.4 I think Tim is there.

6.5 I know what the book is about.

6.6 Let's see if we can find it.

It should be mentioned here that sentential complement clauses produced by children at early age, and those used in spontaneous speech (even by adults) generally do not contain the complementizer *that*. Examples of complement clause markers in Dutch are *dat, om, voor, of, wat* (that, to, for, what). In Berber, the following conjunctions can be used: *h2uma*, *blli/billa, umi* and *bash* (in order to, that, to). The following examples illustrate Berber and Dutch complement clauses.

6.7 het is leuk dat de juffrouw een nieuw boek gaat lezen
(it is nice that the teacher is going to read a new book)

6.8 t-nna-yi ymma-s belli wa ghar-es ca lkitab nni
3S.F-tell.PERF mother-his/her that NEG have-CL.her NEG book that
(his mother told me that she didn't have that book)

Relative clauses

A relative clause is a subordinate clause that modifies a noun or a noun phrase. In other words, a relative clause is about two states of affairs where one is independent and provides specific modification of the other (Tomasello, 2003). Unlike complement clauses, relative clauses are used to modify and specify a noun phrase in more detail. Keenan and Comrie (1977: 63-64) consider:

> *any syntactic object to be an RC [relative clause] if it specifies a set of objects [...] in two steps: a larger set is specified, called the domain of relativization, and then restricted to some subset of which a certain sentence, the restricting sentence, is true.*

> *The domain of relativization is expressed in surface structure by the head NP [noun phrase], and the restricting sentence by the restricting clause.*

At the stage when children start using complex sentences, relative clauses are produced without relative pronouns by using empty nouns like *ones*, *kind* or *thing* as in example 6.9:

6.9 Give me the ones (pens) I got from school

After this stage, children start producing sentences by using nouns (such as pen, man, dog) and relative pronouns (who, that, which) as shown in example 6.10:

6.10 Look, these are the pens that I got from school

6.11 ta xzar busje@nl amacnaou tenni n papa t-ameqran-t
COP 2S-look.AOR van like that of papa CL.it-big-3S.F
(look, a van like that big one of my father) (Anisa, Time 1)

Adverbial clauses

An adverbial clause can be defined as a modifier of the associated matrix clause or verb phrase. Adverbial clauses indicate things like temporal (when), causal (why), conditional (under what conditions) and other relations between clauses. The following examples illustrate some of these relations:

6.12 She bought this toy <u>when</u> she was on holiday (temporal)

6.13 She bought the toy <u>because</u> it was cheap (causal)

6.14 <u>If</u> she shows it to the children, they will like it very much (conditional)

In addition, there are other types of adverbial clauses which indicate place, purpose, reason and manner. Consider the following example:

6.15 ura Yasmin tugha tt ghar-es <u>umi</u> tugha-tt d takkuh2-t.
(umi=when)
Also Yasmin was CL.it-3S.F have-3S when was F-young-F
(Yasmin also had that when she was young)

To operationalize syntactic complexity, transcripts were coded by using four levels of clause combining: coordinate, complement, adverbial and relative clauses (see Chapter 4). Altogether, these types of clause combinations constitute the so-called multi-clause sentences. These four categories are subsequently divided into two main categories according to the degree of complexity: (i) coordinate clauses which are classified as non-embedded sentences are considered simple multi-clause sentences and (ii) subordinate clauses (consisting of complement, adverbial and relative clauses) which are considered more complex. Afterwards, two measures are computed by using the proportions: (a) the proportion of multi-clause sentences in all utterances, (b) the proportion of subordinate multi-clause sentences in all utterances.

6.2 Clause combining in mothers' input

The first results we analyze are those of the mothers. In Table 6.1 the absolute numbers of all four clause combining categories produced by each individual mother during the three measurement times are presented, together with the group means and standard deviations.

Table 6.1: Individual scores, means and standard deviations of mothers' multi-clause utterances in three measurement times

	Coordinate			Complement			Adverbial			Relative		
Time	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3
Age child	3;2	4;2	5;10	3;2	4;2	5;10	3;2	4;2	5;10	3;2	4;2	5;10
Mother												
1	10	8	4	21	6	0	5	1	6	0	3	1
2	14	2	2	17	19	5	5	5	6	9	6	0
3	4	7	5	6	2	2	1	0	3	0	1	4
4	3	1	4	12	0	14	1	4	3	4	7	3
5	4	7	3	11	5	5	1	0	1	4	6	2
6	9	7	2	17	11	10	0	2	2	0	2	2
7	9	2	1	8	3	2	1	1	0	4	5	0
8	0	2	2	1	2	2	1	2	4	0	0	3
9	3	1	3	7	0	3	0	1	0	0	4	0
10	4	2	1	7	0	0	1	1	0	6	2	0
11	5	2	1	10	2	8	1	1	2	4	5	2
12	7	1	3	12	1	9	7	2	5	0	9	4
Mean	6.0	3.50	2.58	10.75	4.25	5.00	2.00	1.67	2.67	2.58	4.17	1.75
SD	3.88	2.81	1.31	5.56	5.63	4.39	2.30	1.50	2.23	3.03	2.66	1.54

The first remark about the clause combining strategies used by the mothers is that they have to be related to the total number of utterances at each of the measurement times. The number of utterances at Time 1 ranges from 38 (mother 8) to 163 (mother 2), at Time 2 from 56 (mother 1) to 119 (mother 4), at Time 3 from 27 (mother 10) to 97 (mother 2), and means varying from 105 (Time 1), to 79 (Time 2) and 62 (Time 3). On average, the mothers say gradually less as the children get older, and the variation is high. Mothers show large differences in the use of clause combining strategies, ranging from only a few (mother 8 and 9) to relatively many (mother 1 and 2). Most of the mothers do not seem to use (in absolute numbers) more multi-clause sentences as the child grows older; several in fact use less (for proportions see below). There is a fairly stable pattern in the relative use of each of the categories. On average, the complement clauses are used most at all three measurement times, followed by coordinate clauses, and to a lesser extent adverbial clauses and relative clauses.

This seems to be true for most of the individuals, an exception being mother 12 who relatively more often uses adverbial and relative clauses compared to her use of coordinate and complement clauses.

Further qualitative examination of the data shows that when complex constructions occur in mothers' speech, Berber is the most used language with Dutch used only sporadically. Table 6.2 gives an overview of different types of multi-clause sentences used by the mothers in the three measurement times.

Table 6.2: Overview of multi-clause sentences used by the mothers in 3 measurement times

	Time 1 (Age 3;2 years)	Time 2 (Age 4;2 years)	Time 3 (Age 5;10 years)
Coordinate clauses	Total 72 (including 4 in Dutch) – 53 with "*d*" (and) – 6 with "*nigh*" (or) – 5 with "*maca*" (but) – 4 without markers – 4 in Dutch with "*en"* (and)	Total 42 (including 2 in Dutch) – 25 with "*d*" (and) – 5 with "*maca*" (but) – 2 with "*uca*" (so) – 8 without markers – 2 in Dutch with "*en"* (and)	Total 31 (including 3 in Dutch) – 11 with "*d*" (and) – 3 with "*maca*" (but) – 1 with "*ssenni*" (so) – 3 with "*nigh*" (or) – 10 without markers – 3 in Dutch with "*en"* (and)
Complement clauses	Total 129 (including 6 in Dutch) – 35 with "*bac*" (in order to) – 29 with "*h2uma*" (to, in order to) – 9 with "*billi*" (that) – 19 with "*main*" (what) – 31 without markers	Total 51 (including 3 in Dutch) – 13 with "*bac*" (in order to) – 9 with "*h2uma*" (to, in order to) – 8 with "*billi*" (that) – 1 "*illa*" = "*billi*" (that) – 6 with "*main*" (what) – 14 without markers	Total 60 (including 5 in Dutch) – 21 with "*bac*" (in order to) – 1 "*ataf*"="*bac*" (in order to) – 15 with "*h2uma*" (to, in order to) – 10 with "*billi*" (that) – 1 "*mamec*" (how) – 2 with "*main*" (what) – 11 without markers

	Time 1 (Age 3;2 years)	Time 2 (Age 4;2 years)	Time 3 (Age 5;10 years)
Adverbial clauses	Total 24 (including 2 Dutch) – 8 of place "*x*" (on), "*ttarf*" (close to), "*zzati*" (next to), "*daniti*" (here), "*awarni*" (behind), "*mani*" (where) – 7 manner: "*am*" (like), amacnaou (like, as), axmi (as) – 2 conditional "*mara*" (if) – 2 time: "*umi*" (when) – 1 reason: "*umi*" (since) – 2 Dutch: conditional "als" (if); causal/consequential "*want*" (because) – 2 unmarked (time)	Total 20 (including 2 in Dutch) – 2 time: "*umi*" (when), "*ruxa*" (now) – 7 reason: "*umi*" (since), "*lah2eqqac*" (because), "*arami*" (since) – 8 conditional "*mara*" (if); "*mri*" (if); "*maolik*" (if) – 2 Dutch: causal/ consequential "*want*" (because) – 1 result clause: "*uca*" (so)	Total 32 (including 5 in Dutch) – 19 time: "*umi*" (when), "*mani*" (when), "*xmi*" (when), "*ghir… t ..*" (as soon as), "*di rweqt*" (at that moment, "*umbaod*" (after), "*ruxa, rexxu*" (now), "*qber*" (before) – 2 reason: "*minzi*" (because) – 1 conditional: "*mara*" (if); – 3 manner: "*am*" (like), – 1 purpose: "*bac*" (so that) – 1 place: "*mani mma*" (wherever) – 5 Dutch: time "*toen*" (when), reason "*want*" (because), conditional "*als*" (if)
Relative clause	Total 31 – "*…i...*" (which is), "*…i…*" (the one who), "*tenni i*", "*wenni i*" (the one who), "*nni*" (the one which), "*mana ya umi*" (one which)	Total 50 (including 2 in Dutch) – "*...i...*" (which is), "*…i…*" (the one who), "*tenni i*", "*wenni i*", "*yinni*" (the one(s) who), "*nni*" (the one which), "*mana ya umi*" (one which). – without markers	Total 21 – "*...i…*" (which is), "*…i…*" (the one who), "*tenni i*", "*wenni i*", (the one who), "*nni*" (the one which)

Regarding coordinate clauses, mothers used in total 72, 42 and 31 coordinate clauses in Time 1, Time 2 and Time 3 respectively. These coordinate clauses included 7 Dutch clauses in the three measurement times. The most frequent coordinating conjunctions are "d" (and), "maca" (but) and "nigh" (or). The following sentences are examples of coordinate clauses used by the mothers in different measurement times:

6.16 (Mother of Hatim, Time 3)
t-tt-wari-d tittawin dduqz-en-t dayes *d* i-qemmum-en d-imeqran-en
2S-see-IMPERF-2S eyes bulge.PERF-3P.F in him and mouths are big
(you see the eyes are bulging out and the mouths are so big)

6.17 (Mother of Ibtisam,Time 1)
ad t-sh2ess-ed ghari *nigh* qa t-tt-gg-ed ca h2aja nneghni?
FUT 2S-listen-AOR-2S to me or T.A. 2S-do-IMPERF any something else
(are you listening to me or you are doing something else?)

6.18 (Mother of Isra, Time 3)
t-ttwara tighmas d t-imeqran-in aqemmum d-ameqran *maca* war t-zemmar ad t-oqeb safi
3S-see.IMPERF teeth COP big-PL mouth COP-big but NEG 3S.F-can.AOR FUT 3S.F-return.AOR
(she sees the big mouth and teeth but she cannot go back)

In addition to explicit use of conjunctions, we also coded unmarked coordinate sentences used by mothers. In three measurement times, mothers used 21 unmarked coordinate clauses as example (6.19) shows.

6.19 (Mother of Saliha, Time 1)
yemma-s t-ssagh-d frisa temghar ø ta t-uca-s tbanant i Tijn
Mother-his 3S.F-buy.IMPERF-directional particle strawberries F-big this 3S.F-give.AOR banana to Tijn
(his mother bought big strawberries and that woman gave a banana to Tijn)

Next to Berber sentences, mothers used also Dutch coordinate clauses, although not frequently. The total number of these sentences is 8 and the conjunction used is "en" (and) as in example (6.20).

6.20 (Mother of Selma, Time 1)
daar@nl is@nl opa@nl buiten@nl *en*@nl oma@nl komt@nl hier t-nna-s yemma-s nni n Tijn
There is grandfather outside and grandmother is coming here 3S.F-tell.PERF-3M mother her that of Tijn
(there is grandpa outside and granny is coming here, Tijn's mother told him)

The second category and the most used one by the mothers is the complement clause. Mothers used in total 129, 51 and 60 complement sentences in measurement time 1, 2 and 3 respectively. The most frequent complementizers are "bac" (to, in order to), "huma" (in order to), "billi, billa" (that), "main" (what) and "mamec" (how). Examples 6.21 - 6.29 illustrate the types of

sentences that frequently were used by the mothers while talking with their children about the book.

6.21 (Mother of Anissa, Time 1)
xzar t-qqar-as deghya qqen tisira *bac* ad n-rah2 ghar h2enna d jeddi
2S-look.AOR 3S.F-tell.IMPERF-him quickly 2S-put on.AOR shoes in order to FUT 3P.M/F-go.AOR to grandpa and granny
(look she is telling him put on your shoes quickly to go to grandpa and granny)

6.22 (Mother of Abdelilah, Time 2)
i-tebo-it baba-s i-rezzu *bac* ad y-arr mmi-s nni
3S.M-follow.PERF-him father-his 3S.M-want.IMPERF in order to FUT 3S.M-get back.AOR son-his that
(his father followed him to get his son back)

6.23 (Mother of Ibtisam, Time 1)
egg tisira qqen tisira h2uma ad n-rah2 ghar jedd-ec d henna-c
2S-do.AOR shoes 2S-put on.AOR shoes in order FUT 1P-go.AOR to grandpa-your and granny-your
(put on your shoes to go to your grandfather and your grandmother)

6.24 (Mother of Selma, Time 3)
rebda ad t-qarri-d lkitab nni *h2uma* ad t-dured mlih2
always FUT 2S-read-AOR-2S book that in order to FUT 2S-become better
(always read that book to become better)

6.25 (Mother of Hatim, Time 1)
nec nnigg-ac *billi* ad n-rah2 rexxu.
I tell.PERF-1S-you that FUT 1PL-go.FUT now
(I told you that we are going now)

6.26 (Mother of Hamza, Time 3)
kijk@nl aqq-ec danitin y-na-s udbib *billi* war din h2etta ci n lmuckil
look P.A-2M here 3S.M-tell.PERF-him/her doctor that NEG there any no a problem
(look, here, the doctor told him that there is no problem)

6.27 (Mother of Hatim, Time 1)
iwa ini *main* da-s t-nna i Tijn
just 2S-say.AOR what COP-to-him 3S.F-tell-PERF to Tijn
(just say what she told Tijn)

6.28 (Mother of Selma, Time 2)
arwah2 ad n-xzar *main* da tt-egg-en
2S-come.AOR FUT 1P-see.AOR what here do.IMPRF-3P.M
(let's see what they are doing here)

6.29 (Mother of Ilyas, Time 3)
xzar t-pinsar *mamec* gha t-suya ig temz2a nni.
Look 3S.F-think.IMPERF how to 3S.F-deal with ogress that
(look she is thinking how to deal with that ogress)

Next to the use of explicit complementizers, mothers used also unmarked complement clauses, especially in Time 1 and Time 2, as examples 6.30 and 6.31 show.

6.30 (Mother of Selma, Time 2) (ø substitutes *that*)
y-nna-s maca t-nni-d ayi ø ad k-ic rah2-gh
3S.M-tell.PERF-him but 2S-tell.PERF-2S to me FUT with you go.FUT-1S
(he told him but you told me I will go with you)

6.31 (Mother of Hatim, Time 3) (ø substitutes *in order to*)
qa t-qqar d-ays di buki nni ø ad t-ssen min ghar t-egg.
T.A 3S.F-read.IMPERF in it in book that FUT 3S.F-know.AOR what FUT 3S.F-do.AOR
(she is reading in that book to know what to do)

The total number of Dutch complement clauses in three measurement times was 14. The complementizers used are "dat" (that) and "wat" (what). Examples of Dutch complement clauses are 6.32 and 6.33.

6.32 (Mother of Isra, Time 3)
t-nna-s *dat*@nl ze@nl niet@nl bang@ moet@nl zijn@nl
3S.F-tell-PERF- her that she not afraid should to be
(she told her that she should not be afraid)

6.33 (Mother of Anissa, Time 2)
kan@nl je@nl nou@nl aan@nl mama@nl vertellen@nl *wat*@nl ik@nl gezegd@nl heb@nl?
can you now to mommy to tell what I said have?
(can you tell mommy what I have said?)

Another multi-clause category used by the mothers is adverbial clauses. They used 76 adverbial clauses in three measurement times (24 in Time 1, 20 in Time 2 and 32 in Time 3), including 9 Dutch sentences. The most frequent types are time, place, reason, purpose, manner, conditional, causal and consequential adverbial clauses. The frequent conjunctions are "umi" (when), "mani" (where), "mri, meolak" (if), "am" (like), "mainzi" (because) and "lah2eqqac" (because). These categories are illustrated in examples 6.34 - 6.42.

6.34 (Mother of Yasmin, Time1)
ura d Yasmin tugha-t ghar-es *umi* t tugha d ttakkuh2t
also CL Yasmin 3S.F-be-PERF 3S.F-have when 3S.F 3S.F-to be-PERF little
(also Yasmin had that when she was little)

6.35 (Mother of Hamza, Time 2)
umi y-wda papa-s daniti i-ruh2 Tim
when 3S.M-fall.AOR papa-his here 3S.M-go.AOR Tim
(Tim went when his father fell here)

6.36 (Mother of Abdelilah, Time 3)
Emma *ghir* t-xecc-ed ghar taddart *t*-bda t-qqar di lkitab nni
Emma as soon as 3S.F-arrive.AOR at home 3S.F-start.AOR 3S.F-read.IMPERF in book that
(as soon as Emma arrived home she started reading in that book)

6.37 (Mother of Anissa, Time 1)
mara war i-qqen tisira nnes war zemmar-en ad uyur-en
If NEG 3S.M-put on.AOR shoes his NEG can.AOR-3P.M FUT walk-3P.M
(if he does not put on his shoes they cannot walk)

6.38 (Mother of Ilyas, Time 2)
y-nna-s *meolik* ula ccek t-qqim-ed x lusa n ssira iri war d ac i-t-wqio walu
3S.M-tell.PERF- him if also you 2S-stay.AOR on paving stones then NEG P.A you PASS-happen nothing

(he told him if you also stayed on the paving stones nothing would have happened to you)

6.39 (Mother of Yasmin, Time 3)
i-oejb-as rh2ar ad t-en t-ssigwed ura d nnettat *minzi* ra-d nitni rebda ssagwad-en-t tteji-n-t war t-ttes.
3S.M-like.AOR-her it FUT them 3S.F-frighten.AOR also her because also P.A they always 3P.M-frighten.IMPERF-her let.IMPERF-P.M-her NEG 3S.F-sleep.IMPERF
(she also liked it to frighten them because they always frighten her and let her without sleep)

6.40 (Mother of Abdelilah, Time 3)
wi d ayi ghar i-oawen-en *bac* war tt-arji-gh di manaynni war t-gwd-egh?
who P.A me FUT 3S.M-help.AOR in order NEG dream.IMPERF-1S in that NEG be scared-IMPERF-1S
(who can help me in order not to dream about that and not to be scared?)

6.41 (Mother of Yasmine, Time 3)
mermi mma t-rah2 ad t-ttes t-tt-arja tirja tioeffan-in
whenever 3S.F-go.AOR FUT 3S.F-sleep.AOR 3S.F-dream.AOR dreams bad-P
(whenever she goes to sleep she dreams about bad things)

The Dutch adverbial clauses used by the mothers are conditional, causal or temporal. For example:

6.42 (Mother of Abdelilah, Time 3)
toen@nl je@nl nog@nl een@nl heel@nl klein@nl jongetje@nl was@nl hadden@ we@ dat@ ook@
(when you were still a very young boy we also had that)

The last category of multi-clause sentences used by the mothers is the relative clause. These were used 102 times in all measurement times, including only 2 in Dutch. The most frequent relative pronouns used by the mothers are "nni, wenni" (the one, that one), "...i..." (who, which), and "umi" (to whom, which). The following examples illustrate how mothers use the relative clauses in their speech with their children.

6.43 (Mother of Ilyas, Time 1)
h2ima ad ac t-uc Senna lbula *nni* isebh2n
so that FUT you 3S.F-give.AOR Senna marble which nice
(so that Senna can give you the marble which is nice)

6.44 (Mother Selma, Time 3)
t-arzja uca t-ruh2 ghar dokter@nl *nni* i-toawan-en
3S.F-dream.AOR then 3S.F-go.PERF to doctor who help.IMPERF
(she dreamed and then she went to the doctor who helps)

6.45 (Mother of Saliha, Time 1)
t-nna-s ini cukran *i ta da*-c y-uc-in tbanant
3S.F-say.PERF say.AOR thank you to S.F-this who COP-you-M 3S.F-give.PERF banana
(she told her say thank you to this [woman] who gave you the banana)

Among the Dutch relative clauses used by the mothers, there were:

6.46 (Mother of Hatim, Time 2)
mensen@nl *die*@nl de@nl straat@nl maken@nl zijn@ weg@nl
(paviors who work on the street are gone)

6.47 (Mother Saliha, Time 3)
t-ruh2 ghar dokter@nl *die*@nl haar kan helpen
3S.F-go.PERF to doctor who her can to help
(she went to the doctor who can help her)

Based on the distinction we made in dealing with the morpho-syntactic level between two main categories of clause combining (coordinate clauses and subordinate clauses), the average scores of the mothers are reported in Table 6.3. Both the absolute numbers and the proportions of the total number of utterances will be presented.

Table 6.3: Means and standard deviations for the total numbers and proportions of multi-clause sentences and subordinate clauses in mothers' input

N=12	Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	M	SD	M	SD	M	SD
Number of utterances	105.42	39.75	79.33	19.48	62.42	23.16
Total number of multi-clause sentences	21.33	11.46	13.58	7.82	12.0	6.66
Total number of subordinate multi-clause sentences	15.33	7.99	10.08	7.35	9.42	6.10
Proportion of multi-clause sentences over all utterances	.19	.07	.18	.11	.18	.08
Proportion of subordinate multi-clause sentences over all utterances	.14	.05	.13	.09	.14	.08

Concerning the total number of multi-clause sentences, the mothers scored on average the highest in Time 1 (mean=21.33), followed by Time 2 (mean=13.58) and finally by Time 3 (mean=12). During each data-gathering round, the variation was high, ranging from 1 multi-clause sentence (mother 10, Time 3) to 45 (mother 2, Time 1). The same holds true for the total number of subordinate multi-clause sentences. The average total scores decreased from 15.33 in Time 1 to 10.08 in Time 2 to 9.42 in Time 3, with a range from 0 (mother 10, Time 3) to 31 (mother 2, Time 1). On average, a decrease in the total number of subordinate clauses used by the mothers is noticed as the children get older. In the last measurement time, when children were at the age of 5;10, mothers used, on average, the lowest number of subordinate clauses in comparison with Time 1 and Time 2. On average mothers produced the highest number of utterances in Time 1. It is also of note that the standard deviations for the total number of multi-clause sentences as well as the total number of subordinate clause sentences are high, indicating that there is a high variation between the mothers taking part in this study.

If we consider the proportions of multi-clause sentences in all utterances, 19% of all the utterances in Time 1 and 18% in Time 2 and Time 3 produced by the mothers during the conversations with their children contained multi-clause sentences; these ranged from 4% (mother 10, Time 3) to 39% (mother 2,

Time 2). ANOVA with repeated measures shows that there was no significant change over time ($F(2,22)=.11$, $p=.89$, $\eta_p^2=.010$). A pattern almost the same can be seen at the level of the proportions of subordinate sentences over all utterances (14%, 13%, 14% respectively for Time 1, Time 2 and Time 3), ranging from 0% (for mother 10, Time 3), to 37% (mother 2, Time 2). ANOVA revealed no main effects of time ($F(2,22)=.10$, $p=.90$, $\eta_p^2=.010$).

When we look at the changes over time for each of the mothers, there are only two mothers who seem to increase the proportion of subordinate sentences over time (mother 6 from 11% to 24% and mother 12 from 12% to 21%); two mothers show a decrease over time (mother 7, from 11% to 6%, and mother 10, from 12% to 0%); all other mothers show variable patterns. This clear lack of increase might be explained by the fact that mothers do not take the age of their children into account and subsequently do not adjust their language use to the child's language competences. Another factor that may account for this behavior is the increased participation of children in the conversation. As children get older, they usually initiate the conversations, which might lead to more simple utterances in the reactions of mothers. The content of the book at hand might also cause differences.

6.2.1 Relationship of mothers' input with SES and literacy

The second research question was whether there is any impact of the family's socio-economic background and literacy activities at home on the use of different complex constructions by the mothers. To find out, a correlation analysis was conducted over the total number and proportions of multi-clause sentences and subordinate clauses, SES and literacy. The results are shown in Table 6.4.

Table 6.4: Kendall Tau correlations of SES and literacy with complex constructions used by the mothers

N=12	Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	SES	Literacy	SES	Literacy	SES	Literacy
Multi-clause sentences	.23	.35	.37	.33	.66**	.45*
Subordinate clauses	.36	.50*	.42*	.47*	.66**	.48*
Prop. multi-clause/utterances	-.03	.21	.22	.08	.38τ	.34
Prop. subordinate/utterances	-.03	.21	.25	.30	.51*	.34

* p<.05; ** p<.01; τ tendency

The general picture that emerges from Table 6.4 is that all correlations of the number of multi-clause sentences and subordinate clauses with both SES and home literacy are positive and that they were all significant in measurement time 3, when children were at age 5;10. Considering the total number of subordinate clauses, the results show significant positive correlations with the literacy background in all measurement times. As for SES, the correlation analysis indicated a significant relationship in Time 2 and 3. The correlations between the proportions of multi-clause sentences and subordinate clauses over all utterances are lower, become more positive as the child grows older, and only reach significance with SES at Time 3 for the proportion of subordinate clauses. For multi-clause sentences there is a tendency.

6.3 Clause combining in teachers' input

In this section the use of clause combining in teachers' input is examined. Table 6.5 gives the individual scores, mean scores and standard deviations for the teachers concerning different types of complex constructions.

Table 6.5: Individual scores, means and standard deviations of teachers' multi-clause utterances in two measurement times

	Coordinate		Complement		Adverbial		Relative	
Time	T 2	T 3	T 2	T 3	T 2	T 3	T 2	T 3
Age child	4;2	5;10	4;2	5;10	4;2	5;10	4;2	5;10
Teacher								
1	8	10	15	11	8	3	3	3
2	0	4	2	5	0	2	2	2
3	1	5	8	8	0	0	1	2
4	13	5	9	5	4	2	2	0
5	2	10	6	5	2	1	2	3
6	5	6	3	7	4	0	1	2
7	0	13	12	6	5	2	1	3
8	6	13	14	9	6	1	0	0
9	12	15	14	0	6	2	3	0
10	1	7	9	2	0	3	0	4
11	13	12	5	5	5	4	0	4
12	19	7	9	8	6	3	0	2
Mean	6.67	8.92	8.83	5.92	3.83	1.92	1.25	2.08
SD	6.33	3.73	4.32	3.0	2.73	1.24	1.14	1.44

As expected, the coordinate and complement clauses were the most used at both measurement times. The average of coordinate clauses used during Time 1 was 6.67 and in Time 2 was 8.92, ranging from 0 (teacher 7, Time 2) to 15 (teacher 9, Time 3). We also see that there is on average an increase over time in the use of coordinate clauses. This holds for most teachers, teacher 4 being a clear exception. The use of complement clauses ranged from 0 (teacher 9, Time 3) to 15 (teacher 1, Time 2). As a matter of fact, teachers used, on average, more complement clauses in Time 1 (8.83) than in Time 2 (5.92). Relative and adverbial clauses were the types used least by the teachers at both measurement times, ranging from 0 to 8. If we look at the relative clauses, the teachers produced the least number in comparison with other types of complex sentences. With regard to adverbial clauses, teachers used, on average, 3.83 during Time 1 and 1.92 during Time 2. As for relative clauses, there is on

average an increase over the measurement times (1.25 during Time 1 and 2.08 during Time 2).

In Table 6.6, an overview is provided of different multi-clause sentences used by the teachers in both school measurement times.

Table 6.6: Overview of multi-clause sentences used by the teachers in 2 measurement times

	Time 2 (Age 4;2 years)	Time 3 (Age 5;10 years)
Coordinate clauses	Total 80 – 67 en (and) – 9 maar (but) – 2 dus (so) – 2 of (or)	Total 107 – 81 en (and) – 13 maar (but) – 3 dus (so) – 10 of (or)
Complement clauses	Total 106 – 60 dat (that) – 20 wat (what) – 12 hoe (how) – 3 of of (whether) – 3 om (in order to) – 8 unmarked	Total 71 – 29 dat (that) – 24 wat (what) – 4 hoe (how) – 2 of of (whether) – 5 om (in order to) – 7 unmarked
Adverbial clauses	Total 46 – 16 conditional: als (if), anders (otherwise) – 15 time/sequential: toen (when), terwijl (while, whereas), dan (than), eerst (first), even (for a while), na (after), nu (now) – 10 reason: want (because) – 2 manner: als (as, like) – 2 place: waar (where) – 1 consequential: zodat (so that)	Total 23 – 7 conditional: als (if) – 8 temporal/sequential: daarna (after that), toen (when), als (when), meteen nadat (immediately after) – 6 reason: want (because) – 2 manner: net als (as, just like)
Relative clauses	Total 15 – 2 wat (which) – 4 dat (that) – 7 die (that, who) – 2 wie (who)	Total 25 – 4 wat (which) – 3 dat (that) – 11 die (that, who) – 7 wie (who)

Coordinate clauses are the most used category by the teachers (80 in Time 2 and 107 in Time 3). The most frequent coordinating conjunctions are "en" (and), "maar" (but), "of" (or) and "dus" (so). The following sentences are

examples of coordinate clauses used by the teachers during the book reading task.

6.48 (Teacher of Abdelilah, Time 3)
hij kauwt met z'n grote bek *en* hij werkt het hele blad naar binnen
(he is chewing with his big beak and he is swallowing the whole leaf)

6.49 (Teacher of Selma, Time 2)
hij gaat weer op z'n rode scooter *maar* hij heeft wat achterop
(he leaves again on his red scooter but he has something back on the luggage rack)

6.50 (Teacher of Isra, Time 3)
en hij is heel klein *dus* het is geen olifant
(and he is very small so it is not an elephant)

The second multi-clause category is the complement clause. The teachers used 106 complement clauses during Time 2 and 71 during Time 3. The complementizers used are "dat" (that), "wat" (what, that), "hoe" (how) and "of" (whether). Examples of complement clauses used by the teachers are:

6.51 (Teacher of Abdelilah, Time 2)
Abdelilah zei *dat* het een schaap is
(Abdelilah said that it was a sheep)

6.52 (Teacher of Ouiam, Time 2)
zullen we eens gaan kijken *wat* het geworden is
(let's see what it has become)

6.53 (Teacher of Anissa, Time 2)
om te kijken of hij echt zo zwaar is geworden
(in order to see if he really has become so heavy)

6.54 Tteacher of Ilyas, Time 3)
hij wil ook laten zien *dat* ie dapper is
(he also wants to show that he is brave)

6.55 (Teacher of Yasmine, Time 3)
kijk eens *wat* ik allemaal kan
(look what I can do)

6.56 (Teacher of Selma, Time 3)
nou gaan we eens kijken *of* muis dat wel kan
(let's see if the mouse can do that)

6.57 (Teacher of Yasmine, Time 3)
maar is het heel dapper *om* te zeggen ik doe niet mee
(but it is very brave to say I do not participate)

The third category of multi-clause sentences used by the teachers is the adverbial clause. In the two measurement times, teachers used 69 adverbial clauses (46 in Time 2 and 23 in Time 3). The most frequent types are reason, manner, place, conditional, temporal, sequential and consequential clauses. Among the conjunctions are: "als" (if), "anders" (otherwise), "toen" (when), "terwijl" (while), "want" (because), "net als" (like, as) and "zodat" (in order to). Examples of adverbial clauses used by the teachers are:

6.58 (Teacher of Ilyas, Time 2)
als je nu twee flessen verf koopt krijg je eentje gratis
(if you buy two bottles of paint now you'll get one for free)

6.59 (Teacher of Isra, Time 2)
terwijl ze dit zegt maakt ze een beweging waarmee ze een rij verbeeldt
(while she is saying this she is making a movement with which she portrays a row)

6.60 (Teacher of Ilyas, Time 2)
hij gaat zijn wollen vacht verven *zodat* schaap niet meer wit is
(he is going to paint his woolen hair so that sheep no longer is white)

6.61 (Teacher of Selma, Time 3)
dan is het heel dapper *als* je eens een keertje niet stoer doet
(then it is very brave if you once don't act cool)

6.62 (Teacher of Abdelilah, Time 3)
ik kroop meteen *nadat* ik geboren ben
(I crawled right after I was born)

The last category of multi-clause sentences is the relative clause. This is the one least frequently used by the teachers (15 in Time 2 and 25 in Time 3). The relative pronouns used by the teachers are "dat" (that), "die" (which, that, who), "wie" (who, whom) and "wat" (what). Examples of relative clauses are:

6.63 (Teacher of Ouiam, Time 2)
daar zie ik het kind *dat* iets wil zeggen
(there I see the child who wants to say something)

6.64 (Teacher of Abdelilah, Time 2)
ik ken ook een dier die draadjes kan maken
(I also know an animal which can make threads)

6.65 (Teacher of Abdelilah, Time 3
hier zie je muisje dat niet mee doet
(here you see the mouse that does not participate)

6.66 (Teacher of Selma, Time 3)
dat is een bloem die op het water drijft
(that's a flower which is floating on the water)

After this overview of specific multi-clause categories, the total number of multi-clause sentences and subordinate clauses are summarized in Table 6.7. In addition, proportions of these categories vis-à-vis the number of utterances are presented.

Table 6.7: Descriptives (means and standard deviations) for the total number and proportions of subordinate and multi-clause sentences in the teachers' input

N=12	Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	M	SD	M	SD
Number of utterances	106.83	32.25	128.33	27.19
Total number of multi-clause sentences	20.58	10.85	18.83	4.97
Total number of subordinate multi-clause sentences	13.92	6.72	9.92	3.63
Proportion of multi-clause sentences over all utterances	.22	.13	.15	.05
Proportion of subordinate multi-clause sentences over all utterances	.14	.08	.08	.03

As shown in Table 6.7, the average scores of the teachers at the level of both the total number of multi-clause sentences and subordinate clauses decrease as

the children get older. When children are 4;2 years old (Time 2), teachers use an average of 20.58 multi-clause sentences but 18.83 as the age of the children reaches 5;10. The same pattern can be seen concerning subordinate sentences (13.92 at Time 2 and 9.92 at Time 3). Note that the relatively high standard deviations indicate considerable inter-individual differences between teachers in both measurement times, as seen before.

Considering the whole transcripts, the proportion of multi-clause and subordinate sentences was calculated. With regard to the first category, the proportion of multi-clause sentences relative to all utterances was .22 in Time 2, ranging from .06 (teacher 10) to .41 (teacher 12), and .15 in Time 3, ranging from .09 (teacher 10) to .21 (teacher 5). The ANOVA analysis indicates that the main effect of time was found not to be significant ($F(1,11)=2.95$, $p=.11$, $\eta_p^2=.212$). As to subordinate sentences, the proportion over all utterances was .14 in the second measurement time, ranging from .04 (teacher 2) to .30 (teacher 1) and .08 in the third measurement time, ranging from .02 (teacher 9) to .11 (teacher 1). The main effect of time shows a significant decrease in the proportion of subordinate clauses between Time 2 and Time 3 ($F(1,11)=8.03$, $p=.016$, $\eta_p^2=.42$).

6.4 Input in two settings: mothers and teachers compared

In order to understand the use of syntactic complex constructions in the input children are exposed to at home and at school, a comparison was made between mothers and teachers. A Mann-Whitney U test was conducted using transcripts resulting from two comparable tasks, namely the book task. Mothers and teachers were compared as to the proportions of multi-clause sentences and subordinate clauses. Table 6.8 gives an overview of the mean values of the proportions together with the significance level in two measurement times.

Table 6.8: Overview of the Mann-Whitney U results on the proportions of multi-clause and subordinate clause sentences of the mothers and teachers in two measurement times

		Age 4;2 years		Significance level	Age 5;10 years		Significance level
		M	SD		M	SD	
Proportion of multi-clause sentences over all utterances	Mother	.18	.11	U=63.50 (p=.63)	.18	.08	U=54.00 (p=.32)
	Teacher	.22	.13		.15	.05	
Proportion of subordinate clause sentences over all utterances	Mother	.13	.09	U=64.00 (p=.67)	.14	.08	U=27.00* (p=.008)
	Teacher	.14	.08		.08	.03	

* p<.01

As for multi-clause sentences, no significant differences were found between mothers and teachers during both measurement times. Teachers' means at Time 2 were higher than those of the mothers, while the opposite can be seen at Time 3. Regarding the proportions of subordinate clauses, almost the same rates of mothers and teachers were found in Time 2, while a significant difference was found in Time 3. Mothers' proportions were significantly higher than those of teachers as children reached age 5;10.

6.5 Clause combining in children's output

6.5.1 Home data

In this section, children's use of complex multi-clause utterances in the home setting is examined. In Table 6.9 the number of utterances and multi-clause utterances of each individual child are presented for each of the measurement times, including the means and standard deviations.

Table 6.9: Individual scores, means and standard deviations of children's multi-clause utterances in three measurement times at home

	Coordinate			Complement			Adverbial			Relative		
Time	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3
Age child	3;2	4;2	5;10	3;2	4;2	5;10	3;2	4;2	5;10	3;2	4;2	5;10
Child												
1	1	2	1	1	0	0	0	1	1	0	1	1
2	1	0	3	0	1	2	0	0	1	0	0	1
3	0	0	1	0	0	0	0	0	0	0	0	0
4	0	2	2	0	0	3	0	0	1	1	0	1
5	0	2	0	0	1	0	0	0	0	0	1	0
6	0	1	0	0	1	0	1	0	1	0	0	0
7	1	0	1	1	1	0	0	0	0	0	0	0
8	1	1	2	0	1	1	0	3	1	0	1	0
9	1	0	1	0	0	2	0	0	0	0	0	2
10	0	0	0	0	0	1	0	0	1	0	0	0
11	0	0	1	0	0	0	0	0	0	0	0	0
12	1	0	0	0	0	0	0	0	0	0	0	0
Mean	.50	.67	1.0	.17	.42	.75	.08	.33	.50	.08	.25	.42
SD	.52	.89	.95	.39	.52	1.05	.89	.52	.29	.29	.45	.67

Examination of the figures in Table 6.9 shows that children, on the whole, produced a low number of multi-clause sentences in the three measurement times with 0 and 1 being the most common numbers. Nevertheless, there are differences among the children: child 4, for example, uses seven multi-clause utterances during Time 3, and child 11 only one. As explained in previous chapters (see Chapter 4), the results dealt with here are the outcome of data from book reading activities. In other words, no specific experimental tasks were constructed explicitly to trigger any kind of subordinate clause. Therefore, the limited use of multi-clause utterances by children does not mean that they are unable to produce such sentences. It can be seen from Table 6.9 that coordinate clauses are the most used category in all measurement times. As regards the progress over time concerning this category, we can see an increase over the three measurement times. The mean scores obtained by children ranged between .50 in Time 1 and 1.0 in Time 3. Regarding complement

clauses, we find that the children's mean scores are the highest compared to the other two subordinate clauses (relative and adverbial clauses). Likewise, an increase can be seen across the three measurement times; the children's mean score in Time 3 was about 4 times as large as in Time 1. Concerning the last two categories, relative and adverbial clauses, four children did not use them at all and only two children used them more than once. The means obtained by children show a similar trend to the other two categories. An increase of the mean scores can be seen with the increasing age of children, due to the fact that some of the children started using them. Notably, notwithstanding the decrease of children's number of utterances across the three measurement times, we see an increase of children's production of all multi-clause types.

In Table 6.10, an overview is given of different multi-clause sentences used by the children during home interaction tasks in three measurement times.

Table 6.10: Overview of multi-clause sentences used by the children in 3 measurement times at home

	Time 1 (Age 3;2 years)	Time 2 (Age 4;2 years)	Time 3 (Age 5;10 years)
Coordinate clauses	Total 6 (3 in Berber, 3 in Dutch)	Total 8 (in Dutch)	Total 12 (3 in Berber, 9 in Dutch)
Complement clauses	Total 2 (1 in Berber, 1 in Dutch)	Total 5 (2 in Berber, 3 in Dutch)	Total 9 (2 in Berber, 7 in Dutch)
Adverbial clauses	Total 1 (in Dutch)	Total 4 (in Dutch)	Total 6 (1 in Berber, 5 in Dutch)
Relative clause	Total 1 (in Dutch)	Total 3 (in Dutch)	Total 5 (3 in Berber, 2 in Dutch)

As Table 6.10 shows, most of multi-clause sentences used by the children during home interaction tasks are in Dutch. Coordinate clauses are the most used ones (26 in the three measurement times, including 20 sentences in Ducth). Examples of coordinate clauses in Berber are:

6.67 (Anissa, Time 1)
h2enna-s t-kkes-as tisira *d* jeddi-s naar@ binnen@ gaan@
granny-his 3S.F-take off.IMPERF-him shoes and grandpa-his to inside go
(his granny is taking off his shoes and his grandpa is going inside)

6.68 (Fatima, Time 3)
Emma t-arezzu ad t-kkar *u* menbeod Emma t-kkar
Emma 3S.F-want.IMPERF FUT 3S.F-stand up.AOR and afterwards Emma 3S.F-stand up.AOR
(Emma wants to stand up and afterwards she stood up)

Examples of coordinate clauses in Dutch are:

6.69 (Isra, Time 1)
dit is poppetje *en* dit is mandje
(this is a doll and this is a basket)

6.70 (Abdelilah, Time 2)
en toen ging ie ze optillen *maar* het was veel te zwaar
(and at that time he wanted to lift it but it was too heavy)

6.71 (Abdelilah, Time 3)
hier is ze bang *en* hier zit ze kopje thee te drinken
(here she is afraid and here she is drinking a cup of tea)

Complement clauses are the second category of the frequently used multi-clause sentences by the children in home interactions (16 in three measurement times, including 11 in Dutch). Examples of complement clauses in Berber are:

6.72 (Ilyas, Time 1)
ad as ini-gh i Senna arwah safi
FUT her tell.AOR-1S to Senna come.AOR it's OK
(I will tell Senna to come, it is done)

6.73 (Ilyas, Time 2) (to)
i-nna-yi baba *ad* qqim-egh x lusa n ssira
3S.M-tell.PERF-me papa-POSS CL 1S.M/F-stay on stones of sidewalk
(my father told me to stay on the stones of the pavement)

Examples of complement clauses in Dutch are:

6.74 (Saliha, Time 2)
Tim zei tegen papa ik ga *om* mijn step te halen
(Tim told his father I am going to collect my scooter)

6.75 (Isra, Time 3)
want zij denkt *dat* het echt dieren zijn
(because she thinks that they are real animals)

The third category of multi-clause sentences used by the children is the adverbial clause. Except for one Berber sentence, all other adverbial clauses are in Dutch (11 in the three measurement times). Examples in both Berber and Dutch are:

6.76 (Hamza, Time 3)
t-tt2es meskina *amacnau* baby
3S.F-sleep.IMPERF poor like baby
(the poor thing is sleeping like a baby)

6.77 (Ibtisam, Time 1)
als ik groot ben dan ga jij een hele grote kopen
(when I grow up then you are going to buy a big one)

6.78 (Isra, Time 2)
en *toen* hem vader uit het raam kijken zei de jongetje zo
(and when his father looked from the window the boy said so)

The least used type of multi-clause sentences is the relative clause. Children used 9 clauses in the three measurement times (3 in Berber and 6 in Dutch). For example:

6.79 (Anissa, Time 3)
zo@nl tamza *nni* t-ameqran-t ging@nl weg@nl
so ogress that F-big-F went away
(so the big ogre went away)

6.80 (Abdelilah, Time 3)
alle auto's *die* een taxi is hebben zo'n ding boven
(all cars which are a taxi have such a thing above)

Figure 6.1 graphically manifests the development of the four categories over three measurement times.

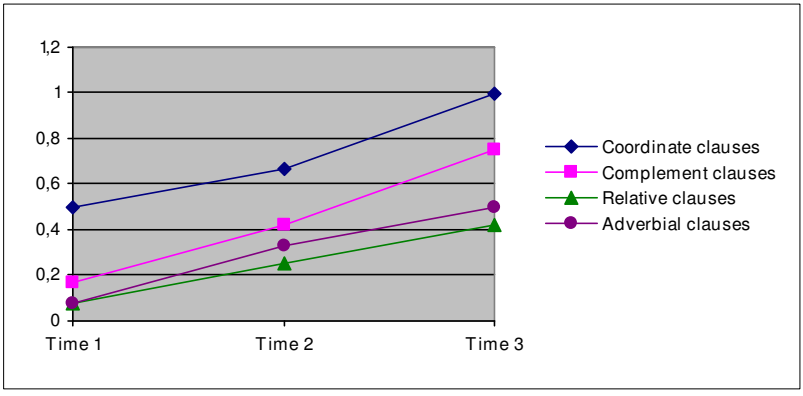

Figure 6.1: development of children's multi-clause sentences in three measurement times at home

After treating the four types of clause combinations separately, we now consider the proportions of both multi-clause and subordinate multi-clause sentences. Table 6.11 lists the descriptive statistics of the multi-clause sentences and subordinate clauses, in addition to the proportions of multi-clause and subordinate sentences relative to all utterances.

Table 6.11: Children's means and standard deviations of multi-clause utterances in three measurement times at home

N=12	Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	M	SD	M	SD	M	SD
Number of utterances	39.08	20.38	32.58	14.24	25.08	16.65
Total number of multi-clause sentences	.83	.72	1.67	2.02	2.67	2.53
Total number of subordinate multi-clause sentences	.33	.49	1.0	1.48	1.67	1.83
Proportion of multi-clause sentences over all utterances	.02	.02	.05	.06	.11	.10
Proportion of subordinate multi-clause sentences over all utterances	.01	.01	.03	.04	.08	.10

On average, the children use between .83 and 2.67 multi-clause sentences, ranging from none (several children at Time 1 and 2), to 7 (child 2 and 4, Time 3). The average use of subordinate sentences (means respectively .33, 1.0 and 1.67) ranges from none (several children) to 5 (child 8, Time 2 and child 4, Time 3).

Regarding multi-clause sentences in relation to all utterances, Table 6.11 shows that the children produce very low proportions of multi-clause utterances in measurement time 1 (2%), ranging from none (several children) to 5% (child 1 and 7). A slight increase can be seen in measurement time 2 (5%), ranging from none to 15% (child 1). In measurement time 3, as children on average reach age 5;10, the proportion was 11%, ranging from none (child 5 and 12) to 33% (child 10). Concerning the proportions of subordinate multi-clause sentences over all utterances, almost the same trend was found as in the previous category. Children produced only 1% in measurement time 1. In measurement time 2 and 3, the proportions were 3% and 8% respectively. The proportions ranged from none (several children) to .33 (child 10, Time 3).

To illustrate the growth, the values for proportions portrayed in Table 6.8 are plotted in Figure 6.2. In order to see if there is any time effect, ANOVA analyses with repeated measures were conducted.

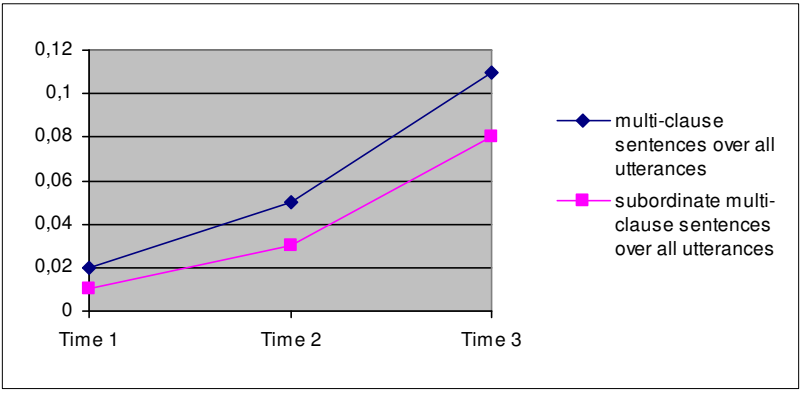

Figure 6.2: Children's proportions of (complex) multi-clause utterances over time at home

As Figure 6.2 depicts, an increase in both categories can be seen across the three measurement times. To verify this, ANOVA with repeated measures were performed to see whether children's use of multi-clause sentences and subordinate multi-clause sentences increased over time. Concerning the

proportions of multi-clause sentences, the results revealed a significant main effect of time ($F(2,22)=6.14$, $p=.02$, $\eta_p^2=.358$).

In the analysis for the proportions of subordinate multi-clause sentences, the amount of progress was substantial and a significant main effect was found for the factor time ($F(2,22)=4.19$, $p=.05$, $\eta_p^2=.361$). The repeated contrasts indicate a non-significant increase from Time 1 to Time 2 ($F(1,11)=2.10$, $p=.17$, $\eta_p^2=.161$), and a significant increase from Time 2 to Time 3 ($F(1,11)=4.07$, $p=.06$, $\eta_p^2=.270$).

6.5.2 Relationship of children's output with family SES and literacy

Another question addressed here is whether the children's use of subordination is associated with parental SES and literacy background. To answer this, a Kendall Tau correlation analysis was conducted between the background characteristics of the family and the different multi-clause types. Table 6.12 summarizes the outcomes of the correlation analysis.

Table 6.12: Kendall Tau correlations of SES and literacy with complex constructions used by the children

N=12	Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	SES	Literacy	SES	Literacy	SES	Literacy
Multi-clause sentences	.21	.43τ	.36	.45τ	.09	.14
Subordinate sentences	.25	.41	.15	.25	.20	.29
Prop. multi-clause sentences	.04	.22	.35	.51*	.08	.13
Prop. subordinate sentences	.15	.34	.15	.24	.22	.21

* $p<.05$; τ tendency

The number of multi-clause sentences and the number of subordinate clauses that children use correlate positively with both SES and the level of literacy practices at home, but there is only a tendency for multi-clauses to correlate with literacy at Time 1 and Time 2. The association of the total number of subordinate clauses with SES and literacy variables, as expected, did not reach statistical significance. As for the proportion of multi-clause sentences over all utterances, only at Time 2 were the correlations with SES and literacy rather

high and significant for literacy. The correlations of the proportion of subordinate clauses with SES and literacy are positive, but not very high.

6.5.3 School data

Analogous to the analysis conducted in the previous section on children's home data, we now present the results of the school data. Table 6.13 gives an overview of means and standard deviations of the absolute numbers of the four categories of multi-clause sentences used by children during the book task at school.

Table 6.13: Children's means and standard deviations of multi-clause utterances in two measurement times at school

N=12	Age 4;2 years		Age 5;10 years	
	M	SD	M	SD
Number of utterances	14.17	11.70	16.75	12.07
Coordinate clauses	.58	1.0	.83	1.03
Complement clauses	.25	.45	.83	1.19
Relative clauses	.25	.45	.42	.67
Adverbial clauses	.17	.39	.50	.67

Since the average number of utterances recorded of the children at school is much lower than at home, ranging from 3 (child 1 and 3) to 39 (child 10) at Time 2 and from 2 (child 3) to 43 (child 6) at Time 3, the data of the individual children are not presented here. From Table 6.13 it can be read that children, as in the home data treated above, generally produced a low number of multi-clause utterances. Coordinate clauses are the most used category in two measurement times, ranging from none (several children) to 3 (child 7 and 9) at both measurement times. With regard to subordinate clauses and considering both measurement times, complement clauses are more used compared to relative and adverbial clauses, ranging from none to 1 at Time 2 and from none to 3 (child 4 and 6) at Time 3. Adverbial and relative clauses are hardly ever used, one child (child 4 at Time 2) using 2 adverbial clauses and one child using 2 relative clauses (child 6 at Time 3). As regards progress over time, it turns out that on average children's use of all multi-clause categories increases from

Time 2 to Time 3. It is noteworthy that the standard deviations listed in Table 6.13 suggest a particularly large variation between children.

An overview of different types of multi-clause sentences used by the children in the two measurement times at school is presented in Table 6.14.

Table 6.14: Overview of multi-clause sentences used by the children in 2 measurement times at school

	Time 2 (Age 4;2 years)	Time 3 (Age 5;10 years)
Coordinate clauses	Total 7 – en (and)	Total 10 – 9 en (and) – 2 of (or)
Complement clauses	Total 3 – dat (that) – of (whether)	Total 10 – 5 dat (that); 3 wat (what); 1 wie (who); 1 die (who/which)
Adverbial clauses	Total 2 – 1 conditional: anders (otherwise) – 1 temporal: toen (when)	Total 6 – 4 temporal/sequential: terwijl (while), toen (when) – 2 reason: omdat (because) – 1 conditional: anders (otherwise)
Relative clauses	Total 3 – 1 dat (which) – 2 die (which, who)	Total 5 – 4 die (which, who) – 1 wie (who, whom)

As with the home data, the most frequently used multi-clause category at school is the coordinate clause (17 in two measurement times). The most used conjunction is "en". Examples of coordinate clauses are:

6.81 (Abdelilah, Time 3)
dit is slak *en* dat is water
(this is snail and that is water)

6.82 (Isra, Time 3)
een muis is beter *of* een andere dier
(a mouse is better or another animal)

The second type of multi-clauses is the complement clause. Children produced 13 complement sentences in two measurement times, using complementizers

such as "dat" (that), "wie" (who), "die" (who, which) and "of" (whether). For example:

6.83 (Hamza, Time 2)
hij zei *dat* het goed is
(he said that it is alright)

6.84 (Hamza, Time 3)
dan denken mensen *dat* spoken bestaan maar dat is niet waar
(then people think that ghosts exist but that is not true)

The third category of multi-clauses are the adverbials. This type was used 8 times by children in two measurement times. Most clauses are temporal, followed by conditional and reason adverbial clauses. Examples of adverbial clauses are:

6.85 (Ouiam, Time 2)
mag ik een keer aaien *anders* gaat ie schoppen geven
(may I stroke it once otherwise he is going to kick)

6.86 (Hamza, Time 3)
je bent onder water *terwijl* je ademt
(you are under water while you are breathing)

The last category is the relative clause. The total number of these sentences used by the children is 8. The pronouns which children used are "dat" (that), "die" (which, who) and "wie" (who, whom). Examples of relative clauses used by the children during school interactions are:

6.87 (Anissa, Time 2)
voor die lange dingen die drie voetjes heeft
(for those long things which have three feet)

6.88 (Abdelilah, Time 3)
hij durft niet de slak die daar is
(it dares not the snail which is there)

Having presented the outcomes of four types of clause combining categories separately, we now examine the proportions of multi-clause sentences in

relation to all utterances and the proportions of subordinate multi-clause sentences over all utterances. Table 6.15 reports the descriptive statistics for these categories during two measurement times.

Table 6.15: Descriptives (means and standard deviations) for the total numbers and proportions of subordinate and multi-clause sentences in children's output at school

N=12	Age 4;2 years		Age 5;10 years	
	M	SD	M	SD
Number of utterances	14.17	11.70	16.75	12.07
Number of multi-clause sentences	1.25	1.14	2.58	2.61
Number of subordinate multi-clause sentences	.67	.89	1.75	2.01
Proportion of multi-clause sentences over all utterances	.15	.16	.20	.20
Proportion of subordinate multi-clause sentences over all utterances	.07	.10	.10	.10

Examination of the figures in Table 6.15 shows that on average the children used about one multi-clause sentence when they were four years old, ranging from none to 3 and about two-and-a-half at measurement time 3, ranging from none to 8 (child 6). The average use of subordinate clauses was even less, with a mean of .67 at Time 2, ranging from none to 3, and 1.75 at Time 3, ranging from none to 6 (child 6). Further examination of the figures in Table 6.15 shows that 15% of all utterances produced by children in measurement time 2 contained multi-clause sentences, ranging from 0% (child 7 and 11) to 43% (child 4). In measurement time 3, the proportion reached 20%, ranging from none to 50%. As to subordinate multi-clause sentences, children produced 7% (range 0-33) in measurement time 2 and 10% in measurement time 3 (range 0-25%).

The proportions of multi-clause utterances and subordinate clauses were also considered in a longitudinal perspective. Therefore, ANOVA with repeated measures was carried out for the school data. For the proportion of multi-clause sentences, no time effect was found through an increase from Time 2 to Time 3 ($F(1,11)=.66$, $p=.43$, $\eta_p^2=.057$). It also emerged that the increase in the proportions of subordinate multi-clause sentences from measurement time 2 to 3 was not significant ($F(1,11)=.53$, $p=.48$, $\eta_p^2=.046$).

6.5.4 Children's home and school data compared

To find out whether the children's output shows a resemblance in the home and the school setting, a Wilcoxon test has been performed on the proportions of multi-clause and subordinate clauses. Table 6.16 presents an overview.

Table 6.16: Overview of the Wilcoxon test results of the proportions of multi-clause and subordinate clauses during the home and the school book tasks

		Time 2 (Age 4;2 years)		Significance level	Time 3 (Age 5;10 years)		Significance level
		M	SD		M	SD	
Multi-clauses	Home	.05	.06	Z=-2.05*	.11	.10	Z=-.87
	School	.15	.16	(p=.04)	.20	.20	(p=.39)
Subordinate clauses	Home	.03	.04	Z=-1.12	.08	.10	Z=-.65
	School	.07	.10	(p=.26)	.10	.10	(p=.51)

* p<.05

As far as the proportions of multi-clause sentences are concerned, a significant difference was found between the home and school data in measurement time 2 (p=.04). Children's proportions at school were three times as high as at home. In measurement time 3, children's rates at school were also higher than at home. This difference, however, misses significance (p=.39). As regards the proportions of subordinate clauses, it turned out that children's mean scores obtained at school are higher than at home in both measurement times, but none of them was statistically significant (Time 2, p=.26; Time 3, p=.51).

6.6 Input-output relationships

This section focuses on the relationship between the caretakers' input in terms of subordinate and multi-clause sentences and children's language production. Table 6.17 lists the outcomes of Kendalls tau correlations between the proportions of subordinate clauses of the mothers and teachers on the one hand and the children on the other.

Table 6.17: Correlations (Kendalls Tau) among mothers' and children's proportion of subordinate clauses during the book task and between teachers' and children's proportion of subordinate clauses (N=12)

Mothers' proportion of subordinate clauses	Children's proportion of subordinate clauses		
	Time 1 (3;2 years)	Time 2 (4;2 years)	Time 3 (5;10 years)
Time 1	.04	.09	-.03
Time 2		.29	-.03
Time 3			-.06
Teachers' proportion of subordinate clauses			
Time 2		-.04	.16
Time 3			-.14

Table 6.17 reveals no consistent correlations between the use of subordinate multi-clause sentences by the mothers and the children nor the teachers and the children. With regard to the home setting, in Time 1 and 2, a positive but not significant relationship was found between subordinate clause utterances in mothers' input and children's use of these clauses. When children reach the age of 5;10, no relationship could be found between mothers' input and children's output. As for the school setting, teachers' use of subordinate clauses also failed to correlate with children's use of subordinate clauses during the book task.

In addition to subordinate clause sentences, we ran correlational analyses between mothers' and teachers' use of multi-clause sentences and children's use of multi-clause sentences. The results of these analyses are displayed in Table 6.18.

Table 6.18: Correlations (Kendalls Tau) among mothers' and children's proportion of multi-clauses during the book task and between teachers' and children's proportion of multi-clauses (N=12)

Mothers' proportion of multi-clauses	Children's proportion of multi-clauses		
	Time 1 (3;2 years)	Time 2 (4;2 years)	Time 3 (5;10 years)
Time 1	.21	.28	-.02
Time 2		.12	-.20
Time 3			-.23
Teachers' proportion of multi-clauses			
Time 2		.40τ	.08
Time 3			-.24

Correlations of mothers' and children's use of multi-clause utterances revealed the same patterns as in the use of subordinate clause utterances. We found positive but not significant correlations in measurement time 1 and 2. However, in measurement time 3, even negative correlations are established between mothers' and children's use of multi-clause sentences. With respect to the school data, the outcomes in Time 2 indicated a trend towards a relationship between teachers' input and children's use of multi-clause utterances. In Time 3, results indicated no significant associations between the input of teachers and the output of children.

6.7 Discussion

In this section the findings presented above will be discussed on the basis of the earlier mentioned research questions. Keeping in mind the possible relationship between children's skills and the input they are exposed to, our first question addressed parental use of complex constructions. More specifically, we examined mothers' use of complex structures during a book reading task. It turned out that the most frequent category of subordinate clauses in mothers' speech was complement sentences across all measurement times. Further examination showed that the use of mothers' multi-clause sentences was the highest during measurement time 1. It must be noted, however, that this might

be due to the fact that mothers produced the highest number of utterances during measurement time 1. The analysis of proportions showed slight differences over the three measurement times, which lends support to the aforementioned explanation. The same patterns can be seen at the level of subordinate clause sentences. The number of subordinate sentences produced by mothers decreased over the three measurement times. In terms of proportions, almost similar patterns were found and the proportions were nearly identical over the three measurement times. This finding leads us to conclude that the frequency of syntactically complex constructions in mothers' input does not change as children get older. This contrasts with our expectation that mothers would use more complex sentence after children reach 4 and 5 years of age. At this age, children are acquiring complex syntactic constructions and need to fully and effectively develop this process by being involved in verbal interactions with adults in their environment. It should be stressed that in all measurement times, high standard deviations of mothers' scores were found. If we consider the multi-clause sentences, ranges of mothers' proportions were 5%-34%, 6%-39% and 4%-30% in Time 1, Time 2 and Time 3 respectively. Thus, no stable pattern could be found among the mothers.

We also investigated whether there is a relationship between observed mothers' input and SES and literacy backgrounds. Different studies seem to disagree with regard to the relationship between family's SES and syntactic complexity in mothers' input. Many cross-cultural studies on SES and parenting show evidence that higher SES mothers use more complex language than do lower SES mothers (Hoff, Laursen, & Tardif, 2002). In contrast, in their examination of parents' linguistic complexity, Huttenlocher *et al.* (2002) found only a marginally significant relation between SES background and the use of complex constructions. Our data show that SES and literacy background were partially related to the composite variables of mothers' syntactic complexity. Especially in measurement time 3 when children reach the age of 5, SES level and literacy practices correlated significantly with the syntactic complexity in mothers' speech.

With respect to teachers' input, the data treated above point out that the teachers in both measurement times tend to use coordinate and complement clauses more than relative and adverbial clauses. It also emerged that the proportions of (complex) multi-clause sentences decreased from Time 2 to

Time 3. This is surprising if one bears in mind that the total number of teachers' utterances in Time 3 was higher than in Time 2.

To locate differences and/or similarities between the input to which children are exposed, we compared the home and school settings. It turned out that teachers, almost at the level of all types of multi-clause sentences, produced higher numbers than the mothers. Proportionally, however, slight differences can be seen between mothers and teachers. Rather unexpectedly, mothers' proportion of subordinate clauses in measurement time 3 is significantly higher than the teachers'. Initially, we assumed that teachers' input would contain higher proportions of (complex) multi-clauses than mothers.

Next to insight into the language input provided by mothers and teachers, we looked at children's use of syntactic complex constructions. Generally speaking, our data point out that the children produced a low number of complex sentences at different ages at home as well as at school. It also emerged that, among subordinate clauses, complement clauses were the most frequently used category by children at home and at school. A plausible explanation is that complementation is the first type of complex constructions that emerges in children's speech after the acquisition of simple sentences (Bloom *et al.*, 1980). Many studies on children's spontaneous speech show that they acquire different complex constructions by age 5. Although children's morpho-syntactic knowledge is essentially complete by the age of 4 or 5 years, they do not necessarily display it in their spontaneous speech at this age. Only between the age of 6 and 10 do children develop the skill to actively produce texts and stories containing multiple and embedded events (Hoff, 2009). The analysis of children's use of (complex) multi-clause sentences at home demonstrates a statistically significant increase over the three measurement times. By the age of 5, children's utterances clearly become more complex. On the other hand, although school data show also an increase of the use of (complex) multi-clause sentences by children, the difference between results in Time 2 and Time 3 was not significant. Comparing the home and school, children tend to use more complex constructions at school than at home. This fits in with our expectations that children would be using more complex talk in school as a setting where academic language is fostered.

CHAPTER 7

The textual level of academic language: decontextualized language

While the previous chapters dealt with the development of academic language features at the lexical level (single words) and the morpho-syntactic level (sentences and complex constructions), the central theme of this chapter is the organization of the text or the discourse. The main focus is on the level of abstraction both in the input provided by mothers and teachers and in the language used by the children. We chose to consider abstraction levels since many studies stress the importance of the use of abstract language as an important bridge to literacy development and school success (Heath, 1982; Snow, 1983; Dickinson & Tabors, 1991; Reese, 1995; Snow, Burns & Griffin, 1998). In section 7.1, textual features of academic language and the operationalization of decontextualized language are addressed. Section 7.2 covers decontextualized language in the input of the mothers. Section 7.3 provides the results of decontextualized language in the input of the teachers. In section 7.4, mothers' and teachers' input are compared. In section 7.5, the results of the children in the home and the school settings are presented. Section 7.6 deals with the input-output relationship. Finally, the findings are discussed in section 7.7.

The objective of this chapter is to shed some light on the use of decontextualized language both by Moroccan Berber mothers and by teachers when talking to young children and by the children. To do so, the following research questions are addressed:

1 To what extent do Moroccan Berber mothers use decontextualized language in their input to their preschool children during a book reading activity?
2 Is the use of decontextualized language during book reading by the mothers related to SES and home literacy practices?

3 To what extent do teachers use decontextualized language in their input to Moroccan Berber preschool children while reading books?
4 To what extent do these children use decontextualized language during a book reading activity at home and at school?
5 Is there a difference between mothers and teachers regarding the use of decontextualized language features during book reading?
6 Is there a relationship between the parents' and teachers' input and the children's use of decontextualized language (output) during a book reading activity?

7.1 Textual features of academic language operationalized

Studies dealing with decontextualized discourse use different measures to assess the degree of decontextualization. Some of these studies focus on grammatical devices, which reinforce the explicitness and specificity during language conversations. These linguistic characteristics are referred to as *literate language features* (Curenton & Justice, 2004). Such devices are conjunctions, adverbs, mental and linguistic verbs and elaborated noun phrases (Greenhalgh & Strong, 2001). Each of these linguistic features is used to enable participants in a conversation to convey meaning in restricted situations with a minimum of contextual cues. For example, conjunctions that are used both for coordination (e.g., *and*, *but*, *or*) and subordination (e.g., *because*, *before*, *after*, *when*) provide additional information about connection of elements in a discourse, in addition to causal/temporal and spatial sequences. Adverbs increase the explicitness of reference to manner, time and place. Mental and linguistic verbs such as *know, think, believe* and *consider* give more information about cognitive and mental states, emotions and activities during a talk or story (Pellegrini, Galda, Bartini & Charak, 1998). Elaborated noun phrases, where other nouns are modified by adding quantifiers, articles and demonstratives increase the specificity of a certain object or event.

Other studies which examined decontextualized language and cognitively challenging talks considered utterances and texts as a whole, instead of separate words, to examine the degree of abstractness. The operationalization of decontextualized language in the current study diverges from that of Curenton *et al.* (2008) in not considering vocabulary and grammatical structures but

instead focusing on the utterance as a whole by using different abstraction levels, as will be explained in the following section. This type of decontextualized language can be defined as an "expression of thinking that is increasingly disembedded from the supportive perceptual context" (Van Kleeck, Gillam, Hamilton & McGrath, 1997: 1262). The levels of abstraction, as used in this study, range from the lowest level, which only requires one to focus on items or events available in the shared context, to the highest level, which goes beyond immediate perception.

In this section, an account of various studies dealing with abstraction as an operationalization of decontextualized language at the textual level will be presented. Subsequently, the concept of abstraction adopted in this study, with its different levels, will be explained.

Heath (1982) distinguishes two levels of abstraction in middle-class mothers' speech used in talking to their children. Starting from the lowest category "What-explanations" (e.g., what happened in the book), which appear at a very early age, followed by "reason-explanations" (e.g., why and how it happened) and "affective-commentaries" (evaluation and judgment). Sigel and McGillicuddy-Delisi (1984) analyzed parental speech by using *distancing* theory. In this model, three levels of abstraction are distinguished. First there is the lower level, which includes naming and labeling objects. The medium level concerns comparing, classifying and contrasting. The higher level involves judging, inferring, and reasoning. DeTemple & Beals (1991) used the *type of talk* approach to measure the degree of abstraction, where dyadic mother-child utterances were coded in terms of immediate versus non-immediate language.

The concept of abstract or decontextualized language in the current study is operationalized by using the "perceptual-language distance" model developed by Blank, Rose & Berlin (1978), where four levels of abstraction are distinguished. This framework has been used ever since in many studies (see also Sorsby & Martlew, 1991; Groenendaal, Goudena & Olthof, 1992; Van Kleeck *et al*., 1997; Hammet, Van Kleeck & Huberty, 2003; Massey, 2004; Nap-Kolhoff & Van Steensel, 2005). The model has four levels of abstraction resting on the assumption that at each level 'language moves increasingly away from direct perception' (Blank, Rose & Berlin, 2003: 1). Table 7.1 shows the four levels of abstractions illustrated with examples.

Table 7.1: The levels of abstraction based on Blank, Rose & Berlin (1978)

I	Matching Perception	Reporting and responding to salient information	– What do you see on the table? – Where is the fork?
II	Selective Analysis of Perception	Reporting and responding to delineated and less salient cues	– Find one that is long and yellow. – What is the boy doing?
III	Reordering Perception	Using language to restructure perceptual input and inhibit predisposing responses	– Show me the part of the egg that we don't eat – Did you go to the zoo?
IV	Reasoning Perception	Using language to predict, reflect on, and integrate ideas and relationships	– What will happen to the cookies when we put them in the oven? – What made the ball move?

In the present study, utterances of children and adults at home and at school are assigned to one of the four abstraction levels described below. For each level, examples are provided from both home sessions (Tarifit Berber and Dutch) and school sessions (Dutch) to illustrate each category.

Level I (Matching perception) is the lowest level of abstraction which focuses solely on concrete items and available information. Within this category, there is a minimal distance between language and perception. Utterances at this level of abstraction deal with visible things in the shared context and they involve *labeling*, that is to say telling the name of a person or an object, e.g., 'what is that? That is a book', *locating* an object or a character, e.g. 'where do you see a horse?', or *noticing* by directing the child's attention to find and understand things without naming them, e.g. 'look here'.

Example 7.1

Berber: (Mother of Ouiam 3;2)
MOT: mayn i-ona wa, d tunubil?
what 3S.M-be.PERF this COP car
(what's this, is it a car?)

Dutch: (Teacher of Abdelilah 4;2)
TEA: ik heb hier een boek.
(I have a book here)

Level II (Selective analysis of perception) concentrates on particular aspects of objects and items in a conversation and draws relationships between these items, in addition to integrating separate entities present in the immediate context. This requires the child to be able to *describe* characteristics of objects, such as number, size, colour, and actions that are going on (e.g., the cat is playing with the ball), *recalling information* previously talked about in a book (e.g., where did Tijn and his mother go?).

Example 7.2

Berber: (Mother with Ouiam 3;2)
MOT: aqq-em ta d ymma-s.
P.A-CL.3S.F this COP mother-his
(you see this is her mother)

Dutch: (Teacher with Anisa, 4;2)
TEA: welke kleur is de auto?
(which color is the car?)

Level III (Reordering perception) represents a major change in the relationship between language and perception. It requires processing information that is not perceptually present in the immediate context or to restructure the ideas and concepts. According to Blank, Rose & Berlin (1978: 16) "at this point the language is no longer mirroring but is now serving to control or reorganize the perception and action". This category includes defining, summarizing, comparing and contrasting, providing point of view, making inferences and judgments and recalling one's own experiences. For example, after a book reading session, a mother may say, "is Tijn happy to visit his grandmother?" or "did we ever travel by train too?".

Example 7.3

Berber: (Mother with Isra, 4;2)
MOT: lah2eqqac i-gga mayn i-nna baba-s.
because 3S.M-do.PERF what 3S.M-say.PERF father-his
(because he did what his father had said)

Dutch: (Teacher with Fatima, 4;2)
TEA: hebben jullie dat wel eens in de winkel gezien?
(have you ever seen that in the shop?)

Level IV (Reasoning about perception) is the most complex one. Utterances corresponding to this category require the child to think about what may, might, could, or would happen to materials and events. This involves prediction, problem-solving and explaining. At this level of reasoning, children need to perform mental operations on decontextualized situations and events. For example, "Why do they need to be more cautious in the street?" or "Do you think they will be on time?".

Example 7.4

Berber: (Mother with Abdelilah, 4;2)
MOT: xatar mara t-uyur-d x lusa nni
dangerous if 2S.M-walk.AOR-2M on paving stone that
(it is dangerous if you walk on that pavement)

Dutch: (Teacher with Saliha, 4;2)
TEA: en wat gaat hij met de verf doen?
(and what is he going to do with the paint?)

We used a book reading session to be able to study the degree of de-contextualization in parents' and teachers' input. Many preschool children are introduced to decontextualized language through a variety of home activities (Davidson & Snow, 1995; Snow, 1991). However, shared book reading is the most common situation for analyzing and operationalizing decontextualized language because it creates a chance for parents and children to link the information provided in the book with the child's everyday experiences, objects and events that are removed from the immediate context (Rosenquest, 2002).

In addition, reading books has been associated with early literacy and school success (Bus, Van IJzerendoorn & Pellegrini, 1995; Teale, 1984). Evidence from many studies reveals that book reading scaffolds the development of children's ability to comprehend stories; it also helps to foster a love for books and reading (Dickinson & Tabors, 2001; Curenton, Craig & Flanigan, 2008). Accordingly, the use of a book reading task in this chapter is appropriate in making a valid and a reliable comparison between the home and school situations. It should be noted that most Moroccan Berber mothers indicated that they regularly read books to their children, and that when they did, they would do so in Dutch and not in Berber. This is due to the oral nature of Berber, which has little or no tradition of book reading and literacy-related activities. The fact that mothers were asked in this study to read to their children in Berber or tell the story in Berber while most of them are used to doing so in Dutch (or not at all) may influence reading strategies and mothers' input as a whole. However, since this was the only way to elicit Berber talk and enhance the naturalness of the setting, mothers were instructed to talk to their children as they would normally do.

As explained in Chapter 4, during the home visits, three different books were used in the three measurement moments. In the first measurement time, the book "*Tijn op de fiets*" (Tijn on the bicycle) was used. In the second and the third times, the books used were "*Tim op de tegels*" (Tim on the tiles) and "*Emma in het spookhuis*" (Emma in the haunted house), respectively. In the school tasks, two different books were used in the two measurement times. The first one was '*De verrassing'* (The surprise) and the second was "*Wie is er hier het dapperst?*" (Who is the bravest here?). For abstraction levels, only task-related utterances (on-task) were coded. Procedural utterances such as *let's get started*, which serve to draw the children's attention, or any other off-task utterances were not coded for abstraction levels. Also, if the utterance was an exact repetition of the speaker's preceding utterance it was not coded. Finally, it should be noted that a number of short utterances could not be coded for abstraction level. For example: *niet dan? (right?/isn't it?), ja (yes), weet ik niet (don't know), goed zo (well done).*

7.2 Decontextualized language in mothers' input

In this section, the results of the mothers' input at the level of decontextualized language are reported. In Table 7.2, the means of the number of utterances, on-task utterances and the means of proportions of the four levels of abstraction are presented. The absolute numbers of the four abstraction levels for each individual mother, teacher and child are presented in Appendix 5.

Table 7.2: Abstraction levels in mothers' input in three measurement times during the book task

N=12	Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	M	SD	M	SD	M	SD
Utterances	105.42	39.75	79.33	19.48	62.42	23.16
% on-task utterances	77%		68%		67%	
Matching (number)	27.00	4,95	12.92	9.52	15.75	13.00
Matching (proportion)	.35	.12	.24	.15	.34	.21
Selecting (number)	36.42	6.95	33.75	10.81	18.17	7.41
Selecting (proportion)	.45	.07	.65	.18	.44	.12
Reordering (number)	12.50	11.04	4.83	5.47	5.17	5.91
Reordering (proportion)	.13	.11	.08	.07	.14	.16
Reasoning (number)	5.92	5.70	2.17	2.52	3.08	3.45
Reasoning (proportion)	.06	.06	.03	.04	.08	.10

During the first measurement time, when the children's age was 3;2 years, the mothers operated mostly at the lower levels of abstraction: matching (M=27; 35%) and selecting (M=36.42; 45%). The mean utterances at the level of reordering and reasoning were 12.50 (13%) and 5.92 (7%), respectively. It should be noted that there is more variation between the mothers at the higher levels (reordering and reasoning) than at the lower ones (matching and selecting). Furthermore, we can see that selective analysis is used most frequently by the mothers (at a ratio of 45% to the total). In the second measurement time, the general finding is that the mother's total number of utterances and on-task utterances is smaller than in Time 1. Again, mothers used the selective analysis level the most (M=33.75; 65%), proportionally even more than in Time 1, followed by matching (M=12.92; 24%). As for the higher

categories of abstraction, the averages of the mothers were 4.83 (8%) at the level of reordering and 2.17 (4%) at the level of reasoning. In the third measurement time, the mothers' strategy was essentially the same as in the first and second measurement times. Compared to other levels of abstraction, they predominantly used selective analysis (M=18.17; 44%). The strategy most frequently used after that was the matching strategy (M=15.75; 34%). The higher levels of abstraction were devices that were used least by the mothers, and the results for reordering and reasoning were 5.17 (14%) and 3.08 (8%), respectively. Note that mothers produced the lowest number of utterances in measurement time 3. It is clear that of all the levels of abstraction used by the mothers, selective analysis was the most frequent by far, in all three measurement times.

After treating each level of abstraction separately, all four categories are transformed into two variables to distinguish contextualized from decontextualized language. Contextualized language consists of the low levels of abstraction (proportions of matching level and selecting level added up), while decontextualized language contains the high levels of abstraction (proportions of reordering level and reasoning level added up). Only the proportions of the decontextualized category will be reported here. The contextualized category will not be considered as the two categories complement each other (and together make up 100%).

Table 7.3: Proportions of decontextualized language in the input of Berber mothers during the book reading task

N=12	Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	M	SD	M	SD	M	SD
Decontextualized language	.20	.16	.11	.10	.22	.26

As can be seen from Table 7.3, the mothers predominantly used contextualized language in all home visits. The proportion of utterances that have been coded as decontextualized did not exceed 25%. In measurement time 1, mothers used 20% decontextualized language, consisting mainly of the reordering level of abstraction (see Table 7.2). In measurement time 2, the proportion of decontextualized language used by the mothers decreased and constituted only 11% of the total number of coded utterances. This may be due to the fact that most

mothers did not really understand the irony of the story of the book and consequently they could not produce cognitively challenging talk. In Time 3, however, mothers used more abstract language and the composite variable of decontextualized talk was 22% of the total number of coded utterances. So there is a relative increase in the use of decontextualized language by the mothers as the children reach the age of 5;10.

The use of decontextualized language by the mothers was considered from a longitudinal perspective by means of ANOVAs with repeated measures. A comparison of the three measurement times revealed a significant main effect of time ($F(2,22)=4.66$, $p=.02$, $\eta_p^2=.298$. Additional analyses with pairwise comparisons using the Bonferroni correction indicated that the decrease of the use of decontextualized language from Time 1 to Time 2 was significant ($p=.04$). The decrease or increase in decontextualized language was not significant from Time 2 to Time 3 ($p=.12$) and from Time 1 to Time 3 ($p=1.0$). The proportions of the mothers' (de)contextualized language over three measurement times are plotted in Figure 7.1.

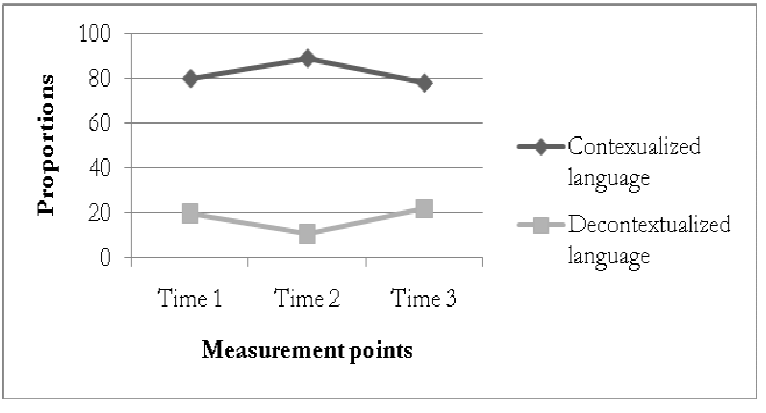

Figure 7.1: Mothers' proportions of (de)contextualized language during the book task

7.2.1 Relationship of mothers' input with SES and literacy

The second research question concerns the extent to which the mothers' use of decontextualized language is related to SES and literacy practices. These two

background variables were correlated with the mothers' scores on each of the abstraction levels. The results of these correlations are displayed in Table 7.4.

Table 7.4: Kendall Tau Correlations of SES and literacy with the levels of abstraction used by the mothers (proportions)

N=12	Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	SES	Literacy	SES	Literacy	SES	Literacy
Matching	-.83**	-.66**	.19	.24	-.83**	-.59**
Selecting	-.57*	-.43	-.32	-.34	-.19	-.27
Reordering	.89**	.66**	.56**	.58*	.88**	.81**
Reasoning	.85**	.67**	.71**	.63**	.86**	.83**

* p<.05; ** p<.01

Notably, in all three measurement times, it turns out that the higher levels of abstraction (reordering and reasoning) are significantly and positively associated with both SES and home literacy measures, suggesting that mothers with a high SES and more frequent reading behavior tend to use more decontextualized language.

More specifically, in Times 1 and 3, a significant negative correlation was found between the background variables and the matching category, which means that the higher the value of the mothers' SES and literacy, the lower the use of the matching category. Additionally, in all measurement times, the selective level did not correlate significantly with SES and literacy. During home measurement time 2, family SES and literacy background were not associated with the lower levels of abstraction. In contrast, the socio-economic status and literacy background of the mothers was found to correlate significantly with the higher levels of abstraction. The difference between measurement time 1 and 3 compared to time 2 with regard to contextualized language and its relation to SES and home literacy activities, might be explained by the nature of the books used. Apparently, the books used in Time 1 and Time 3 contained scenes and pictures which triggered more decontextualized talk than in the book used in Time 2.

In the following examples, we consider two passages from transcripts of two mothers with different SES and literacy backgrounds. Excerpt 1 is taken from a conversation between Ilyas and his mother, whose SES-level and family literacy

practices are at the lower end of the seven point scale that was used (1 and 2 respectively). Excerpt 2 is taken from a conversation between Yasmin and her mother, whose SES and literacy level are among the highest in the group (4.67 and 4.67 respectively). Both transcripts are from conversations between the mothers and their children during the book reading task in the second measurement time at home. The children's utterances have been left out since the focus here is on the mothers.

Excerpt 7.1: (Ilyas's mother)

MOT: y-qqim.
3S.M-sit.AOR
(he is sitting)

[…]

MOT: usi-n-d yinni i-oeddr-en sira nni nna-n-as min da t-ttgg-ed ?
Come.PERF-PL-directional particle those 3P.M-repair.IMPERF-PL pavement that 3P.M-tell.PERF-PL what here do.IMPERF-2M
(those who repair the pavement came and asked him: what are you doing here?)

MOT: y-enna-sen qa qimm-egh x lusa.
3S.M-tell.PERF-them T.A-sit.PERF-1S on pavement
(he told them I am sitting on the pavement)

MOT: y-enna-yi baba qim da x lusa.
3S.M-tell.PERF-CL-me my father sit-AOR here on paving stone
(my father told me to sit on the paving stone)

MOT: haqqa ksi-n-t di kamyun.
P.A take.PERF-3P.M-him in truck
(there they took him in the truck)

MOT: haqqa baba-s i-xezzar day-es.
P.A father-his 3S.M-look.IMPERF in-him
(there his father is looking at him)

MOT: haqqa ksi-n-t.
P.A take.PERF-3P.M-him
(there they took him)

MOT: haqa-t.
PA-he
(there is he)

In Excerpt 7.1, the mother is focusing mainly on items that are immediately available. She calls the child's attention to events in the book by describing activities going on (selecting), and labeling things that are visible in the direct context (matching), in addition to locating persons and objects involved in the activities. The mother repeats the adverb or presentative auxiliary "*haqqa*" (there) four times to direct the child's attention to what is going on.

By contrast, excerpt 7.2 illustrates the high amount of decontextualized language used by Yasmin's mother. In this episode, the mother frequently uses the higher levels of abstraction, reordering and reasoning. An example of reordering is when the mother comments on one of the characters of the story and says that he is obeying his father. In another episode where reasoning is used, the mother not only uses the term "captain" but explains the notion as well to make the story understandable for the child. These patterns are realised by referring to factual knowledge and giving general information to help her child understand the story. In another utterance, the mother explains the causes of certain events.

Excerpt 7.2: (Yasmin's mother)

MOT: y-nna-s waxa mani d-ac y-nna baba-c qim.
3S.M-tell.PERF-CL.him ok since to-you 3S.M-tell.PERF father-your sit-AOR
(he told him OK since your father told you to sit there)

[…]

MOT: baba-s meskin ghar i-ttazzer.
Father-his poor just 3S.M-run.IMPERF
(his poor father is running and running)

MOT: ghar barra dini kapitein@nl.
to outside in there captain
(outside there is the captain)

MOT: kapitein@nl qqar-en-t i wenni i-neddh-en agharrabu.
captain call.IMPERF to the one 3S.M-drive.IMPERF ship
(the captain is the one who is steering a ship)

MOT: y-gga array i baba-s.
3S.M-do.PERF opinion for father-his
(he obeyed his father)

MOT: y-arzu ad t-id i-fekk zi agharabu yin.
3S.M-want.IMPERF FUT it directional particle 3S.M-save.AOR from ship that
(he wants to save him from this ship)

MOT: umi war t-qqim-d x sira haqa xxx xatar.
since NEG 2S.M-stay-PERF on pavement T.A xxx dangerous
(since you did not stay on the pavement, you see xxx it is dangerous)
MOT: min d-as y-enna baba-s uh2armuc a?
what COP-him 3S.M-tell.PERF father-his child this
(what did this child's father say?)

The above passages were taken from complete transcripts of two mothers. The results of the two mothers at all abstraction levels are shown in Table 7.5.

Table 7.5: Distribution of abstraction levels for the mothers of Ilyas and Yasmin

	Ilyas' Mother (SES=1; Literacy=2)	Yasmin's mother (SES=4.67; Literacy=4.67)
Matching	34 (53%)	9 (16%)
Selecting	29 (45%)	33 (57%)
Reordering	1 (2%)	11 (23%)
Reasoning	0 (0%)	5 (9%)

Ilyas's mother, with the lower SES and home literacy practices, uses mainly matching and selective levels and almost none of the higher levels of abstraction. By contrast, Yasmin's mother uses lower as well as higher levels of abstraction.

7.3 Decontextualized language in teachers' input

In this section, the results of the teachers on the abstraction levels will be presented. The distribution of the four abstraction levels are reported in Table 7.6.

Table 7.6: Abstraction levels in the teachers' input at two measurement times

N=12	Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	M	SD	M	SD
Number of utterances	106.8	32.2	128.3	27.2
% on-task utterances	79.5%		76.8%	
Matching (number)	16.83	9.74	19.67	4.21
Matching (proportion)	.19	.06	.20	.03
Selecting (number)	33.67	10.66	39.17	4.69
Selecting (proportion)	.39	.06	.40	.03
Reordering (number)	20.42	1.68	24.00	3.46
Reordering (proportion)	.25	.05	.24	.04
Reasoning (number)	14.00	2.92	15.75	2.96
Reasoning (proportion)	.17	.05	.16	.03

In both school measurement times, the teachers mainly used the selective level of abstraction (39% and 40% at measurement times 2 and 3 respectively), i.e., talking about objects and events without going beyond what is perceptually available. The second category used most by the teachers in both measurement times was reordering. As can be seen from Table 7.6, on average 25% of the utterances at measurement time 2 and 24% at measurement time 3 were coded as reordering. Teachers tend to go beyond the immediate perception and their utterances include elements such as summarizing, identifying similarities and differences between things, defining concepts and providing a point of view. Matching, the lowest level of abstraction, was used by the teachers during both times, 19% and 20% respectively. Finally, the fourth level of abstraction, reasoning, was the least used by the teachers (17% and 16% respectively in the second and third measurement times). It should be noted that the results on the abstraction levels revealed less variation than among the mothers, as is illustrated by the low standard deviations.

After considering every level of abstraction used by the teachers in turn, a distinction is made between contextualized and decontextualized language in teachers' utterances. Table 7.7 displays the proportions of decontextualized language of the teachers in the two measurement rounds.

Table 7.7: Proportions (means and standard deviations) of decontextualized language in the input of teachers during the book reading task

N=12	Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	M	SD	M	SD
Decontextualized language	.42	.09	.40	.05

On average, 42% of the utterances in Time 2 and 40% in Time 3 were characterized as decontextualized language (reordering and reasoning). Testing for the time effect, repeated measures ANOVA shows no significant difference in the teachers' use of abstract language between Time 2 and Time 3: ($F(1,11)=.606$, $p=.45$, $\eta_p^2=.05$).

7.4 Input in two settings: mothers and teachers compared

To get more insight into the input that children receive in home and school contexts, we compared the results regarding decontextualized language use of the mothers with those of the teachers. To this end, a Mann-Whitney U test was conducted using the results of the mothers and the teachers on abstraction levels during the book reading tasks. We consider only two of the measurement times because the first home visit took place before the children entered school. Table 7.8 displays the means and standard deviations of the proportions of the various abstraction levels employed by the mothers and the teachers, in addition to the significance of differences between mothers and teachers at each level of abstraction.

Table 7.8: Overview of the Mann-Whitney U-test results on the proportions of different abstraction levels employed by the mothers and the teachers in two measurement times

N=12		Time 2 (Age 4;2 years)		Mann-Whitney U (Exact)	Time 3 (Age 5;10 years)		Mann-Whitney U (Exact)
		M	SD		M	SD	
Matching	Mother	.24	.15	U=68.0	.34	.21	U=39.0τ
	Teacher	.19	.06	(p=.84)	.20	.03	(p=.06)
Selecting	Mother	.65	.18	U=7.0**	.44	.12	U=49.0
	Teacher	.39	.06	(p=.000)	.40	.03	(p=.20)
Reordering	Mother	.08	.07	U=4**	.14	.16	U=43.0τ
	Teacher	.25	.05	(p=.000)	.24	.04	(p=.096)
Reasoning	Mother	.03	.04	U=0.0**	.08	.10	U=33.0*
	Teacher	.17	.05	(p=.000)	.16	.03	(p=.02)
Decontex-tualized	Mother	.11	.10	U=0.0**	.22	.25	U=39τ
	Teacher	.42	.09	(p=.000)	.40	.05	(p=.057)

* p<.05; ** p<.01; τ tendency

During the second measurement time, with the exception of the matching level, the Mann-Whitney U test reveals significant differences between mothers and teachers on all the other levels of abstraction during book sharing conversations with children. It turns out that mothers used significantly more utterances belonging to the selective category than the teachers did. As for the higher levels of abstraction, apparent in decontextualized language, teachers in both measurement times tended to make use of non-immediate talk more often than the mothers did. With regard to reordering as well as reasoning, a significant difference was found between the mothers and the teachers. On the reordering level, only 8% of the mothers' total number of utterances was found to be of this type of abstraction, whereas the teachers used it in 25% of the utterances produced. An almost similar tendency can be seen at the level of reasoning. Only 3% of the utterances of the mothers were at the level of reasoning, while the teachers produced 17% of their utterances at the reasoning level of abstraction.

In the third measurement time, as Table 7.8 displays, except for the selective level, large differences were found between mothers and teachers in terms of abstract language use. Selecting analysis was the most frequent category employed by both groups (44% for the mothers and 40% for the teachers). A

trend can be observed between mothers and teachers with regard to the matching level of abstraction. On average, only 20% of the teachers' utterances were coded as such while the mothers' average was 34%. With respect to the higher levels of abstraction, reordering and reasoning, teachers produced considerably more utterances at these levels than did the mothers. Only 14% of the mothers' utterances can be categorized as being of the reordering type while the teachers produced 24%. Likewise, 8% of the mothers' and 16% of the teachers' utterances were coded as reasoning. These differences are significant for reasoning and show a trend for reordering.

After considering all levels of abstraction and in order to get a clear picture of the language use of mothers and teachers and the development over time, we computed the proportions of the two variables distinguished (contextualized and decontextualized language), which are derived from the four levels, as was explained earlier in this chapter. Figure 7.2 illustrates the differences between mothers and teachers.

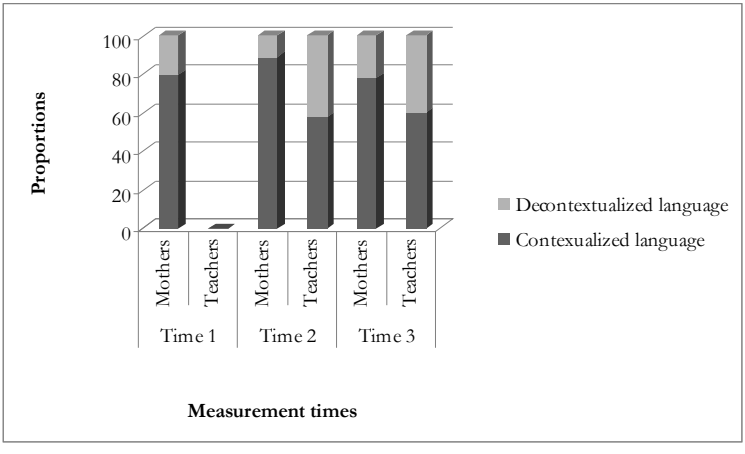

Figure 7.2: Mothers' and teachers' proportions of (de)contextualized language during book reading tasks

As can be seen from the graph, the mothers' conversations are predominantly characterized by contextualized language in both measurement times. Teachers, on the other hand, use more balanced proportions of both types of language as they talk to children. It is worth mentioning that both mothers and teachers do

not show remarkable changes in their use of abstract language over time as children grew older.

In the following passages, we consider the home and school input in the light of two excerpts (7.3 and 7.4) taken from two conversations of the same child, Saliha, talking with her mother and her teacher. Both passages resulted from the book task during the second measurement time. Since we are dealing with language input, in both transcripts, the child's utterances are excluded from the examples below.

Excerpt 7.3: (Mother of Saliha; age 4;2)

*MOT: aqqem y-qqim x lusa.
P.A-2P.F 3S.M-sit.PERF on paving stones
(there he is sitting on the paving stones)

[…]

*MOT: qqar-n as iwa kkar ss-enni!
tell-IMPERF-3P.M him stand up.AOR from-there
(they tell him to stand up)
*MOT: y-ug ad y-kkar
3S.M-refuse.PERF FUT 3S.M-stand up.AOR
(he refuses to stand up)
*MOT: ksi-n t.
take.AOR-3P.M him
(they took him)
*MOT: i-rezzu ad i-rah2 ghar mmi-s.
3S.M-want.IMPERF FUT 3S.M-go.AOR to son-his
(he wants to go to his son)
*MOT: aqqa-t di agharrabu.
P.A-he in boat
(there is he in the boat)
*MOT: gg-in-as optillen@nl.
do.PERF-3P.M-him lift up
(they lifted him up)
*MOT: i da?
and here
(and here?)

*MOT: i wa aqq-et mayn i-tt-eg?
and this T.A-he what 3S.M-do.IMPERF
(and this one what is he doing?)

Excerpt 7.4: (Teacher of Saliha; age 4;2)

*TEA: die gaat kijken hoeveel die weegt op de weegschaal.
(he is going to see how much he weighs on the scale)

[…]

*TEA: en misschien denkt het schaapje wel…
(may be the sheep is thinking…)
*TEA: kijk nou eens hoe dik ik ben!
(look how fat I am!)
*TEA: doe eens een dikke buik.
(try to make a big belly)
*TEA: wat gaat hij doen?
(what is he going to do?)
*TEA: en wat gaat hij met de verf doen?
(and what is he going to do with the paint?)
*TEA: waarom vindt hij zichzelf niet mooi?
(why does he consider himself not pretty?)
*TEA: Saliha, wat voor geluidje maakt die?
(Saliha, what kind of sound does he make?)
*TEA: en dan gaat hij alle wol afscheren en dan is hij helemaal bloot.
(and then he shaves all the wool off and then he is completely naked)
*TEA: heb jij ook wel eens een verrassing gehad?
(have you ever had a surprise like that?)

The talk produced by the mother, as can be seen from excerpt 7.3, focuses on items that can be seen directly in the book. She describes what happens, points to objects and every now and then uses where-questions and yes-no questions that do not go beyond the immediate context. In contrast to the mother, the talk produced by Saliha's teacher (excerpt 7.4) goes beyond the information contained in the book. Using the story, the teacher tries to get children to make predictions, draw inferences and analyze events. In addition, a connection is made with the real world and the child's own experience when the teacher asks the child if she ever had a surprising experience.

7.5 Children's decontextualized language

7.5.1 Home data

Having dealt with the input of the mothers and the teachers, we now present the output of the children. In Table 7.9 children's data in the home environment are presented for three measurement times. It should be noted that children produced most utterances in Time 1 (M=39.1), followed by Time 2 (M=32.6) and Time 3 (M=25.1).

Table 7.9: Descriptive statistics for children's abstraction categories during the book reading task at home

N=12	Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	M	SD	M	SD	M	SD
Utterances	39.08	20.38	32.58	14.24	25.08	16.65
% on-task utterances	78%		75%		81%	
Matching (number)	11.83	2.25	9.50	4.01	7.25	2.14
Matching (proportion)	.39	.09	.39	.17	.37	.11
Selecting (number)	13.67	4.66	9.33	3.68	7.58	2.94
Selecting (proportion)	.44	.10	.40	.16	.39	.14
Reordering (number)	3.50	2.20	3.67	3.85	3.17	3.71
Reordering (proportion)	.11	.07	.14	.13	.13	.14
Reasoning (number)	1.75	1.14	2.00	1.76	2.42	1.83
Reasoning (proportion)	.06	.04	.07	.06	.11	.06

When we look at the data across the three measurement times, we see that the children's (de)contextualized talk remained more or less constant. In the first measurement time, the lower levels of abstraction are the most frequently used ones by children during their talks with the mothers: 39% of the utterances were coded as matching and 44% as selecting. The higher levels, reordering and reasoning, were the least frequent categories. A somewhat similar pattern emerges for the second measurement time. At this stage, the children are at school age (4;2) and they show almost the same results as they do in measurement time 1. One difference is that there is a slight increase in the use of reordering and reasoning categories. During the last measurement time at

age 5;10, children still tend to use, for the most part, the basic levels of abstraction, with a small increase at the level of reasoning.

In order to get a clear picture of the children's use of (de)contextualized language over time during mother-child interactions, the proportions of contextualized and decontextualized language are outlined in Figure 7.3 As can be seen, in most of the utterances produced in the three measurement times at different ages, the children used contextualized language that is not abstracted from the available here-and-now. A slight decrease in the use of contextualized language can be observed as the children get older. On the other hand, the proportions of decontextualized language used by children in all measurement times nowhere exceed 24%. A tiny increase in decontextualized language can be observed between measurement times 1, 2 and 3. However, no significant effect of time could be established. Since Mauchly's test of sphericity was significant (W=.26, Chi-square=13.49, p=.001), the Greenhouse Geisser test statistic was used (F(1.149; 12,640)=1.84, p=.20, η_p^2=.143).

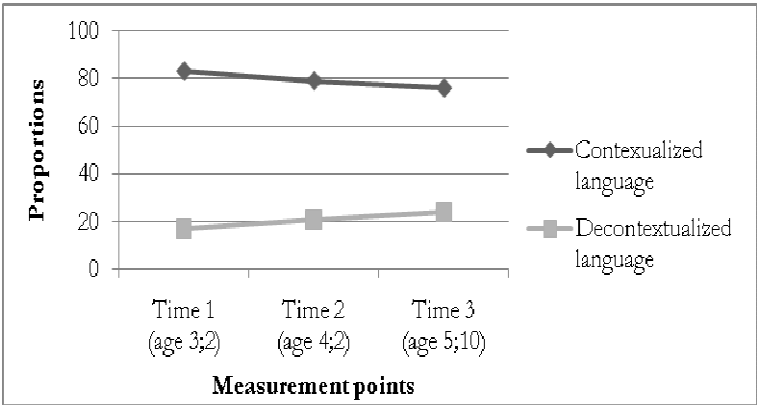

Figure 7.3: Children's proportions of (de)contextualized language during book reading tasks at home

7.5.2 Relationship of children's output with family SES and literacy

Another research question of interest was whether the socio-economic background and literacy activities in the family were important in influencing the type of abstract language used by children. Table 7.10 presents a summary

of the results of the correlation analysis conducted on the children's results at four different levels of abstraction in relation to SES and literacy activities.

Table 7.10: Kendall tau correlations of SES and literacy with the levels of abstraction used by children (proportions)

N=12	Time 1 (Age 3;2 years)		Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	SES	Literacy	SES	Literacy	SES	Literacy
Matching	-.60*	-.50*	-.41τ	-.18	-.53*	-.42τ
Selecting	-.19	-.18	-.40τ	-.48*	-.50*	-.39τ
Reordering	.59**	.55*	.83**	.59**	.79**	.62**
Reasoning	.72**	.61**	.81**	.57*	.59**	.35

* p<.05; ** p<.01; τ tendency

As can be seen from Table 7.10, the level of matching is strongly and negatively related to both SES and literacy in measurement time 1 and 3. That is, children with a high SES and home literacy background produced fewer utterances coded as matching. In measurement times 2, the correlation of matching with literacy is non-significant. As for selective analysis, an examination of correlations between this variable and the family background variables revealed no positive associations over three measurement times. However, in the second and third measurement times, a negative, partially significant, correlation was found between the selective level and SES and literacy. As for the higher levels of abstraction, reordering and reasoning, these were strongly and significantly associated with SES as well as home literacy activities. In all measurement times, the families' literacy practices that the children grow up with at home and their socio-economic background are positively related to their performances at higher levels of abstraction.

7.5.3 School data

In this section we explore the use of abstract language by children in the school setting during a book reading task in two measurement times. Descriptive statistics for abstraction level distributions are reported in Table 7.11.

Table 7.11: Children's means and proportions of four levels of abstraction during a book reading task in school in two measurement times

N=12	Time 2 (Age 4;2 years)		Time 3 (Age 5;10 years)	
	M	SD	M	SD
Utterances	14.17	11.70	16.75	12.07
% on-task utterances	77%		78%	
Matching (number)	3.50	1.09	4.75	1.06
Matching (proportion)	.35	.12	.33	.06
Selecting (number)	4.17	2.48	5	1.41
Selecting (proportion)	.35	.12	.35	.09
Reordering (number)	2.08	.67	2.92	.90
Reordering (proportion)	.20	.05	.20	.06
Reasoning (number)	1.17	.83	1.67	1.07
Reasoning (proportion)	.10	.08	.12	.07

With regard to the first level of abstraction, 35% of children's utterances were coded for matching level in Time 2 (age 4;2) and 33% in Time 3 (age 5;10). The selective level of abstraction was the most frequent category used in both school measurement times: 35% of the utterances produced were coded as selective level during the second and third school visits. For the third level, reordering, almost the same results were found in both measurement times and the proportions were 20% in both measurement times. The highest level of abstraction, reasoning, was used in 10% of the children's utterances at age 4;2 (Time 2) and 12% at age 5;10 (Time 3). As in the home data, it turned out that utterances characterized as reasoning were the least frequent. It is noteworthy that matching and selecting, the two composites of contextualized language, were the most frequent categories by far. Yet, the proportion of level I (matching) utterances slightly decreased if we compare Time 2 and Time 3. By contrast, the frequency of the highest level of abstraction, reasoning, which comprises the decontextualized language, increased slightly over time. In interpreting these outcomes, one should bear in mind that children produced only a small number of utterances at school.

7.5.4 Children's home and school data compared

In order to compare the output of children in the home and school settings, we considered the proportions of contextualized and decontextualized language over two measurement moments. A graphic presentation of the results is displayed in Figure 7.4.

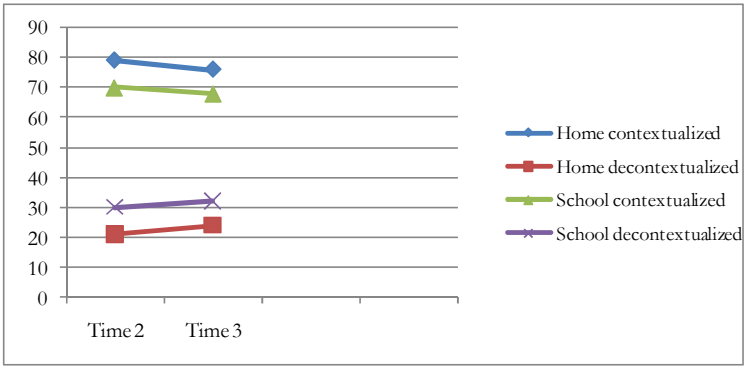

Figure 7.4: Children's (de)contextualized language during book reading tasks at home and at school

In measurement time 2, children used relatively more contextualized language at home (79%) than at school (70%) and more decontextualized language at school (30%) than at home (21%). A similar pattern can be seen at measurement time 3, when children reach the age of 5;10. They used decontextualized language at school more than they did in the home environment, while contextualized language was more frequent at home than at school. The general tendency, as reported above, is for the use of contextualized language in the home and school contexts to decrease over time. Conversely, an improvement can be seen at the level of decontextualized language use as children grow older. At the age of 5;10, children get involved in more abstract conversations which are not related to the perceivable world.

7.6 Input-output relationships

An examination of correlations between the components of decontextualized talk revealed strong and positive relations between the language produced by the mothers and that of the children in all measurement times (see Table 7.12).

Table 7.12: Correlations (Kendalls tau) between mothers' and children's decontextualized language during the picture book reading and between teachers' and children's decontextualized language (N=12)

Mothers' decontextualized language	Children's decontextualized language		
	Time 1 (3;2 years)	Time 2 (4;2 years)	Time 3 (5;10 years)
Time 1	.66**	.73**	.68**
Time 2		.44*	.42τ
Time 3			.68**
Teachers' decontextualized language			
Time 2		.19	.15
Time 3			-.14

* p<.05; ** p<.01; τ tendency

It seems that mothers who involve their children in conversations and talks that are not related to the here-and-now have children who also use significantly more non-immediate language. In the school setting, no significant relationship was found between the input of teachers and the language used by children.

7.7 Discussion

In this section, the findings presented above will be discussed in the light of the research questions and the hypotheses formulated at the beginning of this chapter. So far, we have examined the use of decontextualized language by mothers, teachers and children, as one component of the textual level. We analyzed the input of the mothers in terms of decontextualized speech and examined how children were exposed to language use on different levels of abstraction in the home setting. In addition, the effect of the mothers' SES and

home literacy practices on their use of decontextualized language was examined. In the school context, the input of the teachers was studied. We also compared mothers' and teachers' input with regard to the level of abstraction. Lastly, the children's data in both home and school environment were analysed.

Regarding the mothers' use of abstraction levels, the results of all measurement times revealed that on average mothers mainly used the selective analysis and matching levels of abstraction, which consist of descriptive labeling and talking about actual information and present entities. These results receive confirmation from the findings reported in DeTemple and Beals (1991). This study on low-educated and economically disadvantaged mothers showed that less than 20% of the mothers' talk during shared reading interactions was non-immediate, while more than 80% of their utterances were about immediately available information or irrelevant. In our study, in measurement time 3, when children were at age 5;10, mothers showed slight, though not significant increases in the use of abstraction levels of reordering and reasoning. These results suggest that parental input changes as children grow up. At this age, children receive somewhat more challenging input and are engaged in more abstract talks. These findings are more or less consistent with a longitudinal study carried out on maternal book-sharing (Van Kleeck *et al.*, 1998). This investigation showed a significant increase in the parents' use of higher levels of abstractions as children grew from 2 to 4 years old. Thus, as they get older, children are exposed to language with a higher level of abstraction which is cognitively more demanding.

Furthermore, we found a large amount of variation among the mothers in the number of reordering and reasoning utterances they used during their conversations with their children. As can be seen in Table 7.2, standard deviations of reordering and reasoning variables are almost as large as, and in some cases larger than, the means. These findings are in line with prior research that showed variation in parental language input. Hart and Risley (1995) reported a large amount of variation in the input of low-SES mothers to their preschool children. In another study, Haden *et al.* (1996) found that middle-class parents vary in the amount of abstract language they use during book reading sessions with their children.

Another important finding in this study is the strong relationship between the mothers' use of decontextualized language and their SES and literacy backgrounds. The two components of decontextualized language (reordering

and reasoning) correlated positively with SES and literacy, which suggest that there is a significant relationship between the mothers' socio-economic background and their input. These results confirm earlier studies about low-income mothers and their decontextualized language skills by Curenton, Craig and Flanigan (2008). In their study where the use of decontextualized discourse among mothers from various socio-economic backgrounds is examined, they found that mothers with higher literacy levels were more likely to use some features of decontextualized language.

As for the teachers, the results show that the most frequent level of abstraction is selective analysis, which constitutes contextualized language (39% in Time 2 and 40% in Time 3). Yet, the lowest level of abstraction, matching, is less frequent and the proportions of this category did not exceed 20% in either measurement times. In addition, teachers also tend to engage children in cognitively more challenging talks by using reordering (25% in Time 2 and 24% in Time 3). The highest level of abstraction is the least frequent category but it is still used a fair amount by the teachers. The proportions were 17% and 16% respectively in Time 2 and Time 3. If we consider the composite variables of (de)contextualized language, it turns out that nearly 40% of the teachers' utterances were characterized as decontextualized speech in both measurement times. These results replicate previous findings on conversations between children and their teachers. Dickinson, McCabe and Anastasopoulos (2003) note that during joint book reading sessions, a great deal of teachers' speech is about the organization of the task, asking simple questions and labelling and locating objects. By contrast, only a small amount of teachers' talk involved children in cognitively challenging language.

Considering the importance of home and classroom settings for fostering decontextualized language, a comparison between the mothers' and the teachers' input was made. It has been shown in the previous sections that mothers and teachers differ mainly at the higher levels of abstraction in both measurement times. Mothers principally used context-bound utterances at the levels of matching and selecting. They made minimal use of decontextualized language strategies. These findings lead us to conclude that the language input to which children are exposed in the home setting differs substantially from that of the school setting. The school environment promotes and provides more decontextualized language than does the home environment of bilingual Berber children.

The analysis of children's data at home showed that the most frequently used categories are matching and selective analyses, which are the lowest abstraction levels. Proportionally, in all measurement times, some 80% of the children's utterances were coded as matching or selecting. The distribution of the four abstraction levels shows that the selective category is the most frequently used, followed by the matching category. There was an interesting slight increase of the use of the reordering-level and reasoning-level utterances in measurement time 3 as children reach the age of 5;10. Children produced relatively more decontextualized utterances in Time 3 than they did in Times 1 and 2. Another interesting finding was the correlation of the children's use of decontextualized language with the family's socio-economic status and literacy experiences in the home. The results of all three measurement times provide evidence that higher SES children tend to use non-immediate language more frequently than children with lower SES. This is also true for children who have parents that engage their children in literacy activities more often. Children from high-SES households and whose families have high rates of home literacy practices have more access to academic language through, for example, maternal input, books and other educational activities.

At the level of school data, the distribution of the four abstraction levels is rather different from the home data. While the children's utterances during home tasks were dominated by the matching and selective analysis levels of abstraction, the four levels in school data were somewhat more fairly distributed. In the school setting, children tended to produce decontextualized speech relatively more frequently than they did at home.

CHAPTER 8

Conclusions and discussion

This dissertation deals with the emergence and development of academic language of Moroccan Berber bilingual children in the Netherlands. It explores how these children are exposed to academic language features in communicative contexts in both home and school settings. Using a longitudinal design, the language input of mothers and teachers as well as the language production of children was examined across a period of three years. By means of three empirical studies, three different features, considered as characteristics of academic language, were explored: the lexical, morpho-syntactic and textual level of academic language. This final chapter reflects upon the main findings by giving answers to the research questions formulated in chapter 4. The chapter consists of four sections. In section 8.1, the input of the mothers and the academic language at home is discussed. The input of the teachers and the academic language at school is treated in section 8.2. Section 8.3 deals with the use of academic language by children in both settings. Section 8.4 treats the relationships between input and output. In the last section, 8.5, the implications of this study, future directions for research and limitations of the study will be discussed.

8.1 Academic language input at home

The home setting is the first place where children are expected to encounter the type of language that prepares them for a school career. In the three empirical studies presented in this dissertation, we examined academic language features in the input of the mothers in a period starting when the children were 3, before they entered school, up to the age of 6. The mothers mainly used Tarifit Berber with more or less code-switching to Dutch. The results presented in the above chapters indicate that there were occurrences of academic language in

the mothers' input. However, many differences were found between the mothers with respect to the three domains under study (lexical, morpho-syntactic and textual). Other variables were the age of the children, the socioeconomic status and literacy background of the families.

In the first empirical study dealing with the lexical features of academic language, we were able to compare the data resulting from two different interaction tasks, namely a book task and a picture task. Analyses of these data revealed that the mothers talked more and used lexically richer language during the book task than during the picture task. Moroccan Berber mothers, notwithstanding their lack of book-reading experience in their mother language, turned out to produce a dense and varied lexicon in Tarifit Berber. In contrast, in her study on monolingual Dutch children of the same age as the children in our study, Henrichs (2010) reported that mothers used higher lexical density and lexical diversity during the spontaneous activity of mealtime conversation, than during the structured book reading task. A plausible explanation for the differences between the book task and the picture task is that book reading for the Berber mothers was the only activity associated with literacy, which might have triggered them to talk more and use richer vocabulary.

Considering the production of academic language features, the most significant finding at the lexical level was the increase of most variables from Time 1 to Time 2 and a decrease from Time 2 to Time 3. In the period prior to kindergarten, mothers showed a reasonable and stable increase until children reached school age (4 years old). However, after the children started school, mothers produced on average shorter utterances, used less word tokens and types and used a less rich and varied vocabulary. This pattern was found during both home interaction tasks. A factor that may have caused this decrease is the age of children. As the children reached the age of 5, they played a more active role in the conversation and consequently influenced the language behavior of the mothers. Interestingly, mothers not only did talk less after Time 2, they also used less content words, which are related to lexical density and less word types, which are related to lexical diversity. These outcomes, however, should be interpreted with caution as the group of the mothers reported on here showed considerable variability on the different variables investigated in this study.

In our study, we also focused on the link between SES, literacy practices and lexical academic language features in the mothers' input. In line with our

expectations, the most significant correlations between linguistic variables and family backgrounds were found in the book task. The number of tokens produced by the mothers correlated significantly with SES and literacy practices only in Time 3. With regard to the quality measures, lexical diversity showed the strongest correlation with both SES and literacy in both tasks. Lexical density, on the other hand, correlated with SES and literacy only during the book task. These findings suggest that high-SES and high-literate Berber mothers use more different words than low-SES mothers irrespective of situational context. They also use more content words than low-SES mothers when reading a book to their child, while there is less difference when talking about a picture. The outcomes regarding the SES effect are consistent with Hart and Risley (1995) and Hoff (2003) who found that the quality and quantity of the language used by mothers to their children is associated with SES. These studies showed that mothers with low SES used less language, and when they talked they used a less rich lexicon.

In the second study we considered syntactic complexity as another indicator of academic language. The data collected in the home as well as in the school setting were based on the book task. Our results showed that the type of multi-clause sentence most used by the mothers is the complement clause irrespective of children's age. It is important to note that on the basis of our data, we can only speculate about why mothers do hardly use other forms of complex sentences. Obviously, further investigation is required, in which different tasks and a larger sample than ours are used, to arrive at a clear explanation of the ways mothers use different syntactic structures in conversations with their children at different ages. In considering the relationship between the mothers' use of complex sentences and their family background, we could not find consistent relationships attributable to SES and literacy in all measurement moments. Only when the children reach the age of 5, positive and significant correlations could be established between complex speech in mothers and their SES and literacy. This may be because mothers tend to simplify their use of clause-combining strategies to the children's level when they are younger than 4 years old.

Referring to the mothers' use of abstract language, the results demonstrate that the mothers show almost similar patterns in all measurement times. They mainly use the lowest levels of abstraction, the selective analysis and matching level, which are typically used in talking about objects present in the immediate

context. A slight difference appears in measurement time 3 where mothers begin to use more utterances which belong to the reordering and reasoning categories. This finding suggests that mothers do not use much decontextualized and cognitively demanding language with their children during the first five years of their lives. Another interesting finding is that the mothers' use of decontextualized language was found to be significantly associated with their SES and literacy backgrounds. Mothers with higher SES and literacy levels were more likely to use non-immediate and challenging talk with their (pre)school-age children than mothers with lower SES and literacy levels during the book task. These results are consistent with earlier studies that investigated the relationship between socioeconomic status and literacy and mothers' use of decontextualized language (DeTemple & Beals, 1991; Curenton, Craig & Flanigan, 2008).

When we consider the three aspects of academic language and their existence in the mothers' language use, we note that there is a considerable variation between the mothers. To illustrate, we consider the averages of all mothers on four aspects and the averages of two mothers who, respectively had high and low averages on most academic language features. For each aspect, we looked at Time 1 and Time 3 (for the lexical level we calculated the average of the two tasks) as shown in Table 8.1.

Table 8.1: Variation among mothers in the use of different components of academic register

	Lexical density		Lexical diversity		Proportion of subordinate clauses		Proportion of decontextualized language	
	T 1	T 3	T 1	T 3	T 1	T 3	T 1	T 3
All mothers	1.92	1.99	72.69	68.49	.14	.14	.20	.22
Mother of Ibtisam	1.92	2.52	67.25	87.88	.11	.24	.38	.49
Mother of Fatima	1.11	1.57	45.84	46.60	.08	.11	.08	.00

As can be seen from Table 8.1, there is a large difference between the mean scores of all the mothers and the two mothers at the higher and lower end. In particular, there are huge differences in the use of decontextualized language by the mother at the lower and the mother at the higher end. The differences on lexical density, lexical diversity and the use of subordinate clauses are also large,

especially at Time 3. Besides, while the mother of Ibtisam clearly grows in the use of academic features, Fatima's mother does not grow on all features.

When comparing the results of the Berber mothers with those of monolingual Dutch mothers in the study of Henrichs (2010), it turned out that the main differences are at the level of complex multi-clause sentences and decontextualized language. It concerns the results based on the book task in both studies. In our study with the Berber mothers, the proportions of complex multi-clause sentences over all utterances were 14%, 13% and 14% respectively at Time 1, 2 and 3. On the other hand, the Dutch mothers' proportions were 5%, 9% and 9% respectively at Time 1, 2 and 3. This finding may be due to the relatively high frequency of complement clauses in Berber mothers' speech while the Dutch mothers mainly used coordinate clauses. Regarding the use of decontextualized language, our study showed that the proportion of decontextualized language of Berber mothers was 20%, 11% and 22% respectively at Time 1, 2 and 3. In contrast, the proportions of decontextualized language in the Dutch mothers' speech were 25%, 29% and 37% respectively at Time 1, 2 and 3. We note here that the term used by Henrichs (2010) for decontextualized language was 'nonpresent talk'. A possible explanation of these differences is the fact that the group of Dutch mothers was on average higher educated and had higher literacy levels.

As was discussed in Chapters 5, 6 and 7, the relationship between family background characteristics (SES and literacy) and academic language features in maternal input showed many differences. The most notable finding is the strong relationship between the use of decontextualized language and SES and literacy. The lexical and morpho-syntactic aspects showed low or unstable correlations with mothers' socio-economic and literacy background. In other words, the impact of SES and literacy on the more inherent properties and linguistic units of the language (lexicon and morpho-syntax) is less than on the abstraction level. We could hypothesize that high-SES Berber mothers might have access to and make use of the higher strategies of decontextualization in Berber, Dutch and maybe in other languages. Languages other than Berber have been used for decades in schools and academic contexts and, thus, have developed a common register in which abstraction is one of the basic elements. Berber mothers may use this register, that is not language-specific, from their knowledge of other languages and apply it in their language input to the child.

In sum, the findings of our empirical studies indicate that the input of Berber mothers to their children contains several features of the academic register. There is a large amount of variation between the degree in which these mothers use the academic register. The variation on the aspect of decontextualization can partly and most strongly be explained by the influence of SES and literacy.

8.2 Academic language input at school

Turning now to what happens within classroom interactions, we consider the teachers' input in the three investigated aspects of the academic register. The results on the lexical features reported in Chapter 5 showed a considerable task effect. We found that teachers used more utterances and more content words during the book task than during circle time. The book reading activity seems to foster the type of language linked to the academic register. However, this is not surprising as many studies demonstrated that book reading conversations in classrooms generate more talk, richer language and more abstract discourse (Cochran-Smith, 1984; Dickinson, De Temple & Smith, 1992). As far as lexical diversity is concerned, no significant differences were found between the two interaction tasks. Another interesting finding with regard to lexical variables is the significant increase of teachers' language use from measurement time 2 to measurement time 3. This suggests that both the size and the composition of the teachers' lexical input changes as children grow older. This should come as no surprise since language input at school is partly structured according to school materials in which more complex language is used when the level (and the age of children) gets higher. At home, parents do not have an instructional textbook prompting them to use more complex lexicon as the child gets older.

Concerning grammatical complexity in the teachers' input, we found that coordinate and complement clauses are the largest categories used in teacher-child conversations. We can say that teachers tend to use the most accessible types of multi-clause sentences. In terms of the degree of complexity, coordinate clauses are considered as the most simple multi-clause sentences. Complement clauses, on the other hand are, as argued in Chapter 6, the first complex constructions which emerge in children's speech. This suggests that teachers may take into account children's linguistic abilities and adjust their language use accordingly. It should be noted that teachers used more frequently

other subordinate clauses, such as adverbial and relative clauses, if compared to the mothers who demonstrated minimal use of these types.

When we look at the (de)contextualized language use of teachers during book reading sessions, we see that teachers use both lower and higher levels of abstraction. During the book reading sessions, teachers spend a considerable time in providing instructions, organizing the reading task and giving simple feedback. At the same time, they attempt to engage the children in cognitively challenging talk. These patterns lead teachers to the use of contextualized as well as decontextualized language. A comparison between the input that the children receive at home and the input received at school revealed significant differences. While mothers' language during the book task was dominated by immediate utterances dealing with immediately perceivable and present information (naming, labeling and locating), teachers used a high proportion of infrequent, decontextualized and cognitively challenging talk that is removed from the immediate here and now (analyzing, inferring and reasoning). Our results corroborate the findings of Dickinson *et al.,* (1992), who investigated the co-construction of book reading experiences of young children at home and at school. This study showed that teachers used more non-immediate talk to the children than mothers did.

With respect to the three aspects of academic language used by the teachers, we note that there is a variation between teachers. This is illustrated by considering the averages of all teachers on four aspects and the averages of two teachers who had high and low averages on the same aspects. For each aspect, we considered Time 2 and Time 3 as displayed in Table 8.2. For the lexical level we calculated the average of the two tasks.

Table 8.2: Variation among teachers in the use of different components of academic register

	Lexical density		Lexical diversity		Proportion of subordinate clauses		Proportion of decontextualized language	
	T 2	T 3	T 2	T 3	T 2	T 3	T 2	T 3
All teachers	2.29	2.26	73.85	86.77	.14	.08	.42	.40
Teacher of Ibtisam	3.56	2.52	80.54	87.88	.23	.24	.43	.40
Teacher of Saliha	1.58	1.72	52.08	52.36	.04	.00	.29	.39

What can be seen in Table 8.2 is that there is a large difference between the mean scores of all the teachers and the two teachers at the higher and lower end. The major differences between the teacher at the higher and the teacher at the lower end are at the level of lexical density and the proportion of subordinate clauses. For all other individual scores, see Appendix 4.

If we compare the results of the teachers of Moroccan Berber children with those of teachers of monolingual Dutch children in the study of Henrichs (2010), we see that the main differences are at the level of decontextualized language at Time 3. In our study, the proportions of decontextualized language used by teachers were 42% and 40% respectively at Time 2 and 3. In the Dutch group, however, the proportions were 40% and 64% respectively at Time 2 and 3.

8.3 Children's use of academic language

In our view, children's language development is a mixture of the innate linguistic knowledge conceptualized by Chomsky (1965) containing a number of grammatical principles and parameters that enable them to acquire a language, and the degree of exposure to language input in early childhood and language use (Langacker, 2000; Goldberg, 1995; Tomasello, 2003). Although the aim of the current study is not to support or challenge any of these approaches, we consider language input as a crucial factor in the language development process. An extensive body of research within the usage-based approach showed that children should be involved in social interactions to be best able to acquire language, and that language structure emerges from language use (Clark, 2003; Tomasello, 2003; Diessel, 2004). Children, thus, develop, select and generalize grammatical categories by analyzing and systemizing the input to which they are exposed (Diessel, 2004).

Before discussing features of the academic register in children's language, it should be noted that the language behavior of the target children at home was characterized by a mixture of two languages: Berber and Dutch. To put it differently, when children talked to their mothers, they, as it were, often 'violated' the standards of proper speech and the conventions of using each language separately or according to the context in which they find themselves. When interacting with their mothers, children used features from two of the

languages which are at their disposal to achieve their communicative aims. Because of globalization and diasporic communities, individuals in modern societies, and young urban language users in particular, become acquainted with a wide range of languages. They use some of these language features in their speech even though they do not fully command the languages (Jørgensen, 2008). According to Jørgensen (2008), it is completely irrelevant how many languages are involved in an interaction. Instead of bilingualism or multilingualism, Jørgensen (2008) introduced the term poly-lingualism. In bilingualism or multilingualism, speakers should have access to two or more languages or varieties (e.g. "Turkish is my mother language, I speak Dutch fluently and I can understand some Spanish"). In addition, speakers are expected to have full linguistic competence in these languages and are required to use them in different settings and to different speakers. In contrast, in poly-lingualism, "language users use whatever linguistic features are at their disposal to achieve their communicative aims [...] regardless of how well they know the involved languages" (Jørgensen, 2008: 163). As Møller & Jørgensen (2009: 147) put it, "regardless of our social standing vis-à-vis a given code, as human beings, we do not primarily use "a language" or "some languages", we use *language,* linguistic features, and we do so to achieve our aims". Instead of languages, they propose the term *languaging* to describe the human linguistic behavior and the term *languagers* for the speakers. So, instead of saying that these children use Tarifit Berber or Dutch or both languages at the same time, we can say that these early age *languagers* are *languaging* to their mothers. Basically, in the present study, we looked at what these children are using in terms of language structures, isolating the languages they use only as instrumental devices.

To illustrate this, we choose two examples where two mothers and their children use Berber and Dutch during the book task.

Example 8.1: (Isra and her mother, Time 1)

*MOT: iwa zid?
and add-AOR
(and what else)
*CHI: wah ik@nl heb@nl hem@nl.
yes I got it
(yes, I got it)

*MOT: ja hij wijst naar boven he [$] mayn i-ttwara?
yes he points up what 3SM-see-AOR
(yes, he points up, what does he see?)
*CHI: vogel@nl.
(bird)
*MOT: lla <lijkt een beetje daarop> [$nl] y-ttwara vliegtuig@nl.
no looks like a little bit it 3SM-see-IMPERF plane
(no, it looks a little bit like it, he sees a plane)
*CHI: vliegtuig ik xxx [$nl].
(plane, I…)
*MOT: toen@nl y-gga-s ammu &errrr .
then 3SM-do-him like this &errrr
(and then he did like this errrr)

In example 8.1, we can see that the mother uses Berber and Dutch in order to make sure her daughter understands the story. The daughter, Isra, however gives answers predominantly in Dutch and uses only one Berber word.

Example 8.2: (Anissa and her mother, Time 1)

*MOT: het is klaar [$nl] xzar.
it is finished look-AOR
(it is finished, look)
*MOT: y-nna-s Tijn dag@nl.
3SM-tell-AOR-him Tijn bye bye
(Tijn told him: bye bye)
*CHI: beslama dag@nl dag@nl
bye bye bye
(bye, bye, bye)

In example 8.2, we see that both the mother and her daughter, Anissa, combine Berber and Dutch. Again, we can easily recognize which features each speaker uses here, but not why they do so. The linguistic choices and the sets of linguistic resources are less important than the aim behind using these features to communicate. Both examples seem to suggest that mother and child are communicating with each other and they achieve their communicative goals by using words/constructions/aspects out of the two languages at hand.

The children's data in the home setting indicate that features of academic language emerge especially as children reach the age of 5. Taken as a group,

evidently with a considerable variability, we can conclude that Moroccan Berber children were able to use language which can be typified as belonging to the academic register. This finding is, however, not surprising as children's vocabulary size and quality are the most natural and uncontroversial aspects of language acquisition that grow and change over time.

Another important variable in this study was the influence of the children's family's SES and literacy background on the use of lexical features of the academic register. At the lexical level, only one feature, namely lexical diversity, showed a correlation with SES and literacy at measurement time 3. This is contrary to prior research findings which reported strong links between SES and children's vocabulary development (cf. Hart & Risley, 1995; Hoff-Ginsberg, 1998).

Regarding the second distinctive feature of the academic register, syntactic complexity, it was found that coordination was the most frequently used category of clause combining at all measurement times. This is a natural development, which young speakers go through when they start chaining clauses by using conjunctions such as 'and' or 'but'. Other embedded and more complex clause strategies (complement, relative and adverbial clauses) also emerge in a limited way but show a significant increase as children grow up. A remarkable finding is that almost all the complex multi-clause sentences produced by children while talking with their mothers were in Dutch. This means that when these children want to pack information in a dense, organized and embedded structure, they do that mostly in their second language. It is not clear whether these children are unable to use clause-combining strategies in Berber possibly because they do not hear them often enough in Berber to learn them or whether they just prefer to use Dutch sentences because these are easier.

Also at the level of clause combining, we addressed the question as to whether children's use of complex constructions is related to the family's SES and literacy background. In this study, at different ages, positive but not significant correlations have been established between children's language complexity and their parental SES. This is partly in line with the results reported by Huttenlocher *et al.* (2002). This study demonstrated that parental SES partly showed significant relationships with the use of complex constructions by children. Other studies also showed that SES does matter in children's use of complex sentences. The study of Arriaga *et al.* (1998), for

example, showed that children from high-SES families produce more complex utterances in spontaneous speech than children from low-SES families. Hoff (2006) also showed that higher SES children use longer and more complex sentences when involved in conversations with adults than do lower SES children.

Lastly, the study of the children's use of decontextualized language at home revealed interesting outcomes. Generally, up to age 4, the children mainly use the lowest levels of abstraction, which are context-bound, during mother-child interactions. This pattern, however, changes as children reach the age of nearly 6. At this period of development, the amount of decontextualized utterances produced by children increases. And as shown in the previous chapters, the children's use of decontextualized language was significantly associated with the family's SES and literacy background. This leads us to conclude that families of high SES and literacy activities foster the development of decontextualized language skills which are crucial for children's school achievement and academic success.

As said when we discussed the use of academic register by mothers and teachers, we note also that we found considerable variation between children during the home interaction tasks. This can be illustrated by considering the averages of all children on four different aspects and the averages of two children who, respectively had high and low averages on most academic language features. For each aspect, we looked at Time 1, 2 and 3 (for the lexical level we calculated the average of the two tasks) as shown in Table 8.3.

Table 8.3: Variation among children in the use of different components of academic register at home

	Lexical density			Lexical diversity			Proportion of subordinate clauses			Proportion of decontextualized language		
	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3
All children	1.04	1.39	1.51	55.06	55.33	64.01	.01	.03	.08	.17	.21	.24
Hamza	1.20	1.05	1.07	57.05	58.70	87.17	.02	.00	.13	.38	.52	.57
Hatim	.63	1.05	.98	60.44	49.00	53.19	.00	.06	.00	.17	.13	.15

As Table 8.3 shows, there are considerable differences between the mean scores of all children and the two children at the higher and lower end. In

particular, there are large differences with respect to the proportion of decontextualized language in all measurement times and lexical diversity in Time 3. In addition, Hamza clearly grows in the use of decontextualized language and in lexical diversity over the three measurement times while Hatim does not.

When comparing the results of the Berber children at home with those of monolingual Dutch children at home in the study of Henrichs (2010), it turns out that the main differences are at the level of abstraction. It concerns the results based on the book task in both studies. In our study, the proportions of decontextualized language were 17%, 21% and 24% respectively at Time 1, 2 and 3. The proportions of the Dutch children, on the other hand, were 16%, 23% and 30%. As we can see, there is more increase among the Dutch children than the Berber children, especially from Time 2 to Time 3. Regarding the proportions of subordinate clauses, the results of both groups are comparable with an increase among the Berber group from Time 2 to Time 3.

In the transition made when children enter school, not only does their social environment change but so, radically, does their language experience. In this context, it should be acknowledged that there is a major difference between home and school settings with regard to input providers. At home, parents and children are biological relatives. In addition, as noted by Huttenlocher *et al.* (2002), parents are input providers for the children across the entire period of language acquisition. In contrast, teachers are input providers only for a limited period of time. At school, language changes from an informal means of communication to a more formal and demanding instrument of learning. Many studies show that the home language input is clearly associated with children's vocabulary growth and language development in general. However, no comparable evidence has been shown for the influence of school input during the primary school years. One or two years of school experience have no measurable impact on vocabulary growth (Biemiller, 2005). Biemiller argues that when children are at age 5, it is *home* not *school* which determines vocabulary size and language development.

Recent research on Moroccan and Turkish bilingual children in the Netherlands showed a positive cross-language transfer effect of the language skills developed in L1 on the acquisition and development of skills in L2 (Scheele, 2010). However, other studies argue that this language transfer can only be achieved when children's L1 skills are sufficiently developed (August *et*

al., 2006; Butler & Hakuta, 2004). In any case, our primary interest in the current study was not to investigate the transfer from L1 to L2 but to shed some light on the input these children receive at home in their L1 (and L2) and later at school in L2, in addition to relationships with the output.

The results concerning children's language use in the classroom showed that they produced less utterances at school than at home. This difference is mainly caused by the number of speakers involved in each setting. During home tasks, the target child is usually the only interlocutor with his/her mother. In school tasks, on the other hand, other classmates are also involved in conversations with the teacher.

Regarding the lexical features of the academic register, children showed a remarkable increase from Time 2 to Time 3 in their use of dense and diverse vocabulary. A comparison of home and school data did not show many differences at the lexical level. Only during measurement time 2, did children use significantly denser language at home than at school. With regard to the syntax and abstraction level, children clearly tend to use more clause combining strategies and decontextualized talk in school interactions than in home interactions. From age 4, teacher-child interactions become increasingly a means to transmit specific language and literacy skills that are required at school. From these findings, we can conclude that children use more academic language features at school than at home. When we further examine the children's use of academic language features at school, we notice a considerable variation between the children (see Appendix 4). To illustrate this, we consider the averages of all children on four aspects and the averages of two children who, respectively had high and low averages on most academic language features. For each aspect, we looked at Time 2 and Time 3 (for the lexical level we calculated the average of the two tasks) as shown in Table 8.4.

Table 8.4: Variation among children in the use of different components of academic register at school

	Lexical density		Lexical diversity		Proportion of subordinate clauses		Proportion of decontextualized language	
	T 2	T 3	T 2	T 3	T 2	T 3	T 2	T 3
All children	1.27	1.57	45.16	50.31	.07	.10	.30	.32
Fatima	1.71	1.48	50.13	106.16	.33	.50	.29	.23
Selma	.93	1.57	–	–	.00	.20	.25	.29

As can be seen from Table 8.4, there is a large difference between the mean scores of all the children and the two children at the higher and lower end. In particular, there are large differences in the use of subordinate clauses by Selma and Fatima. We note here that we could not compare our results of the Berber children at school with those of Henrichs (2010) as she did not analyze the output of the Dutch children at school.

8.4 Input-output relationships

Concerning the relationship of the mothers' input with the children's output, correlational analyses of different lexical features during the book task revealed a statistically significant relationship between mothers' lexical diversity and that of their children in Time 1 and Time 2 (between the ages of 3 and 4 years old). It seems that pre-school children who receive more diverse language input are also able to produce a more diverse language. This is consistent with Weizman and Snow's (2001) findings in which strong correlations were found between the exposure to sophisticated vocabulary items and children's own vocabulary. In our study, though, from the age of 4 onwards, no direct link was found between the lexical input of the mothers and the children's output. This might be explained by the fact that after these children entered school at the age of four, they received more input in more than one language (mainly Tarifit Berber at home and Dutch at school). For other features, no strong relationships were found. During the picture task, no significant relationships were found between mothers' and children's lexical features except for the number of tokens in Time 3.

In the home setting, we were also interested as to how input-output relationships look like in terms of the use of clause combining. However, correlational analyses between mothers' and children's use of complex constructions revealed positive but not significant relationships in any of the measurement times.

At the abstraction level, we also examined the extent to which decontextualized language in the mothers' input accounts for their children's use of decontextualized language. Unlike the findings at the level of lexical and morpho-syntactic features, we found synchronous and time-lagged, strong and significant correlations between mothers' decontextualized linguistic input and children's language use. In mother-child interactions, children who are exposed to language which is removed from the immediate place and time, develop abstract thinking and are able to go beyond the perceivable context while they speak. These findings in our study parallel those of Heath (1982, 1983), Sigel and McGillicuddy-Delsi (1984) and Van Kleeck *et al.* (1997). All these studies confirmed that parents' abstract linguistic input is related to children's abstract language abilities.

The strong input-output relationships at the abstraction level and the relatively weak relationships at the lexical richness and clause combining levels are open to many explanations. It might be due to our sample size which is relatively small and the number of the measurement times and time intervals in between. If we consider, for example, the use of clause combining by children, a considerably low number of complex multi-clause sentences was produced over the three measurement times (see Table 6.9). The low level or absence of input-out relationships at the lexical and morpho-syntactic level could, thus, be caused by the low number of children's utterances which contained rich vocabulary and complex sentences. The nature of the task used in our study could also play a role in the degree of richness, complexity and abstractness of the language used both by the mothers and the children. Future studies relying on more tasks could confirm or refute the existence of a relationship between parental input and children's output.

With regard to input-output relationships in the school setting, in comparison with the mothers, less significant correlations were found between the input of the teachers and children's language production. Only the number of tokens used by the teachers was related to the number of tokens used by the children during the circle time in Time 3. At the level of clause combining, the

input-output correlations did not reveal any significant relationships. In other words, teachers who used more syntactic complex utterances did not necessarily have pupils who produced syntactically complex sentences. At the level of decontextualized language, we also did not find any significant relationship between teachers input and children's language production. It may be the case that the effects of the teachers' input on the output of children takes longer to become observable.

8.5 Implications, future directions and limitations

What do our results mean for Moroccan Berber children growing up in the Netherlands? And what are the implications of our outcomes for the Dutch schooling system in general? The challenge for policy makers and educators is to create a curriculum and setting which will bridge the gap between home and school. From this perspective, the role of the teacher is not limited to transferring information to children but in addition, perhaps more importantly, to act as a cultural broker who mediates among the different codes in children's own repertoires so that teaching programs build on, rather than devalue, what children learn at home (Gumperz & Gumperz, 2006). The role of school is to offer children a more structured and organized way of learning to foster academic language. Next to the knowledge they receive in everyday and out-of-school experiences, children need to be introduced to a more structured and organized knowledge system offered in the classroom. Or to use Vygotsky's (1986) terms, while children develop "everyday" and spontaneous concepts through the less structured experiences in everyday life, they acquire "scientific" concepts in a systematic, organized and abstract way in the school context. This challenge becomes even bigger if we take into account that Moroccan Berber children have only limited experience with academic language input in their L1.

As many studies have argued, children's academic language skills in L1 should be sufficiently developed in order to facilitate the acquisition process of L2 academic language skills (Cummins, 2000; Elbers, 2010). It is highly recommended that both parents and educators should be aware of the kind of language input they provide to children and how to build on it. It is therefore necessary that a stronger focus should be placed on the type of language activities which enhance academic language acquisition. For instance, in the

framework of language intervention programs in preschools and kindergartens, specific attention should go to language activities which are characterized by lexical richness, syntactic complexity and decontextualized language in order to enhance academic language development.

The results presented in our work should be interpreted with caution, since this study has a number of limitations. In this study, we focused on maternal input in the home situation. We asked which one of the parents spent most time with the young child at home, and in all cases this turned out to be the mother. Fathers were thus not involved as input providers in the family. In many cultures, mothers are considered to be the prime agents of socialization. In addition, it is presumed that the mother has the primary responsibility for childrearing and that cultural norms are transmitted through her to the child (Feagans & Farran, 1982). Nevertheless, additional or alternative sources of input may play a critical role in shaping children's language use in later years. Future research could also consider the input of fathers, siblings and other family members to reach a more comprehensive understanding of the input which stimulates academic language acquisition by children.

Collecting, transcribing and coding the data used in the current study was labor-intensive. During this longitudinal study, we managed to set up a large and in-depth database and speech corpus resulting from home and school interactions. Unfortunately, because of time restrictions we were not able to transcribe and analyze all these data. Undoubtedly, there are questions that stayed unanswered in the present study. Subsequent and follow-up investigations may consider unanalyzed elements in our sample to track the exact factors that contribute to academic language acquisition. Furthermore, it is important to acknowledge that our sample was far from ideal. On the basis of the relatively small number of respondents who participated in this study, it is difficult to make generalizations and firm conclusions. Future studies require the examination of larger samples to provide an accurate description of the factors influencing children's emergent academic language skills. Another restriction is that we used only three interaction tasks. Future designs should consider more spontaneous and more structured tasks, to provide additional information about the relations between language input and the use of academic register by children. Another point, which should be emphasized, is that we focused on a specific age group, 3 to 6 years old. For future studies, it would be desirable to observe children in a wider age range and at more

measurement times to determine the age at which they master the skills related to academic language.

In conclusion, we believe that this study is a first step towards a deeper understanding of how Moroccan Berber children's academic language develops at home and at school. In our opinion, it is imperative to design more studies which can lead to new insights into this register that children need for their school career.

References

Aarts, R., Demir, S., Kurvers, J., Laghzaoui, M., & Henrichs, L. (2006). Ontluikende schooltaalvaardigheid in interacties tussen moeders en driejarigen. Op zoek naar schooltaalkenmerken in Nederlandse, Berber en Turkse gezinnen. *Toegepaste Taalwetenschap in Artikelen*, *76*(2), 21-38.

Aarts, R., Demir, S., & Vallen, T. (forthcoming). Characteristics of academic language register occurring in caretaker-child interaction: Development and validation of a coding scheme. *Language Learning.*

Abbassi, A (1977). *A sociolinguistic analysis of multilingualism in Morocco.* Ph.D. dissertation, Texas University at Austin.

Albino, J. (1859). *Manual del lenguaje vulgar de los moros de la Riff.* Cádiz: Textbook.

Arriaga, R.K., Fenson, L., Cronan, T., & Pethick, S.J. (1998). Scores on the MacArthur Communicative Development Inventory of children from low- and middle-income families. *Applied Psycholinguistics*, *19*, 209-225.

August, D., Carlo, M.S., Calderón, M., & Nuttall, M. (2006). Developing literacy in English-language learners: An examination of the impact of English-only versus bilingual instruction. In: P. McCardle, & E. Hoff (Eds.), *Childhood bilingualism: Research on infancy through school age*. Clevedon, England: Multilingual Matters.

Bach, K. (1994). Conversational implicature. *Mind & Language*, *9*, 124-162.

Bach, K., & Harnish, R.M. (1979). *Linguistic Commuication and Speech Acts.* Cambridge, Mass.: MIT Press.

Basset, A. (1952). *La langue berbère.* London: Oxford University Press.

Basset, H. (1920). *Essai sur la littérature des berbères.* Alger: J. Carbonnel.

Bentahila, A. (1983). *Language attitudes among Arabic-French bilinguals in Morocco.* Clevedon: Multilingual Matters.

Bernstein, B. (1960). Language and social class. *British Journal of Sociology*, *11*, 271-276.

Bernstein, B. (1971). *Class, codes, and control* (Vol. 1). London: Routledge & Kegan Paul.

Bernstein, B. (1972). A critique of the concept of compensatory education. In: C.B. Cazden, V.P. John, & D. Hymes (Eds.), *The functions of language in the classroom*. New York: Teachers College Press.

Bernstein, B. (1975). *Towards a theory of educational transmissions. Class, codes and control* (Vol. 3). London: Routledge & Kegan Paul.

Berry, J.W., Van de Koppel, J., Sénéchal, C., Annis, R., Bahuchet, S., Cavalli-Sforza, L.L., & Witkin, H.A. (1986). *On the edge of the forest: Cultural adaptation and cognitive development in Central Africa*. Lisse: Swets & Zeitlinger.

Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism*, *10*, 210-223.

Biarnay, S. (1917). *Étude sur les dialectes berbères du Rif: Lexique, textes et notes de phonétique*. Paris: Ernest Leroux.

Biemiller, A. (2005). Vocabulary Development and Instruction: A Prerequisite for School Learning. In: S. Neuman, & D. Dickinson (Eds.), *Handbook of Early Literacy Research* (Vol. 2). New York, NY: Guilford Press.

Blank, M., Rose, S.A., & Berlin, L.J. (1978). *The language of learning. The preschool years*. New York: Grune & Stratton.

Blank, M., Rose, S.A., & Berlin, L.J. (2003). *Preschool language assessment instrument* (2nd ed.). Austin, TX: Pro-Ed.

Bloom, L. (1991). *Language development from two to three*. Cambridge: Cambridge University Press.

Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K. (1980). Complex sentences: acquisition of syntactic connectives and the meaning relations they encode. *Journal of Child Language*, *7*, 235-61.

Bloom, L., Lifter, K., & Hafitz, J. (1980). Semantics of verbs and the development of verb inflection in child language. *Language*, *56*(2), 386-412.

Bloom, L., Rispoli, M., Gartner, B., & Hafitz, J. (1989). Acquisition of complementation. *Journal of Child Language*, *16*, 101-120.

Boer, K. de, & Veldkamp, T. (2005). *Tim op de tegels* [Tim on the paving stones]. Houten: Van Goor.

Boukous, A. (1995). *Société, Langue et Culture au Maroc. Enjeux symboliques*. Rabat: Faculté des Lettres et des Sciences Humaines.

Bouwerman, M. (1979). The acquisition of complex sentences. In: P. Fletcher, & M. Garman (Eds.), *Language acquisition: Studies in First Language Development*. Cambridge: Cambridge University Press.

Bradley, R., & Caldwell, B. (1984). The HOME Inventory and family demographics. *Developmental Psychology*, *20*(2), 315-320.

Brent, M., & Siskind, J.M. (2000). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*(2), B33-B44.

Broeder, P., & Extra, E. (1999). *Language, ethnicity and education: case studies on immigrant minority groups and immigrant minority languages*. Clevedon: Multilingual Matters.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Bus, A.G., Van IJzendoorn, M.H., & Pellegrini, A.D. (1995). Joint book reading makes for success in learning to read. *Review of Educational Research*, *65*, 1-21.

Butler, Y., & Hakuta, K. (2004). Bilingualism and second language acquisition. In: T. Bhatia, & W. Ritchie (Eds.), *The handbook of bilingualism* (pp. 114-144). Malden: Blackwell.

Cadi, K. (1987). *Système verbal rifain. Forme et sens*. Paris: SELAF.

Camps, G. (1980). *Les Berbères, aux marges de l'histoire*. Toulouse: Editions des Hespérides.

Camps, G. (1987). *Berbères, mémoire et identité*. Paris: Edition Errance.

CBS (2001). *Allochtonen in Nederland* [Immigrants in The Netherlands]. Voorburg/Heerlen: Centraal Bureau voor de Statistiek.

CBS (2006). *Allochtonen in Nederland*. [Immigrants in The Netherlands]. Voorburg/Heerlen: Centraal Bureau voor de Statistiek.

CBS (2010). *Allochtonen in Nederland* [Immigrants in The Netherlands]. Voorburg/Heerlen: Centraal Bureau voor de Statistiek.

CBS (2011). *Allochtonen in Nederland* [Immigrants in The Netherlands]. Voorburg/Heerlen: Centraal Bureau voor de Statistiek.

Chacha, M. (1995). *Cway zi tibbuhelya ad war tiwi* [Een beetje van onrijpe gekheid]. Amsterdam: Stichting Izaouran.

Chacha, M. (1997). *Rez ttabu ad dd teffegh tfukt* [Breek de taboe en laat de zon schijnen]. Amsterdam: Stichting Izaouran.

Chacha, M. (2000). *Abrid gher Yizran* [Over de geschiedenis van de zang in het Tarifit]. Amsterdam: Stichting Izaouran.

Chafe, W.L. (1982). Integration and involvement in speaking, writing, and oral literature. In: D. Tannen (Ed.), *Spoken and written language: exploring orality and literacy* (pp.35-54). Norwood, N.J.: Ablex.

Chafik, M. (1982). Pour l'élaboration du berbère "classique" à partir du berbère courant. In: *Actes de la 1ère Rencontre de l'Université d'Eté d'Agadir* (pp. 191-197).

Chafik, M. (2004). *Imazighen: de Berbers en hun geschiedenis*. Amsterdam: Bulaaq.

Chaker, S. (1989). *Berbères aujourd'hui*. Paris: Editions L'Harmattan.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge: M.I.T. Press.

Christie, F. (1985). Language and schooling. In: S. Tchudi (Ed.), *Language, Schooling and Society* (pp. 21-40). New Jersey: Boynton Cook.

Clark, E.V. (1970). How young children describe events in time. In: G.B. Flores d'Arcais, & W.J.M. Levelt (Eds.), *Advances in psycholinguistics*. Amsterdam: North-Holland.

Clark, E. (2003). *First Language Acquisition*. Cambridge: Cambridge University Press.

Cochran-Smith, M. (1984). *The making of a reader*. Norwood, NJ: Ablex.

Cox, B.T., Fang, Z., & Otto, B. (1997). Preschoolers's developing ownership of the literate register. *Reading Research Quarterly*, *3*(1), 34-53.

Cristofaro, S. (2003). *Subordination*. Oxford: Oxford University Press.

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, *19*, 121-129.

Cummins, J. (1991). Conversational and academic language proficiency in bilingual contexts. *AILA Review*, *8*, 75-89.

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Clevedon, England: Multilingual Matters.

Curenton, S.M., Craig, M.J., & Flanigan, N. (2008). Use of Decontextualized Talk Across Story Contexts: How Oral Storytelling and Emergent Reading Can Scaffold Children's Development. *Early Education & Development*, *19*(1), 161-187.

Curenton, S.M., & Justice, L.M. (2004). African American and Caucasian preschoolers' use of decontextualized language: Literate language

features in oral narratives. *Language, Speech, and Hearing Services in Schools*, *35*, 240-253.

Dagevos, J., & Gijsberts, M. (2007). *Jaarrapport Integratie 2007*. Den Haag: Sociaal en Cultureel Planbureau.

Dagevos, J., Gijsberts, M., & Van Praag, C. (2003). *Rapportage minderheden 2003. Onderwijs, arbeid en sociaal-culturele integratie*. Den Haag: Sociaal en Cultureel Planbureau.

Daller, H., Van Hout, R., & Treffers-Daller, J. (2003). Lexical richness in spontaneous speech of bilinguals. *Applied linguistics*, *24*, 197-222.

DASH (2006). *Coding Protocol for Functional Linguistic Analysis*. Amsterdam/ Tilburg/Utrecht: DASH Research Group.

Davidson, R.G., & Snow, C.E. (1995). The linguistic environment of early readers. *Journal of Research in Childhood Education*, *10*, 5-21.

De Jong, P., & Leseman, P. (2001). Lasting effects of home literacy on reading achievement in school. *Journal of School Psychology*, *38*(5), 389-414.

Demir, S. (forthcoming). *Academic language development of Turkish children in the Netherlands* (working title). Ph.D. dissertation, Tilburg University.

Denny, J.P. (1991). Rational thought in oral culture. In: D.R. Olson, & N. Torrance (Eds.), *Literacy and Orality* (pp. 66-89). Cambridge: Cambridge University Press.

De Ruiter, J.J. (2001). Analyse (socio-)linguistique de la Charte nationale d'éducation et de formation. *Languages and Linguistics; An International Journal of Linguistics*, *8*, 29-47.

De Ruiter, J.J., & El Aissati, A. (2004). De recente geschiedenis. In: M. Chafik, *Imazighen: de Berbers en hun geschiedenis* (pp. 113-138). Amsterdam: Bulaaq.

De Temple, J.M. (2001). Parents and children reading books together. In: D.K. Dickinson, & P.O. Tabors (Eds.), *Beginning literacy with language*. Baltimore, MD: Paul H. Brookes Publishers.

De Temple, J.M., & Beals, D.E. (1991). Family talk: Sources of support for the development of decontextualized language skills. *Journal of Research in Childhood Education*, *6*, 11-19.

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

Dickinson, D.K. (2001). Large-Group and free-play times: Conversational settings supporting language and literacy development. In: D.K.

Dickinson, & P.O. Tabors (Eds.), *Beginning literacy with language*. Baltimore, MD: Paul H. Brookes Publishers.

Dickinson, D.K., McCabe, A., & Anastasopoulos, L. (2003). A Framework for Examining Book Reading in Early Childhood Classrooms. In: A. Van Kleeck, E.B. Bauer, & S. Stahl (Eds.), *On Reading Books to Children: Parents and Teachers* (pp. 95-113). Hillsdale, NJ: Lawrence Erlbaum.

Dickinson, D.K., De Temple, J.M., Hirschler, J.A., & Smith, M.W. (1992). Book reading with preschoolers: Co-construction of text at home and at school. *Early Childhood Research Quarterly*, *7*, 323-346.

Dickinson, D.K., & Tabors, P.O. (1991). Early literacy: Linkages between home, school and literacy achievement at age five. *Journal of Research in Childhood Education*, *6*, 30-46.

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge: Cambridge University Press.

Diessel, H., & Tomasello, M. (2001). The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics*, *12*(2), 97-141.

Dik, S.C. (1968). *Coordination: its implications for the theory of general linguistics*. Amsterdam: North-Holland Publishing Co.

Donaldson, M. (1978). *Children's minds*. New York: Norton.

Dromi, E., & Berman, R.A. (1986). Language-specific and Language-general in developing Syntax. *Journal of Child Language*, *13*, 371-381.

Dunlop, A.W., & Fabian, H. (Eds.) (2007). *Informing transitions in the early years*. Berkshire, England: Open University Press.

Durán, P., Malvern, D., Richards, B., & Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics*, *25*(2), 220-242.

Egging, S. (1994). *An Introduction to Systemic Functional Linguistics*. London: Pinter Publishers.

El Aissati, A. (2001). Word Order Change in Berber. *FAB Frankfurter Afrikanistische Blaetter*, *13*, 35-44.

El Aissati, A., Karsmakers, S., & Kurvers, J. (2011) "We are all beginners." Amazigh in language policy and educational practice in Morocco. *Compare, a journal of comparative and international education*, *41*(2), 211-227.

Elbers, E. (2010). Learning and social interaction in culturally diverse classrooms. In: K.S. Littleton, C. Wood, & J. Kleine Staarman (Eds.), *The International Handbook of Psychology in Education*. Bingley, UK: Emerald.

El Kirat, Y. (2001). The current status and future of the Amazigh language in Beni Iznassen community. *Languages and Linguistics*, *8*, 81-95.

El Walid, M. (1994). *Zi redjagh n tmurt ghar ruara ujenna* [Reikend naar het licht]. Amsterdam: Stichting Izouran.

El Walid, M. (1996). *Tifardjas* [Zwaluwen]. Utrecht: Stichting Apuleius.

Ennaji, M. (1985). *Contrastive Syntax: English, Moroccan Arabic and Berber Complex Sentences*. Würzburg: Königshausen und Neumann.

Ennaji, M. (1997). The Sociology of Berber. Change and Continuity. *International Journal of the Sociology of Language*, *123*, 23-40.

Ennaji, M., & Sadiqi, F. (2004). *A grammar of Amazigh*. Fès: Faculty of Dhar El Mehraz.

Ensminger, M.E., & Fothergill, K. (2003). A decade of measuring SES: What it tells us and where to go from here. In: M.H. Bornstein, & R.H. Bradley (Eds.), *Socioeconomic status, parenting, and child development*. Mahwah, NJ: Lawrence Erlbaum.

E-rramdani, Y., Raha, R., & De Santos-Tirado, N. (1998). *Iles inu 1. Tasghunt-Lkitab 1: Tarifit*. Tilburg: Tilburg University Press, Syntax Datura.

Essadki, A. (1997). *Reayad n tmurt: tiqessisin tirifiyin* [Strijdkreet van de aarde: Riffijnse gedichten]. Aalsmeer: Dabar-Luyten.

Extra, G., Aarts, R., Van der Avoird, T., Broeder, P., & Yağmur, K. (2002). *De andere talen van Nederland thuis en op school*. Bussum: Coutinho.

Extra, G., & Maartens, J. (1998). *Multilingualism in a multicultural context: case studies on South Africa and Western Europe*. Tilburg: Tilburg University Press.

Fang, Z. (1997). A study of changes and development in children's written discourse potential. *Linguistics and Education*, *9*(1), 341-367.

Ferguson, C.A. (1959). Diglossia. *Word*, *15*, 325-40.

Field, A. (2009). *Discovering statistics using SPSS for Windows*. London: Sage.

Gottfried, A.E., Fleming, J.S., & Gottfried, A.W. (1998). Role of cognitively stimulating home environment in children's academic intrinsic motivation: A longitudinal study. *Child Development*, *69*(5), 1448-1460.

Grandguillaume, G. (1983). *Arabisation et politique linguistique au Maghreb*. Paris: Maisonneuve et Larose.

Grandguillaume, G. (1995). Comment a-t-on pu en arriver là ? *Esprit*, *208*, 12-34.

Greenfield, P.M. (1972). Oral and written language: The consequences for cognitive development in Africa, the United States and England. *Language and Speech*, *15*, 169-178.

Greenhalgh, K.S., & Strong, C.J. (2001). Literate language features in spoken narratives of children with typical language and children with language impairments. *Language, Speech, and Hearing Services in Schools*, *32*, 114-125.

Groenendaal, J.H.A., Goudena, P.P., & Olthof, T. (1992). Effectieve communicatiepatronen tussen opvoeder en kind: Een longitudinaal onderzoek van de verbale opvoeder-kind interactie in taakgerichte situaties bij peuters en kleuters (3-6 jaar). *Tijdschrift voor Ontwikkelingspsychologie*, *19*(1), 1-29.

Guiraud, P. (1960). *Problèmes et méthodes de la statistique linguistique*. Dordrecht: D. Reidel.

Gumperz, J.J., & Gumperz, J.C. (2006). Interactional sociolinguistics in the study of schooling. In: J.C. Gumperz (Ed.), *The Social Construction of Literacy* (2nd edition). Cambridge: Cambridge University Press.

Haden, C., Reese, E., & Fivush, R. (1996). Mother's extra-textual comments during storybook reading: Maternal style over time and across texts. *Discourse Processes*, *21*, 135-170.

Hajer, M., & Meestringa, T. (1995). *Schooltaal als struikelblok. Didactische wenken voor alle docenten*. Bussum: Coutinho.

Halliday, M.A.K. (1994). *An introduction to functional grammar*. London: Edward Arnold.

Halliday, M.A.K., & Hasan, R. (1989). *Language, context and text: aspects of language in a social-semiotic perspective* (2nd ed.). Oxford, England: Oxford University Press.

Hammam, M. (Ed.) (2004). *Le manuscrit Amazighe: son importance et ses domaines*. Rabat: IRCAM.

Hammett, L.A., Van Kleeck, A., & Huberty, C.J. (2003). Patterns of parents' extra-textual interactions during book sharing with preschool children: A cluster analysis study. *Reading Research Quarterly*, *38*(4), 442-468.

Hart, B., & Risley, T. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology*, *28*, 1096-1105.

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

Heath, S.B. (1982). What no bedtime story means: Narrative skills at home and school. *Language and Society*, *2*, 49-76.

Heath, S.B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. Cambridge: Cambridge University Press.

Heath, S.B. (1994). What no bedtime story means: Narrative skills at home and school. In: B.B. Schieffelin, & E. Ochs (Eds.), *Language socialization across cultures* (pp. 97-124). Cambridge, England: Cambridge University Press.

Henrichs, L. (2010). *Academic language in early childhood interactions. A longitudinal study of Dutch monolingual 3- to 6-year-old children*. Ph.D. dissertation, University of Amsterdam.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*, 1368-1378.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*, 55-88.

Hoff, E. (2009). *Language development* (4th ed.). Belmont, CA: Wadsworth Cengage Learning.

Hoff, E., Laursen, B., & Tardif, T. (2002). Socioeconomic status and parenting. In: M.H. Bornstein (Ed.), *Handbook of parenting. Ecology and Biology of Parenting* (Vol. II, pp. 161-188). Mahwah, New Jersey: Lawrence Erlbaum.

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, *73*(2), 418-433.

Hoff-Ginsberg, E. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics*, *19*, 603-629.

Hutchins, E. (1980). *Culture and inference*. Cambridge, Mass.: Harvard University Press.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236-248.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, *45*, 337-374.

Jørgensen, J.N. (2008). Polylingual languaging around and among children and adolescents. *International Journal of Multilingualism*, *5*(3), 161-176.

Keenan, E.L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic inquiry*, *8*, 63-99.

Khalafi, A. (2000). *Allugha al-um wa sultatu al-ma'sasa* [The mother language and the authority of institutionalization]. Rabat: Imperial.

Kloth, S., Janssen, P., Kraaimaat, F., & Brutten, G. (1998). Communicative styles of mothers interacting with their preschool-age children: A factor analytical study. *Child Language*, *25*, 149-168.

Kraal, K., & Heelsum, A. (2002). *Een dynamisch mozaïek: Nieuwe trends bij Marokkaanse organisaties*. Utrecht: SMT.

Kroll, J.F., & De Groot, A.M.B. (2005). *Handbook of bilingualism: Psycholinguistic approaches*. New York: Oxford University Press.

Laaroui, A. (2007). *Mujmal tarikh al-maghrib* [The concise history of Morocco]. Casablanca: CCA.

Laghzaoui, M. (2005). Language attitudes towards Arabic in Islamic elementary schools in the Netherlands. *Languages and Linguistics; an International Journal of Linguistics*, *15&16*, 111-132.

Laghzaoui, M. (2007). Berber in the Netherlands: A minority language in a multilingual society. In: A. Abel, M. Stuflesser, & M. Putz (Eds.), *Online Proceedings of the 2006 conference on multilingualism across Europe* (pp. 69-67). Bolzano: EURAC Research.

Laghzaoui, M. (2008a). On discourse conventions in academic and interactional Tarifit-Berber texts. In: A. El Aissati (Ed.), *The Amazigh Language at Home and at School: Perspectives on Oral Discourse Structure and Academic Language Skills* (Berber Studies, Volume 21). Koln: Rudiger Koppe Verlag.

Laghzaoui, M. (2008b). Développement de la langue académique chez des enfants marocains amazighes aux Pays-Bas. In J.J. de Ruiter (Ed.), *Langues et cultures en contact: Le cas des langues et cultures arabe et turque en France et aux Pays-Bas* (pp.123-147). Paris: L'Harmattan, Espaces discursifs.

Laghzaoui, M., & Kurvers, J. (2006). De ontwikkeling van schooltaalvaardigheid in het Tarifit en het Nederlands. Een eerste beschrijving van codering in het Tarifit. In: T. Koole, J. Nortier, & B. Tahitu (Eds.), *Bijdragen aan de 5e Sociolinguistische Conferentie* (pp. 330-341). Delft: Eburon.

Landry, S.H., Smith, K.E., Swank, P.R., & Miller-Loncar, C.L. (2000). Early maternal and child influences on children's later independent cognitive and social functioning. *Child Development*, *71*, 358-375.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, *16*, 307-22.

Leiber, L.L. (1999). *Wat gebeurt er in de stad?* [What happens in the city?]. Oosterhout: Aartselaar.

Leseman, P.P.M. (2000). Bilingual vocabulary development of Turkish-Dutch preschoolers. *Journal of Multilingual and Multicultural Development*, *21*(2), 93-112.

Leseman, P.P.M., Scheele, A.F., Mayo, A.Y., & Messer. M.H. (2007). Home literacy as special language environment to prepare children for school. *Zeitschrift Für Erziehungswissenschaft*, *10*(3), 334-355.

Leseman, P.P.M., & Van den Boom, D.C. (1999). Effects of quantity and quality of home proximal processes on Dutch, Surinamese-Dutch, and Turkish-Dutch preschoolers' cognitive development. *Infant and Child Development*, *8*, 19-38.

Lieven, E., Behrens, H., Spears, J., & Tomasello, M. (2003). Early syntactic creativity: A usage-based approach. *Journal of Child Language*, *30*, 333-370.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk* (3rd edition). Mahwah, NJ: Lawrence Erlbaum Associates.

Malvern, D.D., & Richards, B.J. (2000). Validation of a new measure of lexical diversity. In: M. Beers, B. Bogaerde, G. Bol, J. de Jong, & C. Rooijmans (Eds.), *From sound to sentence: Studies on first language acquisition*. Groningen: Centre for Language and Cognition, University of Groningen.

Malvern, D.D., & Richards, B.J. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language testing*, *9*(1), 85-104.

Malvern, D.D., Richards, B.J., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development: quantification and assessment*. Houndmills: Palgrave Macmillan.

Massey, S.L. (2004). Teacher child conversations in the preschool classroom. *Early childhood education journal*, *31*(4), 227-231.

McClelland, C. (2000). *The Interrelations of Syntax, Narrative Structure, and Prosody in a Berber Language* (Studies in Linguistics and Semiotics, Volume 8). Lewiston, NY: Edwin Mellen Press.

McIntyre, L.A. (1976). An investigation of the effect of culture and urbanization on three cognitive styles and their relationship to school performance. In: G.E. Kearney, & D.W. McElwain (Eds.), *Aboriginal*

cognition (pp. 231-256). Canberra: Australian Institute of Aboriginal Studies.

McKee, G., Malvern, D., & Richards, B.J. (2000). Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing*, *15*, 323-338.

Michaels, S. (1981). "Sharing time": Children's narrative styles and differential access to literacy. *Language in Society*, *10*, 423-442.

Møller, J.S., & Jørgensen, J.N. (2009). From language to languaging: Changing relations between humans and linguistic features. *Acta Linguistica Hafniensia*, *41*, 143-166.

Mommers, C., & Verhoeven, L. (1989). *Schatkist* [The treasury]. Tilburg: Zwijsen.

Nap-Kolhoff, E., & Van Steensel, R. (2005). Second language acquisition in preschool playgroups and its relation to later school success. *European Educational Research Journal*, *4*(3), 243-255.

Ninio, A. (1980). Ostensive definition in vocabulary learning. *Journal of Child Language*, *7*, 565-573.

Ninio, A. (1993). On the fringes of the system: Children's acquisition of the syntactically isolated forms at the onset of speech. *First Language*, *13*, 291-314.

Nortier, J. (1989). *Dutch and Moroccan Arabic in contact. Code switching among Moroccans in the Netherlands*. Dordrecht: Foris.

Obdeijn, H., De Mas, P., & Hermans, P. (1999). *Geschiedenis van Marokko* [The history of Morocco]. Amsterdam: Bulaq.

O'Loughlin, K. (1995). Lexical density in candidate output on direct and semi-indirect versions of an oral proficiency test. *Language Testing*, *12*(2), 217-237.

Opstap Opnieuw (1992). *Opstap Opnieuw programma*. Amsterdam: Averroès.

Pauli, L., & Schraerer, K. (2006). *Wie is er hier het dapperst?* [Who is the bravest?] (E. van de Vendel, Dutch Trans). Haarlem: Gottmer Uitgevers groep.

Pellegrini, A.D. (1985). The relations between symbolic play and literate behavior: A review and critique of the empirical literature. *Review of Educational Research*, *55*, 107-121.

Penner, Z., & Mueller, N. (1992). On the early stages in the acquisition of finite subordinate clauses. The syntax of the so-called preconjunctional

subordinate clauses in German, Swiss German and French. *Geneva Generative Papers*, *1*(2), 163-181.

Penninx, R., Münstermann, H., & Entzinger, H. (Eds.) (1998). *Etnische minderheden en de multiculturele samenleving*. Groningen: Wolters-Noordhoff.

Perkins, R.V. (1980). The evolution of culture and grammar. Ph.D. dissertation, State University of New York at Buffalo (University Microfilms 8016230).

Raviv, T.K., Kessenich, M., & Morrison, F.J. (2004). A mediational model of the association between socioeconomic status and three-year-old language abilities: The role of parenting factors. *Early Childhood Research Quarterly*, *19*, 528-547.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Reuning, H., & Wortley, W. (1973). Psychological studies of Bushmen. *Psychologia Africana*, monograph supplement, *7*.

Rosenquest, B.B. (2002). Literacy-based planning and pedagogy that supports toddler language development. *Early Childhood Education Journal*, *29*, 241-249.

Royaume du Maroc (1999). *Charte Nationale d'Education et de Formation. Commission Spéciale Education et Formation.*

Sachs, J. (1983). Talking about the there and then: The emergence of displaced reference in parent-child discourse. In: K.E. Nelson (Ed.), *Children's langua*ge (Volume 4). Hillsdale, NJ: Erlbaum.

Sarriondandia, P. (1905). *Gramática de la lengua rifeña*. Tangier: Imprenta Hispano-Arábiga.

Schachter, F.F. (1979). *Everyday mother talk to toddlers: Early intervention.* San Diego, CA: Academic Press.

Scheele, A. (2010). *Home language and mono- and bilingual children's emergent academic language: A longitudinal study of Dutch, Moroccan-Dutch, and Turkish-Dutch 3- to 6-year-olds*. Ph.D. dissertation, Utrecht University.

Schleppegrell, M. (2001). Linguistic features of the language of schooling. *Linguistics and Education*, *12*(4), 431-459.

Schleppegrell, M. (2004). *The language of schooling: A functional linguistic perspective*. London: Lawrence Erlbaum Associates.

Scribner, S. (1977). Modes of thinking and ways of speaking: Culture and logic reconsidered. In: P.N. Jonson-Laird, & P.C. Wason (Eds.), *Thinking:*

reading in cognitive science (pp. 483-500). Cambridge: Cambridge University Press.

Searle, J. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Shapiro, L.R., & Hudson, J.A. (1991). Tell me a make-believe story: Coherence and cohesion in young children's picture-elicited narratives. *Developmental Psychology*, *27*(6), 960-974.

Shapiro, L.R., & Hudson, J.A. (1997). Coherence and cohesion in children's stories. In: J. Costermans, & M. Fayol (Eds.), *Processing interclausal relationships: Studies on the production and comprehension of texts* (pp. 23-48). Mahwah, NJ: Erlbaum.

Sigel, I.E., & McGillicuddy-Delisi, A.V. (1984). Parents as teachers of their children: A distancing behavior model. In: A. Pelligrini, & T. Yawkey (Eds.), *The development of oral and written language in social contexts* (pp. 71-94). Norwood, NJ: Ablex.

Skinner, B.F. (1957). *Verbal Behavior*. Acton Massachusetts: Copley Publishing Group.

Sluyzer, B., & Oud, P. (1996). *Tijn op de fiets* [Tijn on the bicycle]. Rotterdam: 't Hooft.

Smith, M.W., & Dickinson, D.K. (1994). Describing oral language opportunities and environments in Head Start and other preschool classrooms. *Early Childhood Research Quarterly*, *9*, 345-366.

Smolak, L., & Weinraub, M. (1983). Maternal speech: Strategy or response? *Journal of Child Language*, *10*, 369-380.

Snow, C.E. (1977). Mothers' speech research: From input to interaction. In: C.E. Snow, & C.A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 31-49). London: Cambridge University Press.

Snow, C.E. (1983). Literacy and language: Relationships during the preschool years. *Harvard Educational Review*, *53*, 165-189.

Snow, C.E. (1991). The theoretical basis for the relationships between language and literacy in development. *Journal of Research in Childhood Education*, *6*, 5-10.

Snow, C.E. (1999). Social perspectives on the emergence of language. In: B. MacWhinney (Ed.), *The emergence of language* (pp.257-276). Mahwah, New Jersey: Lawrence Erlbaum.

Snow, C., Burn, M.S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Snow, C.E., Cancini, H., Gonzalez, P., & Shriberg, E. (1989). Giving formal definitions: An oral language correlate of school literacy. In: D. Bloome (Ed.), *Classrooms and literacy* (pp. 233-249). Norwood: Ablex Publishing Corporation.

Sokolov, J., & Snow, C.E. (1994). The changing role of negative evidence in theories of language acquisition. In: C. Galloway, & B. Richards (Eds.), *Input and interaction in language acquisition* (pp. 38-55). London: Cambridge University Press.

Sorsby, A.J., & Martlew, M. (1991). Representational demands in mothers' talk to preschool children in two contexts: Picture book reading and a modeling task. *Journal of Child Language*, *18*, 373-395.

Tabors, P., Snow, C., & Dickinson, D. (2001). Homes and schools together: Supporting language and literacy development. In: D. Dickinson, & P. Tabors (Eds.), *Beginning literacy with language*. Baltimore, MD: Brookes Publishing.

Teale, W.H. (1984). Reading to young children: Its significance for literacy development. In: H. Goelman, A.A. Oberg, & F. Smith (Eds.), *Awakening to literacy: The University of Victoria Symposium on Children's Response to a Literate Environment: Literacy before schooling* (pp. 110-122). Exeter, NH: Heinemann Educational Books.

Tesser, P.T.M., & Iedema, J. (2001). *Rapportage Minderheden 2001*. Den Haag: Sociaal en Cultureel Planbureau.

Tomasello, M. (2000). First steps in a usage based theory of language acquisition. *Cognitive Linguistics*, *11*, 61-82.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2004). What kind of evidence could refute the UG hypothesis? *Studies in Language*, *28*, 642-44.

Tomasello, M. (2006). Acquiring Linguistic Constructions. In: R. Siegler, & D. Kuhn (Eds.), *Handbook of Child Psychology*: *Cognitive Developmen*t. New Jersey: John Wiley & Sons, Inc.

Tomasello, M., Mannle, S., & Kruger, A. (1986). Linguistic environment of 1- to 2-year-old twins. *Developmental Psychology*, *22*, 169-176.

Tough, J. (1982). Language, poverty, and disadvantage in school. In: L. Feagans, & D.C. Farran (Eds.), *The language of children reared in poverty: Implications for evaluation and intervention* (pp. 3-18). New York: Academic Press.

Ure. J. (1971). Lexical density and register differentiation. In: G.E. Perren, & J. L.M. Trim (Eds.), *Applications of linguistics*. Cambridge: Cambridge University Press.

Van Amersfoort, H. (1982). *Immigration and the formation of minority groups: The Dutch experience 1945-1975*. Cambridge: Cambridge University Press.

Van den Boogert, N. (1995). *Catalogue des manuscrits arabes et berbères du Fonds Roux*. Aix-en-Provence: IREMAM.

Van den Boogert, N. (1998). *La révélation des énigmes. Lexiques arabo-berbères des XVIIe et XVIIIe siècles*. Aix-en-Provence: IREMAM.

Van Hout, R., & Vermeer, A. (2007). Rich and happy? Reliability and validity of lexical richness measures. In: H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 93-116). Cambridge: Cambridge University Press.

Van Kleeck, A., Gillam, R., Hamilton, L., & McGrath, C. (1997). The relationship between middle-class parents' book-sharing discussion and their preschoolers' abstract language development. *Journal of Speech, Language and Hearing Research*, *40*(6), 1261-1271.

Van Kleeck, A., Vigil, A., & Beer, N. (1998). *A longitudinal study of maternal book-sharing emphasis on print form and print meaning with preschoolers*. Scientific Technical Presentation at the American, Speech, Language, hearing Association Annual convention. San Antonio, TX.

Van Ommen, S. (2003). *De verrassing* [The Surprise]. Rotterdam: Lemniscaat.

Veldkamp, T., & Boer, K. (2004). *Tim op de tegels* [Tim on the paving stones]. Houten: Van Goor.

Verhoeven, L., Narain, G., Extra, G., Konak, O.A., & Zerrouk, R. (1995). *Diagnostische Toets Tweetaligheid* [Test for bilingual Development]. Arnhem, The Netherlands: CITO.

Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language testing*, *17*(1), 65-83.

Vernon-Feagans, L.V. (1996). *Children's talk in communities and classrooms*. Cambridge, MA: Blackwell Publishers.

Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: M.I.T Press.

Waechter, P. (2007). *Emma in het spookhuis* [Emma in the ghost train]. Amsterdam: Em. Querido's Uitgeverij B.V.

Ward, M.C. (1971). *Them Children: A study in language learning*. New York: Holt, Rinehart & Winston.

Weizman, Z.O., & Snow, C.E. (2001). Lexical input as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology*, *37*(2), 265-279.

Wetenschappelijke Raad voor het Regeringsbeleid (1979). *Rapport Etnische minderheden*. Den Haag.

Yusuf Ali, A. (1998). *Translation of the Koran*. Holly NJ: Islamic Educational Services.

Zafrani, H. (1998). *Deux mille ans de vie juive au Maroc: Histoire et culture, religion et magie*. Paris: Maisonneuve & Larose.

Ziani, A. (1993). *Ad arigh deg wezru* [In steen zal ik schrijven]. Amsterdam: Stichting Izaourane.

Ziani, A. (1997). *Ij n tliwliwt i mulay* [Jubelzang voor de bruidegom]. Amsterdam: Stichting Al-Hizjra.

APPENDIX 1

Example of a home transcript

@Begin
@Languages: Berber
@Participants: CHI fatima zohra Target_Child, MOT mother Target_Mother
@ID: be|dash: 340|CHI|3;1.27|female|berber|||Target_Child|||
@Birth of CHI: 05-APR-2002
@Language of CHI: be, nl
@Language of MOT: be, nl
@Session: 1
@Transcriber: mohammadi
@Date of recording: 01-JUN-2005
@Location: Utrecht, the Netherlands
@Activities: book reading
@Situation: mother and child sitting around a table reading the book "Tijn op de Fiets"
@Time Duration: 00:00:00-00:09:30
@Filename: fatimazohra1.book.cha
*MOT: &ehm aqqem a fatima zohra fiets@nl.
*MOT: ja@nl &mic ja@nl mamec da s qqaren?
*MOT: aqqem a fatima zohra tijn.
*MOT: qa y-nnufar awarni &ehm ## awarni kapstok@nl.
%act: mother points to picture
*MOT: da toch@nl?
*MOT: &ehm qa i-rezzu oawed x tisira.
*MOT: qa t-rezzu xa s yemma s.
*CHI: h2enna s.
*MOT: eh?
*CHI: h2enna s.
*MOT: nee@nl yemma s.
*CHI: yemma s.
*MOT: t-rezzu xa s yemma s.
*CHI: tisira.
*MOT: ja@nl.
*MOT: qa i-rezzu x tsira nnes.
*MOT: aqqa t-tfit id ad as t-qqen tisira nnes.
*MOT: ad rah2-en ghar h2enna s.
*CHI: wah.
*MOT: t-qn-as tisira ad rah2en ghar h2enna s.
*CHI: ghar h2enna s.
*MOT: ja@nl.
*MOT: aqqem oawed.
*MOT: aqqem <t-ggi-t-id> [/] tggit x fiets@nl.

*MOT: t-gg-as wahit tnayen trata.
*MOT: aqqem t-sseny-it.
*MOT: aqqa pop@nl nnes # t-rezzu x pop@nl nnes?
*MOT: mani y-lla pop@nl nnes.
%act: child nods.
*CHI: xxx.
*MOT: <qa t-nya> [/] ny-in x fiets@nl qa tt-eg-en tren [x 3].
*MOT: aqqem da.
*MOT: t-egg-en tren [x 3] s fiets@nl.
*CHI: tren [x 2].
*MOT: ja@nl.
*CHI: fiets@nl.
*MOT: ja@nl.
*CHI: <ik ook tren [x 2] fiets> [$nl] da.
*MOT: ja@nl.
*MOT: aqqem t-nna s oh mech2ar n tunubil-at.
*MOT: bezzaf n tunubil-at.
*MOT: aqqem wa t-egg-n as &ehm beslama.
*MOT: beslama.
*MOT: i tunubil iwin.
*MOT: t-egg-en as beslama.
*CHI: t-egg-en beslama.
*MOT: ja@nl.
%com: mother turns the page
*MOT: aqqem aq ssagh-en-t &ehm aardbei@nl.
*CHI: welk@nl.
*MOT: aqqa-t.
*MOT: ssaghen-t id wanita.
*CHI: waar@nl?
*MOT: ssaghen-t id toch@nl!
*CHI: die@nl!
*MOT: ja@nl.
*MOT: uca <t-uca-s &ehm> [//] t-uca-s tbanant.
*MOT: t-uca-s tbanant uca i-nna-s cucran.
*CHI: cukran.
*MOT: ja@nl i-nna-s cukran.
*CHI: aqqa da macina.
*MOT: ja@nl qqar-en as beslama i &ehm i trein@nl.
*CHI: beslama.
*CHI: t-ufa &eh aqzin.
*MOT: ja@nl.
@End

APPENDIX 2

Example of a school transcript

@Begin
@Languages: nl
@Participants: TEA Eva Teacher, CHI Anisa Target_Child, CLA Class_mate Child
@ID: nl|dash:school code|CHI|04;02.26|female|berber|Target_Child
@Birth of CHI: 30-JAN-2002
@Language of CHI: be, nl
@Language of TEA: nl
@Session: 1
@Transcriber: Véronique Verhagen
@Date of recording: 25-APR-2006
@Location: Utrecht, the Netherlands
@Activities: Book reading
@Situation: The teacher sits togethers with all toddlers in a book reading session
@Time Duration: 00:00:00-00:08:00
*TEA: wat zien we hier?
*CHI: konijn.
*TEA: een.
*CHI: schaap.
*CHI: een cadeautje.
*TEA: &oh, waar zie je +/.
*CHI: daar bij achter.
*TEA: waar zie je een cadeautje?
*CHI: hier.
*TEA: ja.
*TEA: ew gaan kijken.
*TEA: schaap op een motor.
*TEA: dat is spannend.
*CHI: &xxx rijden.
*TEA: hier is schaap.
*TEA: nou.
*CHI: schaap gaat op de wegen.
*CHI: hoe dik is hij.
*TEA: ja.
*TEA: wat goed van jou!
*TEA: schaap gaat wegen.
*TEA: want hij denkt +/.
*TEA: ik heb zoveel dikke wol.
*TEA: ik ben misschien nu wel heel zwaar.
*CHI: &oh maar optellen.
*TEA: en dan gaat hij op de weegschaal staan.
*TEA: om te kijken hoe zwaar hij is.

*TEA: en wat gaat schaap dan doen?
*CLA: spiegel kijken.
*TEA: goed zo, van jou.
*TEA: goed Wafik.
*TEA: ja.
*TEA: want hij wil in de spiegel zien.
*TEA: om te kijken of hij echt zo dik is.
*CHI: op &xxx ook.
*CHI: &xxx buik is ook &xxx .
*TEA: ja.
*TEA: en schaap gaat ook nog meten.
*TEA: met streepjes.
*TEA: om te kijken of hij echt zo zwaar is geworden.
*TEA: en wat gaat schaap dan doen?
*CHI: motorrijden.
*CHI: ja om naar school te gaan.
*CHI: rijden.
*TEA: zou kunnen.
*TEA: als je twee flessen haalt.
*TEA: hoef je er nog maar eentje te betalen.
*TEA: en eentje hoef je maar centjes voor te betalen.
*TEA: nou dat is wel heel erg fijn.
*CHI: <moet> [/] moet hij even weten.
*TEA: en welke kleur zou <ze> [///] schaap nu kopen?
*CHI: oranje.
*TEA: is een beetje naar rood toe &he .
*TEA: dat ze die gaat kopen.
*TEA: en dan met de scooter +...
*CHI: gaat hij &xxx deze [/] deze wielen &xxx +/.
*TEA: hij heeft de verf &mee naar huis genomen [: meenemen].
*TEA: want die had ze gekocht in de verfwinkel.
*CHI: gaat ze haar buik verven.
*TEA: en wat gaan ze verven?
*CHI: op haar buik.
*TEA: en wat nog meer?
@End

APPENDIX 3

List of abbreviations and symbols used in the glosses

1	first person
2	second person
3	third person
S	singular
P	plural
M	masculine gender
F	feminine gender
M/F	masculine or feminine
PERF	perfective
IMPERF	imperfective
AOR	aorist
FUT	future
CL	clitic
COP	copula
IND	indirect
OBJ	object
NEG	negative
P.A.	predicative presentative auxilary
T.A.	predicative temporal auxilary
PAST	past tense
PRES	present
PRO	pronoun
PROG	progressive
POSS	possessive

APPENDIX 4

Table 1: Individual scores, means and standard deviations of mothers' tokens, lexical density and lexical diversity during the book task in three measurement times

	Tokens			Density			Diversity		
Time	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3
Age child	3;2	4;2	5;10	3;2	4;2	5;10	3;2	4;2	5;10
Mother									
1	1331	612	935	2.81	2.88	3.29	87.23	91.55	76.55
2	2183	563	1164	3.34	1.71	3.14	66.86	48.00	119.30
3	822	732	672	1.75	2.91	1.92	73.72	64.43	62.94
4	978	1619	703	3.04	3.44	1.78	119.25	122.18	91.18
5	982	910	649	2.35	3.25	2.90	58.34	113.26	85.74
6	1432	1335	784	2.36	5.71	3.81	101.52	115.13	91.10
7	1488	1007	274	2.61	2.16	1.94	70.55	34.86	47.55
8	363	1001	860	2.11	2.54	2.80	125.30	82.21	61.25
9	574	690	252	2.05	3.07	1.86	63.44	59.70	71.92
10	1188	890	253	2.35	2.86	2.44	81.02	67.34	87.16
11	993	817	713	3.22	2.22	3.13	71.36	66.60	77.80
12	1826	1233	1333	2.52	2.92	3.84	112.49	76.72	81.55
Mean	1180	950.9	716	2.54	2.97	2.74	85.92	78.50	79.50
SD	511.9	321.9	342	.48	0.99	.74	23.09	27.46	18.23

Table 2: Individual scores, means and standard deviations of mothers' tokens, lexical density and lexical diversity during the picture task in three measurement times

	Tokens			Density			Diversity		
Time	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3
Age child	3;2	4;2	5;10	3;2	4;2	5;10	3;2	4;2	5;10
Mother									
1	406	1239	1073	1.29	1.46	1.76	82.63	109.66	12.05
2	961	746	765	1.96	1.42	1.75	70.75	70.85	70.53
3	261	1076	404	.47	1.67	1.21	17.96	67.11	30.25
4	433	709	299	1.04	2.34	.69	93.73	95.50	70.29
5	466	837	390	.99	1.36	.45	25.57	37.17	28.90
6	495	1210	342	1.48	1.41	1.23	32.98	45.95	84.66
7	1130	462	811	1.39	.66	1.89	63.85	12.09	70.18
8	1038	1426	342	1.28	1.61	1.67	67.05	109.74	59.99
9	671	425	197	1.28	1.48	1.06	52.27	49.63	29.33
10	733	535	207	1.48	1.10	1.00	61.27	36.82	17.57
11	462	589	351	1.39	1.85	1.09	61.80	54.04	43.00
12	976	804	325	1.42	1.48	1.13	83.58	46.84	63.09
Mean	699.3	838.2	458.8	1.29	1.49	1.24	59.45	61.28	57.49
SD	292	329.4	272.5	.35	.40	.45	23.61	30.48	29.53

Table 3: Individual scores, means and standard deviations of teachers' tokens, lexical density and lexical diversity during the book task in two measurement times

	Tokens		Density		Diversity	
Time	T 2	T 3	T 2	T 3	T 2	T 3
Age child	4;2	5;10	4;2	5;10	4;2	5;10
Teacher						
1	1674	1889	4.31	2.89	67.75	91.91
2	900	1493	1.58	2.69	84.65	96.09
3	1064	1144	1.63	2.35	65.83	88.46
4	1345	1527	3.35	2.14	64.02	74.60
5	1045	1230	2.92	2.39	92.41	97.00
6	1201	1482	2.53	2.01	67.13	80.99
7	1494	1511	2.15	2.49	66.43	86.03
8	1192	1587	3.38	2.60	65.78	78.15
9	2009	1342	3.93	3.35	79.74	88.79
10	1566	1566	1.63	2.18	47.74	78.94
11	1328	1219	2.25	2.80	78.87	80.95
12	1690	1012	2.98	3.02	78.28	82.17
Mean	1375.7	1416.8	2.72	2.58	71.55	85.34
SD	321.8	238	.91	.40	11.76	7.22

Table 4: Individual scores, means and standard deviations of teachers' tokens, lexical density and lexical diversity during the circle time in two measurement times

	Tokens		Density		Diversity	
Time	T 2	T 3	T 2	T 3	T 2	T 3
Age child	4;2	5;10	4;2	5;10	4;2	5;10
Teacher						
1	915	1049	1.15	2.43	62.38	81.10
2	1003	1371	1.70	2.07	76.46	90.94
3	790	992	1.83	1.99	75.17	83.10
4	677	948	1.82	1.23	65.51	67.67
5	1326	1105	2.22	1.99	74.88	78.13
6	1317	760	2.34	.83	83.78	97.90
7	980	1339	1.90	2.14	78.47	87.20
8	1376	1449	1.82	2.02	76.11	92.69
9	966	1398	1.52	2.17	79.08	97.70
10	965	1078	1.82	2.07	52.24	90.78
11	1444	1489	2.37	2.17	85.01	89.32
12	1080	1224	1.77	2.27	104.81	101.93
Mean	1069.9	1183.5	1.86	1.95	76.16	88.21
SD	243.4	229.2	.34	.46	12.97	9.58

Table 5: Individual scores, means and standard deviations of children's tokens, lexical density and lexical diversity during the book task in three measurement times at home

	Tokens			Density			Diversity		
Time	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3
Age child	3;2	4;2	5;10	3;2	4;2	5;10	3;2	4;2	5;10
Child									
1	131	191	310	.84	1.93	1.97	33.46	90.25	54.05
2	141	213	297	1.17	1.80	2.15	26.65	21.82	68.55
3	377	187	326	.84	2.23	1.60	46.74	66.05	98.38
4	234	296	178	1.33	1.44	.62	–	78.69	116.78
5	30	202	67	.77	1.65	1.20	60.44	71.39	45.70
6	268	300	24	1.79	1.48	1.17	53.62	–	–
7	235	326	93	1.73	1.68	1.50	32.99	27.71	–
8	137	482	240	1.14	2.51	1.86	96.59	33.31	135.23
9	95	106	500	1.03	1.56	2.45	–	64.93	62.18
10	50	42	46	.58	1.27	1.17	–	–	–
11	98	279	137	1.36	1.67	1.78	44.86	88.34	37.30
12	266	67	56	1.67	1.12	.77	67.12	–	–
Mean	171.8	224.3	189.5	1.19	1.70	1.52	51.39	60.28	77.27
SD	103.7	122.7	146.7	.40	.39	0.56	21.55	26.15	35.46

Table 6: Individual scores, means and standard deviations of children's tokens, lexical density and lexical diversity during the picture task in three measurement times at home

	Tokens			Density			Diversity		
Time	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3
Age child	3;2	4;2	5;10	3;2	4;2	5;10	3;2	4;2	5;10
Child									
1	112	456	862	.86	1.21	1.72	–	48.69	53.04
2	137	241	600	.72	.92	2.77	–	28.75	64.68
3	487	126	304	.70	1.04	2.46	29.49	–	38.02
4	97	150	343	1.07	.66	1.53	57.05	38.70	57.56
5	71	142	181	.50	.45	.76	–	26.60	60.67
6	182	558	88	1.00	1.10	.77	57.50	63.02	38.40
7	232	449	459	.66	1.43	1.42	38.36	43.92	18.65
8	211	421	266	.78	1.10	1.08	53.55	67.07	42.33
9	366	573	181	1.10	1.70	1.47	67.29	31.20	26.04
10	201	65	207	.78	.76	1.09	81.17	–	23.99
11	411	309	213	1.32	1.43	1.16	85.48	66.82	88.01
12	240	324	268	1.04	1.26	1.89	63.08	65.43	162.84
Mean	228.9	317.8	331	.89	1.09	1.51	59.22	48.02	56.19
SD	130.2	174.7	216	.23	.35	.63	18.06	16.54	38.95

Table 7: Individual scores, means and standard deviations of children's tokens, lexical density and lexical diversity during the book task in two measurement times at school

	Tokens		Density		Diversity	
Time	T 2	T 3	T 2	T 3	T 2	T 3
Age child	4;2	5;10	4;2	5;10	4;2	5;10
Child						
1	12	66	1.33	1.88	–	–
2	145	142	1.31	1.76	56.90	26.05
3	10	12	1.33	1.00	–	–
4	32	213	1.29	1.76	33.10	40.12
5	22	46	1.00	2.00	–	–
6	155	394	1.36	2.12	29.25	65.06
7	70	114	1.00	1.55	27.55	20.74
8	21	68	1.00	1.25	–	–
9	134	122	2.09	1.38	58.46	60.05
10	109	67	.74	.88	55.63	57.85
11	41	86	1.00	1.30	–	–
12	16	34	1.00	1.33	–	–
Mean	63.92	112.17	1.20	1.52	43.48	44.98
SD	56.32	104.09	.34	.39	14.94	18.79

Table 8: Individual scores, means and standard deviations of children's tokens, lexical density and lexical diversity during the circle time in two measurement times at school

	Tokens		Density		Diversity	
Time	T 2	T 3	T 2	T 3	T 2	T 3
Age child	4;2	5;10	4;2	5;10	4;2	5;10
Child						
1	28	88	1.25	1.44	–	35.50
2	77	187	1.36	1.81	–	30.91
3	97	337	2.08	1.97	50.13	106.16
4	16	144	1.50	1.38	–	57.13
5	82	157	.55	1.39	–	48.38
6	111	38	.94	1.27	51.00	49.16
7	101	201	2.00	1.61	–	–
8	172	86	2.27	4.00	39.80	–
9	187	76	1.20	1.50	65.17	52.15
10	62	54	.93	.33	48.11	45.54
11	24	46	.86	1.83	–	–
12	55	6	1.17	1.00	–	–
Mean	84.33	118.33	1.34	1.63	50.84	53.12
SD	54.17	92.30	.53	.86	9.15	23.08

APPENDIX 5

Table 1: Individual scores, means and standard deviations of mothers' abstraction levels during the book task in three measurement times at home

	Matching			Selecting			Reordering			Reasoning		
Time	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3
Age child	3;2	4;2	5;10	3;2	4;2	5;10	3;2	4;2	5;10	3;2	4;2	5;10
Mother												
1	28	15	19	35	37	20	13	1	3	6	1	3
2	27	9	15	44	43	26	19	4	5	10	3	3
3	33	5	9	33	29	16	5	0	0	1	0	0
4	18	26	3	33	37	6	29	18	15	12	7	10
5	30	3	5	31	28	12	9	0	2	3	0	2
6	22	21	4	41	34	16	24	9	12	14	4	7
7	25	34	19	37	29	17	1	1	0	0	0	0
8	27	12	9	33	34	15	22	5	10	11	5	5
9	32	9	35	35	8	17	1	1	1	0	0	0
10	33	4	34	32	39	21	1	2	0	1	0	0
11	29	8	36	29	54	36	0	6	0	0	1	0
12	20	9	1	54	33	16	26	11	14	13	5	7
Mean	27	12.92	15.75	36.42	33.75	18.17	12.50	4.83	5.17	5.92	2.17	3.08
SD	4.95	9.52	13	6.95	10.81	7.41	11.04	5.47	5.91	5.70	2.52	3.45

Table 2: Individual scores, means and standard deviations of teachers' abstraction levels during the book task in two measurement times

	Matching		Selecting		Reordering		Reasoning	
Time	T 2	T 3	T 2	T 3	T 2	T 3	T 2	T 3
Age child	4;2	5;10	4;2	5;10	4;2	5;10	4;2	5;10
Teacher								
1	14	18	41	45	20	22	11	13
2	10	14	27	33	19	26	10	11
3	19	22	37	40	19	25	16	15
4	12	18	31	39	22	27	10	17
5	16	23	32	38	18	19	12	14
6	19	18	31	36	22	20	15	16
7	31	25	52	45	19	23	17	19
8	13	20	38	41	20	25	12	14
9	14	17	21	37	20	27	16	14
10	40	28	51	44	23	30	15	16
11	8	14	24	30	23	25	19	22
12	6	19	19	42	20	19	15	18
Mean	16.83	16.67	33.67	39.17	20.42	24	14	15.75
SD	9.74	4.21	10.66	4.69	1.68	3.46	2.92	2.96

Table 3: Individual scores, means and standard deviations of children's abstraction levels during the book task in three measurement times at home

	Matching			Selecting			Reordering			Reasoning		
Time	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3	T 1	T 2	T 3
Age child	3;2	4;2	5;10	3;2	4;2	5;10	3;2	4;2	5;10	3;2	4;2	5;10
Child												
1	10	15	11	12	12	9	4	2	1	2	2	1
2	14	13	6	11	13	10	5	3	2	3	3	2
3	12	3	6	16	9	6	2	1	1	1	1	2
4	8	7	5	10	8	7	8	11	11	3	5	5
5	12	4	4	13	9	7	3	1	1	2	1	1
6	9	13	8	17	6	3	4	7	7	3	3	4
7	11	8	9	12	17	13	0	1	2	0	1	2
8	13	11	9	9	12	9	6	6	3	3	4	5
9	12	10	7	16	9	11	1	1	1	0	0	1
10	14	12	10	10	4	7	3	0	0	1	0	1
11	16	13	6	12	8	5	2	1	0	1	0	0
12	11	5	6	26	5	4	4	10	9	2	4	5
Mean	11.83	9.50	7.25	13.67	9.33	7.58	3.50	3.67	3.17	1.75	2.00	2.42
SD	2.25	4.01	2.14	4.66	3.68	2.94	2.20	3.85	3.71	1.14	1.76	1.83

Table 4: Individual scores, means and standard deviations of children's abstraction levels during the book task in two measurement times at school

	Matching		Selecting		Reordering		Reasoning	
Time	T 2	T 3	T 2	T 3	T 2	T 3	T 2	T 3
Age child	4;2	5;10	4;2	5;10	4;2	5;10	4;2	5;10
Child								
1	5	4	8	6	2	3	1	3
2	4	5	7	5	3	3	2	2
3	4	5	6	5	2	2	2	1
4	4	5	5	4	2	2	1	3
5	2	4	3	3	2	3	1	2
6	3	4	7	6	3	3	2	2
7	4	7	2	6	1	2	0	0
8	2	3	1	3	2	4	2	2
9	4	5	2	7	2	2	0	1
10	3	5	5	6	2	3	1	0
11	2	4	1	6	1	3	0	1
12	5	6	3	3	3	5	2	3
Mean	3.50	3.50	4.17	5	2.08	2.92	1.17	1.67
SD	1.09	1.06	2.48	1.41	.67	.90	.83	1.07

SUMMARY IN ENGLISH

Emergent academic language at home and at school

This dissertation reports on a longitudinal study on the development of academic language of 3-6-year-old Moroccan Berber children in communicative contexts at home and at school in the Netherlands. It is one of the in-depth studies forming part of a larger project entitled the "Development of Academic Language at School and at Home" (DASH).

This book consists of eight chapters. Chapter 1 describes the overall DASH project. This is a large scale collaborative research program carried out at three universities in the Netherlands: Tilburg University, Amsterdam University and Utrecht University. The project consisted of one comprehensive study and three in-depth studies. The comprehensive study involved three ethno-linguistic groups, namely Dutch, Turkish and Moroccan Berber. The in-depth studies each focused on a subsample of these groups. This book reports on the in-depth study of the Moroccan Berber group. In addition, this chapter deals with the scientific and societal relevance of the study. Over the last decades, many studies have concentrated on different aspects of school language development, such as narrative skills, technical vocabulary development, and emerging literacy. However, the topic of the acquisition and early development of specific linguistic structures typical of school language, in a first mother tongue has been barely looked at, let alone in a language acquired and used as a second language. This concerns the acquisition of specific lexical, morphological and grammatical structures, which make complex language constructions possible, and accordingly facilitate the development of the academic language register. Different studies have shown that Moroccan and Turkish pupils are lagging far behind their native Dutch peers in school achievements in Dutch primary education. This unfavourable educational position is usually associated with the fact that the children are exposed to a language other than Dutch at home. We assume, however, that there are other factors that might explain educational

underperformance, such as the way they acquire and command academic language skills. This study is an attempt to build on existing findings to investigate the social co-construction of the academic language register of Moroccan Berber children in the Netherlands.

Chapter 2 gives an overview of the position of Berber and its speakers in Morocco and in the Netherlands. First, a description is given of the Berber language structure itself, with a focus on Tarifit. Then, the attitudes towards the languages spoken in Morocco are briefly described and some information provided regarding language policy and language use in Morocco. After that, the position of the Berber language and the Berber community in the Netherlands are described. Two dimensions are distinguished in this respect: the status of the speakers of Berber in the Netherlands and the status of Berber as a language in Dutch society. The first part sheds light on the demographic and immigration background of the Berber community and its members' position in the Dutch educational system and labour market. The second part considers the *status quo* of the Berber language itself in the Netherlands. Aspects examined are the position of Berber in social and cultural life of Moroccans, mass media, religion and education. Finally, Chapter 2 deals with language acquisition, language proficiency, language choice, language dominance, and language preference within the Berber community in the Netherlands.

Chapter 3 deals with the theoretical framework of the study. Firstly, the role of parent-child interaction and input are discussed. Children learn language in a social context in which they interact with their parents. The content of these interactions and more specifically the input parents provide to their children, is crucial in the language learning process that children go through. Children whose social experiences provide more communicative opportunities and richer input build their vocabulary at a faster rate than children with less communicative experiences and poorer input. Thereafter, the notion of academic language is discussed. Academic language can be defined as the use of language for specific, decontextualized and cognitively demanding communication in school-like tasks. For many children, this register is very different from the language used at home. It involves communicating about subjects and things that do not belong to here-and-now reality, mostly without the support of concrete contextual or immediate elements. Academic language encompasses lexical, morpho-syntactic and textual characteristics, which enable a structurally

more elaborated and more abstract communication, whereas informal interactional language, such as at home, is more context-dependent.

Chapter 4 deals with the research design and methodology. The objective of the current study is to examine the co-construction of academic language of Moroccan Berber children in communicative contexts at home and at school. The following main research questions are dealt with:

1 To what extent do Moroccan Berber mothers use features of academic language in their input to their preschool children between the age of 3 and 6?
2 Is the use of academic language features by the mothers related to SES and home literacy practices?
3 To what extent do teachers use features of academic language in their input to Moroccan Berber school children?
4 To what extent do children use features of academic language between the ages of 3 and 6 in the home and school settings?
5 Is there a difference between mothers and teachers regarding the use of academic language features?
6 Is there a relationship between the parents' and teachers' input and the children's use of academic language features?

This study involved 12 Moroccan Berber children (four boys and eight girls) and their caretakers at home and at school, living in the Netherlands in the cities of Utrecht, Ede, Tilburg and Oosterhout. The children were followed during a period of 3 years, from the age of 3 to nearly 6. Different instruments were used in this study – interaction tasks designed for mother-child and teacher-child conversations, a detailed questionnaire dealing with family background, language use and literacy activities in the home context, a short questionnaire for the teachers and vocabulary tests in the first (Tarifit Berber) and second (Dutch) language. The data collection consisted of three measurement times at home and two at school. Concerning the home visits, the first measurement time took place when the children were 3;2 years old, the second measurement time at the age of 4;2, and the third measurement time at the age of 5;10. A total of 60 videotaped sessions of mother-child and teacher-child conversations resulted in a language corpus (120 transcripts and 15,509 utterances) which has been transcribed for the current study. All the home and school data were transcribed verbatim. The interactions have been transcribed

using the *Child Language Data Exchange System* programme (CHILDES) and coded according a coding scheme developed by the DASH research group. The DASH coding scheme contains four main levels: the lexical level, the morpho-syntactic level, the textual level and the socio-pragmatic level.

Chapter 5 deals with the lexical components of academic language. Quantity and quality variables in the language input of mothers and teachers as well as of the children were studied during two different interaction tasks at different times at home and at school. It emerged that mothers produced more talk during the structured task, book reading, than during the picture task. In addition, in the three measurement times, mothers showed a higher lexical density and lexical diversity during the book task than when involved in the picture task. Concerning the relationship between the academic language features in mothers' input and their family background, lexical density and diversity showed a significant correlation with SES and literacy in both tasks in almost all measurement times. As regards the input at school, teachers tended to talk more when they were involved in book reading than when involved in the circle time. School data also demonstrated that teachers talk more and use a richer vocabulary as children grow older. As far as the children are concerned, they seem to use more dense and more diverse language during the book task than during the picture task. As expected, the book tasks, usually associated with schooling, facilitate the use of a richer vocabulary. Looking at the influence of SES and literacy, a positive and sometimes significant impact of both family background factors were found during the book task. Lexical diversity in particular was strongly correlated with the socio-economic status of the family in all measurement times. In the school setting, children produced less utterances than in the home setting. On the basis of these findings, it can be concluded that the language input of the caretakers at home and at school contained many aspects of the lexical features of academic language. Furthermore, children showed a rapid increase on many lexical aspects in the home as well as in the school context particularly as they start their school career.

Chapter 6 deals with morpho-syntax and its relation to the development of the academic language of Moroccan Berber children. The focus is on clause combining strategies in the input of the mothers and teachers as well as in the language production of the children. Concerning the input at home, it turned out that the most frequent category of subordinate clauses in mothers' speech was complement clauses, across all measurement times. We also found that the

frequency of syntactically complex constructions in mothers' input does not change as children get older. Furthermore, Our data showed that SES and literacy background were partially related to the composite variables of mothers' syntactic complexity. Especially in measurement time 3 when children reach the age of 5, SES level and literacy practices correlated significantly with the syntactic complexity of the mothers' speech. With respect to the teachers' input, the data showed that they, in both measurement times, tended to use coordinate and complement clauses more than relative and adverbial clauses. It also turned out that teachers, almost at the level of all types of multi-clause sentences, produced higher numbers than the mothers. Concerning children, our data showed that they produced a low number of complex sentences at different ages at home as well as at school. However, by the age of 5, children's utterances clearly become more complex. Comparing the home and school contexts, children tend to use more complex constructions at school than at home.

The central theme of Chapter 7 is the level of abstraction or decontextualized language both in the input provided by mothers and teachers and in the language used by the children. Regarding the mothers' use of abstraction levels, the results revealed that mothers mainly used the selective analysis and matching levels of abstraction, which consist of descriptive labeling and talking about actual information. Furthermore, we found a large amount of variation among the mothers in the number of reordering and reasoning utterances they used during their conversations with their children. The two components of decontextualized language (reordering and reasoning) correlated positively with SES and literacy, which suggest that there is a significant relationship between the mothers' socio-economic background and their input. As for the teachers, the results show that the most frequent level of abstraction is selective analysis. A comparison between the mothers' and the teachers' input showed that teachers tend to make use of decontextualized talk more often than did mothers. The analysis of children's data showed that they tended to produce decontextualized speech relatively more frequently at school than at home.

Chapter 8 reflects upon the main findings by giving answers to the research questions formulated in Chapter 4. In addition, the implications of this study, future directions for research and the limitations of the study are discussed. The results presented indicate that there were occurrences of academic language in the mothers' and teachers' input. However, many differences were found

between the mothers with respect to the three domains under study (lexical, morpho-syntactic and textual). The relationship between family background characteristics and academic language features in maternal input showed many differences. The most notable finding is the strong relationship between the use of decontextualized language and SES and literacy. With regard to the input-output relationship in the school setting, in comparison with the mothers in the home setting, less significant correlations were found between the input of the teachers and children's language production. It should be noted that the language behavior of the target children at home was characterized by a mixture of two languages, Berber and Dutch. Instead of saying that these children use Tarifit Berber or Dutch or both languages at the same time, we can say that these early age *languagers* are *languaging* to their mothers. An important implication of the study is that both parents and educators should be aware of the kind of language input they provide to children and how to build on it. A stronger focus should be placed on the type of language activities which enhance academic language acquisition. The role of school is to offer children a more structured and organized way of learning to foster academic language. Next to the knowledge they receive in everyday and out-of-school experiences, children need to be introduced to a more structured and organized knowledge system offered in the classroom.

SUMMARY IN DUTCH
(samenvatting in het Nederlands)

Ontluikende academische taal thuis en op school

Een longitudinale studie van 3-6 jarige Marokkaans-Berberse kinderen in Nederland

In dit proefschrift wordt verslag gedaan van een longitudinaal onderzoek naar de wijze waarop de ontluikende academische taal (of 'schooltaalvaardigheid) van 3-6 jarige Marokkaans-Berberse kinderen in Nederland zich ontwikkelt en hoe die in communicatieve contexten thuis (zowel in het Berbers als ook in het Nederlands) en op school (in het Nederlands) worden gevormd. Dit onderzoek is een van de dieptestudies van een groter project met als titel 'the Development of Academic Language at School and at Home' (DASH) (de ontwikkeling van academische taalvaardigheid op school en thuis).

Deze dissertatie bestaat uit 8 hoofdstukken waarvan 3 empirische hoofdstukken. Hoofdstuk 1 beschrijft het DASH project: een grootschalig onderzoeksprogramma in samenwerkingsverband uitgevoerd op drie universiteiten in Nederland: Universiteit van Tilburg, Universiteit van Amsterdam en Universiteit van Utrecht. Het DASH project is een longitudinaal en interdisciplinair onderzoek naar de ontwikkeling van academische taal van 3-6 jarige Nederlandse, Marokkaans-Berberse en Turkse kinderen. Het project bestaat uit één brede studie en drie dieptestudies. De brede studie omvat alle drie de etnisch-linguïstische groepen. De dieptestudies richten zich elk op een steekproef van deze groepen.

Deze dissertatie is een verslag van de dieptestudie over de Marokkaans-Berberse groep. In de afgelopen decennia zijn vele studies verricht naar verschillende aspecten van schooltaalontwikkeling, zoals narratieve vaardigheid, technische woordenschat ontwikkeling, en geletterdheid. Echter, relatief weinig studies zijn er gedaan naar de verwerving en vroege ontwikkeling van specifieke taalkundige structuren, typerend voor de schooltaal, in de eerste taal. Nog

minder onderzoek richt zich op de linguïstische aspecten van de verwerving van schooltaalvaardigheid in een tweede taal. Dit betreft de verwerving van specifieke lexicale, morfologische en grammaticale structuren, waardoor complexe taalconstructies mogelijk worden gemaakt, en dus de ontwikkeling van het academische taalregister. Uit recente gegevens van schoolloopbanenonderzoek in het primair onderwijs in Nederland blijkt dat Marokkaanse en Turkse leerlingen als groep de zwakste positie innemen. Hoewel het feit dat thuis een andere taal gesproken wordt dan het Nederlands vaak als een belangrijke verklaring wordt genoemd, ligt de kwestie ingewikkelder. Er zijn aanwijzingen dat het bij die kennelijk hardnekkige achterstanden vooral gaat om de verwerving van een specifiek register van het Nederlands, namelijk de academische taal. Deze studie bouwt voort op bestaande bevindingen om de sociale co-constructie van het academische taalregister van de Marokkaans-Berberse kinderen in Nederland te onderzoeken.

Hoofdstuk 2 geeft een overzicht van de positie van de Berberse taal en de sprekers daarvan in Marokko en in Nederland. Eerst wordt een beschrijving gegeven van de taalstructuur zelf, met een focus op de variant Tarifit. Vervolgens worden de attitudes ten opzichte van talen, die gesproken worden in Marokko, kort beschreven. Tevens wordt informatie verstrekt over het taalbeleid en het taalgebruik in Marokko en over de implicaties van dit beleid voor de positie van het Berbers in het onderwijs en de Marokkaanse samenleving als geheel. Daarna wordt de positie van de Berberse taal en de Berberse gemeenschap in Nederland beschreven. In dit opzicht worden twee dimensies onderscheiden: de status van de sprekers van het Berbers in Nederland en de status van het Berbers als taal in de Nederlandse samenleving. Het eerste deel belicht de demografische en immigratie achtergrond van de Berberse gemeenschap en de positie van haar leden in het Nederlandse onderwijssysteem en de arbeidsmarkt. Het tweede deel behandelt de *status quo* van de Berberse taal in Nederland. Aspecten die worden onderzocht zijn de positie van het Berbers in het sociaal en cultureel leven van Marokkanen, massa media, religie, cultuur en onderwijs. Ten slotte eindigt het hoofdstuk met de taalverwerving, taalvaardigheid, taalkeuze, taaldominantie, en taalpreferentie binnen de Berberse gemeenschap in Nederland.

Hoofdstuk 3 beschrijft het theoretische kader. Eerst wordt de impact van ouder-kind interactie en taalaanbod (input) besproken. Kinderen leren taal in de maatschappelijke context waarin ze omgaan met hun ouders. De inhoud van

deze interacties, in het bijzonder de input die ouders aan hun kinderen bieden, is van cruciaal belang bij het taalleerproces dat kinderen doormaken. Uit onderzoek is gebleken dat kinderen wiens sociale ervaringen meer communicatieve mogelijkheden bevatten en die een rijker taalaanbod aangeboden krijgen, hun woordenschat in een sneller tempo opbouwen dan kinderen met minder communicatieve ervaringen en een armer aanbod. Vervolgens wordt het begrip academische taal toegelicht. Academische taal kan omschreven worden als het taalgebruik voor cognitief complexe doeleinden in cognitief veeleisende communicatiesituaties zoals dat zich (afhankelijk van de talig-cognitieve gezinssocialisatie) in meerdere of mindere mate thuis voordoet, maar dat toch vooral aan de orde is bij het leren en onderwijzen op school. De instructietaal van leraren en het taalgebruik in schoolboeken zijn daar voorbeelden van. Bij academische taal gaat het veelal om communicatie over onderwerpen die het concrete hier-en-nu met de daarbij direct aanwezige referenten overstijgt. Academische taal is een in alle talen voorhanden taalgebruikregister met een specifiek repertoire aan talige middelen om te spreken of te schrijven over zaken die niet direct voorhanden en/of cognitief complex zijn. Daarmee verschilt academische taal in hoge mate van het informele, alledaagse taalgebruikregister in de thuissituatie, op straat of onder vrienden. Om de ontwikkeling van academische taal te onderzoeken, wordt de zogenaamde functionele benadering in de linguïstiek, ook wel aangeduid als usage-based benadering of cognitieve linguïstiek, als uitgangspunt genomen. Binnen de functionele linguïstiek worden drie variabelen onderscheiden: *field* (cognitieve inhoud), *tenor* (relatie spreker-hoorder) en *mode* (tekstniveau, opbouw verhaal of betoog). Daarnaast, worden de specifieke linguïstische variabelen die kenmerkend zijn voor academische taal beschreven. Deze variabelen bevinden zich op verschillende taalniveaus: er zijn variabelen op het woord niveau (lexicaal niveau), variabelen op het niveau van de zin (morfo-syntactisch niveau,) en variabelen op het niveau van de tekst (tekstueel niveau).

In Hoofdstuk 4 wordt de onderzoeksopzet van deze dissertatie besproken. De huidige studie kent de volgende hoofdvragen:

1 In hoeverre gebruiken Marokkaans-Berberse moeders academische taal in hun taalaanbod aan hun kinderen tussen de leeftijd van 3 en 6?
2 Is het gebruik van academische taal door moeders gerelateerd aan sociaaleconomische status (SES) en geletterdheid?

3 In hoeverre gebruiken leerkrachten academische taal in hun taalaanbod aan de Marokkaans-Berberse schoolkinderen?
4 In hoeverre gebruiken kinderen academische taal tussen de leeftijd van 3 en 6 in thuis en op school?
5 Is er een verschil tussen moeders en leerkrachten met betrekking tot het gebruik van academische taal?
6 Is er een relatie tussen taalaanbod van ouders en leerkrachten en het gebruik van academische taal door kinderen?

Aan deze studie hebben 12 Marokkaans-Berberse kinderen (vier jongens en acht meisjes), hun moeders en hun leerkrachten meegewerkt. De gezinnen komen uit de steden Utrecht, Ede, Tilburg en Oosterhout. De kinderen zijn gedurende een periode van drie jaar, van 3 tot 6 jaar oud gevolgd. Verschillende instrumenten zijn in deze studie gebruikt: interactietaken ontworpen voor moeder-kind en leerkracht-kind gesprekken, een gedetailleerde vragenlijst over familieachtergrond, taalgebruik en geletterdheidsactiviteiten, een korte vragenlijst voor de leerkrachten en woordenschatstests in de eerste (Tarifit Berber) en de tweede taal (Nederlands). De dataverzameling bestond uit drie meetmomenten thuis en twee op school. Het eerste meetmoment thuis vond plaats toen de kinderen 3 jaar en 2 maanden oud waren, het tweede meetmoment op de leeftijd van 4 jaar en 2 maanden en het derde meetmoment op de leeftijd van 5 jaar en 10 maanden. De twee meetmomenten op school vonden gelijktijdig plaats met de laatste twee meetmomenten thuis. Een totaal van 60 op video opgenomen sessies van moeder-kind en leerkracht-kind gesprekken resulteerde in een taalcorpus (120 transcripten en 15.509 uitingen), die getranscribeerd is voor de huidige studie. Alle thuis en school data werden letterlijk getranscribeerd. De interacties zijn getranscribeerd met behulp van de *Child Language Data Exchange System* programma (CHILDES) en gecodeerd conform een coderingssysteem dat ontwikkeld is door de DASH onderzoeksgroep. Het DASH coderingsschema bevat vier belangrijke niveaus: het lexicale niveau, het morfo-syntactische niveau, het tekstuele niveau en het socio-pragmatische niveau.

Hoofdstuk 5 behandelt de lexicale onderdelen van de academische taal. Lexicale variabelen in het taalaanbod van moeders en leerkrachten, alsmede van de kinderen werden bestudeerd tijdens twee verschillende interactietaken op verschillende tijdstippen thuis en op school. Het bleek dat moeders meer praten

tijdens de gestructureerde taak, het voorlezen van een prentenboek, dan tijdens de praatplaat. Daarnaast vertoonden moeders, in de drie meetmomenten, een hogere lexicale dichtheid en lexicale diversiteit tijdens het voorlezen dan tijdens de praatplaat. Wat betreft de relatie tussen de academische taal in het taalaanbod van de moeders en hun familie achtergrond, toonde lexicale dichtheid en diversiteit een significante correlatie met de SES en geletterdheid in beide taken in bijna alle meetmomenten. Ten aanzien van het taalaanbod op school, hebben leerkrachten de neiging om meer te praten als ze betrokken zijn bij de boekentaak dan wanneer zij betrokken zijn in het kringgesprek. Schoolresultaten tonen ook aan dat leerkrachten meer praten en een rijkere woordenschat gebruiken wanneer de kinderen ouder worden. De kinderen gebruiken een taal met een hogere densiteit en diversiteit tijdens de boekentaak dan tijdens de praatplaat. Zoals verwacht vergemakkelijken de boekentaken, meestal geassocieerd met scholing, het gebruik van een rijkere woordenschat. Kijkend naar de invloed van SES en geletterdheid, is een positieve en soms aanzienlijke impact van beide factoren gevonden tijdens de boekentaak. In het bijzonder, de lexicale diversiteit was sterk gecorreleerd met de SES van het gezin in alle meetmomenten. In de schoolcontext produceerden kinderen minder uitingen dan in de thuissituatie.

Hoofdstuk 6 gaat over morfo-syntax en de relatie daarvan tot de ontwikkeling van academische taal van Marokkaans-Berberse kinderen. De focus ligt op de samengestelde zinnen in zowel het taalaanbod van de moeders en leerkrachten als ook in de taalproductie van de kinderen. Wat betreft het aanbod thuis, blijkt de meest voorkomende categorie van bijzinnen van moeders was complimentenzinnen in alle meetmomenten. We vonden ook dat de frequentie van syntactisch complexe constructies in het aanbod van moeders niet verandert als kinderen ouder worden. Bovendien laten onze gegevens zien dat SES en geletterdheid gedeeltelijk betrekking hebben op de samengestelde variabelen van syntactische complexiteit van moeders. Vooral in meetmoment 3, wanneer kinderen de leeftijd van 5 bereiken, correleren SES en geletterdheid significant met de syntactische complexiteit in de taal van moeders. Met betrekking tot het aanbod van leerkrachten wijzen de resultaten erop dat zij in beide meetmomenten de neiging hebben om gecoördineerde zinnen te gebruiken, meer dan relatieve en adverbiale zinnen. Ook blijkt dat leerkrachten, op het niveau van bijna alle soorten *multi-clause* zinnen, hogere aantallen produceren dan de moeders. Met betrekking tot kinderen, wijzen onze

gegevens erop dat ze een klein aantal complexe zinnen geproduceerd hebben op verschillende leeftijden zowel thuis als op school. Echter, op de leeftijd van 5 jaar werden de uitingen van kinderen duidelijk complexer. Overigens hebben de kinderen uit deze studie de neiging om meer complexe constructies op school te gebruiken dan thuis.

Het centrale thema van Hoofdstuk 7 is *decontextualized language* ofwel de abstractie, zowel in het aanbod van moeders en leerkrachten als ook in de taal van de kinderen. Ten aanzien van de moeders, blijkt uit de resultaten dat zij voornamelijk de *matching* en *selective analysis* niveaus gebruiken. Die bestaan uit beschrijvende etikettering en praten over actuele informatie. Verder vonden we een grote hoeveelheid variatie tussen de moeders wat betreft *reordering* en *reasoning* uitingen die zij gebruikten tijdens hun gesprekken met hun kinderen. De twee componenten van *decontextualized language* (*reordering* en *reasoning*) correleerden positief met SES en geletterdheid, wat suggereert dat er een significante relatie is tussen de sociaaleconomische achtergrond van moeders en hun taalaanbod. Wat de leerkrachten betreft, laten de resultaten zien dat het meest voorkomende niveau van abstractie *selective analysis* is. Een vergelijking tussen het aanbod van moeders en leerkrachten laat zien dat leerkrachten de neiging hebben om vaker gebruik te maken van *decontextualized language* dan moeders. De analyses van gegevens van de kinderen tonen dat ze op school relatief vaker de *decontextualized language* produceren dan thuis.

Hoofdstuk 8 bespreekt en reflecteert op de belangrijkste bevindingen van deze dieptestudie door het geven van antwoorden op de onderzoeksvragen. Ook wordt een beknopte evaluatie van de studie gegeven en wordt een aantal aanbevelingen genoemd. De resultaten, gepresenteerd in de drie studies, geven aan dat er sprake is van het gebruik van academische taal in het taalaanbod van moeders en leerkrachten. Er worden echter veel variaties gevonden tussen de moeders met betrekking tot de drie onderzochte domeinen (lexicaal, morfo-syntactisch en tekstueel niveau). De relatie tussen gezinskenmerken en academische taal in het taalaanbod van de moeders toont tevens veel verschillen. De meest opvallende bevinding is de sterke relatie tussen het gebruik van *decontextualized language* en SES en geletterdheid. Met betrekking tot de input-output relatie in de schoolcontext, in vergelijking met moeders in thuiscontext, zijn minder significante correlaties gevonden tussen het aanbod van de leerkrachten en de taal van kinderen. Opgemerkt moet worden dat het taalgedrag van de kinderen thuis werd gekenmerkt door een mengeling van

twee talen: Berbers en Nederlands. In plaats van te zeggen dat deze kinderen Tarifit Berbers of Nederlands of beide talen tegelijk gebruiken, kunnen we zeggen dat deze taalgebruikers, oftewel *languagers,* de *getaal* gebruiken met hun moeders. Een belangrijke implicatie van het onderzoek is dat zowel ouders als leerkrachten zich bewust moeten zijn van het soort taal die zij aan kinderen bieden. Er moet meer aandacht worden besteed aan de aard van de taal-activiteiten die verwerving van de academische taal stimuleren. De rol van school is om kinderen een meer gestructureerde en georganiseerde manier van leren te bieden om de academische taal te bevorderen.

Curriculum Vitae

Mohammadi Laghzaoui was born in *Ait Bouyahie* - Nador (Morocco) in 1971. He learned to read and write at *L'école primaire Laqouass - Hassi Berkane*. After he received his Baccalaureate in 1992, he studied English language and literature at Oujda University (Morocco) where he obtained his B.A. in 1996. In 2001 he received his M.A. in Linguistics at Tilburg University - The Netherlands with a thesis on *language attitudes towards Arabic in the Netherlands*. After working for three years as a teacher of English, in 2005 he started his PhD research at the Department of Culture Studies in the same university which resulted in this doctoral dissertation. While conducting his PhD project, he attended the LSA Linguistic institute at Stanford University, USA in 2007. He is currently affiliated to the Department of Culture Studies and member of Babylon, the Center for Studies of the Multicultural Society, Tilburg University, the Netherlands. Mohammadi Laghzaoui is also a sworn Arabic-Dutch translator.

TILBURG SCHOOL OF HUMANITIES

Dissertations in Language and Culture Studies

1 Sander Bax. *De taak van de schrijver. Het poëticale debat in de Nederlandse literatuur (1968-1985).* Supervisors: Jaap Goedegebuure and Odile Heynders, 23 May 2007.

2 Tamara van Schilt-Mol. *Differential item functioning en itembias in de cito-eindtoets basisonderwijs. Oorzaken van onbedoelde moeilijkheden in toetsopgaven voor leerlingen van Turkse en Marokkaanse afkomst.* Supervisors: Ton Vallen and Henny Uiterwijk, 20 June 2007.

3 Mustafa Güleç. *Differences in Similarities: A Comparative Study on Turkish Language Achievement and Proficiency in a Dutch Migration Context.* Supervisors: Guus Extra and Kutlay Yağmur, 25 June 2007.

4 Massimiliano Spotti. *Developing Identities: Identity Construction in Multicultural Primary Classrooms in The Netherlands and Flanders.* Supervisors: Sjaak Kroon and Guus Extra, 23 November 2007.

5 A. Seza Doğruöz. *Synchronic Variation and Diachronic Change in Dutch Turkish: A Corpus Based Analysis.* Supervisors: Guus Extra and Ad Backus, 12 December 2007.

6 Daan van Bel. *Het verklaren van leesgedrag met een impliciete attitudemeting.* Supervisors: Hugo Verdaasdonk, Helma van Lierop and Mia Stokmans, 28 March 2008.

7 Sharda Roelsma-Somer. *De kwaliteit van Hindoescholen.* Supervisors: Ruben Gowricharn and Sjaak Braster, 17 September 2008.

8 Yonas Mesfun Asfaha. *Literacy Acquisition in Multilingual Eritrea: A Comparative Study of Reading across Languages and Scripts.* Supervisors: Sjaak Kroon and Jeanne Kurvers, 4 November 2009.

9 Dong Jie. *The Making of Migrant Identities in Beijing: Scale, Discourse, and Diversity*. Supervisors: Jan Blommaert and Sjaak Kroon, 4 November 2009.

10 Elma Nap-Kolhoff. *Second Language Acquisition in Early Childhood: A Longitudinal Multiple Case Study of Turkish-Dutch Children*. Supervisors: Guus Extra and Kutlay Yağmur, 12 May 2010.

11 Maria Mos. *Complex Lexical Items*. Supervisors: Antal van den Bosch, Ad Backus and Anne Vermeer, 12 May 2010.

12 António da Graça. *Etnische zelforganisaties in het integratieproces. Een case study in de Kaapverdische gemeenschap in Rotterdam*. Supervisor: Ruben Gowricharn, 8 October 2010.

13 Kasper Juffermans. *Local Languaging: Literacy Products and Practices in Gambian Society*. Supervisors: Jan Blommaert and Sjaak Kroon, 13 October 2010.

14 Marja van Knippenberg. *Nederlands in het Middelbaar Beroepsonderwijs. Een casestudy in de opleiding Helpende Zorg*. Supervisors: Sjaak Kroon, Ton Vallen and Jeanne Kurvers, 14 December 2010.

15 Coosje van der Pol. *Prentenboeken lezen* als literatuur. *Een structuralistische benadering van het concept 'literaire competentie' voor kleuters*. Supervisor: Helma van Lierop, 17 December 2010.

16 Nadia Eversteijn-Kluijtmans. *"All at once" – Language choice and codeswitching by Turkish-Dutch teenagers*. Supervisors: Guus Extra and Ad Backus, 14 January 2011.

17 Mohammadi Laghzaoui. *Emergent academic language at home and at school. A longitudinal study of 3- to 6-year-old Moroccan Berber children in the Netherlands*. Supervisors: Sjaak Kroon, Ton Vallen, Abderrahman El Aissati and Jeanne Kurvers, 9 September 2011.