Effects of literacy, typology and frequency on children's language segmentation and processing units

Effects of literacy, typology and frequency on children's language segmentation and processing units

PROEFSCHRIFT

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 woensdag 1 april 2015 om 14.15 uur

door

Theodora Maria Veldhuis,
geboren op 28 april 1981 te Zwolle

Promotor: prof. dr. Ad Backus

Copromotores: dr. Jeanne Kurvers
                dr. Anne Vermeer

Overige leden van de promotiecommissie:
                prof. dr. Ludo Verhoeven
                dr. Bruce Homer
                dr. Daniel Wiechmann
                dr. Petra Bos

For Thomas and Sophie

Contents

Preface

After secondary school, I wanted to travel around the world, and then I wanted to study Japanese. This despite the fact that my father used to argue that he wanted his daughters to be the CEO of a bank one day (chances were very small that I would become one, after studying a language…) and without proper reason, apart from the fact that I found the Japanese written language extremely interesting. This interest in foreign languages and writing systems and the influence of knowing specific languages and being used to specific writing systems has, in addition to my wish to work as a kindergarten teacher one day, with 24 toddlers and a flute, probably been the reason why I have enjoyed doing research for my PhD and writing this thesis so much. Every thought about the influence of children's typological background on their view on language and every idea on the impact of their knowledge of writing, enthused me to further investigate the role of literacy on the way we think.

This current thesis is a result of all those thoughts and ideas, testing of children (without a flute though) and analyses of results from experimental tasks conducted with children. I would like to thank all the children who participated in the tasks – in pilot sessions and for the final study – first of all, as without them, this study would not have been completed. In addition, I would like to thank the teachers and managers at schools who supported me in testing the children and who allowed me to disturb classes, and who also believed in this study: I really enjoyed talking about your experiences with L2-children and literacy acquisition with you, Margriet, Astrid, Stanley, Helma, Janny, Willemien, Ivonne, Anne-Marie, Will, Souad, Manal, Nicole, Christa, and Maria. Your opinions and information from the work floor helped me greatly in shaping my ideas and thoughts!

Shaping those ideas would not have been possible without the help of my colleagues and my supervisors either. I would like to thank my Turkish colleagues, Seza, Derya, Güliz, and Zeynep, for all the information and

being there while I learned Turkish in Izmir and went to LSA in Berkeley, allowing me to re-upholster our sofas for fun while I had never touched a sewing machine, preparing for and running the marathon together, etc. As I said before, without you, I may have finished this book earlier (no distraction, no support in weird time-consuming activities, no badly planned trips, or travels that took longer than planned, no Mali), but it may also not have been finished at all. I think all the distraction we organized together was in the end good and great fun, and it/you made me work harder when I really had to. Thanks for everything. I only hope that I can pay it back to you, one day.

Sophie: you spent quite a lot of time of the first year of your life on a desk, in front of my bookcase, behind my computer screen, between books and articles on the ground and on my lap while I was writing. You always seemed very interested in what I was writing. You were, with your father, the greatest support I could have wished for: and you made sitting behind my computer and writing more interesting. I am curious how your language and literacy skills will further develop – I follow every step you take with the greatest curiosity and pleasure – and I will definitely try to find my flute before you go to kindergarten.

Dorina Veldhuis
Utrecht, February 2015

CHAPTER 1

Introduction

A recurrent question in the field of linguistics is what constitutes the basic building blocks of language. All over the world, people constantly generate new utterances – which they have never heard or produced before – while at the same time language is a system full of conventions. In addition, human memory capacity is not unlimited. The underlying linguistic system apparently allows for efficient storage, retrieval and processing, as well as a creative combination of units. There must be, therefore, some sort of structure or syntax, with the help of which building blocks get combined into utterances. One task of linguistics is to identify what these building blocks are.

This study is an attempt to obtain further insights into the building blocks of language by means of a battery of experimental tasks conducted among monolingual Turkish and Dutch, and bilingual Turkish-Dutch children, aged 4[1] to 10. Selecting children from these groups and this age range allowed for a detailed investigation of the effect of literacy on children's language segmentation and processing units, as children usually learn to read and write when they are around 6 years old. In addition, it allowed for making a contrast between children from different linguistic backgrounds, specifically speaking languages with quite different typological characteristics, about how they form words and regard units in language, as it was expected that the building blocks for people's interpretation and processing of concepts do not come unmediated. In relation to 'linguistic relativity' (cf. Levinson, 1996; Sapir, 1929, also in Cook, 2011; Whorf, 2012), or more specifically, in relation to knowing a specific language and a specific writing system which may influence people's perception of the structure of language and of the units that it consists of (cf.

---

[1] There was one child in this study who was during the first test session in which she participated only 3 years old.

Bugarski, 1993; Olson, 1994, 1996), it was expected that knowing Dutch or Turkish, or knowing both, and knowing or not knowing about their conventions with respect to word breaking or word marking in writing, has an impact on the units that people recognize and segment language into.

In the field of Cognitive Linguistics, it has been suggested that the units that people parse and produce in language do not necessarily correspond to the units that are marked in writing, such as words and letters or phonemes. Instead, it has been argued that units larger than single words, such as frequently co-occurring sequences like 'bake a cake' – this in comparison to, for instance, 'eat a cake', or 'buy a cake', or 'drop a cake' –, might rather be the units that people process. These units, referred to as '(multiword) constructions', '(lexical) chunks', or 'complex lexical items' (see Chapter 2), may, possibly related to their frequency of occurrence, function as the basic building blocks, rather than smaller units like words (see Chapter 2).

Whether literacy and knowing about writing conventions affect children's ideas on units in language, and to what extent those units consist of multiword sequences as mentioned in the field of Cognitive Linguistics, was the first main question that was posed in the current study. To what extent the expected influence from literacy is at the level of our unconscious language processing, an issue about which not much is known so far, became the second main question. Is what you know about language overtly really what you know about its structure, and are the units that you overtly recognize in language also the units that you use in processing?

1.1   Background to the current study[2]

This thesis explores the question of what the building blocks of language are, as used in language segmentation and processing. This will be investigated for children of varying linguistic and typological backgrounds. Knowing with what linguistic units children operate might help in addressing them adequately in educational programs.

In 2011, about 23% of the youth (0-25 years of age) in the Netherlands was non-Dutch. Of these children and young adults, about 14% had a

---

[2] This thesis was written as part of the 'Building Blocks Project' at Tilburg University, in which the following people participated as well: prof. dr. Ad Backus, dr. Seza Doğruöz, dr. Jeanne Kurvers, and dr. Anne Vermeer. As a team, we considered and discussed all tasks prepared for and data obtained in this study. Therefore, whenever I talk about decisions made for this study, the plural form 'we' may be used, even if the ultimate responsibility for any decisions taken is mine completely.

Turkish background, or at least one parent who was born in Turkey (NJI, '*Nederlands Jeugd Instituut*', 2013). With these numbers, the Turkish youth are one of the two largest groups with non-Western background in the Netherlands, together with the Moroccan youth (NJI, 2013).

The Turkish-Dutch children in Dutch primary schools (aged 4-12) are nowadays mostly third generation immigrants, which means that their parents, or at least one of their parents, have often already completed Dutch education and speak both Turkish and Dutch. Despite this, and despite a host of initiatives aimed at integration of minorities and particularly at improving their Dutch proficiency, bilingual children still seem to have a disadvantage at school, and lag behind their monolingual Dutch peers. Teachers and curriculum designers could address these children more adequately if they know more about the way in which these children parse and produce language, and if they are more aware of the fact that what may seem familiar to them might actually be unfamiliar to young children, especially to beginning second language learners who speak a typologically different language.

Also for language and/or literacy teaching to monolingual children, the outcomes of this study promise to have relevant implications. Following the structural linguistic tradition that views language as a modular system, mainstream educational practice – both in the Netherlands and in Turkey – tends to chop sentences into small segments, so as to facilitate comprehension of the whole sentence, especially for children who have difficulties in learning language. Perhaps this can be helpful once learners have developed a metalinguistic awareness that helps them perceive the smaller units of language in the way it is intended. However, as it is as yet unknown to what extent these smaller units – such as words and phonemes – have psychological reality (the contrary has often been claimed, cf. Olson, 1994, 1996), the use of such small units in education may in fact distort children's literacy acquisition and comprehension, rather than support it (see Chapter 2).

1.2 Overview of the current study

For this thesis, a battery of experimental tasks was conducted with monolingual Turkish, monolingual Dutch, and bilingual Turkish-Dutch children. The experiments contained both relatively offline and relatively online tasks – tasks that were in general more or less focused on conscious or unconscious language processing (see Chapter 3). In most of the tasks,

selected stimulus items of different lengths and frequencies of occurrence were incorporated, in order to be able to compare these, as well as to be able to compare between tasks (see Chapter 5).

In relation to these choices for participants and the selection of tasks and stimulus items, the first challenge was that there are only few publications on testing young children with 'online' measures, a term focused on unconscious language processing that will be discussed in more detail in Chapter 3. There is a large number of previous studies with infants in which online measures such as the preferential looking paradigm or sucking experiments have been applied, but studies with young, pre-literate children that use online measures are hard to find. Consequently, one of the first questions addressed in the preparation of this study was which tasks could also be conducted with these children. This will be discussed in Chapter 4, which provides a description of pilot studies that were conducted to answer this question.

Another challenge that was encountered in the preparatory stages of this study was the selection of stimulus items. In order to test to what extent multiword units could be regarded as candidates for units that children parse and produce in language, as proposed in Cognitive Linguistics, all stimulus items consisted of multiple word sequences such as '*handen wassen*' ('wash hands'). In an ideal situation, all the selected multiword sequences would occur in all tasks, and they would be controlled for familiarity among the children, and their frequency of occurrence. Based on these previous studies it was expected that higher frequency and familiarity would lead to better entrenchment and, therefore, faster processing. Unfortunately, the absence of a sufficiently large Dutch child language corpus made it impossible to arrive at such a selection. Chapter 5 discusses the steps that were taken to overcome this problem.

Moreover, the Turkish-Dutch bilingual children who participated in the experiments turned out to be highly dominant in Dutch, and to have limited knowledge of Turkish (see Chapter 4). Therefore, testing these children in Turkish was problematic, and led to unreliable results. Nevertheless, their results obtained in Dutch did provide insightful information on their language processing, especially in comparison to their monolingual Dutch peers. This issue will be discussed in Chapter 4.

This study consists of in total eight chapters in line with the aims mentioned above, because the chapter organization was motivated by the variety of challenges and issues that were to be discussed: the difference between 'offline' and 'online' experimental tasks, the challenge of testing young,

pre-literate children using offline and online tasks, the comparison between monolingual Dutch and bilingual Turkish-Dutch children, and the comparison between monolingual Dutch and monolingual Turkish children.

The first two chapters provide background information to the study, introduce the central research question, and give an overview of the relevant literature. This introduction motivates the empirical approach underlying this study, in which the children who participated carried out a battery of experimental tasks.

A proposal for a new way of classifying experimental tasks as points on a continuum between 'offline' and 'online' is made in Chapter 3. The distinction is usually conceived of as a dichotomy, but this point of view will be called into question.

Chapter 4 zooms in further on the experimental tasks as developed and conducted for this study in a series of pilots. There are two reasons for including a whole chapter on these pilot sessions. First, it triggers a discussion of methodological issues, including details about the decisions that were taken after the pilot sessions, as these form the background to which the tasks that were finally conducted were developed. Second, the pilot sessions were performed by both monolingual and bilingual children. However, in the final design, the bilinguals were not included as participants, because of their limited knowledge of Turkish, a fact that only became clear during the pilot sessions.

In Chapter 5, the methodology and design of the main study will be discussed, including information on the experimental tasks that were ultimately conducted, the procedures followed in the selection of subjects and items for the tasks, and the way in which the data were analysed.

Chapter 6 and 7 present the results. In Chapter 6, a qualitative analysis is provided of the responses that children gave in one of the language segmentation tasks – a dictation task. In this rather free task, the children were asked to select their own wordings and phrases and make segments in stories that they wanted to tell. The analyses of the segments children provided in this task could give a first indication as to the extent to which multiword units that were mentioned in Cognitive Linguistics are feasible as candidates for segmentation units in language, and to what extent literacy may affect children's segmentation units. More converging evidence for these questions is discussed Chapter 7.

In this chapter, quantitative analyses of five more language segmentation and processing tasks are given. In these tasks, carefully selected stimulus items were included, which allowed for an investigation of how such items were segmented and processed, and for a comparison between tasks. The

analyses will focus on the factors identified in previous studies as likely to affect children's language segmentation and processing units: literacy, typological characteristics, and the frequency of occurrence of specific multiword sequences.

The final chapter, Chapter 8, discusses and summarizes the results, as well as the continuum for the classification of tasks proposed in Chapter 3. Finally, the findings are put into a larger perspective, from which more general conclusions and suggestions for further research are derived.

CHAPTER 2

Units in language[3]

The discussion of what can be regarded as the basic units of language segmentation and processing, often with reference to writing, was already prominent in early studies of language and linguistic structure. Kraak (2006) argues that the study of units in language – more specifically, in Greek – began with the introduction of the alphabet. Traditionally, it was assumed that writing began with words as the basic units, which developed into syllabic writing systems, and ultimately into a system in which speech sounds formed the basis. Accordingly, the basic units in language that were defined and recognized were words, syllables, and, after the introduction of the alphabet, phonemes or speech sounds, which corresponded to letters in writing. It may be argued that phonemes were regarded as the basic and smallest units of language until Sapir stressed the psychological reality of words and sentences in 1920 (see also Kraak, 2006; Kurvers, 2002). According to Sapir, the fact that illiterate Indians were well able to segment sentences into words meant that words had psychological reality (Kraak, 2006; Kurvers, 2002; Sapir, 1920), and thus constituted the basic units of language.

Linguistic theory also often seems to assume implicitly that words form the basic building blocks of language. Interpretations or conclusions about the structure of language and the rules by which language can be explained to work, are often based on word units (e.g., Chomsky, 1975; Pinker, 1994; see also Kraak, 2006). Lexical entries are almost always seen as synonymous with individual words.

Numerous studies of children's metalinguistic development, from the 1960s onwards, have also addressed the question to what extent children have to develop an awareness of words (and of other entities such as

---

[3] This chapter elaborates on Veldhuis (2011, 2012), Veldhuis & Backus (2012) and Veldhuis & Kurvers (2012).

phonemes) as units of language (cf. Homer, 2009; Juel, 2009; Melzer & Herse, 1969; Sulzby & Teale, 1991; Tolchinski, 2004). These studies are often based on experimental research, and have focused on whether or not children are able to distinguish units like phonemes and words in ongoing speech, and whether or not they are able to define concepts such as words in their own terms, and how the awareness of units in language is connected to literacy. These studies relate to two parts of the research question posed in the current study, namely to what extent literacy affects children's building blocks of language, and whether this is visible in children's offline segmentation of language (see Chapter 3 and 5).

More recently, the field of Cognitive Linguistics has also raised the question of what the basic units in language are that people process and produce. A number of units – varying from single words to multiword units, (partially) schematic units (i.e., productive constructions with some fixed elements), and schematic units (i.e., patterns, templates) – have been advanced as functional basic units of language processing. These units have been argued to become entrenched in people's mental lexicons based on their frequency of occurrence (Arnon & Snider, 2010; Siyanova-Chanturia, Conklin & Van Heuven, 2011). In this connection, however, the relation between basic units in online processing and literacy has not been considered, despite the fact that there is a growing body of neuro-imaging studies on language processing that suggests that literacy does affect the online processing of spoken language and that it can change the functional organization of the brain (cf. Carreiras et al., 2009; Ostrosky-Solis, Arellano Garcia & Perez, 2004; Pattamadilok et al., 2010; Petersson, Ingvar & Reis, 2009; Schild, Roder & Friedrich, 2011).

The main purpose of the current study is to shed light on what can be regarded as basic building blocks of language, both in children's offline language segmentation and in their online language processing, and to investigate to what extent literacy and the frequency of occurrence of multiword combinations in language affect those building blocks. The perspectives from the area of metalinguistic development and from Cognitive Linguistics will be combined.

In order to better understand these perspectives, this chapter will first, in Section 2.1, reflect on the previous findings on the development of metalinguistic awareness in language. This will be followed by a discussion of the ideas on units proposed in Cognitive Linguistics in Section 2.2. In Section 2.3, these findings will be linked to the current study.

2.1   Metalinguistic awareness and unit segmentation

Since the late 1960s and 1970s, when interest in the mental processes that underlie the acquisition, storage, production and comprehension of speech and writing gained in popularity, various studies have been published on children's development of metalinguistic awareness, or our 'possibility of raising ourselves above language, of abstracting ourselves from it, or contemplating it, whilst making use of it in our reasoning and observations' (Benveniste, 1974 in Gombert, 1992: 2). In this sense, the term 'metalinguistic awareness', which has been defined and described in many alternative ways since the 1970s (cf. Homer, 2000, 2009), refers to the awareness about language that people show via their linguistic productions. As such, behavioural tasks in which children are asked to manipulate language or to make judgments about language – tasks referred to as 'offline' because they involve conscious decision making processes (see Chapter 3 for further discussion) – can give us insights into children's metalinguistic development, including their awareness of specific units, such as words.

In this section, an outline of theories on metalinguistic awareness and its development will be provided in Section 2.1.1. Subsequently, the main results of earlier work will be discussed in Section 2.1.2.

2.1.1   Metalinguistic awareness and its development

It is probably since Piaget's publication (1926, orig. 1923) that researchers have come to combine knowledge of psychology with the observation of how children develop awareness about language. Piaget discussed the way in which children come to argumentation and collaborative communication, which, according to what he saw in children, develops in stages. In the first stage, children do not really communicate or participate in conversation, but this does appear in the second and third stages. In the second stage, children may start making conversations, usually about concrete topics or actions. In the third stage, which children are said to reach when they are around seven, they may also talk about more abstract themes, and better understand other people's utterances. In fact, Piaget claims that there are several stages that children go through in their language development: first, they use language for themselves; later they use it to refer to objects and activities present in the world around them; and only after that, they use language to refer to more abstract matters.

In later studies by researchers such as Sinclair, Jarvella and Levelt (1978), Karmiloff-Smith   (1992)   and   Gombert   (1992),   children's   language

development, or more specifically, their awareness of language and its structure, was once more investigated and similarly related to different stages or phases. In these studies, it was shown that at first children are only able to use language, later they can recognize and repair errors in language use, and only then children start to develop the ability to talk about language and its structure from a metaperspective (cf. Clark, 1978; Homer, 2009; Slobin, 1978; see also Ravid & Hora, 2009). Marshall and Morton (1978) concluded on a more general level that children first learn something, and only later become able to verbalise explicitly what they have learned – although not all learned behaviour can necessarily be explained verbally in the end. Marshall and Morton stress that metaknowledge has to follow the acquisition, and that it cannot be the other way round. In fact, the theoretical proposals developed by Sinclair, Karmiloff-Smith, Gombert and Marshall and Morton match what Piaget had argued before: awareness of language seems to develop gradually in children and in phases, and children only come to be able to express this awareness once it is there.

Karmiloff-Smith (1992) provided a three-phase model very similar to Piaget's (1976) model of the way in which children become aware of processes they encounter – whether these are linguistic or not. In this model, Karmiloff-Smith makes a difference between implicit representations and different levels of progressive representational explication. In her discussion of the model, she first distinguishes a phase in which it is impossible to define the components of a procedure while one is able to run the procedure in its totality, and then two stages in which there is first explicit but unconscious knowledge of the components, and then explicit and conscious knowledge. As maintained by Karmiloff-Smith, children pass these phases for all linguistic forms, so it describes phonological, morphological as well as lexical development, but they do not have to develop simultaneously. Accordingly, the phases that Karmiloff-Smith distinguished are not age-related, and children can be in the first phase for one form, and in the second or third phase for another, depending on their own endogenous processes.

Gombert (1992) further elaborates on the stages that children pass through in their development of metalinguistic awareness. In order to overcome potential confusion between children's declarative knowledge, the *know-that* of processes, and their intentional monitoring, the *know-how* of processes (cf. Homer, 2000), Gombert (1992) introduces a division between epilinguistic abilities and metalinguistic abilities. This division can be related to the second and third phases in Piaget's and Karmiloff-Smith's models. Epilinguistic ability refers to people's unconscious metalinguistic activities, and is characterized by explicit manifestations of functional awareness of the

organization in language. Metalinguistic awareness, on the other hand, relates to people's reflective abilities, or their ability to describe the processes. This conceptual division should make it possible to define more exactly what stage children have reached, and to what extent they are aware of linguistic systems and structures. In practice, however, the categories have not been used much in analysis.

In general, scholars tend to refer to children's development of metalinguistic awareness for both kinds, ignoring the degree of consciousness involved or the ability to express this consciousness verbally. This makes comparing the results from different studies a difficult task.

The fact that there are quite a lot of terms for metalinguistic awareness – such as metalinguistic activities, (meta)linguistic knowledge, (meta)linguistic consciousness, metacognition, metalanguage, meta-processes, metalinguistic ability or skill, and metalexical and metasyntactic development – further complicates comparing the results from different studies (cf. Bialystok, 1986a; Gombert, 1992). The terms have been used in slightly different circumstances, and each has its own connotations. Metalinguistic 'knowledge', 'consciousness', 'cognition' and 'awareness', but also 'metalinguistic ability' and 'metalinguistic skill', all refer to the knowledge that people have about language. Activities that show or require metalinguistic knowledge are instead referred to as 'metalinguistic activities' or 'processes', whereas terms like 'meta-language' merely refer to the terms that can be used to describe the structure of a language.

The many aspects that are involved in metalinguistic awareness makes a simple interpretation of how it develops in children difficult. For instance, phonological awareness, grammatical awareness, print awareness, and word awareness, are all aspects of metalinguistic awareness (Homer, 2000; Bialystok, 2007). Accordingly, the term 'metalinguistic awareness' may be too broad to be 'of much empirical use' (Homer, 2000: 1). This may explain the contradicting findings one sometimes finds regarding children's metalinguistic development. It may be better to focus only on specific aspects of metalinguistic awareness.

The confusion about terms is related to the multitude of backgrounds from which the phenomenon has been studied. As Gombert (1992) points out, linguists seem to apply the term mostly to refer to the use of language to refer to language, whereas psychologists take it as referring to conscious reflection on language or mental control. The definition of metalinguistics as given by Benveniste (mentioned in Gombert, 1992: 2), which says that metalinguistics is the 'possibility of raising ourselves above language, of abstracting ourselves from it, or contemplating it, whilst making use of it in

our reasoning and observations', combines the two stances, and with that,
seems the most satisfying.

Despite the factors which complicate drawing general conclusions about the
development of metalinguistic awareness in children, researchers have
made progress in understanding the prerequisites of the process.

Some researchers, such as Piaget (1926, orig. 1923) and Karmiloff-Smith
(1992), have argued that meta-awareness develops automatically in children,
as a consequence of general abstraction capabilities, which Piaget suggested
children have until they are 11 or 12 years old (Piaget, 1976). As a result of
normal language acquisition and cognitive development, children come to
'distance' themselves from their linguistic product (Homer, 2009; Karmiloff-
Smith, 1992), and ultimately develop conscious explicit metalinguistic
awareness of a large variety of linguistic forms.

Some researchers, including Coulmas (1989), Olson (1994, 1996) and
Homer (2000), claim a major role for literacy in the development of
children's metalinguistic knowledge, as a result of which children – and
adults who acquire literacy later in life (Morais, 1978; Kurvers, 2002;
Ramachandra & Karanth, 2007) – develop a specific awareness of the
structure of languages (Ravid & Hora, 2009). It is argued that children only
develop meta-awareness of language and its structure when learning to read
and write, and that, as Coulmas (1989: 45) states: 'Writing systems are only
rarely the result of conscious linguistic analysis, yet they are the expression
and materialization of linguistic consciousness.' Bugarski (1993) and Kraak
(2006) argued the same point from a more general perspective, regarding the
development of writing.

This question how metalinguistic awareness and literacy are related has
been addressed numerous times since the 1960s in empirical studies in
which children were asked to perform different experimental tasks. In early
tasks, children were tested on their awareness of units such as phonemes
and words, and the positive impact of this awareness on proficiency in
reading and writing was stressed. Some conceptual studies (cf. Bugarski,
1993; Harris, 2009; Kraak, 2006; Olson, 1994, 1996) suggested that the
awareness of words may be a product of literacy. Yet, the fact that
metalinguistic awareness, and especially phoneme awareness which can be
trained at school, was found to be fundamental to literacy acquisition (cf.
Bradley & Bryant, 1983; Goswami, 2009) received much more attention than
the idea that literacy may also enhance awareness. However, after Bialystok
(1986b) tested children on their development of word awareness, the

directionality of the causal relationship between metalinguistic awareness and literacy started to be questioned in earnest. In Kurvers (2002), this directionality of the causal relationship between metalinguistic awareness and literacy was also investigated by including non-literate adults in addition to pre-literate children as participants in several experimental tasks.

2.1.2 Experimental studies on metalinguistic development

Experimental studies on metalinguistic development have first and foremost focused on children's development of phoneme awareness. In such studies children were usually asked to segment words into smaller parts, delete phonemes from words, or to add phonemes to them (see Goswami, 2009: 136-139 for a summary). It was found that children first develop awareness of 'large units of sound, such as syllables, onsets and rimes', and only later become aware of '"small" units of sound', which include phonemes (Goswami, 2009: 138).

Some studies focused on words. In these studies, children were asked for example to segment sentences into words, to manipulate the concept of the word (for instance by asking them 'Which is the longer word, train, or bicycle?'), or to define the concept of words (cf. Fox & Routh, 1975; Gombert, 1992; Holden & MacGinitie, 1972; Homer, 2000; Homer & Olson, 1999; Karpova, 1966, orig. 1955; Kolinsky, Cary & Morais, 1987; Kurvers, 2002; Kurvers & Uri, 2006; Lazo, Pumfrey & Peers, 1997; Melzer & Herse, 1969; Morais et al., 1986; Morais & Kolinsky, 1995; Morris, 1993; Olson, 1994; Ramachandra & Karanth, 2007; Ravid & Tolchinsky, 2002; Roberts, 1992; Tunmer, Bowey & Grieve, 1983). It has often been concluded that children under the age of 7 do not possess awareness of words, called metalexical awareness (Fox & Routh, 1975; Gombert, 1992; Holden & MacGinitie, 1972; Homer, 2000; Homer & Olson, 1999; Karpova, 1966, orig. 1955; Kurvers, 2002; Kurvers & Uri, 2006; Lazo, Pumfrey & Peers, 1997; Morais et al., 1986; Morais & Kolinsky, 1995; Morris, 1993; Olson, 1994; Ramachandra & Karanth, 2007; Ravid & Tolchinsky, 2002; Roberts, 1992; Tunmer, Bowey & Grieve, 1983).

The first experimental study that found evidence for this conclusion was Karpova (1966, orig. 1955). She asked children aged 3 to 7 to repeat sentences, and to tell her how many words they had heard, and which words. The answers fell into three categories, corresponding to the stages that children apparently have to go through. These stages were: (1) a stage in which children took single sentences as unified messages and did not segment words at all; (2) a stage in which children segmented meaningful

components from sentences (e.g., two female names that were mentioned in a sentence); (3) a stage in which children were able to segment sentences into conventional words, or syllables. These results showed that conventional word awareness is not present in very young, pre-literate children.

After this first study by Karpova, quite a lot of experimental studies were conducted in which children between 3 and 8 years old were tested on their ability to segment sentences into words (cf. Chaney, 1992; Ehri, 1975; Fox & Routh, 1975; Holden & MacGinitie, 1972; Papandropoulou & Sinclair, 1974; Tunmer, Bowey & Grieve, 1983). In these segmentation tasks, children were often asked to indicate the number of words they heard in sentences using poker chips on a table, or by clapping their hands for every word, or by moving a puppet on an underground of hopscotch squares (Chaney, 1992; Ehri, 1975; Holden & MacGinitie, 1972). Most studies concluded that the awareness of words develops gradually, as Karpova (1966, orig. 1955) had suggested (cf. Ehri, 1975; Holden & MacGinitie, 1972; Tunmer, Bowey & Grieve, 1983; Papandropoulou & Sinclair, 1974, see also Kraak, 2006, and Olson 1994, 1996). Young children, aged 4 to 5, were found to base their segmentation mostly on acoustic cues (i.e., phrase and syllable stress, cf. Tunmer et al., 1983), and on meaning. This was seen for instance in their better performance on content words than on function words. As Holden and MacGinitie (1972: 554) state 'the greater the proportion of content words in an utterance, the greater the percentage of correct segmentations' among these children. Function words were mostly kept together with content words by young children, represented by clapping their hands just once, or by a single poker chip (see also Ehri, 1975). This confirmed the idea that young children are often not aware of function words, as they do not refer to objects, events or activities that can easily be visualized (Gombert, 1992; Swiney & Cutler, 1979; Van Kleeck, 1984).

Tunmer et al. (1983) looked at children's behaviour when they were only exposed to content words. They tested 4- to 7-year-olds on their word awareness, using several tasks in which children were asked to tap the number of words they heard in sequences. From their results, Tunmer et al. (1983) concluded that '… the basis on which children segment different meaningful word strings undergoes two changes prior to the attainment of a mature word concept. Children aged 4 to 5 appear to respond primarily on the basis of acoustic factors (such as phrase and syllable stress). Somewhat older children adopt a different strategy, tapping once for each unbound morpheme appearing in the string. However, by age 6 most children have begun to abandon the acoustic and morpheme strategies for the more abstract conception of the word as the smallest meaningful, permutable unit

of language' (p. 592). This shows that it is not only semantics which affects children's segmentations, but that acoustic cues may also play a role.

Similar results were found in studies that used tasks focusing on children's objectivation skills, or their awareness of the concept of the word – thus on real metalinguistic knowledge and not on epilinguistic capabilities (Gombert, 1992). Bialystok (1986) asked children about similarities between words (e.g., what is the same meaning/sound as 'dog': 'frog' or 'puppy'?) and to judge words (which word is larger?). Older children scored better than younger children did. Chaney (1992), Papandropolou and Sinclair (1974), and Kurvers (2002), of which the latter also studied illiterate adults, found similar results. In contrast to previous studies in which experiments had been used to assess children's word awareness, and in which these findings were mostly related to cognitive development (Van Kleeck, 1982), or specifically, linguistic development (Chaney, 1989, 1992; Clark, 1978; Doherty & Perner, 1998), Bialystok (1986b) linked children's responses to their literacy skills. Literate children appeared to show a clear awareness of words, whereas pre-literate children were not able to properly perform tasks that required having the word as a linguistic unit.

Karmiloff-Smith et al. (1996), however, claimed that in addition to age, literacy and linguistic development, the nature of the tasks employed also highly influences word awareness in children. These researchers argued that nearly all of the empirical segmentation and definition studies described above (which rely on tasks such as word-by-word dictation, counting words, pointing to a block or a poker chip for every new word, judgment tasks, word deletion, or changing word order) used offline tasks in which metalinguistic judgments and behaviour are rather detached from normal linguistic processing. This was claimed not to provide sufficient evidence for the conclusion that older, literate children have a higher awareness of words than younger, pre-literate children.

To overcome such task effects, Karmiloff-Smith et al. (1996) developed a task in which children were supposedly triggered to process language more naturally. Pre-literate and literate children heard a story that was read out aloud to them and they were asked to repeat the last word they heard before the researcher paused mid-sentence. Since listening to stories is a more common experience for children than segmenting sentences or giving definitions of linguistic concepts, Karmiloff-Smith et al. (1996) argued that this task would provide more reliable information about the extent to which literacy and age affect children's word awareness.

Their surprising results showed – in contrast to previous studies – that young children are very good at isolating and repeating single words from

running speech. It was shown that 4-year-old children scored 75% correct, and 5-year-olds 96%, suggesting that young, pre-literate children are aware of words as linguistic units.

However, since the children tested by Karmiloff-Smith et al. (1996) were all in the British school curriculum, where there is already a lot of attention for reading skills in kindergarten, Kurvers and Uri (2006) suggested in the conclusions of their replication study that literacy may nevertheless play a major role in children's development of word awareness. Kurvers and Uri had tested pre-literate Dutch and Norwegian children from 3 to 6;4 years old, who were split up into a younger (up to 5;4 years old) and an older age group (5;5 and above). Kurvers and Uri found that the children they tested scored between 24.6% and 29% correct in cases in which they had to repeat the last word. This contrasted with the results obtained by Karmiloff-Smith et al. (1996), whose young, pre-literate child participants had scored more or less identical on word repetition tasks, notwithstanding their age. Furthermore, a small-scale follow-up study in which Kurvers and Uri tested three older literate children demonstrated that these older children were very well able to repeat the last word of a sentence correctly (on average 90%). Accordingly, Kurvers and Uri (2006) suggested that it might very well be possible that the results from the children Karmiloff-Smith et al. (1996) had tested reflect that they had been taught what words are, and therefore had experience identifying them. Their metalexical awareness would then have been higher, because of schooling.

Two other studies in which training effects, or schooling effects, were found, are Homer (2000) and Ramachandra and Karanth (2007). Homer (2000) tested among others 4- to 6-year-old Canadian children on their word segmentation ability in a task similar to the one that Karmiloff-Smith et al. (1996) had conducted, but then making use of two texts. The reason why Homer (2000) used two texts was that one was used in a pre-test condition, and one in a post-test condition. In the pre- and post-test conditions, the children did not receive any corrective feedback. Between these two tests, that were conducted after each other, there was a training session, in which the children who participated in the task received corrective feedback for ten example sentences or until they had two consecutive correct responses. In this training, there was stress on the word 'word', as used by the researcher in the instruction of the task. An analysis of results showed an effect of training: in the pre-test condition, 4-year-olds could repeat 18% of the words they were asked to repeat correctly; the 5-year-olds repeated 33% correctly; and the 6-year-olds 65%. In the post-training condition, the correct scores were much higher – 26%, 48% and 81% respectively. These results not only

show that younger pre-literate children perform more poorly than older literate children in their word awareness, but also that literacy training can lead to better scores.

Ramachandra and Karanth (2007) also found a training effect regarding word awareness, again using a 'last word repetition task', like the one Karmiloff-Smith et al. (1996) had developed and Homer (2000) had used as well. Ramachandra and Karanth tested 30 pre-literate and literate Kannada children (aged 4 to 7, divided into three age groups), and ten illiterate adults. Like Homer (2000), Ramachandra and Karanth conducted tests in pre-training and post-training conditions with two stories, with a training phase with a similar form as Homer had used in between them. Ramachandra and Karanth did however not provide ten separate training sentences in the training session, but simply provided the children with feedback after each pause when they read the first story a second time. Ramachandra and Karanth found different percentages of correct word repetition scores before and after the children had received the feedback. In the pre-training condition, the 4- to 5-year-olds scored 19% correct, the 5- to 6-year-olds 49%, and the 6- to 7-year-olds 89%. Post-training, the percentages were respectively 40%, 63% and 100%. As for the illiterate adults, they repeated 17% of the words correctly before the training, and 40% after it. These results support the suggestions raised by Kurvers and Uri (2006) and Homer (2000): literacy training seems to play at least a facilitative role for word awareness and the ability to segment words from connected speech.

Ramachandra and Karanth's (2007) results for the illiterate adults underlined the importance of literacy training: like the pre-literate children, they were not very good at segmenting the last word from a sentence. This is in line with findings by Kurvers (2002) and by several studies conducted by Morais and colleagues, in which illiterate Portuguese adults were tested. Literacy training and speech segmentation, and the awareness of units such as phonemes and words in language, are closely related, and illiterate adults have been found to be unable to segment small units form language, just as pre-literate children (Kurvers, 2002; Morais, Cary, Alegria & Bertelson, 1979; Morais et al., 1986; Morais et al., 1987).

As Correa and Dockrell (2007: 828) concluded from a battery of tests they conducted with Portuguese children to investigate their word segmentation patterns: 'These results suggest that advances across literacy and language contribute to an understanding of language specific word segmentation patterns.' In other words, as can be seen from the above, a large number of studies of children's and adults' language segmentation abilities have shown

that literacy, particularly the training it provides in identifying words, leads to higher scores on metalexical tasks. However, the studies presented above all focused on languages which are written with alphabetic writing systems, in which words are marked by spaces. In the following section, an overview is given of the results of segmentation tasks conducted with languages in which words are not systematically distinguished in writing.

Cross-linguistic studies into word awareness

Olson (1994, 1996) and Bugarski (1993) claim that words are not the natural units that all speakers easily segment from stretches of running speech. The studies described above, which demonstrated that pre-literate children and illiterate adults have difficulties doing so, especially regarding function words, support this hypothesis.

Cross-linguistic studies with languages that do not use an alphabetic writing system, and in which words are not marked by preceding and following spaces, or that are not written at all, provide further evidence in this direction, although there is one study, by Lin, Anderson, Ku, Christianson and Packard (2011) in which the opposite is claimed. Below, the findings from these studies will be discussed.

For researchers or organisations working with unwritten, endangered languages, identifying words – needed to develop a morpho-syntactic sketch of a language – appears to be a real problem. The problem with figuring out how best to write down a language that is as yet unwritten, and with defining where words begin or end in the speech stream, is that it is hard to find relevant clues. The cues that have been used are complex, and at times controversial.

Van Dyken and Kutsch Lojenga (1993) provided an overview of 'rules' that could be used by researchers. These rules are related to semantic, syntactic and phonological cues that can be found in the spoken language. Despite this approach, Van Dyken and Kutsch Lojenga's remarks about combinations of cues that are needed to arrive at sound practice make it clear that word boundaries are not easily identifiable in spoken language. To be sure, perception studies suggest that infants of only 7.5 months old can recognize words in speech streams (cf. Jusczyk & Aslin, 1995; Houston, Santelmann & Jusczyk, 2004), based on rhythmic and statistical cues, and that even 3-day-old infants are sensitive to word boundary cues (Christophe et al., 1994). Nevertheless, indicating where the word boundaries are in full sentences seems to remain a task that, up to a point, only literate speakers or trained researchers can do.

From findings of studies on languages such as Chinese and Japanese that use non-alphabetic scripts, in which words are not marked as separate units, it appears that people only become aware of those units that are marked as such by the writing system (Homer, 2000; Hoosain, 1992; Bassetti, 2005; Chau, 1997; Veldhuis, Li & Kurvers, 2010; Veldhuis, 2011). When literate native speakers of Chinese were asked to segment sentences into words, defining where words begin or end turned out to be a difficult task: there was no consensus among participants, probably because Chinese marks morpho-syllabic units in writing, rather than word boundaries (Bassetti, 2005; Chau, 1997; Homer, 2000; Hoosain, 1992; Veldhuis, Li & Kurvers, 2010). In Japanese, words are not marked in writing either. Japanese literate native speakers who were asked to divide spoken sentences into smaller pieces, showed no consensus in the units they came up with (Veldhuis, 2011). This suggests that words are not necessarily the units that even literates automatically come up with when segmenting sentences.

However, once people have learned a language in which words are distinguished in writing, it seems that they can easily transfer this knowledge onto a second language (L2, see also Geva, 2006) in which words are not distinguished this way (Bassetti, 2005, 2009; Juffermans & Veldhuis, 2012; Veldhuis, Li & Kurvers, 2010; Yao, 2011).

In a study by Bassetti (2005), awareness of words and syllables among 60 native-speakers (L1) of Chinese was compared to the awareness of these units among 60 English-speaking third- and fourth-year-learners of the language. The L1- and L2-speakers applied different segmentation strategies. Specifically, the results showed that metalinguistic awareness, and the strategies people apply when they are asked to conduct a metalexical task, is highly dependent upon the first language and the features of its specific writing system. Bassetti described the differences she found in the strategies used by L1- and L2-speakers of Chinese in the following way:

> '… for most Chinese natives it [a word-DV] is a syntactic unit, while for most English CFL [Chinese as a Foreign Language-DV] learners it is the equivalent of an English word; for both groups a word is a unit of meaning that cannot be further segmented, but for the English CFL learners, this means a mono- or disyllabic unit, while for the Chinese natives it is also a prosodic unit that can be identified by means of pauses and intonation units, a possibility that never occurs to English CFL learners.' (Bassetti, 2005: 249)

It can be concluded from this argument that metalinguistic – or even metalexical – awareness depends on one's mother tongue, and is influenced by the writing system which is used in that native language, if there is one.

Bassetti (2009) and Yao (2011) provide further evidence for this argument. For both studies, the researchers tested whether the segmentation of what can be called words in English, makes reading easier for L1- and L2-speakers of Chinese Hanzi and Pinyin. In Hanzi, each character represents a mono-morphemic and mono-syllabic unit, and in the writing system, words are not usually marked by spaces. In Pinyin, which is the official Romanization system in the People's Republic of China, the letters of the Roman alphabet are used with diacritics for tones. Pinyin is a supplementary writing system, which is used as a pedagogical tool for both Chinese children and L2-learners, and for applications such as bibliographical references and software development. Because Chinese graphemes (Hanzi) map onto the spoken language at the morpho-syllable level, spacing in Pinyin could be used to separate syllables (and with that, morphemes), but usually Pinyin is written with syllables or morphemes grouped in words separated by spacing. Nevertheless, the places where spaces are inserted in Pinyin writing, are not consistent. In her study, Bassetti tested whether a more consistent segmentation of units in Hanzi and Pinyin writing, according to units that corresponded to word units in English, would help Chinese L1- and L2-readers. Bassetti's and Yao's comparisons of the L1- and L2-readers who participated in their studies show that segmentation does positively affect L2-readers of Chinese, who were also used to word boundary marking in their own writing system, but that there were no such effects for L1-readers.

Hsia (1992), who conducted a battery of tasks to investigate American and Chinese monolingual and bilingual children's ability to identify inter- and intraword boundaries, concluded with respect to the bilingual children's segmentation patterns, that with time, bilingual children appeared 'to develop nativelike phonological constraints' (p. 341). This suggests that even if (late, or beginning) L2-learners may mostly rely on the segmentation strategies as they apply in their L1, exposure to the L2, including exposure to the writing conventions of the L2, affects segmentation strategies.

In our own study with Chinese-Dutch children (Veldhuis, Li & Kurvers, 2010) we confirmed that the writing system in which one learns to read and write has an effect on the kind of unit-awareness that one develops. Bilingual children who were only literate in Dutch were not very good at repeating the last character in a Chinese story (especially not when this was

a character that would be translated by a function word in Dutch), whereas children who were literate in Chinese, or in both Chinese and Dutch, almost scored at ceiling.

In another small-scale study, Juffermans and Veldhuis (2012) discussed the writings of two low-literate multilingual adults in The Gambia, who had not received any formal training in writing Mandinka, the main local language. The writings prepared by these men showed hardly any consensus as to where words should begin or end – as indicated by spaces in three texts that they had written. While one wrote 'Nna FAloo FILITA Aga AYINI', for 'my donkey got lost, I searched for it', the other argued in a hand-written rewriting of this sentence that it should have been written as 'NAA – FALOO FEE LEE TAH NYAA NYENE'. In these writings, the letters representing the pronunciation of Mandinka differ, and – more importantly for our purposes – so do the places where word boundaries are marked. Moreover, in a second version, a digital re-spelling, the second participant corrected the sentence again, into 'NAA FALOO FEELE TAH NGA NYENEE', again both changing letters and word boundaries. Juffermans and Veldhuis (2012: 24) concluded that apparently awareness of words is not based on 'automated processes'. For these low-literate speakers, word boundary marking in Mandinka seemed to be mostly a matter of personal intuition and ad-hoc decisions, rather than of pre-existing knowledge. The spoken language did not provide the adults with much information as to where boundaries should be made. In that sense, words seemed not to be psychologically real units in Mandinka.

Comparable conclusions were drawn from a study in which indigenous Mayan children were asked to write down Mayan sentences in the way they should be written, without having had instruction about the correct way of spelling and word boundary marking (Pellicer, 2004). In this study, 135 children who had learned to read and write in Spanish at school, but whose first, and dominant, language was Maya, were asked to write sentences, and they participated in small groups (up to four or five children) in interviews. From the samples that were obtained, Pellicer concluded that the children did not break up their written language randomly, but that they had some rationality behind the places in which they inserted spaces. Nevertheless, Pellicer (2004: 738) states in line with her findings that children did not break up language in places where they should have done so, according to officially accepted rules, and that '… it is one thing to know about the necessity of blank spaces when writing sentences, but it is quite a different matter to know where to put these blank spaces so as to conform to how native linguists would do.' As was mentioned for the study conducted by

Juffermans and Veldhuis (2012), this study showed that the children knew that they had to insert spaces between words, but what the words exactly were, could not be derived from the spoken language, and was therefore not unambiguously clear to them.

Similar considerations apply to language segmentation at smaller levels, such as syllables or phonemes. Selected units appear dependent on one's native language and on the writing system. Tolchinsky and Teberosky (1998), for instance, found for Spanish and Hebrew children, that the native-language of the children they tested and the writing system to which children were used affected the number of syllables and consonants they segmented. The Hebrew children appeared to be better at segmenting and pronouncing single consonants in isolation, which may reflect 'a major typological feature of Hebrew language with respect to the primacy of consonants' (p. 15), which is 'further reinforced by the script', whereas the Spanish children relied more on syllables.

In Japanese, the sub-syllabic unit *mora* is the unit that is stressed, and that is also the unit that is written in *hiragana* and *katakana*, which constitute two of the Japanese writing systems. Accordingly, this might again indicate an influence of the writing system. Inagaki, Hatano and Otake (2000) argue that the *moraic* unit is much more accessible as the basic unit for segmentation in Japanese than the syllable, which is the unit that gets stressed in for instance Dutch and English.

Then again, in a study in which Chinese children and adults performed in a parsing task by Lin et al. (2011), the authors claimed that words do have psychological reality. In this study, it was concluded – as mentioned in the previous section – that children become aware of words through training, or experience. However, in the study by Lin et al. (2011), it was found that the children (second-graders, fifth-graders and college students – the latter in Taiwan), who participated among others in a two- and three-character parsing task, in which they were asked to indicate words and phrases, were all able to indicate words. Therefore, the authors conclude that '… the findings provide evidence for the psychological reality of the concept of word' (Lin et al., 2011: 53). While indeed all children and adults were found successful in the task, the findings showed, as the authors stated as well, that adults have a better-developed concept of word, and that the children were less sensitive to the difference between words, phrases and nonsense-items, that were also included in the task. Lin et al. (2011) also suggest that children come to learn about the identification of words through experience, and that their ability to identify words only gradually becomes more rule-based.

Accordingly, Lin et al.'s (2011) conclusion about the psychological reality of words, that seems to stand in great contrast to other studies, is more refined than it may look at first sight: the authors do suggest that words have psychological reality, because of their finding that even second-grade Chinese children are able to indicate words from strings of characters, but on the other, the authors also mention that experience enhances word awareness. And as the authors tested only second-grades, who were already enrolled in a literacy education program (even if this was in Chinese, in which word boundaries are not marked on paper), but no pre-literate children – for instance in an oral version of the task – their statement about the psychological reality of words can in fact be doubted. The second-graders will after all have come across multiple character units on paper in their literacy classes, in word- or in phrase-units. As the authors also mention themselves that these younger children had difficulties in distin-guishing words from phrases on paper, their findings may not be as different from other studies in the area as the statement about the psychological reality of words may seem to suggest: Also for the youngest Chinese children in Lin et al.'s (2011) study, identifying words from character strings was found to be a hard task.

In addition to the influence from knowing about writing conventions in a specific language, the morphological typology of a language, i.e., the structure of a word, has also been mentioned as a factor affecting recognition and processing. Morphologically complex words have been said to be processed differently than simple words. As Turkish is a synthetic language, many complex words occur, consisting of content and function morphemes. The word '*masada*', for example, consists of two morphemes, '*masa*', 'table' and '*-da*', a general locative marker, together constituting a word that is translated as 'on the table'. As complex words like this are very common in Turkish, it is for its speakers not possible to rely merely on the smallest meaning of sentence parts to come to correct word segmentations, as it is for simple words and for speakers of more analytic languages. Moreover, the frequency of occurrence of morphologically complex words has also been argued to influence their access from the mental lexicon: it has been argued that commonly used complex words are accessed as wholes by adults; less common complex words may be decomposed (Gürel, 1999).

Awareness of words has thus been found to be related to literacy, to conventions in specific writing systems, to morphological typology, and to rhythmic cues in the spoken language. The current study investigates the

awareness of units of literate and pre-literate speakers in two languages that differ in their morphological typology and conventions regarding word boundary marking on paper: this makes it possible to distinguish general literacy effects from typological effects on language segmentation.

2.1.3   Conclusions from tasks focused on metalexical awareness

From the studies described above, it can be concluded that literacy plays a decisive role in children's and adults' metalinguistic, and probably also epilinguistic, awareness of words. Only children and adults who know writing systems in which words are marked by spaces have been shown to be effective at performing segmentation or at identifying content words and function words. Pre-literate children, who have not yet been trained in recognizing and separating single words from speech streams, have been found to be usually only successful in separating meaningful content words. This is in line with the finding that they '… attend more to the meaning of language than to its formal properties …' (Morris, 1993: 134).

Then again, it should be emphasized that the findings mostly come from experimental tasks that were focused on relatively analytic languages, in which function words and content words are separated by spaces in writing. In more synthetic languages, such as Turkish, co-occurring content morphemes and function morphemes are usually kept together as single words. This example illustrates that speakers of Turkish cannot rely only on the function or meaning of sentence parts to come to correct responses in word segmentation or description tasks. The question is how word recognition and word segmentation works for these speakers, whether it differs from what speakers of analytic languages do, and to what extent pre-literate and literate speakers differ in languages like this in what they recognize and process as units.

Another question is to what extent the results reviewed above are representative of the units that people store and process in their mental lexicon. All of the tasks described above, focused as they were on either metalinguistic or epi-linguistic knowledge of words, can be regarded as relatively 'offline' tasks (see Chapter 3), as they allow participants to make conscious decisions about the language sample they had to describe or segment. To what extent this corresponds to people's unconscious processing of language, in perception and production of speech, remains as yet largely unknown.

2.2 Cognitive Linguistics on unit processing

The question of what constitutes the basic units of language has also been raised in the fields of Cognitive Linguistics and Construction Grammar. In contrast to the receptive focus of metalinguistic studies, and its focus on segmentation, Cognitive Linguistic work often zeroes in on the units that people rely on in parsing and producing language. Awareness of units in the metalinguistic sense is less of an issue in these studies.

In this section, proposals in Cognitive Linguistics on what may count as units in language will be outlined in Section 2.2.1. In Section 2.2.2, some major results will be discussed from studies with children, followed by a discussion in Section 2.2.3 of what we know about the impact of literacy on the processing.

2.2.1 Units in processing

Studies in the Cognitive Linguistics and Construction Grammar traditions have claimed that larger units, i.e., units that contain some internal complexity, form the basic planning units in language use rather than single words (cf. Croft, 2001; Goldberg, 2006; Langacker, 2008; Tomasello, 2003). These multiword expressions can sometimes be fully or maximally specific, as in the case of formulaic sequences (e.g., 'kick the bucket', cf. Wray, 2002) or sequences that happen to occur often. One can imagine, for example, that in the speech of some people the clause 'It rains a lot in Holland' is recurrent, or parts of it are (Doğruöz & Backus, 2009). Some multiword expressions instantiate constructions that are at least partially schematic, as in constructions like 'the X-er the Y-er' (cf. Goldberg, 2006), or, as an instance of the sentence mentioned before from Doğruöz and Backus (2009), a combination of words and open slots that can be filled with specific words, such as [*It* $V_{weather}$.pres. ADV *in* N] (See Appendix 12 for a list of abbreviations). Completely schematic constructions, such as ditransitive, causative or passive constructions, or morphological constructions such as past tense [V + -*ed*], or a combination of a verb plus an adverb in [V ADV], may also form part of the inventory of basic units of language (Doğruöz & Backus, 2009; Goldberg, 2006; Langacker, 2008; Tomasello, 2003). This latter type often corresponds to what can be regarded as syntactic templates, while the first type, the maximally specific units, are part of the lexicon. Partially schematic constructions are somewhere in between, and they can often be regarded as a combination of a syntactic pattern and one or more lexical items, containing slots in which words or morphemes can be inserted in a

restricted creative manner – 'restricted' because usually not just any kind of word or morpheme can be inserted in a slot.

In a schematic representation, this leads to a visualization such as provided in Figure 1:

Most specific	Partially schematic	Most schematic
●——●		
Lexicon		Syntax
[*rains a lot*]	[*rain*-pres. ADV], [V *a lot*]	[V ADV]
[*It rains a lot in Holland*]	[*It* Vweather.pres. ADV *in* N]	[SUBJ V PP]
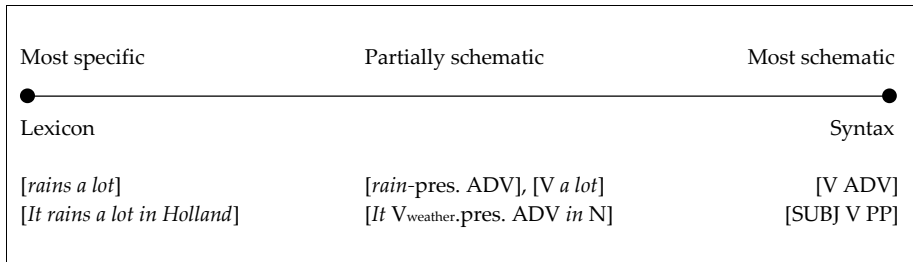
Figure 1: Schematic representation of different types of multiword expressions (adjusted from: Doğruöz & Backus, 2009: 44)

Since various types of constructions can overlap, any utterance is assumed to consist of an accumulation of word and multiword units of different types inserted into overlapping constructional templates (Ellis & Cadierno, 2009). Any sentence then consists of an accumulation of constructions (cf. Mos, 2010), or, as Goldberg announced: 'It is constructions everywhere' (LSA conference, 2009).

For linguists, the question that will be raised when defining language as accumulations of constructions is what makes up a construction, or when a sequence of co-occurring words can be identified as the instantiation of a construction (cf. Smiskova-Gustafsson, 2013). In general, constructions are defined as sequences of words which usually occur together in a similar way or with only limited variation, and that therefore may be stored and processed as wholes. There is thus an underlying usage-based approach to theorizing, which suggests that every instance of usage, or its frequency of occurrence, increases that unit's degree of entrenchment, i.e., the strength of its trace in the memory of the person who has just used or heard the unit (cf. Barlow & Kemmer, 2000; Taylor, 2002).

This idea that higher usage leads to better entrenchment of multiword sequences and generalizations of more schematic templates, implies that frequency data should tell us much about what is and what is not well entrenched, and to what degree a unit is entrenched. The last decade has produced an impressive body of work that attempts to find evidence for this approach, in which usually large databases are taken as a basis for frequency

information about particular units (cf. Barlow & Kemmer, 2000; Conklin & Schmitt, 2008; Gürel, 1999; Langacker, 2008; Lehtonen & Laine, 2003; Soveri, Lehtonen & Laine, 2007).

Langacker (2002) was among the first to argue that there is a correlation between the frequency of occurrence of a unit – as an example Langacker provided the past tense of 'to drive' – and its entrenchment. Though the regular past tense of 'to drive' could be predicted to be 'drived', based upon frequency speakers know that this would be ill-formed. For the relatively infrequent verb 'to thrive', however, Langacker argues that the past tense form 'thrived' might be accepted more easily. This is explained by the fact that its past tense form does not occur as often as 'drove' or 'driven', while the regular past tense schematic construction [V + -ed] does. What Langacker thus suggests is that the token frequency of a unit affects its entrenchment.

Experimental research extrapolates from this hypothesis and predicts that entrenchment will affect the ease and speed of processing. This hypothesis has been tested repeatedly. Adult participants have for instance been asked to perform a cloze test in which they had to replace function words (cf. Conklin & Schmitt, 2008), or they were tested on whether or not they processed complex words or inflected nouns holistically or not (Gürel, 1999; Lehtonen & Laine, 2003; Soveri, Lehtonen & Laine, 2007). Other studies tested the way in which frequency of occurrence enhanced the memorization of constructions occurring in stimulus sentences (Ehrismann, 2009; Mos, 2010). All of these studies confirmed that formulaic sequences, frequently occurring constructions and multi-morphemic (complex) words may well be processed as wholes, whereas complex words with lower frequencies tend to be decomposed in processing (Gürel, 1999).

Barlow and Kemmer (2000) also stress the importance of frequency. They argue, as Taylor (2002), that though more frequent occurrence of a linguistic unit will lead to better entrenchment, linguistic representations should properly be seen as emergent, rather than stored fixed entities. They have to be regarded as cognitive routines, and the linguistic representations are therefore '… nothing more than recurrent patterns of mental (ultimately neural) activation …' (Barlow & Kemmer, 2000: xii). What Barlow and Kemmer argue is in fact that constructions are not automatically stored as wholes to begin with, but need to get entrenched before they can be processed as wholes.

Bybee (2002) and Dąbrowska (2004) also argued that frequency contributes to the use and production of multiword sequences. Bybee (2002) showed that frequently occurring multiword sequences, such as 'do not know', are usually reduced in spoken language, in this case to 'dunno'. This

phonological reduction would suggest that specific multiword phrases are likely to be perceived and produced as one single block, and not as a combination of separate words, even by young children. Dąbrowska (2004) argued that the frequency of especially different types of words, and the phonological similarity of, in her study, different complex Polish words, contributes to their use and the production by adults and children. As Dąbrowska argues, children and adults can make generalizations over the inflection of complex words both on the basis of regularity – according to type frequency – as well as to phonological similarity – which would also work for irregular complex words, and the production of new irregular words.

2.2.2   Cognitive Linguistics on children's processing units

In 2000, Tomasello argued that children's syntactic competence may differ from adults'. In line with this idea, Cognitive Linguists such as Tomasello (2006), Bannard and Matthews (2008), Bannard and Lieven (2009), and Arnon (2010), have focused on the units that children use and produce in language. As the Cognitive Linguists discussed above in Section 2.2.1 in relation to mostly adult language use, these researchers also suggested that toddlers base their utterances on units larger than words, rather than on individual words. In his work on how children acquire language, Tomasello (2006) stresses that they usually start out imitating their parents or caretakers, not only in language, but in behaviour more generally. Only later do they start to use language in a more creative way and do they manipulate the language they know. As children are usually not exposed to utterances that consist of single words, nor to explanations of grammar, but rather to connected speech streams, Tomasello argues that children often acquire multiword phrases early on, and only later realize what the parts are of which such phrases are composed.

The view that children first imitate their parents' or caretakers' language and only later apply their knowledge in a more creative way was empirically confirmed by Bannard and Lieven (2009). They compared the speech production of 2-year-old toddlers to their caregivers' language input, and found that only 3% to 14% of all children's utterances were not directly derived from speech strings they had used or heard before. Imitation of language the children had been exposed to, specifically of strings of multiple words, was found to be highly important in shaping children's utterances. Children simply copy the speech streams they hear, relatively independent of how many words such streams consist of.

In an experimental study into the way in which L1- and L2-children and adults use compositional multiword phrases of higher or lower frequencies of occurrence according to corpus analyses, how they complete them or produce them, Arnon (2010) found support for the Starting Big Hypothesis: children start out with units that are usually bigger than words, and only later accumulate an understanding of the relations between linguistic units at smaller levels (Arnon, 2010).

Further experimental support was provided by Bannard and Matthews (2008). In contrast to Arnon (2010), they did not use an adult corpus as the baseline for frequency-data of specific multiword sequences, but rather used a corpus of language produced by and directed to a 2-year-old boy. Against this baseline, Bannard and Matthews tested how well children aged 2 and 3 were able to repeat four-word-sentences that either had a high frequency of occurrence (e.g., 'sit in your chair') or a low frequency of occurrence (e.g., 'sit in your truck'). Their results show that both the 2- and the 3-year-olds were more likely to correctly repeat the first three words of the high frequency multiword units and that 3-year-olds were in general better and faster in repeating the sentences than the 2-year-olds. Furthermore, all children pronounced the first three words of high frequency items faster than those of the low frequency items. This supports the hypothesis that children use concrete, exemplar-based multiword items as units in their language production, just like adults.

A study by Brandt, Verhagen, Lieven and Tomasello (2011) also provides evidence for the idea that frequency may affect children's language processing. In a production task conducted in German, the children were in this study prompted to change high and low frequently occurring verbs of different types (transitives, mental-state complement-taking verbs, and communication complement-taking verbs) in sentences that they had to produce. The results from this study showed that the children were better at changing low frequency transitive verbs into high frequency transitive verbs than the other way round, but no clear effects, or even reverse effects, were found for verbs of the other types.

Accordingly, the question of how frequency affects children's language processing remains as yet unanswered. Does the fact that children produced high frequency multiword sentences better and faster (Bannard & Matthews, 2008) really mean that children process these sequences as wholes? Does the time that it takes children to repeat phrases tell us something about processing? And does the effect of frequency diminish over time, as children become better at decomposing and analysing language as they get older, as Arnon's study suggested?

These are questions that are also touched upon in the current study, especially with regard to literacy. After all, literacy was found by the metalinguistic studies reviewed above to have a decisive effect on the extent to which children are able to decompose language into smaller units.

2.2.3 Literacy effects on online language processing

The fact that literacy has so far played no role in Cognitive Linguistic studies that aim to explain the units in children's and adults' processing, does not mean that there are no studies at all in the field of Cognitive Linguistics that have investigated literacy effects on language processing. On the contrary, there are quite a number of studies in which ERP- and fMRI-techniques, as well as eye-tracking methodologies, have been applied to examine how language processing in illiterates and pre-literates differs from that of literates.

A growing body of evidence suggests that literacy affects the online processing of spoken language and that it changes the functional organization of the brain, or more specifically, the density of white matter adjacent to regions involved in the processing of language when producing or listening to language (cf. Carreiras et al., 2009; Ostrosky-Solis, Arellano Garcia & Perez, 2004; Pattamadilok et al., 2010; Petersson, Ingvar & Reis, 2009; Schild, Roder & Friedrich, 2011).

Studies by Morais and Kolinsky (1995), Reis and Castro-Caldas (1997) and Kosmidis, Tsapkini and Folia (2006), for instance, have also shown that knowledge of phonemes, as acquired in literacy education, seems to influence the processing of pseudo-words, but not of existing words. Portuguese (Reis & Castro-Caldas, 1997) and Greek (Kosmidis, Tsapkini & Folia, 2006) literates and illiterates had no different responses when repeating existing words, but literates were better at repeating pseudo-words, probably as a result of the phonological processing route that literates, who are trained in recognizing phoneme-letter correspondences, may use (cf. Reis & Castro-Caldas, 1997; Kosmidis, Tsapkini & Folia, 2006). An ERP-study among the same participants in Reis and Castro-Caldas' study reinforced this idea, as the brain activation in literates also differed from that in the illiterates.

Another study that compared language processing in pre-literates and literates is Schild, Roder & Friedrich (2011). They investigated to what extent pre-literate and beginning readers differed in their lexical access. Their results confirmed the suggestion by Reis and Castro-Caldas (1997), that

lexical access in pre-literate children is less based on acoustic detail than in beginning readers.

Eye-tracking studies have so far not often dealt with differences in language processing between pre-literates and literates, presumably because most eye-tracking studies rely on written stimulus material. There are a number of studies on literacy activities with pre-literates, but these are mostly focused on where young children look in books and whether they look at print or illustrations (cf. Evans & Saint-Aubin, 2005). Eye-tracking studies with literates, on the other hand, are more numerous, including all kinds of studies of language processing, such as reading strategies in different languages (cf. Kemper & Liu, 2007; Peng, Orchard & Stern, 1983; Rayner, 1998), the processing of garden-path sentences (cf. Trueswell et al., 1999) and anticipation in sentence processing (cf. Kamide, Altmann & Haywood, 2003).

Eye-tracking studies that focus on reading and reading strategies are the ones most closely related to the current study, as they involve an interest in the kinds of units that people rely on in reading, e.g., the number of letters they can maximally take into account when reading, or the effect of placing spaces between characters in written Chinese, a language in which usually no spaces occur in text. However, these tasks cannot be performed by pre-literates, and, moreover, they are only focused on reading as one specific kind of language processing.

Thus, despite the findings in these studies, and even though a relationship between literacy and the recognition of words has been identified in offline studies of children's metalinguistic development, it has not been established yet from the perspective of Cognitive Linguistics how literacy affects the units that people utilize in online language processing.

2.2.4 Conclusions from studies into unit processing in Cognitive Linguistics

It is clear that also within Cognitive Linguistics, attention is paid to the way in which people parse and produce units in language, and to identifying what those units are. So far, studies with children have indicated that they usually start imitating their parents' linguistic productions, and this includes many multiword units, as claimed by Arnon's (2010) Starting Big Hypothesis. The offline studies into children's metalinguistic development discussed earlier suggested something very similar: children appear to treat larger sequences as units, before figuring out that they consist of smaller parts.

Studies with adults have shown that the frequency of occurrence of multiword sequences plays a crucial role in their processing and production of utterances. Bannard and Matthews (2008) and Brandt et al. (2011) investigated possible frequency effects with young children based on frequency data from a single child and with different types of verb-forms. To date, however, more elaborate studies in which children are tested on frequency effects in their language processing are lacking, although children's language processing might, as their syntactic processing (cf. Tomasello, 2000), differ from adults'.

The role of literacy, found to be a critical factor in children's metalinguistic development, has not been given much attention in Cognitive Linguistic studies. Whenever literacy is explored, it has not been related to the processing of units in language.

What all the cited studies confirm is that single words are not necessarily the basic units of language, in terms of what people process when they listen to or produce utterances. Instead, various kinds of multiword units, which can be based on more or less schematic templates, are more probable candidates for the status of basic unit in processing.

2.3   Connection to the current study and summary of research questions

In the current study, pre-literate and literate children took part in experiments in which they had to segment units in relatively offline tasks and use units in processing language in relatively online tasks (see Chapter 3). This way this study connects with the studies of metalinguistic development, with their focus on the role of literacy, and with Cognitive Linguistic studies, and their focus on unconscious language processing.

In addition, the possible impact of the frequency of multiword sequences, as is also often mentioned by Cognitive Linguists, will be examined as well, also in order to see whether this impact overrules the expected literacy effect. Testing for this impact is made possible by the inclusion of the stimulus items consisting of multiword sequences with higher and with lower frequencies of occurrence.

Finally, the possible impact of typology is included in the design. Typology is usually accepted to involve all structural and functional features of a language. This also includes the way in which a language is organized, thus for instance whether or not morphemes are bound or free. If one language's convention is synthetic or agglutinative, which means that it

tends to bind morphemes, whereas another is more analytic, and distinguishes more free morphemes – both in writing and in speech –, then this may affect the ideas people have on what constitutes the basic units of language. Moreover, for literates, the conventions they know about word breaking or word marking will also show in the respective languages on paper. There will be typological spelling differences in places where word breaks occur between the languages.[4]

Since findings of language segmentation tasks suggest that languages with their specific conventions on paper in word marking, and, for instance, binding morphemes or not, affect people's knowledge of units, and as it has also been argued that phonological features of a language, such as phonological reduction, may affect unit recognition, the tasks developed for this study were prepared in two typologically unrelated languages – Turkish and Dutch. They were performed by Turkish, Dutch and some bilingual Turkish-Dutch speakers. Turkish is an agglutinative language, in which most words are multi-morphemic (often combining function morphemes with content morphemes), and Dutch is a more isolating language, in which function and content morphemes are usually separate words. Including both pre-literate and literate speakers of Turkish and Dutch in this study, would make it possible to distinguish typological effects of a language on the units that children recognized and processed from effects of knowledge of the conventions in writing and word marking of the languages on paper.

Chapter 4 will go into further detail about the tasks that were developed for this study; before that, Chapter 3 will provide a discussion of what can be regarded as 'relatively offline' and 'relatively online' tasks, as this distinction is more intricate than the dichotomy used so far between 'segmentation' and 'processing' suggests. First, however, the research questions posed for this study will be outlined:

Main research question:
What is the influence of literacy and typological background, and of the frequency of occurrence of specific multiword sequences, on the way in which children segment and on how they process spoken language?

---

[4] Orthography, which is a related term, is the methodology of a written language. Orthography includes rules of spelling, with a focus on the grapheme-sound relation in a written language, and aspects such as capitalization, punctuation, and word break marking. Word breaking, or binding words, is however not only related to writing, but also to the typology of a language itself – and whether or not words can for instance be separated by other words, or put in an inverse order (cf. Coulmas, 1989).

Sub-questions:

(1)   How do offline and online tasks differ from each other? (Chapter 3)

(2)   How can offline and online tasks be applied to investigate language segmentation and processing units among young, pre-literate children? (Chapter 5)

(3)   What is the effect of typological background (being a speaker of Turkish or Dutch) on the units that children segment and process in language? (Chapter 6-8) And what is the effect of being bilingual in Turkish and Dutch? (Chapter 4)

(4)   What is the effect of literacy on the units that children segment and process in language? (Chapter 6-8)

(5)   Does the entrenchment of multiword sequences, as measured by frequency of occurrence of such sequences, affect the segmentation and processing of such sequences in young children? (Chapter 6-8)

(6)   Is there an interaction between the effects of literacy, typology and the frequency of multiword sequences on the way in which children segment and process language? (Chapter 6-8)

As can be seen from these research questions, this study in fact investigates three dimensions of children's language segmentation and processing.

Offline versus online tasks: Not dichotomous but continuum?[5]

The main question asked in this study is whether the units that children recognize and process in language differ in relation to their literacy skills and language background, and the frequency of occurrence of individual multiword sequences. In order to investigate these questions, several experimental tasks were developed. One could perhaps also observe spontaneous language use, but comparing utterances from completely natural language behaviour and conversation on pauses that may indicate processing units in language, is hard, because of the large variation that people show in their utterances.

From the literature on unit awareness and unit processing, as discussed in Chapter 2, it is clear that many experimental tasks have been developed for investigating how adults and children decompose language into building blocks, or use them to compose an utterance. Generally speaking, this range of tasks can be divided into offline and online tasks. 'Offline' refers to conscious or explicit knowledge, or metaknowledge, related to what Homer (2009), following Piaget, referred to as (one's) *know-that*; 'online' relates to the unconscious, or implicit knowledge or processing, which can be described as one's *know-how* (cf. Homer, 2009: 488; Marinis, 2008; Piaget, 1976).

Using both types of tasks, it should be possible to answer the main research question of the current study, as offline data have been argued to zoom in on the conscious decisions made in language segmentation, and online data on the unconscious processes that underlie natural language processing.

Since, however, it has been found that there are degrees of awareness (see Chapter 2), rather than a dichotomous distinction between aware and

---

[5] This chapter elaborates on Veldhuis & Kurvers (2012).

unaware, the definition of tasks as simply offline or online needs to be reviewed. Offline and online tasks do not rely clearly on either conscious or unconscious language processing, but on the whole seem to range from completely unconscious to entirely conscious modes of using language. The two-way division of tasks does not match this range.

Instead of a two-way division between offline and online tasks, this chapter proposes a continuum, ranging from relatively offline to relatively online tasks. First, the differences between offline and online tasks will be examined in detail (Section 3.1). Doubt about the accuracy of the dichotomy leads to a survey of issues and characteristics that should be taken into consideration to come to an alternative characterization as a continuum (Section 3.2). An evaluation of this proposal, in the light of the theories discussed in Chapter 2 and the degree to which they match the empirical results of the present study (to be presented in Chapter 5-7) will be undertaken in Chapter 8.

3.1   Differences between offline and online tasks

In the 1960s, the question was raised what the difference is between language competence and performance (cf. Chomsky, 1964, in Sekerina, Fernández & Clahsen, 2008). Generativists believed that intuitions about grammaticality tapped into competence, and that this referred to the idealized linguistic capacity or knowledge people have about language. Performance was instead assumed to indicate how the processing mechanisms worked that put that knowledge to work. Still, most experimental studies into children's language development up until the 1990s focused only on competence, which was tested with the help of behavioural and judgment tasks. These tasks can be referred to as 'offline' tasks: they provide information on explicit knowledge about language.

Since the 1990s, there has been a shift in experimental studies towards a focus on performance. Such studies usually work with 'online' tasks, which measure language processing in real time. These tasks concentrate on implicit knowledge, as reflected in the unconscious production of language, and they allow inferences about processing routes and mechanisms. In this sense, offline and online tasks can be contrasted, and seem dichotomous.

With the introduction of online tasks, scholars have shown increased interest in exactly how tasks should be defined as either offline or online (cf. Karmiloff-Smith et al., 1996). It turns out that studies in applied linguistics and in cognitive psychology tend to differ in their definitions: some for

instance refer to a task as online only if it involves real time measuring, whereas others focus more on whether the task taps into implicit knowledge, or the presence or absence of any conscious reflection on linguistic performance (cf. Fernández & Smith Cairns, 2011; Sekerina, Fernandéz & Clashen, 2008). Accordingly, there is disagreement on whether or not a task should be called offline or online (see Karmiloff-Smith et al., 1996, and Kurvers & Uri, 2006).

In general, however, in studies in which offline and online tasks have been used side-by-side, offline tasks are usually those that elicit explicit reflection on language. Accordingly, offline experimental tasks – often used in studies on children's (meta)linguistic development – include tasks in which participants have to apply their knowledge, such as answering comprehension questions or providing judgments (cf. Fernández & Smith Cairns, 2011). These tasks are focused on explicit knowledge and usually allow participants time for consideration and reflection.

Online tasks, in contrast, try to capture in some way mental processing loads or brain activity (cf. Marinis, 2003), usually as participants use natural language, productively or receptively (cf. Clahsen, 2008; Marinis, 2003; Sekerina, Fernández & Clashen, 2008; Fernald et al., 2008; Fernández & Smith Cairns, 2011; Männel & Friederici, 2002; Snedeker & Thothathiri, 2008). Online tasks may focus on neural mechanisms, or give insights into real time procedural operations (Fernández & Smith Cairns, 2011; Männel & Friederici, 2002; Sekerina, Fernández & Clashen, 2008).

The methods used for investigating neural mechanisms, directly or indirectly (Männel & Friederici, 2002), usually include technologically advanced techniques and devices, such as ERPs, fMRIs, or eye tracking. Self-paced reading or self-paced listening tasks have also been claimed to have an online nature. The same holds for click tasks (cf. Cohen & Mehler, 1996; Ladefoged & Broadbent, 1960) in which participants are asked to detect clicks which are superimposed on orally presented sentences, as they tap into real time language processing (see Chapter 5 for a more elaborate discussion).

The distinction between offline and online tasks thus relates not only to the often mentioned time lag between the stimulus and the participant's response (Theodore Marinis, Nel De Jong, personal conversations), which should indeed make a difference in whether or not participants are conscious of their answers, but also to whether their processing is distorted or not, and whether the stimulus material is manipulated or not. In general, the distinction hinges on the performance or behaviour that is required from

the participants in a language task:[6] as long as a task provides insights into real time processes in which natural, unconscious, language processing or production is not obstructed, a task can be regarded as an online task (cf. Rayner et al., 2007; Sekerina, Fernández & Clashen, 2008; Trueswell, 2008). This is the case for instance in tasks in which eye movements are tracked. Whenever metalinguistic awareness and explicit commentary on language is involved, participants have time to consider their language use (or their judgments), or natural processing is disturbed, a task can be regarded as an offline task. The next section, however, argues that it is better to order tasks on a continuum than to categorize them into two discrete categories.

3.2   Offline and online tasks: A proposal for a continuum

Since the characterization of a task as offline or online thus highly depends on whether participants make use of their metalinguistic knowledge, and since studies of metalinguistic development argue that consciousness is gradable (see Chapter 2), a simple two-way distinction between offline tasks that tap into metalinguistic knowledge and online tasks that do not focus on such knowledge, seems insufficient. At least three stages have been claimed to exist in children's development of linguistic awareness: complete lack of any metalinguistic awareness, some control of linguistic operations, and clear explicit awareness (see Chapter 2). Therefore, three different types of tasks would seem to be needed to investigate these stages. However, as the cut-off points of these stages cannot be defined precisely, even a three-way distinction seems not to suffice. It cannot be known for sure when children move from one stage to another, let alone that specific tasks can be focused so precisely that they only target one of these stages of children's metalinguistic awareness.

What is more, since there are also several other factors that affect whether language can be processed unconsciously or not, and which are not all dichotomous either, a more subtle categorization of tasks seems desirable. Some experiments focus on people's judgments, whereas others require participants to manipulate language. These types of tasks are now both regarded as offline in nature, but they differ considerably in the linguistic knowledge they tap into. Accordingly, the term offline seems too broad. In addition, it seems hard to define tasks as either completely offline or

---

[6]   But see Guilqin and Gries (2009: 5) who relate online language processing to the naturalness of language data collection.

completely online: some tasks may focus to a large extent on unconscious processing, but still require some use of metalinguistic knowledge. Calling such tasks online, as has been done so far, is imprecise.

The lack of a clear definition of the extent to which the different factors may or should be present in the tasks in order for them to be regarded as either offline or online, makes the dichotomy problematic. There is no clear cut-off point for the degree to which a task may include manipulated language materials to still be called an online task, or the extent to which it may involve metaknowledge. Some experiments might test real time, natural, unobstructed language processing, but still use modified language, whereas others may take natural language use as a basis, but call for metalinguistic skills or conscious judgments. Rather than a strict two-way distinction between entirely offline and entirely online tasks, a continuum between those poles will be suggested.

3.2.1 Criteria for the placement of tasks on the continuum

In order to see to what extent a task should be regarded as offline or online, all factors or criteria that are involved have to be considered. Their presence or absence should be checked to see whether a task should be regarded as more on the offline or more on the online side of the continuum.

The factors involved relate to the focus, design and materials that are used, as well as to what is required of the participants. Below, questions that will need to be considered will be dealt with separately. The questions all focus on aspects which have previously been mentioned as important in characterizing tasks as offline or online (cf. Karmiloff-Smith et al., 1996; Kurvers & Uri, 2006; Marinis, 2003; Sekerina, Fernández & Clashen, 2008).

Design of the task
Since the main difference between offline and online tasks seems to lie in the degree to which participants rely on conscious thought about language, this aspect should be investigated before any additional criteria. Therefore, the first question may be:

(1)   Are the participants aware of the fact that a language task is at stake?

If the answer to this question is 'yes', then it can be suggested that the task is located more on the offline side of the continuum, as it can be expected that participants will try to, consciously or not, control their linguistic productions and decisions more if they know that they participate in a language task than if they are unaware of this. If, in contrast, the question is

answered negatively, then the task will appear more on the online side, as any data obtained from participants will then only reflect their unconscious dealing with language. Tasks can then be ranked on the continuum as in Figure 2. This will provide a first global indication of the offline or online nature of the task.
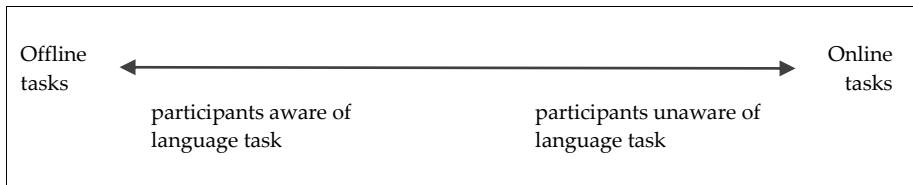
Figure 2: First division of tasks over proposed continuum

It should be noted that in Figure 2 the extremes of the continuum are not actual endpoints, as indicated by the arrow tips. The reason for this is that, to the best of my knowledge, there are no purely offline or purely online experimental tasks. A task that only taps into unconscious language processing, blocking any conscious thoughts about language, is hard to imagine, since participants in experiments usually at least know that they are going to do some kind of language-related task. Hence, they may pay more attention towards language, at least some of the time, than during regular language processing in everyday conversations. On the other hand, it is also impossible for a task to induce a complete focus on conscious thinking about language, blocking all unconscious processing.

Second, it should be emphasized that the placement of the tasks is not precise, as it is not known where and how cut-off points should be placed between regions of the continuum. This first localisation is meant only as an indication as to what the continuum could look like, but it is impossible to say at this stage where the relatively offline or online side begins or ends.

Degree of consciousness is not the only factor affecting the nature of the task. In order to come to a more refined placement on the continuum, further characteristics of the tasks need to be evaluated. Most of these were discussed in Veldhuis and Kurvers (2012), and can be summarized through the following questions:

(2) Does the task consist of elicited material, such as utterances prepared by researchers focusing on specific aspects of language? Does it include pseudo-language, or pre-selected sentences (thus not language participants produced themselves)?

(3) Does the task include unnatural language, such as sentences in which noise is inserted, that disturbs normal, regular language processing?

(4) Does the task involve answering a question after an utterance or sentence has been completed?

(5) Does the task allow participants as much time as they like before they respond (i.e., is there no time constraint)?

Question (2) looks at the type of language that is used in a task, specifically in the stimulus items. If a task includes artificial language, natural language processing is of course disturbed. In such cases, participants will need to apply their (meta)linguistic knowledge, which will make the task more offline. If a task includes regular language, processing will be more natural, as it will resemble everyday language use more. If, however, a researcher has prepared materials for a task, such as elicited sentences, then there may not be completely normal language processing, even if the sentences used in a task might occur in real world conversations as well. As long as sentences are given one by one to a participant, the resemblance to normal conversation is minimal. In such a case, the participant will need to pay attention to language more and in a different way than in regular conversation, as the sentences do not occur in a context. The attention this requires from the participant will make the task relatively offline.

The question raised in (3) is related to that in (2), but focuses on whether or not the language materials are presented in a natural way. In psycholinguistic experiments, the standard way of providing stimulus material is often unnatural (e.g., sentence by sentence, or selected text). If regular language processing is disturbed, for instance by presenting participants with sentences in which noise is included on purpose (especially if participants have to manipulate or react to that noise), then participants will be hindered in their efforts to process the contents of the language they are presented with. Therefore, the task will not be a completely online task, perhaps even relatively more offline.

Question (4) is also related to this idea, as it concentrates on whether or not the language presented to the participant is complete. If participants are asked to answer questions about the language they have been exposed to in the task after the sentences have unfolded, then participants will need to use their memory to come to a decision about their response. If, on the other hand, participants only briefly go 'offline', as Karmiloff-Smith et al. (1996) called it, in the middle of more regular language processing, then the task will be placed more on the online-side of the continuum.

Finally, if there is no time constraint in a task, as mentioned in question (5), allowing participants to consider the stimulus material for as long as they like, the task has to be considered to be relatively offline. If, however, prompt answers are required, allowing participants little time to reflect, then the task will be more online in nature.

Role of the participants
The above mentioned questions all focus on characteristics of the task and the materials. What the participants have to do while carrying out the task also contributes to it being more offline or more online in nature. Questions focused on the role of participants that should also be asked, are the following:

(6) Do participants have to make judgments about language, or provide definitions of linguistic elements?

(7) Do participants have to manipulate the language in a behavioural task?

(8) Do participants have to conduct tasks that provide insight into their procedural operations, or do they have to do 'nothing special' (i.e., blood pressure may be taken, or eye movements are tracked, without the participants knowing that these reflect their language processing)?

(9) Is the specific goal of the task known to the participants?

(10) Do the participants have to read and/or write (instead of listen and/or speak)?

If the answer to question (6) is 'yes', then the task will be of an offline nature – as linguistic judgments clearly require metalinguistic knowledge. If question (7) applies, then the task will also be relatively offline – as it will reflect metalinguistic, or, in the terminology of Gombert (1992), 'epi-linguistic', knowledge (see Chapter 2). However, in behavioural tasks in which the participants manipulate language but are not asked to provide metalinguistic commentary, the researchers will have to interpret in some way to what extent, and which, metalinguistic skills were employed by the participants. Accordingly, these tasks will be more online than judgment tasks in which the use of metalanguage, such as 'this sentence consists of three words', more clearly reflects participants' metalinguistic knowledge or awareness. If the task allows a researcher to interpret participants' automated procedural operations in dealing with language, as touched on in (8), then the task will be more online in nature, since the participants produce language without explicitly being asked to monitor it. Tasks in

which participants do not have to do anything special with the language to which they are exposed, while the researcher might still investigate and examine their processing mechanisms or their neuronal organization, as mentioned in question (8) as well, are the most online: these tasks clearly allow a relatively uninhibited focus on the unconscious linguistic processes that take place in participants.

Question (9) asks what participants know about the task. If the specific goal of the task has been explained to them, they will know what to focus on, or they can guess which behaviour will lead to the highest (correct) scores in a task. In such cases, participants may (try to) adjust their linguistic behaviour accordingly. This will lead to some degree of consciousness regarding linguistic choices, which is a feature typical of offline tasks.

In (10), the question is raised which modality participants use to complete the task. The modality seems to affect the offline or online nature of a task. As Havelock (1986) (see also Kurvers, 2002) has observed, writing does not express concepts directly. Accordingly, writers have to take some distance from the ideas or things that they want to convey, and reflect on the language. This reflection will make a task in which writing is involved more offline than a task in which speaking or listening is involved. In reading, a more or less similar process is at stake. Reading and writing can be seen in some respects as 'poorer' language than listening and speaking: there is no prosody (Harris, 2009, following McLuhan, 1964), no elaborations on the topic under discussion can be called for, and the writer of a text cannot clarify himself by repeating the point he wishes to make in other words (see Kurvers, 2002), or by adjusting his tone – all things people usually do in spoken conversation. Accordingly, in the words of Olson (1994: 93) (see also Kurvers, 2002: 8), writing 'captures only a privileged aspect of the utterance, namely "what is said" and not "how it is to be taken"'. Readers have a harder time than listeners to infer connotations. Therefore, they need to reflect on the language used on paper more than on the language they might hear in a conversation. This makes a task that relies on written language more offline than one in which listening and speaking are involved.

After these questions are answered, a more subtle indication of where the tasks belong on the continuum can be given. This is represented in Figure 3.

Conclusions about questions

Depending on the answer to the ten questions posed above, tasks can be located on the continuum. They distinguish between features of the task, and requirements of the participants.
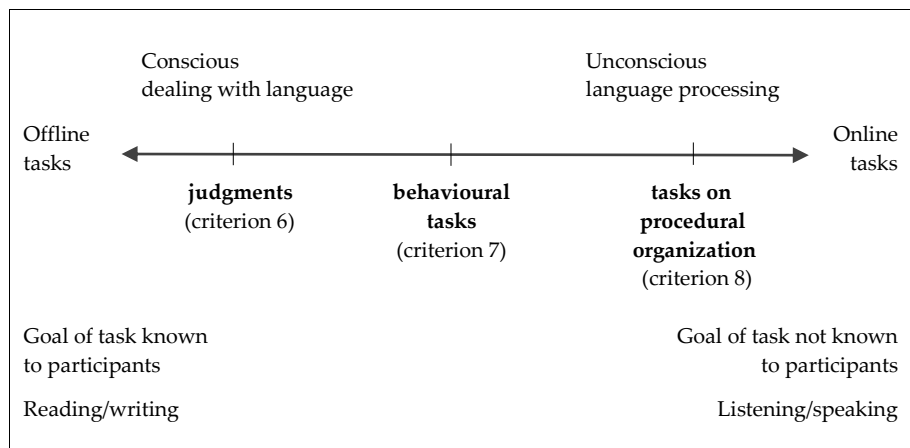
Figure 3: More detailed division of tasks over proposed continuum

The following section discusses in more detail how the answers to the questions should be interpreted.

3.2.2 Defining tasks as offline or online after answering the questions: A preliminary proposal

As mentioned before in Section 3.1, time management in a task is often regarded as the most important factor for the identification of a task as either online or offline (Theodore Marinis; Nel de Jong, personal communication; Fernández & Smith Cairns, 2011). This criterion is reflected in question (5) above. However, some tasks, such as a listening task in which participants have to make grammatical or semantic decisions in the midst of listening to stimulus material, have both offline and online features. In such a task, participants process the text in a natural way while listening to it. This would place the task on the more online side of the continuum. However, since participants will be dealing with elicited material (material criterion, cf. question (2)), the task also has offline features. Hence, the argument that time management is sufficient to indicate whether a task is offline or online, seems inadequate. The other criteria, covered by the other questions, cannot be neglected either: all criteria together reflect whether a task can be called offline or online. The weighting of these criteria for an ultimate decision, however, is a difficult matter (see Chapter 8).

    Therefore, the relative importance of the criteria was not incorporated for our first proposal. In an ideal world, the questions would need to be discussed in relation to the way in which the responses to the questions

should be weighed. The picture that will be obtained after weighing would then reflect the extent to which a task can be called offline or online, and where about it should be placed on the continuum – relative to other tasks that is, since absolute values cannot be assigned.

However, as the world is more complex than ideal with respect to this matter, weighing the responses is not so straightforward. Because of limited research thus far into the features that make tasks more offline or online in nature, assigning weight to the responses to the questions posed above could possibly only be based upon intuition. Therefore, the responses to the questions will in this thesis be dealt with as if all features are equally important, and only positive or negative responses to the questions, which will be translated into plusses (obtains a 'one') or minuses (obtains a 'zero'), will be counted to estimate to what extent tasks fall more on the respective offline or online side of the proposed continuum.

In addition to this methodological decision with respect to the placement of tasks on the continuum, it should be borne in mind that the placement of tasks on the continuum is not absolute but relative. The credits assigned per answer are based on whether or not a question can be answered with an affirmative, but without having absolute clarity regarding the weight of the questions, tasks can only be ranked against each other.

It is important that no questions are left out, as all factors or characteristics covered by the questions may contribute to the offline or online nature of a task. This does not mean that each question is relevant for each specific task – sometimes they may not be applicable or relevant – but not considering the questions may lead to an incomplete and biased picture. Accordingly, it seems best to follow the schema as presented in Figure 4, in which it can be established whether a task has offline characteristics – as touched upon in the questions posed above – or not. If the answers to the points mentioned in the schema are positive, a one (1) can be assigned showing that a task has this more offline characteristic. If not, a zero (0) could be assigned.

After answering the questions with ones, for confirmative answers, and zeroes, if a question is not confirmed for a specific task, counting the zeroes and ones will show per task whether it has to be placed more on the offline or more on the online side of the continuum.

Relation to question	Offline characteristics
1	Conscious of language task
2	Elicited material
3	Non-natural language included in the task
4	Linguistic input finished at moment of observing participant's behaviour
5	Time for consideration of answer (i.e., is there no time constraint)
6	Judgments required
7	Manipulation required
8	Linguistic procedure required
9	Specific goal of task known to participants
10	Reading/writing (instead of listening/speaking)

Figure 4: Schematic overview of offline characteristics of tasks

3.3   Final remarks

In Chapter 5, the tasks conducted for the present study will be presented, and a proposal will be made for their placement on the continuum – based on the criteria described here in Figure 4.

After detailed discussions of the tasks, and presentation of the results they yielded (Chapter 6-7), a conclusion with respect to the proposed continuum will be attempted in Chapter 8. Before this, however, Chapter 4 will provide an overview of the tasks that were prepared and tested in pilot sessions with monolingual Dutch and bilingual Turkish-Dutch children. This chapter will also contain a discussion of the results obtained from these pilots, as they have not only contributed to the final methodological decisions, but also show how the bilingual and monolingual children differed in their responses and behaviour, and to what extent testing bilinguals was informative.

Pilot studies

In order to answer the questions posed in Chapter 1, namely what influence literacy, typological background and frequency have on the units that children recognize and process in speech, several pilot testing sessions were organized in order to assess different testing methods. The tasks were performed by pre-literate and literate monolingual Dutch children and by bilingual Turkish-Dutch children. The reason for including bilingual children was that their data were expected to allow disentangling the effects of typological background from those of bilingualism (see Section 4.1 below).

In this chapter, the methodological issues that arose and the decisions that were taken on the basis of the pilot sessions, such as the ultimate exclusion of a self-paced listening task and the selection of particular stimulus items, will be discussed. Results from the pilot studies showed first and foremost how the tasks worked, but they also provided interesting insights with respect to the effects of bilingualism. In this chapter, the results will therefore be presented rather extensively, and they are used as a first indication of empirical results and an aid in helping to further specify the hypotheses formulated on the basis of literature in Chapter 2. Only the results obtained for Dutch will be discussed, though, as the number of bilingual participants who were also tested in Turkish was too small to allow any quantitative analyses. Moreover, the Turkish researcher who conducted the tasks in Turkish with the bilingual children had the strong feeling that these children were unable to understand her well enough in Turkish, making the validity of the results doubtful for this group of participants.

Before going into the findings in Section 4.2-4.5, Section 4.1 will provide background on why the initial design included bilingual children in the first place. In the final section of this chapter, the implications of the pilot studies for the main study will be summarised. In general, it should be noted that

the pilot studies yielded two decisions: which tasks to keep, and not to test bilingual Turkish-Dutch children. For the latter reason, there are fewer participants in the main study than in the pilot (see Chapter 5).

4.1   Testing Turkish-Dutch bilinguals

The initial idea was to compare language segmentation and language processing in monolingual Dutch, monolingual Turkish and bilingual Turkish-Dutch pre-literate and literate children. The reason for selecting these three groups was that effects of literacy, typology and bilingualism could be teased apart. It is to be expected that in addition to literacy, the typological profile of the language involved may affect which building blocks children use in processing language (see Chapter 2). The inclusion of bilinguals would additionally allow for distinguishing between the impact of literacy, bilingualism and typological background, even if it was already known that while adult literate Turkish-Dutch bilinguals in the Netherlands tend to be orally proficient in Turkish and Dutch, many only read and write well in Dutch.

The awareness of units in Dutch may differ between bilingual and monolingual children. Several researchers have investigated the effects of bilingualism on metalinguistic development since the 1970s. So far, the findings have been ambiguous: some studies have found positive effects of bilingualism, e.g., on the measure of control over linguistic processes (Bialystok, 1986b; Bialystok, Craik & Luk, 2008; Francis, 1999; Narain & Verhoeven, 1994); others have found no relation between bilingualism and metalinguistic development (Bialystok, 1987).

Bialystok (1987) suggests that the reason for such contradictory findings is that there is no clear definition of 'metalinguistic skills'. Indeed, the multi-faceted nature of these skills is not well understood, as was discussed in Chapter 2. Homer (2000) emphasizes that metalinguistic awareness involves various components of language, such as phonology, grammar, orthography and words. Accordingly, it should not come as a surprise that investigating the effects of bilingualism on general metalinguistic development leads to controversial conclusions.

However, even if the focus is restricted to one particular metalinguistic skill, there is no conclusive evidence for whether being bilingual has a positive effect or not. Several studies of phoneme recognition have shown that there are no significant differences between monolingual and bilingual children's abilities to segment speech into phonemes (cf. Nicholadis &

Genesee, 1996; Schwartz, Leikin & Share, 2005). However, Cutler, Mehler, Norris and Segui (1992) concluded that French-English bilinguals behaved differently in their segmentation of words from either French or English monolinguals. In this study, participants had to press a response key when they heard a word which started with a particular sound sequence, for instance *ba-* or *bal-* (in words such as 'balance' and 'balcony'). It was found that the bilinguals '… as a group did not simply mimic the performance of English monolinguals with English materials and of French monolinguals with French language materials' (p. 381). In this sense, differences between monolinguals and bilinguals have been found in language segmentation studies as well, even if this study did not mention a positive or negative effect of bilingualism explicitly.

In contrast, Narain and Verhoeven (1994) show that the difference between monolingual and bilingual children in their development of metalinguistic awareness is limited. These authors did find better metalinguistic knowledge in bilingual than in monolingual children, as Uri (2001) and Bialystok (2007), but Narain and Verhoeven claim that this finding was probably more related to the proficiency levels of bilinguals than to being bilingual in itself. Veldhuis, Li and Kurvers (2010) also found that Chinese-Dutch biliterate children's responses on a unit segmentation task in Chinese differed from bilingual children who were monoliterate in Dutch. In this study, a relation between unit recognition and literacy was found. Children who were literate in either only Chinese, or in both Chinese and Dutch, were able to correctly segment character units in spoken Chinese, whereas bilingual Chinese-Dutch children who were monoliterate in Dutch and pre-literate bilingual Chinese-Dutch children, were unable to do so. These findings are in line with Bassetti (2007), who found that both Chinese L1-speakers and Japanese L2-speakers of Chinese, who had first learned the writing system of Japanese, which is rather similar to that of Chinese, distinguish segments in Chinese that are more similar to each other than to those of English L2-speakers of Chinese. In fact, the positive effect of bilingualism on language segmentation tasks was in these studies thus found to be closely related to literacy.

There are few studies that have focused on the effects of bilingualism on online processing of lexical units in language. Rayner et al. (2007) compared eye movements of readers from different language backgrounds, and found that monolingual readers of English and Chinese apply different strategies in looking, and that bilingual Chinese-English readers switched between the strategies applied by their monolingual Chinese and monolingual English peers. Accordingly, it is assumed that bilinguals can activate their languages

simultaneously while reading or while doing visual experiments. Bilinguals also seem to be able to access their two languages simultaneously in spoken word recognition (cf. Blumenfield & Marian, 2007; Weber & Cutler, 2004). These findings suggest that bilinguals may partially process different units from monolinguals, who do not have to deal with the processing of two languages when using or listening to language. So far, however, this has not been confirmed: most studies with bilinguals focus on how they access their lexicon or how they process words or visual scenes; there are no studies as yet that investigate whether, and how, the units that bilinguals parse and produce in spoken language differ from those that monolinguals use.

Including Turkish-Dutch children in the tasks conducted for this study permitted us to investigate this matter. It also allowed us to examine to what extent bilinguals exhibited the same behaviour as monolingual Turkish children. The typological background in itself was expected to affect children's language segmentation and processing units (see Chapter 2). Turkish is, in contrast to Dutch, an agglutinative language, in which multiple morphemes are written together in single words (see Chapter 2). This means that Turkish words are often longer than Dutch words, and morphologically more complex.

The comparison between monolingual and bilingual children on their Dutch or Turkish language segmentation and processing behaviour, as well as between literate and pre-literate children, would in theory allow for distinguishing between influences from typological background, bilingualism and literacy (see also Veldhuis, Li & Kurvers, 2010). Accordingly, pre-literate and literate monolingual Dutch and monolingual Turkish children, as well as pre-literate and literate bilingual Turkish-Dutch children were originally envisioned to participate in the study.

4.2   Tasks conducted in the pilot studies

Seven tasks were prepared and tried out. These included relatively offline language segmentation tasks and more online language processing tasks (see Chapter 3). The relatively offline tasks included a dictation task, a sentence segmentation task,[7] and a last-part repetition task. The more online tasks were a click task, a self-paced listening (SPL) task, a picture-naming task focused on compound nouns and adjective+noun-combinations, and a

---

[7] This task has in previous publications been referred to as a 'tapping task'. See Veldhuis (2011) and Veldhuis, Vermeer and Yan (2012).

production task focused on labelling prepositional phrases. All tasks were conducted in Dutch, and the sentence segmentation, last-part repetition and picture-naming tasks were also conducted in Turkish.

The SPL-task and the production task were deemed especially suitable for investigating frequency effects on children's language production and processing (see Chapter 2). The dictation, sentence segmentation and last-part repetition tasks were developed to provide insights into the way in which children segment spoken language, the click task to examine the way in which children process multiword units in language, and the picture naming task to compare pre-literate and literate children's knowledge of units used in writing.

Section 4.2.1 and 4.2.2 provide details with respect to the two tasks that were eventually discarded. The other tasks will be discussed in Section 4.3-4.5.

4.2.1   The production task and the SPL-task: Two tasks especially focused on frequency effects

The SPL-task and the production task were specifically developed to measure effects of the frequency of multiword units. As mentioned in Chapter 2, previous research has suggested that the frequency of occurrence of such chunks, or of the co-occurrence of specific words, leads to their entrenchment. To what extent frequency has an effect on the units that children use, has, however, not been investigated so far in an elaborated way.

The SPL-task and the production task were developed to investigate this supposed influence of frequency. In the production task, children were triggered to name frequent and less frequent prepositional phrases after seeing pictures which represented these phrases. More specifically, on a computer screen children saw one-by-one 19 pictures in which a smiley was located in, on, or under a specific object, and they were asked to answer the question 'Where is the smiley?' ('*Waar is de smiley?*'). An example, adjusted to black and white, is given in Figure 5.
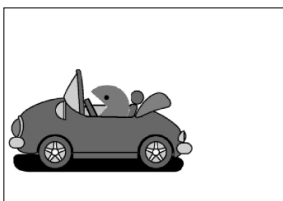
Figure 5: Example of screen used in the production task (representing 'in the car')

In the self-paced listening task, children had to press a button after they were exposed to high and low frequency multiword targets in sentences that were cut into parts consisting of one or several words. An example was: '*ik heb – mijn handen – onder de kraan – gewassen*' ('I have washed my hands under the tap'). Children were presented with this sentence part by part, as represented by hyphens in the example sentence, and they were only given the next part when they pressed the button. In this particular example, the sequence '*onder de kraan*' ('under the tap') was the target item. The children's reaction times of pressing the button were measured from the moment that the target had been completely pronounced. In this SPL-task, the critical sequences consisted of prepositional phrases and noun+verb-combinations, whereas in the production task, only prepositional multiword sequences were included.

The multiword sequences that occurred in these tasks, either in the form of pictures (in the production task) or as pre-recorded phrases (in the SPL-task), were selected on the basis of frequency data obtained from the Corpus of Spoken Dutch (*Corpus Gesproken Nederlands*, *CGN*, Nederlandse Taalunie, 2004) and Google. The item '*onder de kraan*' occurs quite often according to the CGN and Google corpora, whereas its counterpart in the task – starting with the same phoneme and including the same preposition – '*onder de krant*' ('under the newspapers') does not occur often at all.

For the analyses of children's responses, the first measure in the production task was how often children provided the target locative phrase that was shown in the drawings. Secondly, it was logged how quickly children produced the prepositional phrase that was depicted. In the SPL-task, it was analysed how quickly children pressed the button after having heard the target. It was expected that children would produce and process the frequent multiword sequences faster than those that occurred less often, as was found in studies with adults (see Chapter 2). For both tasks, hypotheses about the effects of literacy, bilingualism and typological background were not formulated, but the tasks were provided both to monolingual and bilingual pre-literate and literate children, so that any effects from these factors could be investigated in an explorative way.

The pilot sessions showed that the production task was very difficult for young, pre-literate children aged 4 to 5 (N=34). Some of these children provided incorrect prepositions in their answers (saying 'on the car' instead of 'in the car'); one bilingual child who was tested in Dutch provided the Turkish words for 'newspaper', 'tent' and 'tap'; one pre-literate girl also described the object, by saying that the smiley was 'in the house for babies',

with which she referred to the picture of a tent. Because of answers like these, children's responses were not analysed further statistically, as for the youngest children the task was probably too hard to perform. This would make the results unreliable. Accordingly, even though this task could possibly be used to measure the impact of frequency among older children, it was not retained for the main study.

The results of the SPL-task, which was conducted with 69 pre-literate and literate children, aged 4 to 10, were not encouraging for the hypothesis that a higher frequency would lead to shorter processing times. The children's reaction times for pressing the button were on average slightly longer after high frequency targets (M=522.4ms, SD=286.8ms) than after low frequency targets (M=501.6ms, SD=276.7ms), as is shown in Table 1 below.

	Pre-literates (N=25)		Literates (N=44)		Total	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
High frequency targets	674.2	(345.2)	436.1	(206.3)	522.4	(286.8)
Low frequency targets	630.8	(378.2)	428.3	(161.5)	501.6	(276.7)

Table 1: Reaction times in ms in SPL-task after high and low frequency targets

A Repeated Measures analysis, with literacy as between factor, frequency as within factor and age as a co-variate, revealed that the frequency of the multiword sequences was not a significant factor (p>.05). The effects of literacy and age were not significant either, nor was there a significant interaction (p>.05).

Just like the production task, the SPL-task also turned out to be rather complex for young children: among the pre-literate children, more than 60% only pressed the button at the end of the sentence (after 1000ms), or did not press the button at all. Due to this behaviour, the results from the SPL-task were not reliable. Therefore, it was decided not to retain this task for the main study either.

4.2.2 A picture-naming task focused on children's knowledge of units in writing

The picture naming task was meant to allow an investigation of what children regard as units in writing, more specifically whether they would regard singular and plural compound nouns and adjective+noun-combinations (such as 'hot bath') as single or multiple units when written.

In order to test this, a task was developed in which children were presented with pictures that depicted compound nouns, such as '*wasmachine*' ('washing machine') or '*boterham*' ('slice of bread') (lit: 'butter+ham') or adjective+noun-combinations such as 'hot bath'. Below these objects, which occurred in the pictures either once or twice, the name of the object or objects were written either as one word, or as two words. This way, below the picture of '*boterham*' would be written '*boter ham*' or '*boterham*', or, in the plural form, '*boter hammen*' or '*boterhammen*'. The reason for presenting either one or two of the objects was that this would allow investigating whether children regard plural nouns more often as two words (perhaps to be expected as they saw two objects). An example is given in Figure 6.
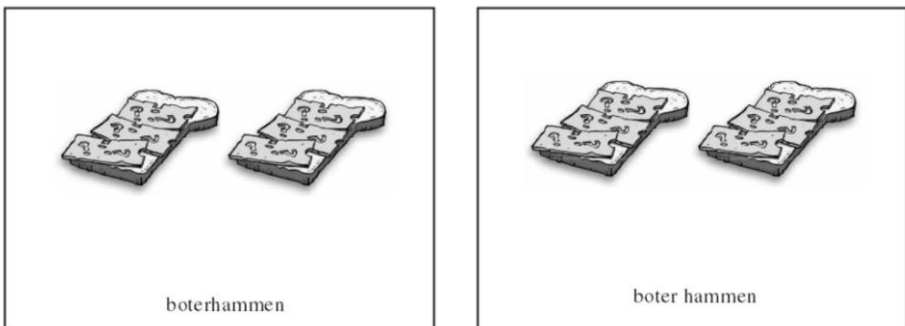
Figure 6: Example of an item used in the picture naming task, adjusted to black and white: the plural form of sandwich ('*boterhammen*')

Children were asked whether the word written below the picture was a correct representation of the picture or not. The picture-naming task was conducted both in Dutch and in Turkish with the bilingual children, so that it could be analysed to what extent typological background and knowledge of conventions in the writing systems of either language would affect the children's responses. In Dutch, compound nouns are usually written as one word, whereas in Turkish most compound nouns, including all the ones of which the children saw pictures, are written as two words. Both languages write adjective+noun-combinations as two words.

This task was conducted in Dutch and in Turkish with monolingual Dutch and bilingual Turkish-Dutch children. Even the youngest children who participated in the task were well able to choose one or the other representation. Still, the decision was taken in the end not to include this task in the main study, or to further analyse the data obtained (see Chapter

5). Although the results obtained from this task may provide reliable evidence with respect to the way in which children think about writing conventions, the task appeared not really to provide information on the way in which children segment or process sentences or multiword chunks of language. It rather elicited children's (meta)knowledge about printed language and their ideas on how the boundaries of written words can be drawn in language (see also Homer & Olson, 1999; Meltzer & Herse, 1969). Therefore, it would not lead to the sort of information looked for in this study.

4.3   Tasks focused on literacy, typology and frequency effects

As mentioned above, the click task, the dictation task, the sentence segmentation task and the last-part repetition task are all designed to investigate the effects of literacy, typology and frequency. These tasks were retained for the main study (see Chapter 5). In this section, the background and set-up of these tasks will be discussed. The results from the pilot sessions will also be given, as they provided interesting insights especially with respect to bilingualism.

4.3.1   Tasks focused on children's language segmentation

Three of the tasks were relatively offline and focused on children's conscious language segmentation (see Chapter 3). These include the dictation task, the sentence segmentation task and the last-part repetition task. All of these were adjusted versions of tasks that had been applied by other researchers before. In the pilot sessions they were conducted in both Dutch and Turkish with the bilingual Turkish-Dutch children, and with the Dutch monolingual children in Dutch.

Since the dictation task was only conducted with 10% of the Dutch monolingual children who participated in the pilot sessions, and with even less than 10% of the bilingual children, the pilot results of this task will not be discussed here. The other two tasks were conducted in Dutch with all monolingual and bilingual Turkish-Dutch participants, at schools in Tilburg, Epe and Rotterdam. Below, the details with respect to the participants, the methodology and the results from these tasks will be provided both for the monolingual and for the bilingual children. This will show that the sentence segmentation and last-part repetition tasks provided us with reliable and

valid information with respect to our research questions. Therefore, these
tasks were retained in the main study.

4.3.2   A task focused on children's online processing: The click task

A version of the click task was adopted to tap into the way in which children
process units in language in a relatively unconscious, i.e., online, way (see
Chapter 3). Children were exposed to pre-recorded sentences to which they
listened on a computer. The idea was to find out whether pre-literate and
literate monolingual and bilingual children would process multiword
sequences in a similar way.

   The details of this task will be discussed in the next section. At this point,
however, it is important to note that the task did indeed provide valid and
reliable information, so this task was also retained for the main study.

4.3.3   Hypotheses for the sentence segmentation, last-part repetition and
          click tasks

On the basis of earlier studies (see Chapter 2), it was expected that the
sentence segmentation, last-part repetition and click tasks would show the
following effects:

– Literacy affects children's language segmentations. This would show up
  as a larger number of word segmentations in the data from literate
  children than in those from pre-literate children. Pre-literate children
  were expected to focus more on meaning, and to come up with more
  varied and larger segmentations, including multiple words, as they
  would not 'break up' conventional multiword sequences (cf. Arnon, 2010;
  Morris, 1993; Tomasello, 2006, see also Chapter 2). Literacy might also
  affect processing. This should show up as a larger proportion of words
  among the units that language is decomposed into in a processing task
  such as the click task.

– Typological background affects children's language segmentation and
  processing in Dutch. This would be shown by more single word
  responses among the monolingual Dutch children, who speak an
  isolating language, and more multiword responses among the bilingual
  Turkish-Dutch children, who in addition to Dutch speak an agglutinative
  language, in which words are usually more complex and multi-
  morphemic than in Dutch (see Chapter 2).

The tasks were furthermore expected to provide some first insights into how literacy and typological background, or knowing yet another language than Dutch, may interact, as both pre-literate and literate monolingual and bilingual children participated in the tasks. If the two factors would interact, then children who were literate in Turkish would be expected to come up with larger units in the tasks than their peers who were literate in Dutch only, as (written) words are usually longer and more complex in Turkish than in Dutch. For bi-literate bilingual children, no clear expectations were defined, as these children would know about word spacing conventions in two languages. For the pre-literate Turkish and Dutch children, or bilingual children, no differences would be expected to occur, as pre-literate children are not yet biased towards specific units.

4.3.4 Participants in the sentence segmentation, last-part repetition and click task

A total of 198 children participated in the sentence segmentation, last-part repetition and click task in the pilot sessions in Dutch. Table 2 presents the division of these children over groups, based on literacy, grade, and linguistic background.

The pre-literate children who participated in the pilot sessions all attended kindergarten. This means that they had not yet received any formal literacy teaching. It was verified that they were indeed not literate by asking them to write down a number of simple words – a task they could not perform. The other children were in grades 1 to 3 (*groep 3-5*). They had received at least seven months of formal literacy education, and could write simple CVC (consonant-vowel-consonant) words that they were provided with, such as '*vis*', '*jas*' and '*kat*' ('fish', 'coat', 'cat'). The children were between 4 and 10 years old (M=6.8, SD=1.6). There were in total 91 (46%) boys and 107 (54%) girls.

Children participated voluntarily in the tasks, and the children who participated were randomly selected per grade. Parental consent for participation of their children was obtained via a letter containing information about the research project and the experiments. This letter was distributed via the children's teachers.

Due to time limitations and unforeseen circumstances (such as illness, obligatory school tests, PE lessons), not all children were included in all tasks. See Section 4.4.1 for the exact number of children who participated in each task.

	Grade	Monolingual Dutch children	Bilingual Turkish-Dutch children
Pre-literate	K1	14	22
	K2	14	31
	Total	28 (37.8%)	53 (42.7%)
Literate	E1	15	24
	E2	16	29
	E3	15	18
	Total	46 (62.2%)	71 (57.3%)
Total		74	124

Table 2: Division of participants over groups

4.3.5 Materials and procedure

Two of the tasks were relatively offline in nature. In the sentence segmentation task, children were asked to show, using little blocks on a table, how many parts (*stukjes*) they heard in 17 sentences that were read out loud to them by the researcher (cf. Holden & MacGinitie, 1972). The second offline task was the last-part repetition task (based on Karmiloff-Smith et al., 1996), in which the researcher read out a story to the children and asked them to repeat the last part (*laatste stukje*) every time she paused somewhere in the middle of a sentence. The story contained 28 pauses.

In both tasks, some of the selected targets consisted of multiword sequences of which the conventional Turkish translation would be a single word: there were nine of such targets in the sentence segmentation task and four in the last-part repetition task. An example is '*Ik heb mijn schouder gestoten*' ('I have hit my shoulder'), with a pause after '*schouder*'. 'My shoulder' is the multiword target; its Turkish equivalent is written as one word: '*omuz-um*',[8] 'shoulder-POSS'.

In the more online click task (cf. Cohen & Mehler, 1996), children had to repeat 13 pre-recorded sentences with clicks of white noise occurring mid-sentence, five times within selected multiword sequences that again would be written as single words in Turkish. An example is: '*Dit is mijn X jas met bloemen*' ('this is my coat with flowers', '*mijn jas*', 'my coat' is the target unit, and the X indicates where the click was). Children were asked to imitate the click in their repetitions.

---

[8] The hyphen is here inserted for clarity and does usually not appear in Turkish writing.

Of each task two versions were constructed, with different stimulus items. Both versions, however, contained multiword units that were comparable in their frequency of occurrence (according to a Google search) and in the level of difficulty of the nouns and verbs occurring in the units (according to *Woordwerken*, Schrooten & Vermeer, 1994). Children only participated in one version of the tasks, and performed them individually in a quiet room at their schools, and in a playful manner. Children participated in several sessions (on several days) in different tasks, which were conducted right after each other with small breaks in between, to make sure that a limited attention span would not negatively affect the results. The Turkish-Dutch children also conducted a quick vocabulary test in Turkish and in Dutch, with the names of body parts as test items (see also Veldhuis, Vermeer & Yan, 2012).

4.3.6 Analyses of the sentence segmentation, last-part repetition and click tasks

Even though this initial study served as a pilot, analyses were carried out to find out whether the monolingual and bilingual pre-literate and literate children differed in the number of times they deconstructed multiword targets into single words and how often they retained them in their responses.

To measure whether it is literacy, bilingualism or both which affected children's segmentation and processing units, the first investigation was focused on how many single word responses the children gave in the last-part repetition task, how many times they segmented the multiword target units word-by-word in the sentence segmentation task, and how many times they were able to repeat the clicks correctly between the words in the multiword targets.

To analyse to what extent the multiword target items were indeed entrenched as processing units, and to examine a possible effect from also knowing an agglutinative language, it was also counted how often the monolingual and bilingual children repeated the target multiword constructions in the last-part repetition task as wholes; left the multiword target units unsegmented in the sentence segmentation task (i.e., segmented those items as wholes); and did not repeat the clicks occurring within multiword target units (but either repeated them just before, or just after the units). In order to see whether age or grade affected the results, scores were also compared across grades.

4.4   Results of the sentence segmentation, last-part repetition and click task

This section will provide the pilot results from the three tasks. The effects of literacy and bilingualism will be investigated in Section 4.4.1, while Section 4.4.2 looks at the possible effect of knowing Turkish (in addition to Dutch).

4.4.1   Effects of literacy and bilingualism

Table 3 presents an overview of the pilot data, for all three tasks. It gives the means of single-word repetitions in percentages made after the four multi-word targets in the last-part repetition task that would be translated as single words in Turkish, the number of word-by-word segmentations for the nine multiword targets in the sentence segmentation task that would be translated as single words in Turkish, and the number of correctly repeated clicks in the click task with the five multiword targets that would be regarded as single words in Turkish. The results are given per group, in percentages of the number of targets in each task.

	Monolingual children		Bilingual children		Total	
	Pre-literate	Literate	Pre-literate	Literate	Pre-literate	Literate
Last-part repetition task (4 targets)	39.29 (38.38) (N=21)	73.81 (37.02) (N=42)	30.87 (34.74) (N=44)	70.64 (33.24) (N=65)	33.59 (35.87)	71.88 (34.64)
Sentence segmentation task (9 targets)	44.73 (28.21) (N=22)	86.78 (20.16) (N=46)	43.56 (25.86) (N=45)	83.33 (19.14) (N=63)	43.94 (26.44)	84.79 (19.56)
Click task (5 targets)	13.13 (17.40) (N=16)	32.69 (22.93) (N=44)	10.83 (19.71) (N=24)	28.45 (24.01) (N=59)	11.75 (18.62)	30.26 (23.54)

Table 3: Single-word responses per task and group in %, M (and SD)

As can be seen from Table 3, overall the literate children produced more single word segmentations in the tasks than the pre-literate children, and the monolingual Dutch children had slightly more single word segmentations than their bilingual peers. Univariate analyses with word responses as dependent variable, and literacy and bilingualism as independent variables showed that there was a significant main effect of literacy on children's

word segmentations in all tasks, with medium effect sizes: $F_{last-part\ rep}(1,167)=12.38$, $p=.001$, $\eta_P^2=.07$; $F_{sentence\ segm}(1,171)=15.98$, $p=.000$, $\eta_P^2=.09$; $F_{click}(1,138)=5.65$, $p=.02$, $\eta_P^2=.04$. Whether children were monolingual or bilingual did not lead to a significant main effect in any of the three tasks ($p>.05$). There was no interaction between literacy and bilingualism in any of the tasks either ($p>.05$). These results suggest that literacy does affect the segmentation strategies that children employ in the tasks, but that being monolingual or bilingual does not.

To assess the possible confound between age or schooling and literacy – literate children were older and in the higher grades – a second analysis was done, in which the effects of the grades which the children attended were investigated among the literate children. In Table 4, the percentages of single word responses per task are presented per grade. In the univariate analyses of these results per task, the single word responses or correct repetitions of clicks between words within multiword target units were again the dependent variable; language background (being monolingual or bilingual) and grade (1, 2 or 3; corresponding to *groep 3-5*) were now the independent variables.

	Monolingual children			Bilingual children			Total		
	Grade 1	Grade 2	Grade 3	Grade 1	Grade 2	Grade 3	Grade 1	Grade 2	Grade 3
Last-part repetition task	70.83	65.00	85.00	65.00	71.91	75.00	67.19	69.44	79.55
	(33.43)	(46.10)	(28.03)	(35.73)	(32.78)	(32.08)	(34.45)	(37.66)	(30.27)
(4 targets)	(N=12)	(N=15)	(N=15)	(N=20)	(N=27)	(N=18)			
Sentence segmentation task	76.48	90.97	92.59	71.30	88.77	90.12	73.45	89.65	91.25
	(25.04)	(13.59)	(17.65)	(24.23)	(13.11)	(11.98)	(24.35)	(13.18)	(14.63)
(9 targets)	(N=15)	(N=16)	(N=15)	(N=21)	(N=24)	(N=18)			
Click task	26.43	35.94	35.24	24.80	30.69	28.89	25.62	32.79	31.67
(5 targets)	(22.05)	(24.44)	(22.41)	(22.79)	(27.39)	(21.04)	(22.10)	(26.06)	(21.53)
	(N=14)	(N=16)	(N=14)	(N=17)	(N=24)	(N=18)			

Table 4: Single-word responses per task and group in %, M and (SD)

Table 4 shows that, among the literate children, the children in the higher grades gave more single word answers and had more correct repetitions of clicks between words in multiword units than the children in grade 1 (*groep 3*). In the click task, children in grade 2 provided on average almost

the same number of single word responses as the children in grade 3. This was also the case for the responses of the monolingual and bilingual children separately. In the other tasks, the children in grade 3 produced more single word responses than children in grade 2, and children in grade 2 had more single word responses than children in grade 1. In the last-part repetition task, however, monolingual children in grade 2 had *fewer* single word segmentations than the children in grade 1.

Even though most of these differences appeared to be small, the difference between grades was significant. This was seen in the univariate analyses: there was a main effect of grade in the last-part repetition task and in the sentence segmentation task: $F_{rep}(1,161)=6.15$, p=.000, $\eta_P{}^2$=.13; $F_{sentence\ segm}(1,165)=6.08$, p=.000, $\eta_P{}^2$=.12; $F_{click}(1,132)=1.72$, p=.15, $\eta_P{}^2$=.01. A pair-wise comparison showed that for the last-part repetition task and the sentence segmentation task, the children in grade 1 had significantly fewer single word segmentations than children in grade 2 or 3, who did not differ from each other.

There were no significant main effects of bilingualism in any of the tasks, and there was no interaction between grade and bilingualism in any of the tasks either (p>.05).

These results suggest that in general, the effects of being bilingual and of years of schooling were limited, and that the differences in children's responses in all three tasks were mainly caused by whether the children were literate or not. Literacy led to main effects in all tasks, whereas bilingualism and grade did not.

4.4.2   Multiword sequences as basic units of language and effects of typological background

In order to assess whether the Turkish-Dutch bilingual children regarded the selected multiword targets as units more often than the Dutch monolingual children, and also to investigate to what extent the multiword sequences that were selected for the pilot study indeed functioned as basic units, a third univariate analysis was done, this time with the multiword target responses as dependent variables. For the click task, this meant that it was investigated how often children in their repetition responses moved the clicks that were inserted within a multiword target to a place just before or just after it. For instance, instead of correctly repeating 'on the X table', the child might say 'X on the table', or 'on the table X' (X representing the click). The dependent variables in this analysis were thus multiword segmentations in the sentence segmentation task and the last-part repetition

task, and the number of times clicks were moved to a position just before or after the target multiword unit in the click task. The independent variables were once again literacy and bilingualism. In Table 5, the results of this analysis are provided.

	Monolingual children		Bilingual children		Total	
	Pre-literate	Literate	Pre-literate	Literate	Pre-literate	Literate
Last-part repetition task (4 targets)	25.00 (28.50) (N=21)	13.69 (22.23) (N=42)	35.04 (33.65) (N=44)	16.67 (23.57) (N=65)	31.79 (32.20)	15.50 (22.99)
Sentence segmentation task (9 targets)	9.33 (14.92) (N=22)	0.48 (3.28) (N=46)	12.86 (18.11) (N=45)	0.35 (1.96) (N=63)	11.70 (17.10)	0.41 (2.59)
Click task (5 targets)	51.45 (25.33) (N=16)	46.74 (23.38) (N=44)	30.21 (33.41) (N=24)	49.86 (24.76) (N=59)	38.71 (31.88)	48.53 (24.11)

Table 5: Multiword responses per task and group in %, M (and SD)

As Table 5 shows, pre-literate children seemed to provide more segmentations that corresponded to the target multiword units than the literate children, but not in the click task. This corresponds well with the literacy hypothesis. The *lower* scores on multiword retention for pre-literate children in the click task are primarily due to the bilingual pre-literate children, who had substantially fewer multiword retentions than their literate peers. In general, however, bilingual children preserved multiword targets more often than monolingual children. This was especially obvious among the pre-literate children, which goes against the idea that pre-literate children speaking different languages would not differ in their responses (see Section 4.3.3). In the click task, however, the pre-literate bilingual children moved the clicks to a place just before or just after the target unit much less often than the monolingual pre-literate children. Instead, they moved the click more often to a place at the beginning or at the end of the sentence in their repetitions, which is not informative regarding the question whether the multiword sequence was a single unit for them or not.

The univariate analyses of the tasks showed a significant main effect of literacy in the last-part repetition task and in the sentence segmentation task, $F_{rep}(1,167)=9.41$, p=.003, $\eta_p^2=.05$; $F_{sentence\ segm}(1,171)=11.00$, p=.001, $\eta_p^2=.06$, but

not in the click task (p>.05). There was no significant main effect of bilingualism in the last-part repetition task and in the sentence segmentation task (p>.05), but there was a trend of an effect of bilingualism in the click task, $F_{click}(1,138)=3.62.$ p=.059, η_P^2=.03. This was also the only task in which there was an interaction between literacy and bilingualism, $F_{click}(1,138)=5.92$, p=.02, η_P^2=.04. As mentioned, this interaction was probably caused by the low number of times that pre-literate bilingual children moved the clicks to a place just before or just after the multiword target in their repetitions, compared to the other groups.

These results suggest that the effect of knowing a more agglutinative language was limited. As was concluded for single word responses above, literacy seemed to be a much more important predictor of the way in which children segmented multiword targets. However, bilingualism did seem to be the most important factor determining the replacement of clicks just before or just after multiword targets.

To see whether and how schooling and age may have affected the results, it was once more investigated how children from grades 1 to 3 differed in their multiword responses. These results are presented in Table 6.

	Monolingual children			Bilingual children			Total		
	Grade 1	Grade 2	Grade 3	Grade 1	Grade 2	Grade 3	Grade 1	Grade 2	Grade 3
Last-part repetition task	12.50 (22.61)	15.00 (20.70)	13.33 (24.76)	13.75 (18.98)	17.90 (24.97)	18.06 (26.75)	13.29 (20.06)	16.86 (23.31)	15.90 (25.63)
(4 targets)	(N=12)	(N=15)	(N=15)	(N=20)	(N=27)	(N=18)			
Sentence segmentation task	1.48 (5.74)	–	–	0.53 (2.42)	0.46 (2.27)	–	0.93 (4.09)	0.28 (1.76)	–
(9 targets)	(N=15)	(N=16)	(N=15)	(N=21)	(N=24)	(N=18)			
Click task	46.43 (27.63)	45.00 (22.95)	49.05 (20.69)	50.20 (26.11)	45.14 (26.75)	55.83 (20.31)	48.49 (26.42)	45.08 (24.99)	52.86 (20.43)
(5 targets)	(N=14)	(N=16)	(N=14)	(N=17)	(N=24)	(N=18)			

Table 6: Multiword responses per task and group in %, M (and SD)

Table 6 reveals that the children in grade 1 provided more multiword responses than children in grades 2 or 3 in the sentence segmentation task. This would correspond with the literacy hypothesis. In the last-part repetition task, however, the children in grade 1 made the fewest multiword repetitions. In the click task, while children in grade 3 repeated the clicks

most often just before or just after the multiword targets, children in grade 2 did this the least often. Despite these differences, which may suggest a difference in scope and interpretation of the tasks, or an unpredictable pattern for multiword decompositions among the literate children, it should be emphasized that the differences between the grades were in fact small.

The differences between monolingual and bilingual children were small as well: in the last-part repetition task and in the click task, the bilingual children seemed to decompose language into multiword units slightly more often than monolingual children, but in the sentence segmentation task, the bilingual children in grade 1 kept the multiword target as a whole less often than their monolingual peers. Again, this suggests an unstable pattern.

From the univariate analyses of these pilot results, it can be concluded that there was only a significant main effect of grade in the sentence segmentation task, $F(1,165)=3.26$, $p=.01$, $\eta_P^2=07$, but not in the other two tasks ($p>.05$). There was no significant effect of being bilingual or not in any of the tasks ($p>.05$), and there was also no interaction between literacy and typological background ($p>.05$).

Surveying the results from all children combined, literacy again appeared to be the most decisive factor in children's multiword segmentations in these tasks – at least in the two relatively offline tasks in which children were specifically asked to segment sentences or repeat specific segments from the input sentences. In the more online click task, this pattern was not found. The implications of these pilot studies for the design of the main study are discussed in the next section.

4.5 Discussion and conclusion of the three tasks conducted among bilingual and monolingual children

The pilot data suggest a major role for literacy, but only minor effects of bilingualism, or at least of knowing both a more analytic and a more agglutinative language. The grade being attended had little effect either. There was a significant main effect of literacy on the number of single word responses given in all three tasks, and on multiword responses in the two relatively offline tasks, but not in the click task.

4.5.1 Proposed improvements for the more offline tasks

The sentence segmentation and last-part repetition tasks were based on previous studies in which these tasks had been used (cf. Homer, 2009;

Karmiloff-Smith et al., 1996; Kurvers & Uri, 2006). The main difference between the tasks prepared for this study and the previous versions was that in the tasks as prepared for this study selected multiword sequences, hypothesized to be fixed units that could be processed holistically, were included as targets. The selection of these targets was based on the probable familiarity to children of the words occurring in them, on comparability on the basis of the initial phonemes to other content words used as stimulus items, and on the number of syllables that the content words consisted of (this was also the case in the task conducted by Kurvers & Uri, 2006). The division between high and low frequency targets which will play a role in the main study was not yet taken into account during item selection. Post-hoc analyses did allow us to make a distinction between frequently and infrequently occurring multiword targets, but as the numbers of high and low frequency targets were unequal, this analysis was not reliable. For the main study, the selection of items was adjusted more carefully to the question whether frequency mattered.

Likewise, there were only a few items that differed in whether they were written as one or as multiple words in Dutch and Turkish (see Appendix 2): having more of these targets, otherwise comparable across the languages, was expected to lead to more reliable information in the main study. Still, it could have been expected that the relatively large number of children who participated in the pilot study would counterbalance the negative effect of a small number of targets.

Apart from this shortcoming in the selection of items, the three relatively offline tasks – the dictation task, the sentence segmentation task and the last-part repetition task – seemed, with some adjustments for item selection, suitable for an investigation of the research questions posed in this study.

4.5.2 Shortcomings of the click task

The click task, which in general worked well with the children, yielded some unexpected results. In retrospect, it was found that the results mainly differed from what was expected as a result of the item selection, and that a lot depended on how the results are interpreted.

It was assumed that children prove that they are able to segment the multiword sequence into smaller segments, i.e., words, if they manage to repeat sentences with a click within a multiword unit correctly. This reasoning follows the logic that in order to be able to repeat clicks within multiword units, children would have to break them up into words. For assessing children's performance on the multiword targets, however, it was

only investigated how often children, incorrectly, placed the clicks at the borders of the targets: i.e., just before or just after the unit. This would indicate that children had retained the target unit. There were many cases, however, in which especially pre-literate children placed the click at the beginning or at the end of a sentence, or at any other location in the sentence, which made it impossible to say whether the target unit was preserved or not. It might well be the case that children who repeated the clicks at any location within the sentence that was not inside the multiword target left the target unit unanalysed. Since the data from the click task would not allow a clear picture of this issue, it was decided to include additional online tasks, to see whether further evidence about how children process multiword sequences online could be found.

Moreover, as in the relatively offline tasks, the limited range of target items that were included – the translation equivalents in Turkish had to be single words – and the fact that two different types of constructions were taken together in the quantitative analysis (possessive phrases and prepositional phrases) may also have corrupted the findings. If it can be assumed that multiword sequences, no matter what underlying schematic constructions they instantiate, are entrenched in similar ways in our mental lexicon just on the basis of frequent occurrence, then taking the two types of constructions together is perfectly acceptable. After all, storage of the whole construction would mean that you do not have to attend to the internal structure anymore. If, however, different schematic templates are handled in different ways during language processing, then collapsing the two types in one analysis may not be warranted. For this study, however, it was impossible to quantitatively analyse the different constructions separately, due to the limited number of relevant items. A qualitative analysis of only five target items would not be telling much either. Therefore, it was decided not to address this issue any further for the pilot data, but to make sure that for the main study, targets were even more carefully selected, so as to be able to investigate any differences between schematic templates (see Chapter 5 and 7).

4.5.3   Encouragement for exploring literacy and bilingualism further

In general, the literacy effects found in the pilot study are so strong that it is very likely that literacy indeed has an effect on children's language segmentation and perhaps on processing; this confirms what was suggested in numerous previous studies (cf. Berthoud-Papandropoulou, 1978; Bugarski, 1993; Ehri, 1975; Harris, 2009; Homer, 2000; Olson, 1994, 1996). No

differences between monolingual Dutch and bilingual Turkish-Dutch children were found in the analyses, which suggests that there was no measurable effect from being mono- or bilingual.

However, the evidence that there is no effect of bilingualism is not necessarily conclusive. The bilingual children tested in the pilot study were mostly from the third generation of immigrants, with parents who were born and raised in the Netherlands. Thus, even though the children were all bilingual, in the sense that they reported to speak both Turkish and Dutch at home (unless they had trouble answering the question clearly), the degree to which they actually used Turkish, especially for book reading and watching TV, was limited. Accordingly, it is possible that their knowledge of Turkish was also limited.

This was confirmed by the bilingual proficiency measures as obtained through a lexical naming task (see Veldhuis, Vermeer & Yan, 2012). The bilingual children who participated in this task were clearly not 'balanced bilinguals', but highly dominant in Dutch, even though their vocabulary in Dutch appeared to be smaller than that of their monolingual peers. Their dominance in Dutch may of course explain why there was no influence from bilingualism in the results. The only significant difference between the bilingual and monolingual children was in the click task, and only for literate children. This may well be the result of a task effect. It may also be the result of a difference in Dutch proficiency between the monolingual Dutch and bilingual Turkish-Dutch children, whose proficiency in Dutch was lower than that of their monolingual peers, and whose Turkish proficiency was low as well (cf. Veldhuis, Vermeer & Yan, 2012). In the latter case, there is an effect of bilingualism, but not of being bilingual per se: differences in vocabulary and/or language proficiency in Dutch were found between the monolingual and bilingual children tested in the pilot sessions as reported in Veldhuis, Vermeer and Yan (2012). The differences in the results in the click task can, therefore, not necessarily be subscribed to knowing two languages or knowing a specific type of language – in this case an agglutinative language.

Language proficiency and being bilingual could not be disentangled in this study as the bilingual participants in the pilot sessions received lower scores on their Dutch language test than their monolingual peers. And even if more balanced Turkish-Dutch bilinguals could be found, the question should be raised how representative those children would be for all Turkish-Dutch bilinguals in the Netherlands or for bilinguals in general.

4.6   Implications of the pilot studies for the main study presented in this thesis

In addition to the fact that the pilot studies stimulated the development of shorter testing sessions in the main study, so as to warrant children's attention in the tasks, as well as dividing the testing sessions over several days, also enhancing children's attention levels, the pilot yielded three main decisions: which tasks to do, and not to test bilingual Turkish children, nor children in grade 3.

The decision not to include a group of bilingual children in the main study – which led to lower numbers of participants in the main study than in the pilot sessions – was based on the fact that it appeared impossible to verify whether any differences between bilingual and monolingual children's responses were caused by their bilingualism, their knowledge of Turkish, or their degree of proficiency in Dutch. Distinguishing between the effects of knowing Turkish and bilingualism would perhaps be possible if a second group of bilinguals speaking a language with a different typological background had been included, but doing that was beyond the scope of this present study.

Since it also appeared that there were no differences between the children in grades 2 and 3, and as children in grade 3 appeared to become easily bored with the tasks, presumably because they must have been easy for them, it was also decided not to include children from grade 3 anymore in the main study. Concentration problems among these children would have affected the reliability of their results. Instead, the focus was going to be on data from children from the first year of kindergarten until grade 2, in a between subjects analysis. Again, excluding children in grade 3 from the experiments, led to a lower number of participants in total in the main study than in the pilot.

The pilot study gave also rise to some adaptations in target item selection. All tasks included multiword sequences. These included transitive noun+verb-combinations, such as 'wash your hands' ('*handen wassen*'), prepositional locative phrases, such as 'in the tree' ('*in de boom*'), and possessive phrases, such as 'my hairbrush' ('*mijn borstel*'). Items had been selected on the basis of their probable familiarity among children, as indicated in Schrooten and Vermeer (1994; see Chapter 5 for further information), and also word length (in syllables) and initial phoneme of the content words in the multiword items. The number of items per

grammatical construction was not controlled for in advance in the pilot sessions, nor was the frequency of occurrence.

Post-hoc analyses of the pilot data revealed that the effects from frequency of occurrence of targets could be investigated, but also that it would be useful to investigate the effects of constructional type and frequency more systematically than was done so far. In order to be able to do this, only transitive and locative constructions were used in the main study, so that the number of items per type could be increased. Specific items were selected based on their relative frequency of occurrence and their probable familiarity among children, both in Dutch and in Turkish.

A final implication from the pilot study was that the picture naming task, the production task and the self-paced listening task (SPL) were not included in the main study. The reason was that these tasks turned out to be too difficult for young children to provide reliable results, or that they did not allow sufficient focus on the research questions posed for this study. Instead, another task that would tap into children's language processing was developed for the main study, which will be referred to as the mixed-words task (see Section 5.3.5). Like the SPL-task, this task provides insight into the way in which children process units in language, but in contrast to the SPL-task, it does not require that children carry out a complex performance (such as pressing a button within a limited time). Instead, this task only asked children to repeat sentences in which the selected multiword units occurred in mixed, or jumbled, order. This has been done previously in tasks focused on phonological analysis and synthesis. The assumption is that words occurring in the wrong order in a running sentence do not usually obstruct a proper understanding of the sentence. By having both pre-literate and literate children perform this task, the mixed-words task could provide converging evidence about the way in which pre-literate and literate children segment and process the selected multiword targets that would be included in the other tasks.

In the next chapter, all of the tasks conducted in the main study will be discussed in detail. The description of the pilot studies was meant as a description of the search for useful tasks, and how the selected ones were adjusted. The pilot clearly showed that the inclusion of bilinguals in the experiments was more problematic than expected, and that item selection needed refinement in order to maximize the chances of finding answers to the research questions posed in this thesis.

Methodology of the main study

To study the research questions, a battery of tasks was compiled that varied in the degree to which they could be regarded to be offline or online (see Chapter 3). These were presented to monolingual Dutch pre-literate and literate children. Two of the relatively offline segmentation tasks were performed by monolingual Turkish pre-literate and literate children as well, so as to be able to investigate whether typological background and knowledge of writing conventions with respect to word spacing affect unit recognition in spoken language.

In this chapter, the methodology employed will be discussed. Section 5.1 presents the participants. Next, Section 5.2 discusses the selection of multiword target items. The background of the tasks and their set-up will be discussed in Section 5.3, also including information on the participants per task, and the method employed in coding the responses. This overview is followed by a discussion of the probable placement of the tasks on the proposed offline-online continuum (see Chapter 3) in Section 5.4. Some final remarks close off the chapter in Section 5.5.

5.1 Participants

A total of 86 monolingual pre-literate and literate Dutch children from a primary school in Utrecht in the Netherlands participated in the study, as well as 45 monolingual Turkish pre-literate and literate children from two elementary state schools in Istanbul. The pre-literate children were in their first or second year of kindergarten (K1 or K2) and they were not able to write or read more than some single letters and their names; the literate children were in grade 1 or 2 of elementary school (E1 or E2). Table 7 presents an overview of the numbers of children per grade.

	Grade	Dutch		Turkish	
		N	%	N	%
Pre-literate	K1	25	29.1	NA	
	K2	21	24.4	20	44.4
	Total	46	53.5	20	44.4
Literate	E1	20	23.3	NA	
	E2	20	23.3	25	55.6
	Total	40	46.6	25	55.6

Table 7: Participants per grade

Of the Dutch children, 44 (51.2%) were boys and 42 (48.8%) were girls. Their ages ranged from 3 to 8 years old (M=5.3, SD=1.1). In the Turkish group, there were 29 (64.4%) boys and 16 (35.6%) girls; they were between 5 and 8 years old (M=6.3, SD=1.1). Of four pre-literate Turkish children, the age was not recorded.

All literate children were older than all pre-literate children, and a T-test revealed that the Turkish monolingual children were on average significantly older than the Dutch monolingual children, t(125)=-4.57, p=.000. To analyse a possible bias caused by an age-effect, age is taken as a co-variable in the analyses where possible (see Chapter 7).

For the Dutch children, parental consent for participation in the tasks was obtained via letters with information about the study distributed to the children's parents, and for the Turkish children consent was obtained via the teachers and directors of the schools. It was made sure that none of the children had any severe physical problems, certainly not concerning hearing or sight, and that none were diagnosed with concentration problems, as that would possibly affect the results.

5.2   Selection of high and low frequency targets

All the tasks conducted for this study were based on experiments used previously by other researchers (see Section 5.3). They were judged suitable for investigating the effects of literacy, typology and the frequency of occurrence of specific multiword combinations on children's segmentation and processing of language. Specific multiword sequences, selected in advance, were included in four of the six tasks, including the ones conducted in Turkish (they were not included in the dictation task and the

eye-tracking task – see Section 5.3). Their frequency of occurrence was analysed for particular syntactic environments. The selection of target items will be discussed below in Section 5.2.1-5.2.3.

5.2.1 The selection of multiword target units

The multiword sequences that were selected as targets consisted of transitive noun+verb-combinations in which the verb occurred in its infinitival form (e.g., '*handen wassen*' in Dutch, 'wash hands', or '*el temizlemek*' in Turkish) and locative prepositional or postpositional phrases (e.g., '*op de stoel*' in Dutch [PREP DET N] or '*koltukta*' [N+GEN.LOC] or '*koltuğun üzerinde*' [N+GEN POST+GEN.LOC] in Turkish, 'on the table'). The reason for including only transitive and locative phrases was that investigating whether or not children maintain multiword units is harder if the sequences differ too much in their syntactic make-up: it is possible, for example that one type of syntactic phrase lends itself to holistic storage easier than another type. Opting for only two types of multiword sequences, which are often argued to constitute holistic units (cf. Goldberg, 2006; and Chapter 2), and which will therefore be referred to as 'hypothetical units' in this thesis, allowed for a bottom-up approach, in which it could be examined whether and how children segmented and processed sequences that only differed from another in frequency, not in syntactic structure.

The reason for selecting particular instances of these schematic units, i.e., for particular specific, or 'filled' units that instantiate the schemas (see Figure 1, page 30), was to allow for making comparisons between two different sequences that contained the same word, such as the noun. An instance of this is the comparison between '*handen wassen*' ('wash hands') and '*handen schoonmaken*' ('clean hands'). Such a comparison would not have been possible if only schematic or partially schematic units (see Figure 1, page 30) had been used.

Opting for transitive and locative constructions furthermore allowed for an interesting comparison between Dutch and Turkish, as well as a comparison within these languages. In Dutch, both the transitive and the locative constructions consist of several words in writing. The transitive construction is considered to consist of at least two words, the object noun and the verb (e.g., '*handen wassen*', 'wash hands'), though determiners often need to be added (e.g., '*je handen wassen*', 'wash your hands'). The locative prepositional phrase consists of three words in Dutch (e.g., '*op de stoel*', 'on the chair'). In Turkish, the transitive constructions also consist of two words (e.g., '*el temizlemek*', 'wash hands'), but the locative phrases consist of either

one word, in which the noun occurs with a general locative case marker in the structure [N+GEN.LOC] (e.g., '*bahçede*', 'in the garden'), or they consist of two words, with the structure [N+GEN POST+GEN.LOC] (e.g. '*bahçenin içinde*', 'in the garden'). The locative constructions are prepositional in Dutch and postpositional in Turkish.

Locative and transitive constructions could be compared between the languages. The handling of the two types of locative constructions can also be compared within Turkish: it could be examined whether the rather general locative morpheme that is directly attached to the noun in the shorter construction was segmented differently from the longer form, which includes the more specific postposition, and it could be investigated to what extent this longer form was comparable to the segmentation and processing of the Dutch locative phrases, in which the preposition and the determiner are separate function words.

The probable familiarity of the multiword stimulus items for the youngest children who participated in the tasks and the frequencies of occurrence of the items were taken as a basis for item selection.

For this, lists with many different locative prepositional phrases and transitive noun+verb-combinations were first prepared in Dutch, on the basis of a brainstorm session and making use of the words that occur in the wordlist in *Woordwerken* (Schrooten & Vermeer, 1994). It was then checked to what extent young children, in the first year of kindergarten in the Netherlands, could be assumed to be familiar with these phrases, again using *Woordwerken* (Schrooten & Vermeer, 1994). This corpus, which is available on the internet, provides information on the order of acquisition of words in Dutch. Words occurring in the first three lists (which include 3,000 words) can be reasonably assumed to be known by Dutch 4-year-olds (Van de Guchte & Vermeer, 2003). Accordingly, locative and transitive constructions that included words that did not occur in these lists were excluded.

5.2.2 The selection of a frequency measure

Of all remaining multiword expressions and the words they contained, frequency information was collected. First, the frequency of the words and the sequences in the Corpus Spoken Dutch (*Corpus Gesproken Nederlands*, *CGN*, Nederlandse Taalunie, 2004) was checked. The reason for choosing CGN was that this corpus is relatively large, and that the sources of the data are clearly described.

However, the raw frequencies obtained from this corpus were very low – too low to be reliable – so a second frequency measure was sought, this time using the Dutch indexed webpages of Google.[9] In retrospect, it appeared from a correlation analysis that the Google-derived data corresponded quite well to the CGN-derived ones: a two-tailed Spearman correlation analysis of 72 multiword units about which frequency information in Google and in CGN (in Dutch) had been obtained, showed a very strong correlation, r_s=.89, p<.01. Google was, however, a more reliable source, simply because of corpus size, and this corpus could also be used for frequency searches in Turkish.

Next, the frequency measure to be applied had to be considered. In the field of Cognitive Linguistics, in which frequency has been argued to play an important role for language processing, various measures have been used. In some studies, raw frequency in a corpus has been taken as the basis, whereas in others log-frequency-scores of the occurrence of specific words, bi-grams or tri-grams, or more complex methods such as Mutual Information (MI)-values or collostructional strength have been used (cf. Arnon, 2010; Mos, 2010; Stefanowitsch & Gries, 2003; Verhagen, 2011).

Raw frequency data and log-frequencies
The disadvantage of taking simple raw frequency data, or even log-frequency-data, of multiword sequences as a basis for further research is that these data do not provide any information with respect to the relative frequency of occurrence in relation to the occurrence of the single words they contain (or of any sequences of words contained within the larger sequence). For instance, the raw frequency or the log-frequency of 'handen wassen' ('wash hands') needs to be related to the frequency of 'hands' and 'wash'. And though it has been argued that log-scores provide more insight into the relative frequency of words than raw frequency data (see also Verhagen, 2011), since log-scores represent normalized frequencies and not the Zipfian distribution that is typical of raw frequencies (cf. Keller & Lapata, 2003), this method does not provide refined information with respect to the relative co-occurrence of two or more words. This means that even though raw frequency information may show that '*handen schudden*'

---

[9] The Google-size was for the Dutch pages determined on the basis of *http://www.worldwideweb size.com/index.php?la*; for the Turkish web-pages in Google, the size was estimated on the basis of the average of the 50 most frequently occurring words in the METU-corpus (of Turkish) and their frequencies, then using the formula used to estimate the size of the Dutch indexed webpages as well: Estimated Google size = (frequency of a word in Google / frequency of that word in corpus B) * total size of corpus B.

('shake hands') occurs more often than for instance '*handen vouwen*' (lit. 'fold hands', 'pray'), the fact that '*handen*' occurs even more often in other instances, for instance with '*wassen*' (in 'wash hands'), and that therefore the relative frequency of co-occurrence of both '*handen schudden*' and '*handen vouwen*' is low, is not taken into account when using raw frequency data.

Bi-gram or n-gram analyses

Bi-gram or n-gram analyses, which are popular in especially Anglo-American studies on frequency effects (Arnon, 2010), do provide information about the relative frequency of occurrence of a specific expression in relation to the total number of times that a word it contains occurs in general, in any construction. In an n-gram analysis, the relative frequency of occurrence of an expression can be calculated.[10]

A disadvantage of n-gram analyses is that whenever a researcher wishes to investigate the relative frequency of a word as occurring within a specific construction, the construction needs to be pre-defined. In other words, in n-gram analyses, one may investigate the relative frequency of for instance the word 'hands' in 'wash hands', but in this case, it needs to be accepted that the construction at stake is [wash NOUN], and that 'hands' is the variable word within this construction. This pre-defining of the structure, and of the 'head' as it were, (in this case 'wash', which occurs with a noun) is not always possible – one could equally well claim that the basic construction behind [wash hands] is [V hands], given the existence of, in Dutch, 'shake hands' and 'fold hands'.

Collostructional analysis

Another method of calculating the relative frequency of words in constructions has been described by Stefanowitsch and Gries (2003). In their approach, called the 'collostructional analysis', the focus is on the relative frequency of the occurrence of words in specific constructions in relation to their total number of occurrences as well as to the total number of occurrences of the multiword sequence and of the schematic template or construction of which the sequence is an instantiation. For the multiword expression [set off], which Stefanowitsch and Gries provide as an example, a collostructional analysis would take into account the frequency of the single

---

10 The formula that can be used to calculate bi-grams (with [w1 w2] as the schematic representation of the multiword target item) is:

$$P[w1\ w2] = \frac{P[w1\ w2]}{P[w1]} \quad \text{or} \quad P[w1\ w2] = \frac{P[w1\ w2]}{P[w2]}$$

words 'set' and 'off', of [set off], and of the schematic template for this expression, i.e., of phrasal verb constructions with the form [V PREP]. This means that in order to calculate the collostructional strength, four frequencies are needed: the frequency of lexeme L in construction C, the frequency of the lexeme L in all other constructions, the frequency of the construction C with lexemes other than L, and the frequency of all other constructions with lexemes other than L (see also Stefanowitsch and Gries' website, or Stefanowitsch & Gries, 2003). These frequencies are then entered into a [2-by-2] matrix and submitted to a Fisher exact test (or any other distributional statistic; including one written by Stefanowitsch and Gries themselves and available for use with the program *R). The p-values then obtained provide information on the association strength between the words and the construction: the smaller the p-value, the stronger the association.

As Stefanowitsch and Gries (2003) conclude, defining what counts as an instance of construction C is subjective and up to the researcher as it is as yet unknown what level of schematicity is the most relevant. In the example above, both [V PREP] or [V off] could be taken as the schematic template for 'set off', with equal justification. Therefore, using collostructional analysis only seems suitable as long as the researcher has strong arguments to take a specific sequence of words and a specific schematic structure as the basic construction, or as the basis for analysis, and frequency information on the schematic templates is available. As yet it is, however, often impossible to collect frequency data from corpora on schematic templates.

Mutual Information (MI) values

Another measurement of association strength that is often used is Mutual Information (MI). For the calculation of this value, the pre-definition of the schematic structure is not needed. The Mutual Information value gives information about the relative frequency of the co-occurrence of two words, in relation to the total number of times that the words occur within a corpus. Of the value that is obtained in such a calculation, the log-quotient is taken, since, as mentioned before, corpus frequencies have a Zipfian distribution, and a normal distribution is needed for statistical analyses.[11]

In this calculation, the number of times that the first word of the construction (w1), such as, for instance, 'set' in [set off] occurs in the corpus

---

[11] The formula that can be used for the calculation of Mutual Information is the following (adjusted from Brown et al., 1992):

$$P[w1\ w2] = Log \frac{P([w1\ w2]/corpus\ size)}{(P([w1]\ /\ corpus)-P([w1\ w2]/corpus\ size)*(P([w2]/corpus)-P([w1\ w2]/corpus\ size)}$$

is multiplied by the number of times that the second word (w2) 'off' in the expression [w1 w2] occurs in the corpus. Church and Hanks (1990) explain that 'the assumption behind this calculation is that the words are distributed at random in a text [at chance, in our terminology]. It is obvious to a linguist that this is not so, and a rough measure of how much *set* and *off* attract each other is to compare the probability with what actually happens …' (p. 25).

The disadvantage of MI-analysis is that only two word-sequences can be taken into account. With Enhanced Mutual Information (EMI), as proposed by Zhang, Yoshida, Tang and Ho (2009), it is possible to calculate the relative frequency of co-occurrence of three or more words.[12]

5.2.3   The selection of targets on the basis of MI and EMI values

For this study, MI and EMI values were computed to measure the relative frequency of the target items included in the tasks. These measurements were chosen because the main focus was to investigate whether the frequency of individual multiple word sequences affects children's segmentation and processing, without having a strong expectation as to whether the transitive and locative phrases used as schematic templates actually form basic schematic units in the children's mental representations. Moreover, MI and EMI values did not require us to define the 'heads' of the (transitive and locative) constructions, as would have been required for some other frequency calculations. See Appendix 4 for the selected items and their MI or EMI values.

The MI and EMI values were used to create two bins: one containing items with a 'high frequency of occurrence' and one with 'low frequencies'. For the creation of these two bins, the ten targets occurring at the extreme ends of the complete lists of potential transitive and locative multiword target items – ranging from items with very high to very low MI and EMI values – were selected. It was then made sure that the MI and EMI values did not vary significantly within these groups, but that they did differ significantly between the high and low frequency groups created. Potential target items that were not near the extreme ends of the frequency range were discarded.

Following the creation of these bins, one more division was made. All bins were randomly divided in two. This led to two sets of items (sets A and

---

[12]   The formula used to calculate this EMI value is:

$$P[w1\ w2\ w3] = \mathrm{Log} \frac{P([w1\ w2\ w3]/\text{corpus size})}{(P([w1]/\text{corpus})-P([w1\ w2\ w3]/\text{corpus size}))*(P([w2]/\text{corpus size})-P([w1\ w2w3]/\text{corpus size}))*(P([w3]/\text{corpus size})-P([w1\ w2\ w3]/\text{corpus size}))}$$

B), both with half the high frequency transitive phrases, half the low frequency transitive phrases, half the high frequency locative phrases and half the low frequency locative phrases (see Appendix 4 for the item sets). The reason for splitting the bins was that not all selected items could be incorporated in the tasks all at once, taking into account that the youngest children participating in the tasks would be only 4 years old, and thus have a limited attention span. By creating two comparable bins, every task could be split in two, and children could be given only half a task in one testing session. With the creation of the two sets of items, it was also possible to give a child one task with item set A in one session, and another task with item set B in the same session, whereas on the next testing day, the child could be given item set B in the first task, and item set A in the second task. This way, children never came across the same target item twice in one testing session.

For the selection of Turkish items, a similar procedure was followed. First of all, lists with possible transitive and locative expressions were prepared. Their suitability in tasks with young Turkish children was informally estimated by two native speakers; there is no Turkish equivalent for *Woordwerken*. Where doubts occurred about the suitability of items for young children, alternatives were chosen. Then, the MI values were calculated for the multiword target items, with Google frequencies as a basis, and bins were created. Again, two sets of items were prepared, including both high and low frequency transitive and locative phrases. In the two tasks that were prepared in Turkish, one set occurred in one of the two versions, so that, as in Dutch, children could perform two tasks on one day without being exposed to the same target twice.

In the Turkish versions of the tasks, there were also locative phrases included that consisted of a single word in writing, such as '*bahçede*' ('in the garden'). For these, relative frequencies could not be calculated, since it was impossible to count the frequencies of bound morphemes. Hence, these locative phrases were merely selected on the basis of their meaning: they always corresponded to a locative postpositional phrase that occurred in the longer and more complex form (e.g., '*bahçenin içinde*', 'in the garden').

Although the shorter and longer forms of a locative target item could both be translated using the same prepositional phrase in English or Dutch, it should be noted that the longer forms are in fact semantically more specific. In the shorter form, with the form [N+GEN.LOC], the general locative morpheme '*-de/-da*' is used. This morpheme is translated by various locative prepositions in English and Dutch: 'in', 'on', 'under', etc. In the longer locative form, which consists of [N+GEN POST+GEN.LOC], the morpheme before the general locative marker – here referred to as a

postposition, although it is generally analysed as a genitive-marked noun ('inside', 'outside', etc.) – has a relatively specific meaning, compared to the general locative marker. Since the more general shorter forms can be expected to occur more often in everyday Turkish than the longer forms, the use of these two forms will probably not just be affected by the meaning of the forms, but also by their frequencies of occurrence. A detailed discussion with respect to these two forms and their comparability will be given in Chapter 7.

For now, it will suffice to say that all target items occurred in all tasks, and that both in Dutch and in Turkish, it was made sure that target items did not occur at the beginning or end of sentences in any of the tasks, so as to avoid wrap-up or attention effects (cf. Field, 2004).

5.3   Tasks

Six tasks were developed for this study. All of these were introduced in Chapter 4: a dictation task, a sentence segmentation task,[13] a last-part repetition task, a click task, a mixed-words task and an eye-tracking task. The Turkish children, all native speakers of Turkish, performed only two of these tasks. In Table 8, the numbers of children who participated are indicated, per task, language, and grade.

		Grade	Dutch		Turkish	
			N	%	N	%
Dictation task	Pre-literate	K1	19	27.5	–	–
		K2	16	23.2	–	–
	Literate	E1	15	21.7	–	–
		E2	19	27.5	–	–
Sentence segmentation task	Pre-literate	K1	20	27.0	–	–
		K2	19	25.7	20	44
	Literate	E1	19	25.7	–	–
		E2	16	21.6	25	56

---

[13] This task has been referred to as a 'tapping task' in previous publications. See Veldhuis (2011) and Veldhuis et al. (2102). Calling the task a 'sentence segmentation task' seems however more accurate. See Section 5.3.2.

		Grade	Dutch		Turkish	
			N	%	N	%
Last-part repetition task	Pre-literate	K1	20	26.3	–	–
		K2	20	26.3	20	44
	Literate	E1	19	25.0	–	–
		E2	17	22.4	25	56
Click task	Pre-literate	K1	22	27.5	–	–
		K2	19	23.8	–	–
	Literate	E1	19	23.8	–	–
		E2	20	25.0	–	–
Mixed-words task	Pre-literate	K1	23	28.0	–	–
		K2	19	23.2	–	–
	Literate	E1	20	24.4	–	–
		E2	20	24.4	–	–
Eye-tracking task	Pre-literate	K1	13	21.3	–	–
		K2	14	23.0	–	–
	Literate	E1	16	26.2	–	–
		E2	18	29.5	–	–

Table 8: Dutch and Turkish participants per task and grade

The Dutch children performed the tasks in sequence, with small breaks in between, over five separate sessions on five days. In each testing session, a child participated in two or three (parts of) tasks. Apart from the dictation and eye-tracking tasks (see Section 5.3.1 and 5.3.6), the tasks were split into halves and presented to the children in such an order that they did not come across the same item twice in one session (see Section 5.2). At most two weeks, and at least two days, passed between participation in the first and the second half of a task. The dictation task was either conducted at the beginning or at the end of a session. The Dutch children also performed a cognition test, namely Raven's (1962) Coloured Progressive Matrices (sets A, A_B, B). This task was conducted in one session, together with the eye-tracking task. In addition, background information was collected and information about the Dutch children's reading proficiency skills was collected from formal school tests (the 'CITO' tests).

Below, the background and the set-up of the tasks, the coding of the results, and the questions that were asked to obtain background information will be discussed in separate subsections.

5.3.1   Dictation task

A task using dictation in order to obtain insights into children's language segmentation was first developed by Chaney (1989). Chaney asked primary school children to dictate a well-known story (The Pledge, or the Stegosaurus) word-by-word so that she could investigate the units that children seemed to work with, and investigate whether they were able to segment words equally well for different types of words: content words and function words, and words of different syllable lengths.

In a BA-thesis, Chau (1997) also asked children to dictate a story. In contrast to Chaney (1989), however, children were now asked to recite sentences provided by the researcher like a robot. Again, the focus of this study was to investigate how many words children could segment correctly from the sentences.

The dictation task developed for the present study is based on these two sources but differed in two ways from these predecessors.

First, children were not asked to recite or dictate a story that they knew well, but rather to tell the researcher a story that they came up with themselves. They were asked to tell the researcher for instance about a nice experience they had had (e.g. a trip or a vacation). The reason for asking the children to come up with their own stories was to avoid children making segmentations according to well-rehearsed rhythmic patterns, which are often stressed in children's stories and nursery rhymes. Not having elicited materials in the task probably also led to more natural language processing (see Chapter 3). The multiword target items that were selected for the other tasks (see Section 5.2) thus did not occur in this task.

A second difference with previous incarnations of the task concerned the unit that children were requested to segment in their dictations. Since the focus in this study was on investigating what segments children would come up with unprompted, children were not asked to tell the story or give the sentences word-by-word, as Chaney (1989) and Chau (1997) instructed, but rather to tell the story 'part by part' (*stukje voor stukje*), so that the researcher could write it down. Children were given one or two examples containing both single words and multiword sequences as parts ('*Vorige … week is … er … iets vreemds … gebeurd*', lit. 'Last … week is … there … something strange … happened', dots indicating pauses made by the researcher), to avoid a bias towards a particular style of segmentation.

The dictation task was only conducted with the Dutch children.

Coding the dictation task

To analyse the results of the dictation task, it was first investigated how many segments children had distinguished in their stories, and how many of those were conventional words. The segments were assigned to one of the following categories:

(a) Responses smaller than words: a segment was assigned to this category if it consisted of a syllable or a sound, e.g., − *ka − b − ou − ter − s* − ('gnomes', hyphens indicating pauses children made) consisted of five sub-word segments.

(b) Single-word reactions: a segment would be assigned to this category if it was a single word, e.g., − *elfjes − en* − ('− fairies − and −'). Monosyllabic words that were segmented along word boundaries were also assigned to this category, even though it cannot be known for sure whether children based their segmentation in these cases on words or on syllables.

(c) Multiple-word reactions: a segment was assigned to this category if it consisted of more than one word, e.g., − *dit boek − gaat over* − ('− this book − is about −').

Within these categories, several sub-categories could be distinguished. Category (a) for instance includes 'syllables' and 'phonemes', and multiple-word reactions (category (c)) could be 'determiner+noun-combinations', 'prepositional combinations', 'combinations with the schematic form [and X]', X standing for 'any possible word', or verb+noun-combinations. A qualitative analysis of these categories will be provided in Chapter 6. For the quantitative analysis of the task, which will be provided in Chapter 7, the abovementioned three categories were used. Specifically, the total numbers and percentages were calculated of each type of response (as a ratio proportion of the total number of segments that a child had produced).

5.3.2 Sentence segmentation task

This task was conducted in Turkish and in Dutch. It was based on a word-count task that was described by Karpova (1966, orig. 1955) and a tapping task developed by Holden and MacGinitie (1972). These tasks were later also applied by Ehri (1975), Berthoud-Papandropoulou (1978), Tunmer, Bowey & Grieve (1983), Bialystok (1986b), Durgunoğlu and Öney (1999), and Kurvers (2002). In these studies, children, and sometimes adults (Kurvers, 2002), were asked to indicate how many words they heard in grammatical or

ungrammatical sentences they were provided with, either by clapping their hands, by tapping on the table, by counting, or by moving plastic chips on a table. Alternatively, they were asked to count words in sentences (Karpova, 1966, orig, 1955; Berthoud-Papandropoulou, 1978; Durgunoğlu & Öney, 1999).

In a progressive segmentation task, related to the task as developed for this study, Fox and Routh (1975) first asked illiterate and literate children to segment sentences that they provided them with into smaller segments, and then asked them to give 'a little bit' of the first word of the sentences they had segmented. By avoiding the term 'word' in the instructions to the tasks, it could be analysed how often participants provided word units more or less automatically (Kurvers, 2002), and how natural such units apparently are for participants. By asking participants furthermore to also segment a word into even smaller parts or bits, a progressive segmentation task cannot only tap into children's ability to segment words from speech, but also explore their abilities to distinguish other units in language.

Since the current study is more explorative, in the sense that it aims to find out what kinds of building blocks pre-literate and literate Dutch and Turkish children operate with, rather than whether they are able to recognize single words, children were not asked to indicate word units, but rather how many parts or bits ('*stukjes*' in Dutch, or '*bölüm*' or '*parça*' in Turkish) they perceived in the stimulus sentences (as in Kurvers, 2002). The Dutch children were asked to repeat 29 sentences per testing session, read out to them by the researcher, and the Turkish children were given 30 sentences this way. All children were asked to indicate with small blocks on a table of how many parts they thought the sentence consisted, while they repeated it. To avoid effects from rhythm and stress, the task did not involve clapping one's hands or tapping on a table.

There were two testing sessions, in which different sentences with different target items (from set A or B, see Section 5.2) were included. All sentences contained both mono- and multi-syllabic words, open and closed class words, and 20 of the sentences that a child heard per day contained a multiword sequence consisting of a prepositional locative phrase (e.g., 'in the kitchen') or a transitive noun+verb-combination (e.g., 'wash hands'). In addition, there were nine filler sentences in the Dutch version of the task in which no such multiword targets occurred. In the Turkish version, there were five filler sentences in addition to the 20 sentences that contained selected multiword targets. The other five sentences in the Turkish version

of the task included the single-word locatives, such as '*bahçede*' ('garden' + GEN.LOC).

All children were given three example sentences before they performed the task, in which the researcher used both single words and multiword sequences as examples of answers to avoid any bias towards a specific segmentation style (e.g., '*Ik vind – groene – appels erg lekker*', 'I really like green apples', hyphens indicating segment boundaries). Appendix 5 contains the stimulus sentences.

Coding the sentence segmentation task

For the analysis of the sentence segmentation task, it was investigated whether and how children had segmented the multiword target sequences that occurred in the sentences, notwithstanding the segmentations they made in the rest of the sentence. The obtained answers were assigned to one of the following five categories:

(a) Responses smaller than words: an answer was assigned to this category if the target item was only or mostly (as could be counted from the number of parts made in a target) segmented into syllables or sounds, e.g.'… – *han* – *den* – *wasse* – *n* – …'or '… – *e* – *l* – *te* –*mi* – *zle* – *mek* –…' ('… – wa – sh – hand – s – …'; hyphens indicating segmentations made by the children, dots indicating other parts of the sentence in which this target item occurred).

(b) Single-word reactions: an answer would be assigned to this category if the multiword target item was completely segmented at word boundaries, e.g., '… – *handen* – *wassen* – …' in Dutch or '… – *el* – *temizlemek* – …' in Turkish ('… – wash – hands – …').

(c) Complete target construction reactions: an answer was assigned to this category if the target item was left unanalyzed, and segmented as a whole, e.g., '*je moet* – *je* – *handen wassen* – *voor* – *het eten*' in Dutch, or '… – *el temizlemek* – …' in Turkish ('you have to – wash hands – before – dinner').

(d) Other multiple-word reactions: an answer was assigned to this category if the child had used multiword sequences that nevertheless did not correspond to the target item, e.g., '… – *op de* – *stoel staan*' ('… – on the – chair standing').

(e) Other: all other responses, such as deictic reactions, or non-verbal or mixed responses, were coded as 'other', e.g., a reaction to a name: 'But Bram was not on the swing today!', or 'I know Fleur too!'.

The total numbers and percentages for each response type were calculated for all segmentations of the target items. These percentages were later used in the analyses in which the effects of literacy and of the frequency of occurrence of specific expressions were assessed (see Chapter 7).

5.3.3 Last-part repetition task

The last-part repetition task was developed on the basis of Karmiloff-Smith et al. (1996). In that study, the researchers read a story out loud to the children, with pauses occurring in the middle of sentences. Children were then asked to repeat the last single word that the researcher had said, before she made the pause. The task is intended to provide evidence on how well children are able to segment single words, of various syllable lengths, from ongoing speech. Homer (2000), Kurvers (2002), Kurvers and Uri (2006) and Ramachandra and Karanth (2007) also used this repetition task, and apart from translating it into other languages (Chinese, Dutch, Norwegian and Kannada respectively), and using a different story, no major adjustments were made to the original task. In all tasks, participants (children and, in Kurvers, 2002, and in Ramachandra & Karanth, 2007 also adults) were asked to repeat the last single word that the researcher had said before a pause, and in all tasks, it was investigated whether pre-literate children and illiterate adults differed from literate children in their segmentation of content and function words.

As an alternative to asking children to repeat the last word, Karmiloff-Smith et al. (1996) in a follow-up study also asked children to repeat the last 'thing'. This resulted in longer responses, which was interpreted as showing that children do make a distinction between 'last word' and 'last thing' when carrying out this task.

A possible explanation for the differences in responses after 'word' and 'thing' may however also lie in the fact that in regular language use, 'thing' will more often refer to a message than to a linguistic unit (as in: 'The only *thing* we did was …'). In that sense, the choice for 'thing' may indeed be why children gave longer answers than when asked to repeat a 'word', but only because the two words do not have the same connotation in everyday language use.

In the last-part repetition task as it was developed for the present study, children listened to the researcher reading a short story out loud, in which 30 pauses were inserted in mid-sentence position. In contrast to the earlier studies, the children were not asked to repeat the last single word or thing that they had heard before the pause, but rather to repeat the last piece or bit ('*laatste stukje*' in Dutch or '*bölüm*' or '*parça*' in Turkish). This way, the task was designed not to focus on children's awareness of words, or on their ability to distinguish single words from ongoing connected speech, but rather on the units or building blocks they recognized in speech. After a pause, backtracking was used to continue the story, which meant that the researcher would repeat more of the sentence than just the last word or target unit (cf. Karmiloff-Smith et al., 1996; Levelt, 1989).

Before the task began, all children were given three example sentences with pauses, like '*Er was eens een klein meisje …*' in Dutch ('Once upon a time there was a little girl …'). Children were told they could choose what they thought was the last part they had heard, and the answers that were given by the researcher for these examples consisted both of single words and multiword parts (like '*meisje*', or '*een klein meisje*', 'girl' or 'a little girl'). In the Dutch version of the task, pauses were made after 20 selected multiword targets and after ten filler items, which were either function or content words. In the Turkish version of the task, there were also 20 pauses after the selected multiword targets. In contrast to the Dutch version, however, there were also five pauses after single-word locative targets (e.g., '*bahçede*', 'garden'+gen.loc) in addition to five pauses after filler items, which were all function words.

Like the sentence segmentation task, the last-part repetition task was prepared and conducted in Turkish and in Dutch. Appendix 6 contains the stories that were used.

Coding the last-part repetition task
The analysis of the answers obtained largely corresponded to the method applied for the sentence segmentation task, except for one additional category ('anticipations', see below). All responses were assigned to one of the following categories:

(a) Responses smaller than words: a sentence was assigned to this category if only a repetition of the last syllable or sound was presented, e.g. '*l*', after the pause in the sentence '*Zelfs op de stoel …*' ('Even on the chair …'), or in Turkish '*e*' or '*de*' after a pause in the sentence ending in the target item '*bahçenin içinde*' before the pause ('in the garden').

(b) Single-word reactions: single-word reactions were those where a child
     had repeated only the last word from the utterance that he/she had just
     heard, e.g., '*stoel*' after the pause in the sentence '*Zelfs op de stoel …*'
     ('Even on the chair …'), or in Turkish '*içinde*' after hearing '*bahçenin
     içinde*' before the pause ('in the garden').

(c) Complete target construction reactions: an answer was assigned to this
     category if a child repeated a multiple-word target item after the pause,
     e.g., '*op de stoel*', after '*Zelfs op de stoel …*' ('Even on the chair …') or in
     Turkish '*bahçenin içinde*' after hearing '*bahçenin içinde*' before the pause
     ('in the garden').

(d) Other multiple-word responses: an answer was assigned to this
     category if a child repeated more than just the multiple-word target
     item after the pause. This could for example be the repetition of the
     whole sentence.

(e) Anticipations: an answer was assigned to this category if a child did not
     (merely) repeat from the sentence he/she had heard, but also
     anticipated what would come up next, e.g., a child said '*is het cadeautje
     niet!*' ('is the present not') after hearing '*Zelfs op de stoel …*'.

(f) Other: all other responses, such as deictic reactions or non-verbal
     responses, were coded as 'other', e.g., after hearing about a dog given
     as a birthday present: 'A real dog?! Did she get a real dog for her birth-
     day?!', or 'Owww, I would also love to get a dog for my birthday!', or,
     after a sentence about dough that fell to the floor: 'Oops, my mother
     always gets really angry if I make a mess in our kitchen …'.

As for the sentence segmentation task, the total numbers and percentages of
each response type were calculated so that they could be incorporated in the
analyses for literacy and frequency effects (see Chapter 7).

5.3.4 Click task

The click task was an adjusted version of earlier click tasks (cf. Abrams &
Bever, 1969; Cohen & Mehler, 1996; Cutler et al., 1993; Delogu, Conte &
Sementina, 1998; Fukuda, 1983; Green, 1977) in which participants heard a
sentence with a click in it and were either requested to indicate with a pen
on paper where they had heard the click (this method was called *click
localization*) or to press a button as soon as they heard it (*click detection*). In
the latter case, reaction times were used to draw conclusions about linguistic
processing (cf. Abrams & Bever, 1969; Cohen & Mehler, 2004; Cutler et al.,

1993; Green, 1977). Since the end of the 1960s click localization tasks have been criticized as they might not really reflect linguistic processing, but rather sentence reproduction in writing (cf. Abrams & Bever, 1969; Fukuda, 1983). Click detection tasks, likewise, have been criticized, for instance because natural listening and linguistic processing might differ from listening in click detection (cf. Mitchell, 2004), or because click detection reaction times may be the result of acoustic effects rather than of processing features (Cutler et al., 1993; Fukuda, 1983). Still, the method was applied quite frequently to investigate linguistic processing until the end of the 1980s, when it seemed to fall out of favour, – or maybe it was simply surpassed by more advanced online methods such as ERPs, fMRI-scans and eye tracking. Even so, the click task may be a good technique for investigating language processing, if precautions are taken to circumvent the problems that the task has met with in the past.

The criticism that the click task largely depends on acoustic effects has been investigated several times. Indeed, acoustic effects should be taken into account when conducting a click detection task, as acoustic features of language, such as vowel-length and other prosodic features of words and sentences, have been found to affect click detection (Cutler et al., 1993; Fukuda, 1983).

As long as clicks are, however, not imposed within a word, but only after a word, and if the clicks are placed after one class of words only (thus only after open or only after closed class words) or after both classes of words equally often, then the problem bothering Cutler et al. (1993) could be avoided. In that case, different word classes no longer lead to bias. Moreover, by only presenting the task to native speakers, any influence from not understanding the language and not being able to process the language as such, which was also mentioned as a point of critique (Cutler et al., 1993), can be avoided as well.

Green (1977) and Delogu, Conte and Sementina (1998) tried to ensure that subjects actually processed the language to which they were exposed by asking them either to repeat the sentence they heard (after performing a click detection task) or by asking them to come up with a suitable follow-up sentence to the sentence in which they had been requested to detect the click. If they could repeat or complete the sentence, Green (1977) and Delogu, Conte and Sementina (1998) argued that participants had heard and processed the sentence in a natural way. These additions, however, make the task complex. Pressing a button, as is requested in the click detection task, may already intervene with natural processing; the additional tasks exacerbate that effect. This does not mean that comprehension should not be

controlled for, but the way in which it is tested should be as unintrusive as possible. Simple comprehension questions may be the most reliable means to do this.

Apart from these studies, in which it was investigated to what extent responses to click tasks depend on acoustic features, click detection and click localization tasks have been applied to investigate sensitivity to clause boundaries (cf. Abrams & Bever, 1969; Chapin, Smith & Abrahamson, 1972; Fodor & Bever, 1965; Fukuda, 1983). All these studies, in which clicks were put in the middle of clauses or right after them, suggested that participants perceived clicks better or faster when they occurred at clause boundaries than when they occurred within clauses. Chapin, Smith and Abrahamson (1972) and Fodor and Bever (1965) furthermore showed that the subjective click localization, as indicated by participants, tended to be at syntactic boundaries, or at the clause boundaries that were closest to the actual location of the click. This suggests that people process language not along word boundaries but in bigger units.

In previous studies, researchers only focused on clause boundaries. Since the purpose of the current study is to investigate whether and to what extent children process language along word boundaries or along larger selected multiword chunks, and to investigate how the frequency of occurrence of such chunks influences processing, clicks were placed after and within selected multiword target items (see Section 5.2).

We expected that highly frequent expressions would be more entrenched as single units than expressions with lower frequencies (see also Chapter 2 and Section 5.2). It would therefore be harder for participants to hear clicks within frequent expressions than within an expression with a lower use frequency. Regarding literacy, it was expected that literate children would be better at locating the clicks that occurred within multiword target items than pre-literate children, as literate children are trained to decompose language, including multiword units, into the smaller parts (words), between which the clicks occurred.

In order to investigate these hypotheses, and taking into account experiences of previous studies that used click tasks, pre-literate and literate children were asked to repeat pre-recorded sentences with a click inserted into them, and to make a click sound to indicate their perception of the location of the click while repeating the sentences. The clicks occurred in 40 target sentences within selected multiword expressions, for which the frequency of occurrence was controlled, e.g., '*je moet je **handen X wassen** voor het eten*' ('you have to wash your hands before dinner", bold script

representing the target unit, X representing the click). In 18-filler sentences, the click occurred after such units, e.g., '*zullen we **film kijken X** vanavond?*' ('shall we watch a film tonight?'), or elsewhere if no specific target unit was included at all. The children obtained comprehension questions after most repetitions, in order to check whether they had processed the meaning of the sentences as they would do in regular language processing.

In total 58 sentences were created and recorded in Dutch using the computer program Praat (Boersma & Weenink, 2012). Clicks of white noise (-10dB, 50ms) were inserted in Audacity in mid-sentence position within or right after the target expressions. The clicks were preceded and followed by 40ms of silence (also created in Audacity), as pilots showed that the clicks were otherwise extremely hard to perceive. The multiword target items consisted of ten transitive noun+verb-combinations and ten locative prepositional phrases per session. Frequency was controlled for. There were also nine fillers per session, in which clicks either occurred right after selected multiword targets, or in which no selected expressions were included. In these sentences, there were also clicks inserted in mid-sentence, but not after or within specific multiword expressions.

In a single session, the children first heard three examples and were then asked to repeat the first 29 sentences, and to indicate where they had heard the click by making a click sound themselves in their repetitions of the sentences. They were tested on the next 29 sentences a couple of days, or at most two weeks, later in a second session. Appendix 7 lists the stimulus sentences.

Coding click task
For the analyses of the click task, the number of correctly repeated clicks was first calculated for all sentences. It was then counted how often children placed the click erroneously right before or after the multiword target items when they in fact had occurred within such items. By comparing the score to correctly repeated clicks that had occurred within or after multiword target units, it would be possible to analyse whether children indeed replaced the clicks right after or before the multiword targets, as was expected on the basis of the studies that focused on clause boundaries, and on the literature on the processing of multiword units. It was expected that the literate and the pre-literate children would be equally capable of repeating clicks that occurred after the multiword units, as in this circumstance it was not necessary to 'break up' the sequence (cf. Arnon, 2010; Tomasello, 2006), but that the literates would be better at correctly repeating clicks that occurred

within multiword targets, as this would require decoding the multiword chunks into its constituent words.

Like for the other tasks, the numbers and percentages of the scores per type of answer (correctly repeated / replacement just before or after target units / other position) were included in statistical analyses to investigate the effects of literacy and frequency (see Chapter 7).

5.3.5   Mixed-words task

The mixed-words task developed for this study was a variation on better known tasks that usually tap into children's phonological awareness. Such tasks ask children to switch, or not to switch, phonemes around that occur in a jumbled order in words. This makes it possible to assess subjects' auditive or temporal ordering abilities or the level of phonemic awareness that they have at that moment (cf. Adams, 1990; Huizenga & Robbe, 2013; Huizenga, 2010). These tasks are focused on the idea that children have to become aware of sounds or phonemes in order to manipulate them, and build on the assumption that in general letters or phonemes that occur in jumbled order within words do not (highly) affect the understanding of these words, at least as long as the first and the last letter remain in place.

Tihs lteatr armgenut has aslo been made by rsceaerhres who have swohn taht ppleoe can sitll qtiue esaily raed and unrdsetand seetencns in wchih hadlry any ltetres ocucr in teihr crroect psotinois (cf. Velan & Frost, 2007). The fact that the above sentence can be understood once given some attention suggests that our understanding of written words is not highly dependent on the order in which letters within words occur. Quite the contrary, our literate brain, trained as it is in reading, seems to repair uncommon positioning of letters automatically, suggesting that for a proper understanding of language, a misplacement of letters does not always obstruct regular language processing.

If this 'Cambridge effect' (cf. Velan & Frost, 2007) holds for letters in words, then it might well be possible that for a proper understanding of sentences, the order in which words occur does not always matter either. In fact, slips of the tongue in which people mix up the order of words (e.g., 'the shoe of your heel' instead of 'the heel of your shoe') are not uncommon, and they do not often seem to disturb understanding.

Likewise, 'funny images' in which sentences are presented with mixed-word orders are nowadays popular on social media sites and on T-shirts; see Figure 7 below for some examples. These images suggest that, indeed,

people's brains tend to repair sentences in which words occur in jumbled order more or less automatically.

In a test in which illiterate and literate adults were asked to repeat pseudo-words, Reis and Castro-Caldas (1997) found that illiterates tended to make more repairs in the pseudo-words they obtained for repetition, changing them into real words. Reis and Castro-Caldas suggest that illiterates have difficulties in processing nonsense language (as they cannot dissociate phonological knowledge from actual language to the same extent as literates, who have learned about phonemes), and that they attach greater value to meaning (as Morris, 1993, stated). This, in turn, could suggest that whenever illiterates are exposed to sentences with words occurring in jumbled order, they will make more repairs as well.

This is actually what has been found among children who were asked to repeat sentences that are jumbled up or that contain grammatical errors. Hirsch-Pasek, Gleitman and Gleitman (1978) and Kurvers (2002) report about studies in which it was found that children aged 5 and illiterate adults tend to repair ungrammatical sentences, seemingly without knowing that they do so, when just asked to repeat them. And Bialystok (1986b) concluded that in sentences with regular word order, especially younger children tend to segment meaningful phrases instead of words, whereas in sentences with jumbled order, they rather refer to syllabic units. From these three studies, it thus seemed that illiterates are less able than literates to play around with ungrammatical language or pseudo-language: they process language on the basis of meaning rather than form.
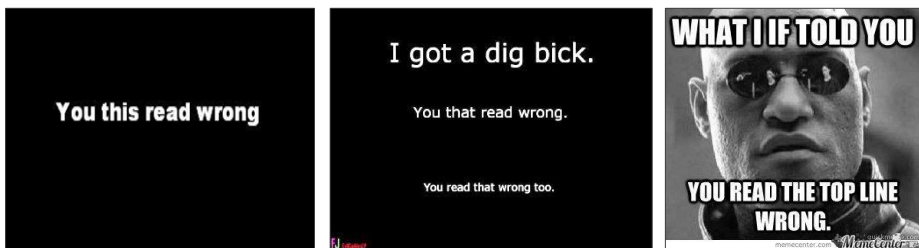
Figure 7a, b, and c: 'Funny sentences' with mixed-word order, adjusted to black and white[14]

---

[14] Figure 7a: Reprinted from Tumblr, retrieved January 21, 2013, from:
http://www.tumblr.com/tagged/you-read-this-wrong (copyright by Cherrybam).
Figure 7b: Reprinted from Funnyjunk, retrieved January 21, 2013, from:
http://www.funnyjunk.com/funny_pictures/3246036/You+read+that+wrong/ (copyright by Evilcookieslv).
Figure 7c: Reprinted from Memecenter, retrieved January 21, 2013, from:
http://www.memecenter.com/fun/855713/did-you-read-it-wrong (copyright by Memecenter).

Based on these findings, the mixed-words task prepared for this study included sentences in which selected multiword expressions occurred with jumbled word order. Children were asked to repeat the sentences verbatim. Fifty-eight sentences that were largely similar to those used in the click task were developed. Of these, 40 sentences contained a selected multiword expression in mid-sentence position in jumbled order. The other 18 sentences were fillers, in which words were not mixed.

The mixed order for the transitive noun+verb-sequences was always a simple swap of noun and verb. Thus instead of '*je moet je* **handen wassen** *voor het eten*' ('you have to wash your hands before dinner'), children heard '*je moet je* **wassen handen** *voor het eten*' ('you have to your wash hands before dinner'). For the prepositional locative phrases, two types of mixed order occurred: in the first type, the correct order such as '*op de stoel*' ('on the chair') would be replaced by '*stoel op de*' [N PREP DET] and in the second type, the mixed order would be '*stoel de op*' [N DET PREP]. A prepositional target item either occurred in the first mixed order, or in the second, but not in both. There were also nine filler sentences, in which words were not presented in mixed order.

Each child heard 29 sentences, including the nine fillers, in the first session, and the next 29 sentences in the second session on another day. Before the task began, children were presented with three examples, out of which two included multiword expressions that were rather similar to the target units, and presented with mixed-word order. See Appendix 8 for the sentences occurring in the mixed-words task.

The hypothesis was that if frequency has an effect on children's verbatim repetition of sentences with mixed words, it was expected to be most visible in the highly frequent multiword expressions: these would be repeated more often in their corrected form. Regarding the literacy effect, there was no clear hypothesis. Despite the fact that previous studies lead one to expect that pre-literates would be worse at verbatim repetition of sentences with ungrammatical word orders (cf. Hirsch-Pasek, Gleitman & Gleitman, 1978; Kurvers, 2002; Reis & Castro-Caldas, 1997), this expectation could not be extended to predict that the target items would be processed holistically if they changed the word order in their repetitions. Not changing the word order of a sentence with mixed words, after all, does not necessarily mean that children processed the sentence they repeated word-by-word (see below).

Coding the mixed-words task

For the analysis of the mixed-words task, it was checked whether the frequency of occurrence of the target expressions affected the results. Accordingly, the repetitions of the children were categorized as follows:

(a) Literally repeated: a response was assigned to this category if a child had literally repeated the sentence he/she was asked to repeat – with mixed-word order.

(b) Not literally repeated, regular word order: a response was assigned to this category if the child had applied the regular (or original) non-mixed-word order in his/her repetition of the sentence.

(c) Not literally repeated, but still mixed-word order: a child's response was assigned to this category if he/she had repeated the sentence with jumbled order, but the sentence was not repeated verbatim.

(d) Not literally repeated: '*de*' ('the') locative multiword target changed in children's repetition into '*d'r*' (which is more or less equal to 'there'): a response was assigned to this category if the child had replaced the '*de*' ('the') which often occurred in locative multiword targets with '*d'r*', which increased the grammaticality and probability of the sentence: e.g., a child would say: '*stoel d'r op*' ('chair there on') instead of the actual stimulus item '*stoel de op*' ('chair the on').

(e) Other responses: these included for example deictic reactions or non-verbal responses, e.g., after a sentence in which bananas and mandarins were mentioned: 'But I don't like mandarins at all!'.

For the analyses, it was assumed that if children literally repeat the sentences, with the mixed-word orders (response category (a)), this shows that they can isolate words from a larger expression. In contrast, the responses to category (b), in which children changed the multiword targets back to their regular forms, was assumed to be evidence for storage of the targets as wholes: mental representation of fixed units is then seen to override the actual input.

Answers assigned to category (c), in which children did repeat the multiword targets with a mixed order, but not the same one as in the stimulus item, were assumed to show that children did realize that the sentence they heard was not in the right word order, but that they could not repeat verbatim what they had heard – a sentence with mixed-word order – or that they did not manage to make a correct and proper repair of the

sentence they had heard. Answers assigned to category (d) were assumed to show that children wanted to make the sentence more plausible.

This categorization of responses allowed for an analysis of the frequency effect: calculating how often children left the multiword targets with mixed-word order untouched in their repetitions and calculating how often they changed the targets back into a regular word order allowed us to investigate whether children more often changed the high frequent targets back than the targets with lower frequencies of occurrence. This would have been expected on the basis of the higher degree of entrenchment of the former (see Section 5.2). The literacy effect could not be investigated as straightforwardly with this task: on the one hand, not changing the words back to their regular order could be interpreted as showing that children could process multiword sequences word by word. However, since young children are very well able to imitate whatever they hear (or see), and since changing the word order back into its regular order also required the manipulation of the language to which the children were exposed – a metalinguistic task – this theory would not necessarily hold. Accordingly, in the analysis of this task (Chapter 7), only frequency effects will be discussed.

5.3.6   Eye-tracking task

In 1998, Rayner published an article in which he discussed '20 years of eye tracking'. Even though eye tracking may still be regarded an innovative way of investigating linguistic processing, it is not really new anymore. Nevertheless, the application of eye trackers has definitely changed since the early 1980s.

The first studies in linguistics using an eye tracker were mainly focused on reading (Anderson, Goldberg & Hidde, 1976; Hyönä, Niemi & Underwood, 1989; Rayner, 1998; Rayner & Pollatsek, 1989; Vitu & O'Regan, 1995). In such studies, the question was usually which words participants focused on while reading sentences that were presented to them on a screen. This is assumed to provide information on how words are processed in reading, and on how people anticipate and process sequential information in sentences (Kamide, Altmann & Haywood, 2003). Cross-linguistic effects on people's reading styles and strategies, or effects of specific writing conventions, such as the inclusion or exclusion of spaces, have also been investigated (cf. Peng, Orchard & Stern, 1983; Pynte & Kennedy, 2006; White, 2008).

In addition to these, Cooper (1974) was probably the first to investigate where people look when presented with various objects on a screen while

listening to spoken language. This paradigm has become known as the 'visual world paradigm' (Altmann, 2004). The basic idea behind the visual world paradigm is that people's eyes make unconscious movements or fixations on different pictures that they are presented with while listening to audio input (cf. Cozijn et al., 2011). This way, researchers have investigated children's and adults' semantic and syntactic processing.

For instance, Trueswell et al. (1999) investigated whether or not children and adults process ambiguous garden-path sentences such as 'put the frog on the napkin in the box' in a similar way. More specifically, Trueswell tested whether children and adults thought about a frog that had to be put on a napkin, and then had to be put in a box, or whether they thought about a frog that was already on a napkin, which had to be put in a box. From the eye movement data Trueswell et al. (1999) collected, and from having children acting out the sentences they were presented with, the authors concluded that adults and children (aged 5) do not process constituents in ambiguous sentences in the same way: While the children tended to bind the prepositional phrase to the verb phrase, which made them look at the empty napkin, and while they showed little urge to repair their interpretation, adults appeared able to do so.

The conclusions of Trueswell et al. (1999) have been disputed in follow up eye-tracking studies and in acting out tasks in which children were confronted with similar ambiguous sentences (cf. Meroni & Crain, 2003; Weighall, 2007). Trueswell et al. (1999) had linked their findings to children's limited processing resources. Meroni and Crain, in contrast, rather explained children's results as an 'order of mention' effect, and they showed that 5-year-olds can show adult-like results in eye-tracking tasks with ambiguous sentences.

While this debate is not particularly relevant for this thesis, the fact that eye-tracking studies using the visual world paradigm have shown to provide reliable information on the way in which even young children, at the age of 4 or 5, process language and units (in the form of constituents in the studies mentioned), was the reason why the eye-tracking task as developed for this study also involved a visual world paradigm. The task was largely explorative, as, to the best of my knowledge, no eye-tracking study has been used before to investigate whether pre-literate children process single words in language.

For the current eye-tracking task, the question was in what units children process the input when they are listening to language. In line with the other tasks developed for this study, the eye-tracking task focused on

investigating whether children process specific selected multiword units (consisting of noun+verb-combinations), that varied in the frequency of occurrence, as wholes, or whether they would break them up into smaller segments. The expectation was that if the frequency of occurrence of multiword units affects people's visual processing, then the children would look quicker at more frequent target pictures than at less frequent counter targets. Secondly, it was expected that if literacy affected children's processing of the multiword targets, then literate children would show processing of all separate parts of the target noun+verb-combinations in their eye movements, their gaze directions, and their fixations.

In order to investigate these hypotheses, a large number of line drawings were created, all similar in style, that represented different parts of selected multiword noun+verb-combinations or the complete multiword target to see how children would process such targets: word by word or as wholes. Children saw four pictures on a screen: one representing a complete target noun+verb-combination such as '*handen wassen*' ('wash hands'), two pictures representing the single words of which that construction was composed ('wash' and 'hands'), and one counter target picture which represented a multiword noun+verb-combination related to the noun in the target unit, such as '*handen vouwen*' (lit: 'fold hands', i.e., 'pray'). It was then investigated where and how quickly children fixated their gaze while listening to sentences in which the target constructions, the single words, or the distractors were mentioned.

As mentioned in Chapter 5, the noun+verb-combinations in this task were different from the ones occurring in the item sets A and B that were used in the other tasks. The reason for choosing different target items for the eye-tracking task was that it seemed hardly possible to visualize locative phrases such as 'in the bus' or 'under the cupboard' without making use of some object, animal or other referent indicating the location. This would make the picture more complex (see Scheurwater, 2010, for a pilot in which depicting prepositional phrases was tested). For the same reason, not all of the transitive constructions that were used in the other tasks were suitable for the eye-tracking task. An expression like '*bladzijde omslaan*' ('turn the page'), for example, would have been very hard to depict. Such sequences were thus excluded from the eye-tracking task. Finally, since all test items needed distractors consisting of other transitive combinations with the same noun (e.g., '*handen wassen*', 'wash hands' had '*handen vouwen*', 'fold hands' as a distractor), a further subset of the previously used items seemed unsuitable. Both the target unit and the counter target needed to be

depictable. See Appendix 9 for the targets that were regarded as easily recognizable by four colleagues to whom the drawings were shown.

For the task itself, all four line-drawings – depicting the target unit, the counter target, the noun and the verb – were presented at the same time to the children on a grey screen, since the contrast between black (line drawings) and white (background) had been found to be rather blindening (Scheurwater, 2010). An example of a screenshot of a set of drawings is given in Figure 8.
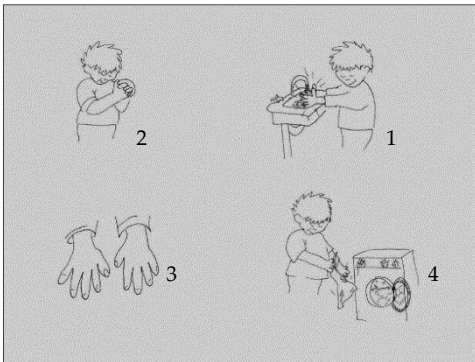
Figure 8: Screenshot of target-item '*handen wassen*'

As can be seen from Figure 8, the four pictures were each placed in one quarter of the screen. The numbers in this figure did not occur in the task; they are just included here to indicate the above-mentioned types:

(1) '*Handen wassen*' ('wash hands') – target (for which both a noun-picture and a verb-picture also occurred on the screen)

(2) '*Handen vouwen*' ('fold' hands = 'praying', in English) – counter target (for which no verb-picture occurred on the screen)

(3) '*Handen*' ('hands') – distractor 1 (picture of the noun of the target and counter target)

(4) '*Wassen*' ('wash') – distractor 2 (picture of the verb of the target)

The location of the four types of pictures varied between the screens, which were prepared in PowerPoint. In the example screen presented in Figure 8, the counter target (number 2 in this example) is presented in the first quadrant (upper left), the target (number 1) in the second (upper right), distractor 1 (number 3) in the third (lower left) and distractor 2 (number 4) in the fourth (lower right). The quadrants were – for the analyses of results –

created, but not shown to the children on the screens, and they were of equal size. The position of the different types of pictures varied over stimuli.

For the auditory input, a native speaker of Dutch, not part of the research team, read out the basic sentence '*Je ziet hier nu …*' ('Now you see here …*')* filled out with either the target noun+verb-combination (condition 1), the counter target (condition 2), or, in the case of fillers, the name of the object (noun) or the verb. All sentences were recorded in Praat (Boersma & Weenink, 2012). To make sure that the natural breaks between the '*nu*' ('now') in the basic sentence and the object mentioned after that were similar in all sentences, natural silences were added to the sentences – when needed – or deleted from the sentences that were recorded in Audacity. The length of the pause between '*nu*' and the next word was set at 49ms, which corresponded to the median of the time that the speaker paused in all the sentences she had pronounced when recording the sentences.

During the task, the children were sitting in front of a computer screen (60 cm distance at most) to which an SMI-eye tracker RED 250 at 250Hz was connected. Before the task began, calibration was performed using a nine-point calibration procedure. A practice session of three trials followed, in which the children saw three screens on which four pictures appeared. These pictures were not test material; in fact, the four pictures on these three screens were not related to each other. As in the real test, however, children did hear the input over loudspeakers with a relevant basic sentence, which was '*Je ziet hier nu … X*' ('Now you see here … X'), with X representing one of the pictures, representing either objects or the simplex or complex verbs that were shown on the screen. Children were told simply to look at the picture that corresponded to what they heard.

After three seconds of fixating, the children would see the next screen, with only a fixation cross in the middle of it, to make sure that children's eyes were not fixated on one corner or quarter of the screen before they went to the next screen, with the next four pictures, and a new spoken sentence. After the examples, the real test began. A similar procedure was applied as for the examples, except that in the real test, the pictures were related to each other in the way described above.

After each screen, the children had to answer a comprehension question, to keep their attention and to check whether they had heard the sentence correctly. These questions were related to the pictures and were asked by the researcher who was present during the experiment. The task included twelve screens with target pictures and ten screens included as fillers. These fillers involved a similar basic sentence as for the target sentences, with

some object included in it. On the filler screens, however, no counter targets or distractor pictures were shown: children saw four unrelated pictures, just as in the examples.

Following the hypotheses, it was expected that children would focus quicker on frequent targets than on less frequent counter targets, and it was expected that literate children would show more single-word processing in their fixations (e.g., from the pictures (1) 'hands' to (2) 'wash' to (3) 'wash hands').

Coding the eye-tracking task

While the sentences with the target items or filler items were played to them, children's eye movements, their gaze direction, and their fixations were recorded. For the analysis of these data, the sentences were divided into segments of words and pauses, as occurring in the spoken sentences, with an additional 0.5 seconds at the end of each sentence before the tier was broken off (in NoteTab Light). The critical multiword sequence occurred after segment or word number 5, after the basic sentence '*Je ziet hier nu*' plus a silence, thus from number 6 onwards. Accordingly, only fixations from segment number 6 and onwards were informative for the analyses. Any looks or gazes before would not be relevant for our research question.

The eye movements were parsed into saccades and fixations by the Eyelink software using the default settings. There were four areas of interest or quadrants that the children could direct their attention to when looking at the screen (see Figure 8). The analysis of the paths of fixation involved assigning the fixations to these areas or quadrants. Children's fixations were automatically assigned to a quadrant in the program Fixation (Cozijn, 2006). These were manually checked and adjusted when necessary (see Appendix 3 for the steps taken during adjustments).

The manual adjustments yielded a data loss of 6.8% of all fixations by all children on the target screens. Of these, 0.7% were lost as a result of indeterminacy as to where the fixations had fallen, and 0.2% were losses as a result of a too limited number of fixations on the screen (<5 fixations). Finally, 5.8% of the data were discarded because of the fact that it was impossible to interpret four or more of the ten target screens that a child had seen (>33%) as a result of unclear fixations, or technical problems.

For the data that remained, the following points were further analysed:

– What was the average first fixation number (after segment 5, see above) when the children looked at the target or countertarget pictures upon hearing either the target (condition 1) or the countertarget (condition 2) pronounced?

The answer to this question would show to what extent frequency of occurrence of targets and countertargets affected children's gazes. Literacy, which was defined on the basis of the grade that children attended (K1 and K2 were pre-literates; E1 and E2 were literates) was also taken as an independent variable in the analysis, to investigate a possible literacy effect.

– Which route did the pre-literate and literate children take to get to the target picture for which they heard audio input? This question was analysed by counting how often the pre-literate and literate children showed one of the following patterns in their eye-movement streams when the target or counter target began to be pronounced in the audio input:
(1)   from distractor 1 (noun picture) to distractor 2 (verb picture) to target or counter target;
(2)   from distractor 1 to target or counter target

It was expected that if children decomposed the multiword target or counter target into smaller parts, then they would show either route (1) or (2) in their fixations.

   Again, a comparison was made between pre-literate and literate children's data, to assess whether there was any influence of literacy on how the audio input was processed.

Since coding and analysis of the eye-tracking data differed considerably from the methods applied to the other tasks, the results will not be compared to those from the other tasks directly. In the discussion of the results (in Chapter 7), the eye-tracking task will therefore be dealt with separately, and only its general outcomes will be integrated into the general conclusions.

5.3.7   Other instruments used for the main study

As mentioned in Section 5.3, the Dutch monolingual children who participated in this study also carried out a non-verbal intelligence test. The test chosen was Raven's Coloured Progressive Matrices (1962). Having these

data made it possible to check whether there was a relation between children's responses in the tasks and their cognitive development (Dellatolas et al., 2003). The test consists of 36 puzzles in which participants are requested to identify a missing element that would complete a certain pattern or shape (cf. Smith, 2012). Of these, 18 puzzles were selected. The reason for not using all puzzles was that this would have been rather time-consuming, and that conducting only half of the puzzles has been argued to give reliable results as well (see the manual provided with the task book, Raven, 1962). It was made sure that in the puzzles that were selected, the correct piece occurred in different places an equal number of times.

In addition, the Dutch children's language proficiency was assessed on the basis of scores they had obtained in the official school test (CITO-test), which they performed by the end of the school year in which the testing sessions had taken place. The raw scores on vocabulary knowledge (*woordenschat*) were collected for children in E1 and E2, as vocabulary has been found to be representative for children's general language proficiency (cf. Veldhuis, Vermeer & Yan, 2012). For the children at kindergarten, scores on the tests for 'kindergarten language' ('*Toets kleutertaal*') were collected.

Due to time limitations, the Turkish children did not participate in the Raven's Coloured Progressive Matrices (1962), and for these children, reading proficiency could not be assessed on the basis of official school test results.

Both the Dutch and the Turkish children were given short semi-structured surveys that had been attached to their response forms. They were filled out by the researcher who conducted the tasks with the children and who asked the questions in an informal way upon meeting a child. The questions regarded children's age and class, their hobbies, whether they liked reading books and watching TV, the frequency with which they watched TV and read books at home (or were read to), and about their musical skills. Notes from children's responses were taken by the researchers, in the form of a five-point-scale for questions dealing with frequency (e.g., 'How often do you watch TV?'), with zero indicating 'never', and four indicating 'on a daily basis'. As for the other questions, the information obtained from the children (such as their age) was simply written down, or yes/no was checked. See Figure 9 for the answer sheet used in the survey.

Name child: _____   recording: _____

- Gender:
- Age:
- Birthday:
- Grade:
- Literate / illiterate:
- Brothers:
- Sisters:
- Music skills: yes / no
- Hobbies:
- Home language:
- TV: language _____ + 0 1 2 3 4 (never – daily)
- Friends language:
- Reading out aloud (home): language _____ + 0 1 2 3 4 (never – daily)
- Reading – self (home): language _____ + 0 1 2 3 4 (never – daily)

Figure 9: Answer sheet for survey on background factors

Information about the background factors was obtained because previous studies (for instance Laghzaoui, 2011) have shown that they may vary a lot and may influence the outcomes in experimental tasks. An additional advantage was that since the children were asked these questions right after meeting the researcher in the first session, children who were about to be tested were already chatting to the researcher – in a more or less informal way – and that way they may have felt more at ease during the experimental tasks that followed and that required more concentration from the child.

The correlation coefficients from the information obtained from the informal surveys and from the school tests and the intelligence test with the number of single-word responses that children provided in the tasks will be discussed in Section 7.1. The reason for correlating the single-word responses to the background factors was that, although single-word responses could not be regarded as the correct responses in the tasks, they were expected to be the kind of responses that literate monolingual Dutch and Turkish children (as well as adults), who have fully[15] developed in their literacy and metalinguistic skills in Dutch or Turkish, would provide.

---

[15] 'Fully' here does not mean that one ever stops learning in the field of literacy, but it refers to literate children and adults who are already fully aware of word-units in language.

5.4   Placement of tasks on the offline-online continuum

Based on the set of criteria outlined in Chapter 3, each experimental task conducted in this study can be regarded as relatively offline or relatively online in nature. Accordingly, the ten questions posed in Chapter 3 were answered for each of the tasks, and points were assigned to the questions, depending on whether they were answered with 'yes' (1 point) or 'no' (0 points). Table 9 shows all tasks and the number of points they received, in the form of a ratio between 0 and 1. This ratio was later used as to define the degree of offline- or onlineness of a task.

Offline characteristics	Sentence segm. task	Last-part repetition task	Dictation task	Click task	Mixed-words task	Eye-tracking task
Conscious of the task being a linguistic task	1	1	1	1	1	0
Elicited material	1	1	0	1	1	1
Non-natural language included in the task	1	1	1	1	1	0
Linguistic input finished at moment of observing participant's behaviour	1	0	1	1	1	0
Time for consideration of answer (i.e., is there no time constraint)	1	1	1	1	1	0
Judgments required	0	0	0	0	0	0
Manipulation required	1	1	1	0	0	0
Linguistic procedure required	1	1	1	1	1	0
Specific goal of task known to participants	1	1	1	0	0	0
Reading/writing (instead of listening/ speaking)	0	0	0	0	0	0
Ratio of points/10 questions	0.8	0.7	0.7	0.6	0.6	0.1

Table 9: Points assigned per task and question

The relative positions of the tasks along the continuum are presented in Figure 10.

OFFLINE ◄━━━━━━┿━━━━━━┿━━━━━━┿━━━━━━━━━━┿━━━► ONLINE

      Sent. Segm. Last-part rep./Dict. Mixed words/Click   Eye tracking
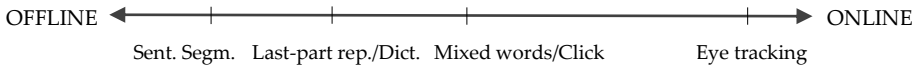
Figure 10: Overview of tasks along the offline-online continuum

Figure 10 shows that the sentence segmentation task was more offline in nature than the last-part repetition and dictation tasks, which were in turn more offline than the mixed-words and click tasks, and that the eye-tracking task was the most online. It should be remarked here that the placement of the tasks on the continuum is not absolute and that the distances between tasks are not indicative for the mutual differences. The continuum merely represents whether a given task is more offline or online than another, and it shows that none of the tasks were purely offline or online.

All tasks possessed a combination of offline and online characteristics. In relation to previous studies in which the offline or online nature of a task was discussed, it can be concluded that the last-part repetition task, which Karmiloff-Smith et al. (1996) regarded as an online task, is located more towards the offline end of the continuum. Children had time to consider their responses, and they knew the purpose of the study, in the sense that they were aware of the fact that they had to manipulate selected sentences in stories. The click task and mixed-words task appeared to be more offline than the eye-tracking task; the latter was not placed at the most extreme end of the continuum, however, as it does not directly provide information regarding neural processing of natural, regular language. There was no task that can be regarded as purely offline either, as children never had to make verbal judgments regarding (meta)linguistic properties of the language used in the tasks.

In Chapter 8, the placement of tasks on this continuum will be discussed once again. In that discussion, the results presented in Chapter 6 and 7 will be considered. This will show whether more offline tasks indeed lead to different types of responses than more online tasks, and whether the continuum that was outlined in Chapter 3 is reflected in the results.

5.5   Final notes

The tasks used in this study are all focused on investigating children's language segmentation and language processing, and they were all intended

to provide information on whether and how children are affected by literacy, typological background and the frequency of occurrence of specific multiword combinations.

Before turning to a quantitative analysis of the results, the following chapter will present a qualitative analysis of the dictation task. This was a rather free task, which could provide information on whether children indeed would come up with the types of multiword combinations that were included in other tasks. Moreover, the dictation task offered a first insight into how pre-literate and literate children differed in how they segmented their own sentences and stories.

Chapter 7 will give a more direct answer to the question of how literacy, typology and the frequency of occurrence of specific multiword combinations influence children's language segmentation and processing. The results discussed will show how the tasks differed in what they required from the children – more offline thinking or more online processing.

An analysis of the dictation task [16]

In this chapter, the results of the dictation task will be discussed. This task was only carried out in Dutch. The participants told the researcher a story in bits, so that she could write it down conveniently. The children were free to talk about whatever they wanted, although the researcher did prompt them by asking about any nice experiences they had had over the weekend or on holiday, just to get them talking (see Chapter 5 for more information on the method).

Because of the free nature of the task, the results cannot be compared to the treatment of individual words and multiword units in the other tasks. However, precisely because of the free nature of the dictation task, it does allow for analyses into a possible literacy effect on segmentation strategies, and on the kinds of multiword sequences that were maintained in the segmentations. Attention was paid to whether any schematic templates recurred in the multiword sequences that children used, and if these were in evidence, their relation to frequency (see also Chapter 5).

Section 6.1 provides a short overview of the participants and the materials that were used. Section 6.2 contains general information on the dictations the children produced, followed by an overview of the way in which the dictations were analysed in Section 6.3 and 6.4. In Section 6.5 and 6.6, literacy and frequency effects are investigated, and in Section 6.7, the findings from the dictation task are interpreted; some general conclusions will be drawn.

---

[16]   This chapter elaborates on Veldhuis and Backus (2012).

6.1 Participants and materials

As mentioned in Chapter 5, a total of 69 Dutch monolingual children carried out the task. Table 10 provides an overview of the number of children per grade.

	Grade	Number of participants per grade	
		N	%
Pre-literate	K1	19	27.5
	K2	16	23.2
Literate	E1	15	21.7
	E2	19	27.5
Total		69	100.0

Table 10: Participants in the dictation task per grade

The task was based on Chaney (1989) and Chau (1997) (see Chapter 5). No specific materials needed to be prepared. Instead, the children were asked to come up with their own output, and tell the researcher a story about some nice experience they had had. The children were asked to tell this story to the researcher bit by bit ('*stukje voor stukje*') so that she could write it down – as if they were dictating the story to the researcher. Children were free to choose their own topic. One of the implications was that children's stories varied in length and in the number of segmentations that were made in the telling. In Section 6.2, a general overview is given of the results.

6.2 General remarks about the dictations

Before going into details, some general remarks are in order about the dictations and the language used. These form the background to which the results need to be interpreted.

First, it is noteworthy that the pre-literate children in the first year of kindergarten were not always able to dictate exactly what they wanted to say. While during natural conversation with the researcher most of these children had no problem narrating their chosen experience, they often seemed to get confused about their choice of words when carrying out the dictation task in which they were asked to pause after every 'bit' of the story. An example of this is provided in (1), hyphens marking the segmentations made.

(1)   (30, Mo., K2)

  Ik – ging – v – oetballen – vrienden.

  I – went – p – lay football – friends.

  'I went playing football friends.'

This example, which was provided by a boy in K2, is ungrammatical, as the preposition 'with', indicating with whom the boy went playing football, is missing. Before starting the dictation task, the child had told the researcher about his weekend, including how he had played football with friends. In his dictation, however, he experienced difficulties saying this with a grammatical sentence, using only the content words. Having to segment the sentences apparently imposed a cognitive load heavy enough to cause mistakes sometimes.

Ungrammatical phrases occurred in ten of the dictations of the children in K1 and in six of those by children in K2. For children in E1 and E2, telling a story in segments did not appear to be a problem: they did not have any ungrammatical sentences in their dictations.

Elliptical constructions, however, in which it was usually the verb that was omitted, occurred in all grades. The 35 pre-literate children in K1 and K2 provided 14 ellipses, and the 24 children in E1 and E2 13. The occurrence of ellipses may be related to the task: as children were asked to talk about some nice experience, some felt encouraged to just sum up the things they had done, as in the following example of a child in K2, who talked about the stupidity of lambs, unable as they are to draw (compared to herself).

(2)   (26, Ei., K2)

  Een schaapje – gaat drinken
  Het – is – een – lammetje – hij – geen – tekening maken.

  A sheep – goes drinking
  It – is – a – lamb – he – no – picture drawing.

  'A sheep goes drinking. It is a lamb, he no picture drawing.'

There were also two children in K1 who provided what could be regarded as exclamatory phrases. In one case, a boy commented on how seawater tasted (he was telling a story in which he had accidentally swallowed the water) by saying '*vie-ie-s!*' ('disgusting!'). In the other case, another boy exclaimed with enthusiasm that his holidays had been fun: '*leu – ke – va – kan – tie!*' ('nice

holidays!'). Apart from these two cases, however, children did not comment on their own stories.

Another notable feature of the dictations as provided by all children in all grades was that multiword sequences of more than four words, which were mostly complete clauses, occurred mostly at the end of sentences (e.g., *'waarin je kan verdwalen'*, 'in which you can get lost'). This suggests a task-related wrap-up effect. It may also be a consequence of the cognitive load required when children have to remember their story and the phrases they wanted to produce while they are dictating, as this is after all a rather unnatural type of language production. The phenomenon could also be related to the information structure of relatively short phrases in Dutch: in such phrases, the meaningful information is usually given at the end of sentences. It may be the case that children construed this information as single units. Since it is impossible to know which of these reasons applies, they are all counted as multiword sequences in the quantitative analysis (see Section 6.3).

The type of language used in the dictations was fairly basic, for all grades, as derived from the length of utterances and the grammatical constructions that were used. As for the latter, there were only ten occasions in which children used complex sentences in their dictation. Among these were two conditional clauses such as '*Het zou vet zijn als ik een doppel maak in de straaljager en een looping*' ('It would be cool if I made a dopple in the jet, and a looping'; 58, Ka.). Five other complex sentences contained relative clauses, such as '*We waren in een soort huis waarin je kan verdwalen en de muren waren zacht*' ('We were in a kind of house, in which you can get lost, and the walls were soft'; 37, Yc.). Of the other three complex sentences, two contained adverbial clauses with time references such as '*Toen ik jarig was ging ik met mijn vrienden naar KidZcity*' ('When it was my birthday, I went to KidZcity with my friends'; 37, Yc.), and one with the subordinating conjunction '*omdat*' ('because'; 10, Je.).

All other sentences were either simple, or contained the coordinating conjunctions 'and' or 'and then'. The use of these conjunctions may have been related to the task: children's narrations generally contain a high number of phrases and sentences starting with 'and then' (cf. Litjens, 2001; Litjens et al., 2001).

The length of the sentences, identified on the basis of pauses or silent breaks, confirmed the simple nature of the language used. The stories consisted of one to six sentences.

The average number of sentences dictated by the children in their stories varied per grade, as did the number of words, which numbered between 7 and 42. The average numbers of units that children used as segments also differed. Some children only provided single-word or sub-word units in their dictations whereas others used multiword sequences, and hence produced a much lower total number of segments.

In Table 11, the average number of sentences and words used in children´s dictations per grade, including information on the standard deviations, is given, as well as information on the average number of units that children used in their dictations.

	Sentences used per dictation		Conventional words used per dictation		Total number of units segmented	
	Mean	SD	Mean	SD	Mean	SD
K1 (N=19)	3.47	(1.17)	21.00	(6.79)	15.47	(7.55)
K2 (N=16)	3.06	(1.12)	17.88	(8.34)	16.31	(8.94)
E1 (N=15)	3.07	(0.88)	20.93	(4.11)	21.80	(6.41)
E2 (N=19)	3.21	(1.08)	25.32	(8.22)	21.05	(6.12)
Total	3.22	(1.07)	21.45	(7.50)	18.60	(7.68)

Table 11: Sentences and words used per grade

The differences in the number of sentences provided were not significant (p>.05), but the differences in the number of words were, $F(1,65)=3.2$, p=.03, η_P^2=.13. Turkey post-hoc analyses revealed a significant difference in the number of words used by the children in K2 and E2 (p=.02). The variation in length of the dictations resulted from the fact that some children were more enthusiastic storytellers than others, and that the researcher did not want to force children to say more if they really did not want to, so as to avoid bias in the results because of disinterest or lack of concentration.

The average number of units that children segmented in their stories also differed per grade. The children in E1 and E2 segmented more units than the children in K1 and K2. This difference appeared to be significant, $F(1,65)=3.4$, p=.02, η_P^2=.13. A higher number of units suggests that children segmented speech into words or even smaller units, while a lower number suggests that they segmented their speech in units that were larger than words. In Section 6.3, this will be discussed in more detail.

6.3   Analysing the dictations

As a first step in the analysis, the numbers of times that specific kinds of segmentations occurred per child were counted. The results were then collapsed for all children from the same grade. The categories were:

(a)   Single-word units: e.g., '*ik – was – jarig*' ('it was my birthday') consists of three single-word units.

(b)   Sub-word units: e.g., there are six sub-word units, as indicated by bold script, in '*ik – **bot** – **ste** – **te** – **gen** – de – **pun** – **t** – van – de – tafel*' ('I bumped into (lit: against) the corner of the table'), namely syllabic units in the words '*botste*' ('hit') and '*tegen*' ('against'), and phonemic segmentation of the 't' in the word '*punt*' ('corner'). '*pun*' was initially categorized as 'other sub-word unit', but for the analyses in this chapter, phonemic, syllabic and other sub-word units were taken together.

(c)   Multiword sequences: These included various multiword sequences that were not further segmented into smaller parts by the children. They could consist of any sequence of multiple words, from complete sentences or clauses, to a sequence of just two words, e.g., '*een deken*' ('a cover') in the following dictation, made by a boy in K2:

> *Dan – vind – ik – **een deken** – en – een – kussen.*
> Then – find – I – a cover – and – a – pillow.
> 'Then I find a cover and a pillow.'

Literate children were expected to segment their language along word boundaries, whereas pre-literate children would not yet be able to apply such knowledge.

    For the computation of the number of single-word units and sub-word units, monosyllabic words that were segmented as single units by the children were assigned to category (a). Thus, in a dictation like (3) below, provided by a boy from K1, the words printed in bold were assigned to category (a), even if it could of course not be known for sure whether this boy made his segmentations along word boundaries or syllabic boundaries.

(3)   (8, Fi., K1)

> **Ik – heb – de** – *blauwe – geskied.*
> **Ik – heel – vaak** – *ge – zwo – mmen.*
> **En – ik – heb – ook – nog – veel – keer** – *friet – jes –* **en** *– pi – zza – ge – ge – ten.*
> *Leu – ke – va – kan – tie!*

I – have – the – blue – skied.
I – many – times – swi – mming.
And – I – have – also – still – many – times – fries – and – pi – zza – ea – ten.
Nice – ho – li – days!

'I have been skiing the blue one.
I (have) swum many times.
And I have also eaten fries and pizza many times.
Nice holidays!'

Assigning the monosyllabic words in the dictation to the single-word category may have led to bias. However, assigning them to the sub-word category often made very little sense, as children generally maintained multisyllabic words as single units. Not incorporating the ambiguous cases would have been possible as well, but this would have led to a great loss of data, especially among the younger children, who provided simpler language than the older children. This showed, amongst other things, in the use of more monosyllabic words. Therefore, it was decided to include monosyllabic words that were segmented as single units in category (a).

Category (c) comprises both multiword sequences that could be regarded as 'hypothesized multiword units' (see Section 5.3.1), and other sequences. This category was included to check whether pre-literate children provide more segments that are longer than single words than literate children do, and to see whether there were specific schematic templates to be recognized in the multiword segments. Pre-literate children were especially expected to maintain multiword units in their dictations as they have been argued not to pay attention to word boundaries yet (see Chapter 2), while literate children would more often stick to words, because of their literacy education.

6.4   General results

To get a general idea of the kinds of segmentations that children produced, we first look at how often a specific category occurred per grade. Table 12 contains the findings, in percentages.

Grade	Single-word units		Sub-word units		Multiword sequences	
	Mean	SD	Mean	SD	Mean	SD
K1 (N=19)	38.93	(22.93)	21.11	(26.02)	39.96	(33.81)
K2 (N=16)	35.15	(27.00)	31.10	(31.05)	33.74	(39.85)
E1 (N=15)	75.70	(17.18)	18.27	(18.94)	6.03	(6.75)
E2 (N=19)	74.40	(19.56)	7.22	(10.06)	18.38	(21.78)
Total	55.81	(28.73)	19.00	(23.81)	25.20	(30.98)

Table 12: Summary of units provided by children per grade and category, in % of total units segmented

The preferred category taking all children together was the single word (55.81%). Sub-word elements and multiword sequences were segmented at approximately half its rate. Differentiated by grade, however, the picture is much more varied. Words were the clearly preferred unit for segmentation in grades E1 and E2. In these grades, around 75% of segments were words. This seems to indicate a literacy effect: children in E1 and E2 were well trained in recognizing written words.

The pre-literate children in grade K1 and K2 used all three categories often. The children in K1 most often provided multiword sequences. They only provided units smaller than words in 21% of their segmentations. This may be related to the fact that meaningful units in Dutch usually consist of either single (content) words, or of multiword units (which usually include combinations of content and function words), but not so often of sub-word units, i.e., Dutch has relatively few bound morphemes. The children in K2 did on average not really show a preferred segmentation unit. This may be related to the fact that the curriculum of these children at school can also be regarded as a period of transition: in K2 in the Netherlands, children often practice with units that are smaller than words (notably syllables), but after this, they learn to analyse and synthesize also smaller units of language (such as phonemes) into single words. The results presented in Table 12 reflect this.

The results indicate a literacy effect. There was also more consistency in how the children in E1 and E2 segmented their stories than among the younger, pre-literate children. The data for these older children moreover indicate a developmental pattern. The specifics of this pattern may be related to educational practice: In E1 language classes, children still often focus on segmenting or analysing and fusing or synthesizing parts or sounds of words ('*hakken en plakken*', auditive synthesis and analysis; cf. Huizenga,

2010; Huizenga & Robbe, 2013). This focus on the parts of language that are smaller than words might have enhanced the recognition of sub-word elements. For the children in E2, the language program is mainly focused on words, phrases and sentences (writing longer sentences and short texts). In class, they do not focus on sub-word units anymore because decoding has now been automatized.

From a Repeated Measures analysis with the average percentages of word, sub-word and multiword segmentations per group as dependent factors, and groups as independent factor, it derived that for all these response types, group significantly affected the results: $F_{words}(3,65)=17.1$, $p=.000$, $\eta_P^2=.44$; $F_{subwords}(3,65)=3.3$, $p=.02$, $\eta_P^2=.13$; $F_{multiword}(3,65)=4.7$, $p=.005$, $\eta_P^2=.18$. Post-hoc analyses revealed that for single words, this difference was caused by the difference in the number of times that children in K1 and K2 versus those in E1 and E2 provided single-word units as segments ($p=.000$).

For the sub-word units the difference was caused by the large difference between K2 and E2 ($p=.02$); and for the multiword responses by the large difference between K1 and E1 ($p=.01$). Between other groups, no significant differences were found ($p>.05$). As argued above, the high incidence of sub-word units in K2 and its demise in E2 is probably in part explainable as an effect of school practice. For the significant difference for multiword units between children in K1 and E1, a similar argument may hold: children in K1 provided many more multiword responses because their education has not biased them towards word-based segmentation yet.

Interestingly, the total number of multiword segments in E2 is relatively high, at 18.38% of all segments provided by this group. Even though this may have to do, as suggested before, with the practice in school to focus on phrases and sentences more, this higher number was also influenced by the dictation of one girl, who provided 12 multiword segmentations in her story, in which she summed up to which attractions she had been in a fun park. Her dictation is presented in (4):

(4)   (55, Do., E2)

In de Python – en Joris en de Draak – en – ehm – de – Hollandse – de Vliegende Hollander – en ook nog de Droomvlucht.
Daar zijn – elfjes en trollen.
De Slakkenbaan – de paardentram – en de bobsleebaan – en de – Piranha – en de – het schommelschip.

In the Python – and Joris and the Dragon – and – er – the – Dutchman – the Flying Dutchman – and also still the Dreamflight.
There are – fairies and trolls.

The Snailroad – the horse tram – and the bobsleigh track − and the – Piranha – and the – the swinging ship.

As can be seen from this dictation (4), this girl provided a single segment for almost every attraction that she mentioned. In listing objects such as diplomas and family members, two other children also provided single segments consisting of determiners+nouns or the word 'and' plus a determiner+noun for every item they mentioned. The determiner+noun-template is a construction that was mentioned in previous studies as well as a construction that people seem to parse and produce holistically (see Chapter 2).

In order to further investigate to what extent multiword segments are of a limited number of schematic templates, a count was made of how often the following multiword constructions occurred:

(a) Multiword sequences that consisted of a determiner(+adjective)+noun: e.g., '*een deken*' ('a cover') in the following dictation, made by a boy in K2:

*Dan – vind – ik – **een deken** – en – een – kussen.*
Then – find − I – a cover – and – a – pillow.
'Then I find a cover and a pillow.'

Determiner+noun sequences that included adjectives were also included in this category.

(b) Multiword sequences that consisted of a prepositional phrase: e.g., '***in de trein***' ('in the train') and '***in de boot***' ('in the boat') in the following dictation, made by a girl from K1:

*We gingen – **in de trein**. – Toen – gingen we – **in de boot**.*
We went – in the train. – Then – we went – in the boat.
'We went by train and then we went by boat.'

(c) Multiword sequences that consisted of the subject + verb: e.g., '*we gingen*', ('we went'), '*ik kan*' ('I can'), and '*het deed*' ('it did'). Forms of the subject and verb with inversion were also assigned to this category.

(d) Multiword sequences that consisted of object + verb, possibly including a determiner: e.g., segmentation of the following multiword sequences as wholes: '*frietjes eten*' ('eat fries'), '*hut bouwen*' ('build a treehouse'), '*tekening maken*' ('make a drawing').

(e) Multiword sequences that were predicates (pred.): e.g., '*heeft een zwembad gewonnen*' ('has won a swimming pool') as segmented by a girl in the sentence: '*Abel − mijn broer – heeft een zwembad gewonnen*' ('Abel, my brother, has won a swimming pool').

(f) Multiword sequences that consisted of complete sentences or clauses (sent./clauses): e.g., modifying clauses such as '*waarin je kan verdwalen*' ('in which you can get lost'), or '*Ik ging met m'n zwempak zwemmen*' ('I went swimming in my swimming suit').

(g) Multiword sequences that consisted of ['en'+X]: These included units like '*en daar*' ('and there') and '*en toen*' ('and then'), which mostly consisted of discourse markers.

(h) Other multiword sequences: these included segments such as '*als je*' ('if you') and '*zwempak aan*' ('swimming suit on') and others that could not be assigned to one of the categories mentioned above.

The categorization of responses is indicated more precisely in Appendix 10. In Table 13, an overview is given of the average number of times that the children in the various grades retained these kinds of multiword (mw) sequences.

Grade	DET (+ADJ)+ N	SUBJ+ verb	PP	(DET+) OBJ+V	Pred.	Sent./ clauses	[EN_X]	Other mw units	Total mw units
K1 (N=19)	0.53 (0.84)	0.79 (0.71)	0.68 (0.95)	0.32 (0.75)	0.32 (0.67)	0.58 (0.96)	0.53 (0.70)	0.32 (0.58)	4.05 (2.76)
K2 (N=16)	0.44 (0.63)	0.63 (0.89)	0.37 (0.62)	0.25 (0.45)	0.19 (0.40)	0.44 (0.63)	0.31 (0.79)	0.06 (0.25)	2.69 (2.58)
E1 (N=15)	0.33 (0.62)	0.07 (0.26)	0.13 (0.35)	0.07 (0.26)	0.00 (0.00)	0.13 (0.35)	0.27 (0.46)	0.07 (0.26)	1.07 (1.22)
E2 (N=19)	0.89 (1.41)	0.50 (0.99)	0.61 (0.85)	0.33 (0.67)	0.00 (0.00)	0.06 (0.24)	0.72 (1.23)	0.33 (1.03)	3.44 (3.60)
Total	0.56 (0.95)	0.51 (0.80)	0.47 (0.76)	0.25 (0.58)	0.13 (0.42)	0.31 (0.65)	0.47 (0.86)	0.21 (0.64)	2.91 (2.89)

Table 13: Mean number (and SD) of different types of multiword sequences per grade

Table 13 shows that the average of 2.91 multiword segments that each child produced, combinations of determiners with nouns, possibly including adjectives, occurred most often, followed by sequences of subjects plus

verbs. The noun phrases were common especially in the data from children in E2, while the highest numbers of [SUBJ V] segmentations were from the children in K1 and K2. Predicates occurred least often; in fact, they did not occur at all in the responses by the children in E1 and E2, suggesting that for them, predicates were not a natural unit.

However, the relatively high standard deviations suggest a lot of individual variation. There were children who did not provide any multiword sequences at all, as well as a child who did not produce any segments that were smaller. This makes it difficult to arrive at any kind of generalization.

To overcome this difficulty, the raw frequencies represented in Table 13 were re-calculated as percentages of the total number of times that children in different groups had provided multiword sequences. This calculation is presented in Table 14.

Grade	DET (+ADJ)+N	PP	SUBJ+ verb	(DET+) OBJ+V	Pred.	Sent./ clauses	[EN_X]	Other mw units
K1 (N=19)	11.37 (15.85)	16.67 (18.21)	22.83 (24.31)	7.71 (15.19)	5.95 (10.88)	13.54 (20.77)	10.76 (12.33)	11.16 (25.57)
K2 (N=16)	15.08 (17.70)	16.27 (28.87)	22.87 (29.48)	11.11 (17.88)	6.75 (13.07)	19.25 (28.97)	6.60 (12.74)	2.08 (7.22)
E1 (N=15)	31.48 (42.85)	8.33 (17.68)	2.78 (8.33)	5.56 (16.67)	0.00 (0.00)	9.26 (18.84)	31.48 (42.85)	11.11 (33.33)
E2 (N=19)	24.32 (30.25)	18.63 (27.83)	10.13 (14.70)	9.70 (17.38)	0.00 (0.00)	2.56 (9.24)	23.55 (28.94)	11.11 (29.40)
Total	19.25 (27.02)	15.58 (23.28)	15.93 (22.54)	8.66 (16.30)	3.52 (9.20)	11.29 (20.91)	16.82 (26.02)	8.96 (24.90)

Table 14: Means number of different types of multiword answers in % (and SD) of the total number of multiword units per grade

Table 14 confirms that the most frequent type of multiword sequence was the sequence in which children combined determiners with (adjectives and) nouns. Predicates occurred least often. There was variation between grades, and the variation between children per grade, as represented by the standard deviations, was large as well.

The average percentages shown in Table 14 do show, however, that the pre-literate children show a preference for complete sentences or clauses in the multiword sequences that they produced, or of the proto/clausal combination of subjects with verbs, whereas the children in E1 and E2 provided more [EN_X]-units and the quintessential phrasal unit [DET (ADJ) N].

Then again, it should be noted that the categories used for the assignment of responses were mutually exclusive, and that the percentages in Table 14 are for the different categories dependent on each other: if an answer was assigned to one category, it could not be assigned to another. Accordingly, the results for one category may be slightly inflated, while the occurrence of another one may be underestimated. Still, the high standard deviations mean that the generalizations cannot be taken as indicative for every single child.

Importantly, especially those multiword segments assigned to the categories 'sentences or clauses' and 'predicates', were not units that occurred frequently in the corpus (see Appendix 10). If these two categories are removed, the picture of the relative occurrence of the different constructional types in children's dictations changes, see Table 15.

Grade	DET (+ADJ)+N	PP	SUBJ+ verb	(DET+) OBJ+V	[EN_X]	Other mw units
K1 (N=16)	13.61 (19.29)	20.45 (21.61)	27.34 (25.89)	8.65 (17.40)	17.89 (26.16)	12.04 (25.85)
K2 (N=11)	26.06 (31.86)	19.55 (30.68)	28.64 (30.97)	15.15 (22.92)	8.33 (16.24)	2.27 (7.54)
E1 (N=9)	33.33 (43.30)	13.89 (33.33)	2.78 (8.33)	5.56 (16.67)	33.33 (43.30)	11.11 (33.33)
E2 (N=14)	22.58 (29.79)	18.49 (28.30)	16.56 (27.86)	9.01 (16.90)	23.06 (29.36)	10.32 (28.40)
Total	22.41 (30.21)	18.52 (27.10)	20.19 (26.66)	9.62 (18.20)	20.01 (29.42)	9.24 (25.02)

Table 15: Relative occurrence of constructional types in % (and SD), excluding multiword sequences from the categories 'sentences/clauses' and 'predicates', as these occurred seldomly in the corpus (see Appendix 10)

As can be seen from Table 15, when the multiword responses assigned to the categories 'sentences/clauses' and 'predicates' were excluded, the responses that were assigned to the other categories occurred between 9.24% to 22.41% of the responses. There were no extremes in these percentages, meaning that there was not one category of multiword sequence that children treated as one unit in a similar way. The pre-literate children in K1 and K2 however most often dictated subject+verb-combinations as single units.

This may be related to the single meaning of such units: usually, such a phrase refers to one activity performed by a specific person – this is easily regarded as one single conceptual unit. For the literate children, the highest

numbers are achieved for the category [EN_X]. This is likely due to their experience with (more complex) storytelling, in which they combine events in chronological order. Intensive use of the phrase 'and then' is a natural ingredient in telling such stories (cf. Litjens, 2001; Litjens et al., 2001). Determiner+noun-combinations also occurred often among the children in E1 and E2, which may also be related to the single concept that such combinations describe.

To what extent these multiword units can indeed be related to their frequency of occurrence will be discussed in Section 6.6, First, Section 6.5 below will provide a further analysis of the effects of literacy.

6.5 Investigating the literacy effect

In order to analyse a possible effect from literacy, due to which literate children would be expected to provide more single-word units in their segmentations than pre-literate children, the number of single-word units that the children in the different grades provided was further examined. Table 11 already suggested that literacy affected children's strategies, as literate children used many more single-word units than pre-literate children.

    For a further investigation, the ratio of single-word segments as divided by the total number of words that occurred in the stories was calculated. This ratio was called the 'single-words ratio'. The interpretation of this ratio is as follows: If the ratio approximates 1, then it means that children segmented almost as many single-word units in their dictation as there are in the story. A lower ratio indicates that relatively few single words were segmented. In Table 16 the average number of single words children segmented in their stories is given, together with the mean number of conventional words that they could have segmented and the single-words ratio per group.

Grade	Number of single words segmented per story		Total number of conventional words in story		Ratio single-word units/ conventional words in story
	Mean	SD	Mean	SD	
K1 (N=19)	7.05	(5.83)	21.00	(6.79)	0.34
K2 (N=16)	6.94	(5.95)	17.88	(8.34)	0.38
E1 (N=15)	15.93	(4.15)	20.93	(4.11)	0.77
E2 (N=19)	16.05	(6.48)	25.32	(8.22)	0.67
Total	11.43	(7.21)	21.45	(7.50)	0.53

Table 16: Average number of single-word segmentations and average total number of conventional words in the story, and ratio of single-word segmentations/ conventional words in story

As mentioned above, the literate children provided single-word units most often (15.9-16.1 single-word segmentations in their stories, which corresponded to 74-76% of all their segmentations made as shown in Table 12). The ratio also suggest a literacy effect: for the literate children, the single-word ratio is between 0.67 and 0.77, much closer to 1 than the ratios of the pre-literate children, which are between 0.34 and 0.38. Surprisingly, the single-word ratio is higher for children in E1 than in E2. This may seem to go against the expectation that older literate children would provide relatively more single words in their dictations than younger literate children, but it may be related to practices in class: there is a focus on segmenting words into smaller parts and synthesizing those smaller parts into words again in E1. This focus on single words may explain the higher single-word ratio for E1 children than for E2 children.

From Univariate analyses with the single-words ratios per grade as the dependent factor, and grade as the independent factor, it derived that there was indeed a main effect for grade, $F(1,65)=12.5$, $p=.000$, $\eta_P^2=.37$. Post-hoc analyses showed no significant differences between children in E1 and E2, nor between K1 and K2 children ($p>.05$). A Univariate analysis with the single-words ratios as the dependent factor, literacy as the independent factor and age as a covariate, confirmed a main effect of literacy, $F(1,66)=6.85$, $p=.01$, $\eta_P^2=.09$, but not of age ($p>.05$)

6.6   Children's segmentations in the dictation task and their relation to frequency

In order to investigate to what extent the frequency of occurrence of the combinations that were preserved as multiword segments in the data may have determined the segmentations, all multiword segments were entered as search terms between quotation marks in the Dutch pages of Google (in October-December, 2012). The individual words in these segments were similarly entered in Google. If a word could be spelled in different ways, as was for instance the case for numbers (e.g., '*negen puppy's*', 'nine puppies') which can be written as words or as numbers ('*negen*', or '9'), both spellings were entered in Google. The frequency results were then added up. The frequency values thus obtained (included in Appendix 10) were used for further analyses. The list of all multiword segments that the children provided, with their raw frequencies, is given in Appendix 10. The relative frequency information as provided by EMI values, which indicates the strength of boundedness within a multiword sequence (see Chapter 5), was not computed for these segments, as they varied to a large extent in structure and as the raw frequency information already showed clearly that the frequencies of occurrence varied highly in range. Calculating EMI values with such big differences between the various multiword segments would in this case not lead to more, or more reliable, information.

For the multiword segments in the K1 data, the range of frequencies was huge: from 0 to 70,700,000. The two most frequent ones are 'lexical bundles', i.e., co-occurring function words ('*en ik*', 'and I', and '*en met*', 'and with'). As to be expected, the longest multiword segments occurred least often in the Google-corpus. These longer phrases, such as '*heb de hangmat opgehangen met papa*' and '*omdat 'ie dan zijn moeder terug kan vinden*' are very specific in meaning, and therefore unlikely to have been used much by others, if at all. Clearly, it would be inefficient if units for such specific meanings were stored.

Six of the ten most frequent multiword segments are prepositional locative phrases. This suggests that such phrases are natural candidates for early multiword units. Other frequent multiword segments are '*en toen*' ('and then') and '*gingen we*' ('went we') (this latter phrase has 'inverted' word order, which occurs in Dutch after a sentence-initial adverb: a very common word order pattern). Unit status for these phrases also seems to be confirmed by the fact that there several different children used these same multiword segments.

For the children in K2, the range in frequencies is also huge: from 52,900,000 for the most often occurring multiword segment ('*en dan*', 'and then') to 0. The most frequent multiword segments were again mostly lexical bundles consisting of function words. Only the tenth segment in the list includes a noun ('my mother'); the first nine were phrases like 'I want', 'I may', 'I find', and 'and then'.

In fact, for 'and then', two forms exist in Dutch: '*en dan*', which refers to a time in the future, and '*en toen*', which refers to a time in the past. While in Google '*en dan*' occurs much more often than '*en toen*', the children produce '*en toen*' more often, probably because they were asked to tell a story about an event or activity that had taken place in the past in the dictation task. Both forms, however, were often segmented as single units by all children.

Interestingly, the data for prepositional phrases yielded less natural candidates for segmentation than among the children in K1 and K2. Only number 17 on the list of multiword segments for K2 is such a phrase ('*in m'n haar*', 'in my hair'). The score for another prepositional phrase that was produced by one child in K2 was contaminated by unrelated occurrences. This concerns the phrase '*naar de bos*', ('to the forest'), in which an ungrammatical determiner is used. Instead of the neuter determiner '*het*', the common gender counterpart '*de*' is used. That this multiword segment still occurred quite often in Google is probably due to the fact that '*naar de bos*' appeared as part of larger phrases, such as '*naar de bos en lommerweg*', in which '*bos en lommerweg*' is a street name, and because of the fact that '*bos*' also occurs in Google as an abbreviation, as BOS.

As with the data of the younger children, the longest forms did not give any hits. The segments '*ik heb parfum spuit*' ('I have perfume spray') and '*zag er blauw met wit uit*' ('looked blue with white') for instance, were not found in Google. These segments were presumably composed from smaller parts, rather than stored as complete units.

Turning to the literate children, it is clear that grade E1 children used fewer multiword segments than the children in kindergarten. Their multiword segments did not vary as much either. The most frequently occurring segment, produced by two children in E1, consisted of the collocation '*opa en oma*' ('granddad and granny'), which occurred 17,500,000 times in Google. Two different children used it as a segment in their dictations, not surprisingly since the children were asked to talk about a nice experience during their holidays or weekends, and visiting grandparents is what many families do on the weekend.

Once more, the longer segments did not occur at all in Google. These consisted mostly of complete sentences, clauses or predicates (e.g., '*gingen we*

in een bunker zitten' and '*heb de hangmat opgehangen met papa*', resp. 'we went to sit in a shelter' and 'have put up the hammock with daddy').

Finally, the multiword segments made by the children in grade E2 varied in frequency from 1 to 360,000,000. The most frequent segment was again a lexical bundle, namely '*met een*' ('with a'). This also holds for most other segments. Only the segments '*in de hand*' ('in the hand'), '*op vakantie*' ('on holidays'), and '*één dag*' ('one day') contained content words.

The segment '*en X*' ('and X'), in which X stands for a variety of words and multiword sequences, also occurred several times in the E2 data. But even if this might indicate that the partially schematic construction ['and' X] is a natural candidate for segmentation, its frequency was highly dependent upon the frequency of occurrence of the full expression. For instance, '*en tv kijken*' ('and watching TV') and '*en ook nog de Droomvlucht*' ('and yet also the Droomvlucht' (which is the name of an attraction in a Dutch fun park), had completely different frequencies of occurrence. Both did, however, occur in the corpus.

The findings indicate that there is a wide range in the frequencies of occurrence of the multiword segments that the children kept together in the dictation task, and that it is hard to draw conclusions on the basis of this task with respect to frequency. Then again, the dictation task does show that in general, whether children were pre-literate or literate, their multiword segments included both long ones produced as novel sequences in the course of the storytelling – sequences with very low or zero occurrence in Google –, as well as segments with a broad meaning that did occur relatively often. This especially holds for lexical bundles.

6.7   Discussion and conclusions

The question what kinds of units pre-literate and literate children distinguish in their language production was not so easy to answer on the basis of the dictation task. Even though several schematic templates could be detected in the segments that children came up with, such as prepositional phrases, object+verb-combinations, and determiner+adjective+noun-combinations, the large variety in kinds and lengths of segments, did not allow for the identification of clear patterns.

Applying another version of this dictation task, for instance asking all children to dictate the same story (see Chaney, 1989), might overcome this problem, and would make it easier to compare data, but on the other hand

children would be affected by what they knew about the story or song they were asked to dictate, for example about rhythm or stress patterns.

Nevertheless, the literacy hypothesis could be confirmed: pre-literate children in K1 and K2 provided significantly fewer single-word units in their dictations than literate children in E1 and in E2. The acquisition of literacy, which makes children become aware of word boundaries in language that are not always clear in connected speech, was reflected in the kinds of responses the children provided, as the literate children gave more single-word responses than the pre-literate children.

The fact that literate children in E1 and the pre-literate children in K1 and K2 also often provided sub-word segments, also suggests an influence from schooling: although it was expected that young, pre-literate children would mostly provide meaningful multiword segments in their dictations (see Chapter 2), the older pre-literate children in K2 also often provided sub-word segments in their dictations, corresponding mostly to syllables or phonemes. This can be related to classroom practices in K2, in which many pre-reading practices involve segmenting spoken words and sentences, for example tapping syllables, and to beginning reading and writing exercises in E1, in which children often have to segment spoken words in phonemes, and synthesize those phonemes into words. The fact that K2 children thus often provided sub-word segments seems to reflect what they know from school. Children in K1 produced fewer sub-word segmentations, and more multiword segmentations, which corresponds to the idea that these children attached great value to meaningful units, and were not yet as well trained as their peers in K2 in segmenting language into smaller parts.

The children in E1 still seemed to show traces of their earlier training: most of the segments these children made corresponded to single words, but 18% corresponded to sub-word units. This result suggests a competition between the effects of their previous training, and from the newly learned focus on combining sub-word units into larger parts, which correspond to words. The fact that the children in this task were tested after Christmas, thus in the second term, means that the children in E1 had come across words often: teachers use single words in dictations in grade 1, and children in E1 are supposed to be able to read and use simple children's books and other teaching materials independently. This way, the mixture of segments that consisted of single words and the sub-word segments the children in E1 produced, seems to reflect their schooling.

The children in E2 usually provided single words in their dictations. However, in contrast to the children in E1, the second most frequent type of segments made by them was the multiword sequence. Many of these

seemed ad-hoc combinations of function words, so-called lexical bundles. There is no single schematic constructional template that could be applied to these multiword segments. This was not only typical for the children in E2: there are no obvious templates among the many multiword segments in any of the data.

In response to the question how pre-literate children, who are unlikely to be biased towards words (cf. Kurvers & Uri, 2006; Veldhuis, Li & Kurvers, 2010), segment sentences, and whether multiword segments are motivated by high frequency, it can be concluded that frequency did not seem to play an important role for children's multiword segmentation in the dictation task: children did produce multiword sequences, but these included cases of low frequency. On the other hand, the high frequency of multiword sequences such as '*en toen*' does influence their use as segments in dictation.

Nevertheless, unambiguous support for the view that language users process frequent multiword sequences as wholes could not be found (cf. Wray, 2008), as several phrases that could be expected to be units on the basis of their corpus frequency were broken up into words during dictation. There may be several reasons for this.

First, the mental lexicon of pre-literate children probably does not correspond well to the Google corpus we worked with for this study, because the inventory of units used in input to young children may differ quite a bit from what is recorded in Internet language use.

Second, the dictation task is an offline task, in which children have time to think about their segmentation styles. The metalinguistic knowledge the task therefore relies on is at odds with the unconscious processing that online tasks tap into, and that most of the entrenchment literature is based on (see Section 2.2). This casts doubt on the ability of other offline tasks, such as grammaticality judgments, to provide reliable evidence about underlying linguistic competence.

In order to investigate this further, and to see whether the findings from this task would correspond to what can be found in more controlled language segmentation tasks, two other relatively offline and three relatively online tasks were conducted with the same children as reported on above (see Chapter 5). The results from these tasks will be discussed in the next chapter.

A quantitative analysis of the tasks: Literacy, typology and frequency effects

In this chapter, the results will be discussed in relation to the hypothesized effects from literacy and frequency. It was expected that literate children would provide more word segmentations in the more offline tasks and process language more along word boundaries in the more online tasks than pre-literate children. Similarly, it was expected that multiword sequences with a high relative frequency of co-occurrence would be processed holistically. Finally, we investigated whether there was an interaction between literacy and frequency.

Two tasks were carried out by Turkish children in Turkish, allowing for a comparison between the Turkish and Dutch children's responses. This comparison could throw light on the question whether there might be an effect of typological background. It was expected that literate Turkish and Dutch children would not provide similar units as responses, as word boundaries are not placed the same way in Turkish and Dutch, due to the fact that Turkish is an agglutinative language and Dutch an isolating one (see Section 2.1.2). Furthermore, by including both pre-literate and literate Turkish and Dutch children in the tasks, it should theoretically be possible to disentangle effects from typological background and from literacy, and see to what extent they interact.

Before we discuss the data in detail, Section 7.1 will provide an overview of the extent to which the (sociolinguistic) background factors, such as home language and language used for book reading, as well as school test scores, affected children's single-word responses in the task. The results of the tasks conducted for this study will then be presented separately for the Dutch and Turkish participants, in Section 7.2 and 7.3. The results from the eye-tracking task, conducted only in Dutch, will be discussed in 7.4. The reason for discussing this task separately from the others is that the results obtained in this task are not directly comparable to those of the other tasks. It is

impossible to categorize responses in this task in a similar way as was done for the other tasks. In Section 7.5, the Dutch and Turkish results are compared. The most important conclusions will be briefly summarized in Section 7.6, before their implications are discussed in Chapter 8.

7.1 Background information

In addition to the experimental data, background information was collected about the children through short survey questions and school test scores (see Section 5.3.7). The following information was gathered and analysed:

(1) For the Dutch children: scores on Raven (1962)'s Progressive Coloured Matrices, which yields information on cognitive skills.

(2) Vocabulary scores for Dutch children in E1 and E2 on a school (CITO) test, and scores on the *Toets kleutertaal* ('kindergarten language test') for children in K1 and K2.

(3) Information about the following sociolinguistic factors:
   − how often children watched TV;
   − how often children read books at home or were read to.

For the Turkish children, only the sociolinguistic background factors as mentioned under number (3) were collected.

As explained in Chapter 5, in order to assess possible effects from these sociolinguistic factors, of intelligence and of general language proficiency, correlations (Pearson's correlation co-efficient, two-tailed) were computed with children's single-word responses in the dictation, sentence segmentation, last-part repetition and click tasks. For the click task, this meant that correlations were calculated with children's correct repetitions of mid-target unit clicks (see Section 5.3.4). First, the analyses were made for the children in kindergarten, in grade K1 and K2, and then the analyses were made for the children in E1 and E2.

    Few correlations were found for the youngest children; the results can be found in Appendix 11. Briefly, only the results from the Raven's intelligence test had a significant correlation with the children in K1 and K2's single-word responses in the click task (r^2=.36, p=.04). For all other tasks, the background information did not seem to affect the number of single-word responses.

For the older children, a few more correlations were found. These results are also represented in Appendix 11 (2). There was a significant negative correlation between the scores on Raven's intelligence test and the mean percentage of single-word responses in the sentence segmentation task (r^2=-.41, p=0.02), and a significant positive correlation between the scores on Raven's test and single-word responses in the click task (r^2=.47, p=.003), and between the children's scores on the language test and their single-word responses in the click task (r^2=.36, p=.03). All other factors were not related to the children's results (non-significant correlations).

For the Turkish children, only the sociolinguistic background information was obtained. There were no significant correlations found for these children's single-word responses in the tasks and their answers to the sociolinguistic background questions (see Appendix 11 (3)).

As there were only four significant correlations (three positive, one negative) found between the Dutch children's single-word responses and the factors under investigation, we proceeded to analyse the effects from literacy, typological background and the frequency of occurrence of specific multiword sequences on their language segmentation and processing units.

7.2   Results from the tasks in Dutch

In this section, the results from the tasks performed by the Dutch children will be discussed. Section 7.2.1 will provide a general overview of the types of responses that the Dutch children provided per task (all but the eye-tracking task). In this overview, it will be indicated for all of the tasks to what extent the children produced different kinds of responses (ranging from sub-word units to multiple-word responses, see Section 5.3). After this, the ratio of word segmentations per task and the number of multiword target unit segmentations as obtained from the tasks conducted in Dutch will be discussed in Section 7.2.2 and 7.2.3, as to be able to examine the influence of literacy on children's performance in the tasks.

As a first step, however, the reliability of the tasks was investigated. The reliability (Cronbach's alpha) turned out to range from acceptable to high for all the tasks. For the sentence segmentation task for single-word segmentation (n_{items}=40), α=.94; for the last-part repetition task for single-word repetition (n_{items}=40), α=.91; for the click task for correctly repeated

clicks (n_{items}=40), α=.88; and for the mixed-words task for verbatim repeated sentences, excluding the filler sentences (n_{items}=39), α=.72. For this last task, the greatest increase in alpha would come from deleting the item '*bos het in*' ('forest the in'). Removal of this item would however only increase alpha by .04.

For the dictation task, the reliability of the children's responses could not be calculated, as children had come up with different numbers of sentences in their dictations.

7.2.1   General results from the tasks conducted in Dutch

Table 17 to 22 present, per task, the average percentages of times that the Dutch children in the different grades at school provided a specific response-type. With few exceptions, the categories to which responses could be assigned were similar across all tasks (see Chapter 5). They consisted of:

(a) Sub-word units, which consisted mostly of syllables and phonemes (e.g., '… *hand – den – was – sen* …' ('wash hands') hyphens indicating segment boundaries made).

(b) Single words (e.g. '… *op – de – tafel* ….', 'on the table') ;

(c) Multiword combinations that corresponded to the transitive and locative target units (e.g., a segment consisting of the complete unit '*op de tafel*' in the sentence segmentation task). These units were the hypothetical units included as targets, see Section 5.2.1.

(d) Other multiword combinations, including some that contained target multiword units (e.g., the repetition of the sentence '*ze moet haar handen wassen*' ('she has to wash her hands') in the last-part repetition task).

(e) Other responses, such as deictic reactions or non-verbal responses (see Chapter 5).

For the dictation task, responses could not be assigned to category (c), as children were free to construct their own story and the target units that were included in the other tasks did not occur. No responses were assigned to category (e) in this task either, as children only used sub-word units, single-words and multiword sequences.

	Sub-word responses		Single-word responses		Multiword responses	
	Mean	SD	Mean	SD	Mean	SD
K1	21.11	(26.02)	38.93	(22.93)	39.96	(33.81)
K2	31.10	(31.05)	35.15	(26.97)	33.74	(39.85)
E1	18.27	(18.94)	75.70	(17.18)	6.03	(6.75)
E2	7.22	(10.06)	74.40	(19.56)	18.38	(21.78)
Total	18.99	(23.81)	55.81	(28.73)	25.20	(30.98)

Table 17: Responses in the dictation task[17] in % of the total number of segments made (N=69)

Table 17 illustrates that single-word units occurred most in the dictation task, followed by multiword units. This was already mentioned in Chapter 6. The children in K1 and K2 provided almost equal numbers of single-word and multiword responses, whereas the children in E1 and E2 showed a preference for single words. This might indicate an effect of literacy, as children in E1 and E2, who have learned about words, preferred to segment their sentences into words. The fact that sub-word units were mostly provided by children in K2 underlines this impact of schooling: in this grade, children are often trained in recognizing units smaller than words, such as syllables.

Table 18 shows the results for the last-part repetition task.

For this task, the additional category, of 'anticipation', was included. This refers to cases in which children guessed what might be coming next in the story after the researcher had paused. Guesses consisted of single words (e.g., '*hondje*' ('doggie') when the pause followed 'a very sweet' …) or multiple words (e.g., '*uitpakken en mee spelen*' ('unwrap and play along') after 'she goes immediately' …).

---

[17] NB: For the dictation task, the percentages of sub-word, word and multi-word responses are calculated in comparison to the total number of segmentations that children provided themselves in their stories in this task. For the other tasks, the percentages relate to the total number of responses that children provided per task. See Chapter 5.

	Sub-word responses	Single-words	Target unit resp.	Other multiwords	Anticipa-tions	Other
K1	.60 (.99)	21.19 (16.63)	30.02 (14.80)	35.19 (26.61)	.09 (.41)	12.92 (8.55)
K2	.65 (1.16)	25.03 (19.10)	27.46 (10.84)	35.25 (21.80)	.00 (.00)	11.61 (9.07)
E1	.70 (1.60)	34.34 (22.54)	23.10 (9.10)	33.61 (27.87)	.26 (.84)	8.00 (6.74)
E2	1.08 (1.31)	42.06 (27.12)	32.35 (14.91)	21.37 (17.61)	.00 (.00)	3.14 (3.67)
Total	.75 (1.15)	30.29 (22.57)	28.15 (12.86)	31.62 (24.22)	.09 (.48)	9.10 (8.15)

Table 18: Responses in last-part repetition task in %, M (and SD) (N=74)

In general, Table 18 shows that single-word repetitions, repetitions of the target unit, and other multiword responses occurred most often. Sub-word responses and anticipations were less frequent. Other reactions, however, such as remarks related to children's own experiences with events like the ones occurring in the stories, did occur more often, mostly among the younger children in K1 and K2.

Single-word responses occurred more often among the older children in E1 and E2 than among the children in K1 and K2. The differences were not as big as in the dictation task, however. In general, all children provided multiword responses relatively often, though these certainly did not always coincide with the target units.

Table 19 presents the results of the sentence segmentation.

	Sub-word responses		Single-word responses		Hypothetical target unit responses		Other multiword responses		Other	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
K1	29.40	(27.70)	17.82	(8.74)	13.33	(11.40)	20.48	(23.48)	18.97	(15.44)
K2	32.59	(29.52)	20.77	(9.00)	14.62	(16.68)	17.41	(14.17)	14.61	(9.00)
E1	26.54	(19.31)	50.11	(15.96)	9.34	(11.03)	2.56	(3.67)	11.45	(8.18)
E2	32.34	(29.20)	37.99	(22.97)	13.81	(17.01)	8.02	(12.23)	7.84	(7.44)
Total	30.05	(26.12)	31.52	(19.93)	12.69	(13.92)	12.25	(16.81)	13.49	(11.30)

Table 19: Responses in sentence segmentation task in %, M (and SD) (N=72)

Table 19 shows that in the sentence segmentation task the children provided mostly sub-word units and single-word units. The other categories, including multiword target units, occurred less often, especially among the children in E1 and E2. These children mostly used single-word segments, while the children in K1 and K2 were responsible for most of the sub-word units. The difference suggests a literacy effect, which will be further explored in Section 7.3 below.

Table 20 provides an overview of responses for the click task. All children's responses were taken into account (mid-unit clicks, post-unit clicks, and clicks in fillers). The table shows that children did not re-place the clicks within words in their repetitions: sub-word responses did not occur. Single-word responses, i.e., responses in which children repeated the clicks within the multiword target unit (for instance, '*je moet je handen X wassen voor het eten*' ('you have to wash X your hands before dinner'); X representing the click sound children made), occurred more often among the children in E1 and E2 than among the K1 and K2 children. This was also the case for the repetition of clicks right before or after the target units, here represented in the rows 'target unit responses'. Younger children in K1 and K2 gave more 'other' responses, which included the placement of clicks all the way at the end of the sentence, or at another place in the sentence.

	Sub-word responses		Single-word responses		Hypothetical target unit responses		Other	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
K1	.00	(.00)	8.64	(14.67)	10.71	(15.10)	80.65	(28.32)
K2	.00	(.00)	13.18	(15.20)	13.82	(13.06)	73.00	(25.93)
E1	.00	(.00)	32.41	(16.97)	23.21	(12.92)	44.38	(26.69)
E2	.00	(.00)	52.68	(14.97)	28.08	(11.22)	19.23	(9.48)
Total	.00	(.00)	26.38	(23.30)	18.76	(14.78)	54.87	(34.09)

Table 20: Responses in the click task for all sentences in %, M (and SD) (N=80)

The picture changes somewhat if we look at the correctly repeated target items in which clicks occurred only in mid-unit position (which was assumed to indicate word by word processing). The relevant data are presented in Table 21.

	Sub-word responses		Single-word responses		Hypothetical target unit responses		Other	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
K1	.00	(.00)	7.54	(13.77)	14.24	(19.98)	78.21	(31.42)
K2	.00	(.00)	9.00	(13.25)	19.23	(17.99)	71.77	(28.11)
E1	.00	(.00)	29.08	(16.46)	32.78	(18.70)	38.14	(29.77)
E2	.00	(.00)	48.53	(19.03)	38.78	(15.80)	12.69	(11.53)
Total	.00	(.00)	23.25	(22.95)	25.97	(20.53)	50.78	(37.39)

Table 21: Results click task for mid-unit clicks only in %, M (and SD) (N=80)

As can be seen from Table 21, for clicks occurring in mid-unit position, such as in '*handen X wassen*' ('wash X hands', X representing the click), the children in the higher grades E1 and E2 more often correctly repeated the clicks than the children in the lower grades (K1 and K2). The literate children also seemed to replace the click more often to just before or just after the target unit, responding for instance with '*X handen wassen*' or '*handen wassen X*'. The pre-literate children in K1 and K2 more often repeated the clicks in completely different places (78% and 72% of all their responses).

In Table 22, children's responses in the mixed-words task are represented. For this task, an additional scoring category was needed, called 'DE changed into D'R' (see Chapter 5) while the category of single-word responses was logically impossible in this task.

	Hypothetical unit responses		DE changed into D'R		Other (not including verbatim repetitions)	
	Mean	SD	Mean	SD	Mean	SD
K1	3.37	(3.60)	4.64	(3.55)	8.18	(8.60)
K2	6.33	(9.68)	3.92	(3.79)	8.62	(8.51)
E1	4.82	(8.32)	3.21	(3.78)	5.05	(4.06)
E2	2.67	(7.66)	4.76	(3.01)	2.42	(2.93)
Total	4.24	(7.50)	4.15	(3.53)	6.11	(6.96)

Table 22: Responses in the mixed-words task in %, M (and SD) (N=82)

In general, Table 22 shows first and foremost that most responses could not be assigned to the created categories. The reason for this was that most children repeated the sentences verbatim, thus with the mixed order.[18] Changing back the targets that were presented in mixed order occurred most often among the children in K2, but still this percentage was very low (6%). Changing '*de*' into '*d'r*', creating a semantically meaningful sentence, did not occur very often (3.2%-4.8%), contrary to expectations. The number of 'other' responses, which consisted for the largest part of deictic reactions or references to children's real life experiences, was again higher for the younger children in K1 and K2 than for the children in E1, and was lowest for the children in E2. This suggests that the older children were better able to approach the task as a metalinguistic task, and were able to ignore the wrong word order.

Before going further into the details of the Table 17 to 22 presented above, it should be noted that the responses divided over the different response types were not completely independent from each other: a response that was assigned to one specific category, automatically led to the exclusion of other response categories (see Chapter 6). Despite the bias that may have been caused by this categorization, however, the responses per type or category seemed to differ for most tasks to a large extent.

If we compare Table 17 to 22, it appears that there were differences across tasks. Children provided most single-word responses in the dictation task (55.81%). In the last-part repetition task, on the other hand, children did not show an overall preference for a specific response type: the average percentages of single-words, multiword target units and of other multiword sequences was almost equal, around 30%. The sentence segmentation task featured yet another pattern, children most often providing units smaller than words (30.05%) or single-word responses (31.52%). In the click task, the most often occurring response type was 'other responses' (54.87%), followed far behind by single-word responses (26.38%). The fact that the 'other responses' were so high for the click task may be related, however, to the fact that there were only three categories available for coding this task.

    The variation in the average numbers of response types will partly be caused by the difficulty of interpreting responses in some of the tasks. For

---

18 There is no category for verbatim repetitions of the sentences included in Table 23, as such a category would not be informative with respect to the research questions posed here, and responses assigned to such a category would not be comparable to categories in other tasks.

the dictation task, the sentence segmentation task and the last-part repetition task coding was straightforward, but the interpretation of responses in the click task and mixed-words task required more consideration.

As regards the click task, it was assumed that children were able to process single words within multiword targets if they were able to correctly place a click that had occurred within such a target, between two words that together formed the multiword target: e.g., if children replaced the click as indicated by the X in the following sentence in the correct position, '*je moet je* **handen X wassen** *voor het eten*' ('You have to wash X hands before dinner'). A child who could do this was assumed to be able to break up the multiword target '*handen wassen*' ('wash hands') into single words. If, however, a click that occurred inside a target unit was placed just before or right after the multiword target, as in '*je moet je* **handen wassen X** *voor het eten*' or '*je moet je* **X handen wassen** *voor het eten*', then it was assumed that the child had not broken up the multiword target '*handen wassen*'. Accordingly, an answer of this second type was assigned to category (3), in which the multiword targets were left intact.

For the mixed-words task, a response was assigned to category (3) if a child changed the mixed-word order in the stimulus multiword target item into the regular word order. In such cases, a child would for instance say '*Je moet de* **tafel dekken** *van mama*' ('Mum wants you to set the table') instead of the same sentence with mixed-word order for the multiword target: '*Je moet de* **dekken tafel** *van mama*'.

Because of these coding differences, the percentages found for the different categories across the tasks cannot always be compared directly. To be sure, the sentence segmentation, dictation and last-part repetition tasks seemed more closely related to each other than to the mixed-words and click tasks (see Chapter 5). In these three tasks, children could choose their own segments. The different results across the tasks may, however, also be explained by other differences between the tasks: in the sentence segmentation task, children were given relatively short sentences to divide up, and they provided many units that were smaller than words in this task. In the dictation and last-part repetition tasks, on the other hand, segments were part of a larger context, i.e., a story, and children more often provided multiword segments. It is possible that in these tasks, children were more motivated to pay attention to the overall meaning, and their segmentation behaviour may reflect their higher involvement in the unfolding story.

Interestingly, the numbers of single-word and multiword responses in the click task are fairly close to those in the sentence segmentation and last-part repetition tasks. In the click task, children seemed to preserve the

multiword target units almost as often as in the other two tasks (18.76%), and the percentage of single-word responses was also quite similar (26.38%). In the mixed-words task, children did not often produce the multiword target unit: only in 4.24% of the cases. Since about 85% of the sentences was repeated verbatim in this task, without correcting any deviant word order, it can be suggested that maybe literal repetition is not so hard for children, or at least not harder than changing a salient 'error' into its correct alternative, giving the multiword target unit its regular word order. Below, this issue will be discussed further in relation to the frequency of occurrence of these multiword target units.

If we look at the response types per grade, it is clear that the children in grades E1 and E2 provided more single-word responses than the children in grades K1 or K2. In the dictation and sentence segmentation tasks, the K1 and K2 children rather often came up with sub-word units (21.11%-32.59%) but also with multiword responses. These, however, did not generally correspond to the preselected multiword target units (17.41%-39.96%). As for the sub-word units, it should be noted that the children in K2 provided more such units in their responses than the children in K1. This may be a result from schooling, as classroom practices in Dutch kindergarten that prepare for reading and writing, and to which children in K2 have been exposed more and longer than children in K1, are mostly focused on sub-word units such as syllables.

In the last-part repetition task, on the other hand, children hardly ever provided sub-word units (0.60%-1.08%). Further underlining the differences across tasks, in this task the average percentages of multiword responses and multiword target unit responses was much higher, and occurred almost equally often across all four grades (21.37%-35.25%).

In the click task, the children in E2 again most often provided responses that were coded as 'single words' (52.68%), whereas the children in the other three grades most often provided responses that were assigned to the category 'other' (44.38%-80.65%), though the children in E1 here provided the lowest number of 'other' responses, and the children in K1 the highest. This high number of 'other' responses for these children was mainly caused by the high number of repetitions in which children placed the clicks at the beginning or end of the sentence.

Correctly placing the click within the multiword target unit was mostly done by children in E2, closely followed by the children in E1 (resp. 48.53% and 29.08%).

As mentioned above, in the mixed-words task children did not often change the jumbled word order to the regular word order at all: they only did so in 2.67%-6.33% of the sentences, and it was mostly the children in K2 who did so.

In sum, the overview of results presented in Table 17 to 22 suggests that at least partially the use of the various categories was dependent on the tasks. Overall, though, children in E1 and E2 more often provided single-word units in their responses than the children in K1 and K2, while the preservation of complete multiword target units as segments seemed not to depend on literacy or on the grade in which the children were. The next section will specifically explore whether, how, and to what extent there was an effect of literacy on the children's responses.

7.2.2   The impact of literacy in the Dutch tasks

Table 17 to 22 suggest that pre-literate children differ considerably from literate children in their responses. This suggests a literacy effect. According to the literacy hypothesis, literate children would provide more single-word responses than pre-literate children.

In this section, this hypothesis will be investigated further, by analysing the number of single-word units provided in the tasks (except in the eye-tracking task, see Section 7.4). Hence, Unianova analyses were done for the dictation, sentence segmentation, last-part repetition and click tasks, with the average percentages of single-word responses as dependent variable, literacy as independent factor, and age as a co-variable. The results from the mixed-words task were not included in this analysis, as there was no hypothesis with respect to a literacy effect in this task (see Section 5.3.5). For the click task, only the results obtained from sentences in which the click had been placed mid-unit were taken into account (see Section 7.2.1). Table 23 presents the average numbers of single-word segmentations in the four tasks for the pre-literate and literate children.

	Dictation (N=69)		Last-part rep. (N=74)		Sentence segm. (N=72)		Click (N=80)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
K1	38.93	(22.93)	21.19	(16.63)	17.82	(8.74)	7.54	(13.77)
K2	35.15	(26.97)	25.03	(19.10)	20.77	(9.00)	9.00	(13.25)
Total pre-literate	37.20	(24.55)	23.01	(17.70)	19.18	(8.86)	8.22	(13.38)
E1	75.70	(17.18)	34.34	(22.54)	50.11	(15.96)	29.08	(32.78)
E2	74.40	(19.56)	42.06	(27.12)	37.99	(22.97)	48.53	(38.78)
Total literate	74.97	(18.29)	37.98	(24.76)	44.57	(20.13)	39.05	(20.16)
Total	55.81	(28.73)	30.29	(22.57)	31.52	(19.93)	23.25	(22.95)

Table 23: Single-word reactions by task, literacy, and grade in %, M (and SD)

Table 23 shows that overall, the literate children provided more single-word responses in the tasks than their pre-literate peers. The results from the Unianova analyses per task with single-word responses as dependent variable, literacy as independent factor and age as covariate, are presented in Table 24. There was no main effect for age in any task (all p>.05), but there was a significant main effect of for the dictation, sentence segmentation and click tasks. The effect sizes were medium for the click task and dictation task, and high for the sentence segmentation task. In accordance with our hypotheses, this suggests that literacy indeed generates single-word reactions.

	Age			Literacy		
	F	p	η_P^2	F	p	η_P^2
Dictation task (N=69)	.47	.47	.01	11.32**	.001	.15
Last part rep. (N=74)	.03	.87	.000	2.53	.12	.03
Sent. segm. task (N=72)	2.16	.15	.003	27.70***	.000	.29
Click task (N=80)	1.84	.18	.02	10.86**	.001	.12

Table 24: Anova: single word as dependent, literacy as independent factor and age as covariate

For the analysis of the results from the click task, as presented in Table 23 and 24, only the sentences in which the clicks had occurred in mid-unit position were taken into account. This means that in Table 23, only the clicks occurring within the multiword target units were considered, as they could represent ´single-word´ responses; not clicks that occurred just before or

after target-like units, in fillers, or clicks located in positions unrelated to target-like units. To further investigate the impact of literacy, the children's responses to clicks placed in the middle of multiword units and clicks in filler items were compared. Repeated Measures analyses were carried out, with the percentages of correct repetitions of clicks in either mid-unit or filler positions, which also included post-unit clicks, as dependent variables, and literacy again as an independent variable, and with age as a co-variable. The results from these analyses are shown in Table 25.

	Pre-literates (N=41)		Literates (N=39)		$F_{Literacy}$ p, η_P^2
	Mean	SD	Mean	SD	
Mid-unit click (40 items)	8.22	(13.38)	39.05	(20.16)	F(1,77)=10.3, p=.002, η_P^2=.12
Filler click (18 items)	16.43	(22.52)	51.28	(20.88)	
$F_{location\ click}$, p, η_P^2	F(1,77)=.004, p=.95, η_P^2=.000				

Table 25: Correct scores by target-type and group in %, M (and SD) and RM-analyses

Although there were more correct responses for clicks in fillers than for clicks within multiword target units, the location of the click did not lead to a significant main effect (p=.95). The main effect of literacy (controlled for age), on the other hand, was significant, as shown in Table 25, F(1,77)=10.3, p=.002, η_P^2=.12. There was no significant interaction between click location and literacy, F(1,77)=0.00, p=.99, η_P^2=.000. This suggests that it was easier for literate children than for pre-literate children to place the clicks in the correct positions. The lack of an interaction suggests that though it was harder for both literate and pre-literate children to repeat a click that occurred within a multiword target unit like '*handen wassen*' ('wash hands') than to repeat a click outside such a unit, this effect was not stronger for pre-literate than for literate children. These results suggest that in general the literate children were better at repeating the clicks than the pre-literate children, but that whether or not the clicks occurred within or after the high and low frequency multiword target units did not affect the results.

This, in turn, reinforces the literacy effect. The finding that literate children were better at placing clicks between the words that supposedly form a multiword unit, seems less related to the entrenchment of these units, than to the fact that literate children are better at correctly placing clicks where they had occurred in the audio-recordings: they are better at recognizing separate words. Although the fact that children scored higher on the filler sentences than on the test sentences with mid-unit clicks might indicate that the entrenchment of the selected multiword units did affect

children's repetitions of clicks somewhat, there was no effect of the location in which the clicks occurred.

7.2.3   The effect of entrenchment in the Dutch tasks

The fact that literate children more often provided single-word responses than pre-literate children does not necessarily mean they do not process multiword units holistically. Similarly, that pre-literate children did not come up with single words as often does not automatically mean that they do use multiword sequences as the basis for their language segmentation and processing.

To investigate this issue, two aspects of the data were investigated. First, the influence of the frequency of occurrence of specific multiword units on children's language segmentation and processing was analysed. Secondly, we examined whether different types of multiword units – locative prepositional phrases and transitive noun+verb-combinations – yielded different results.

The underlying hypothesis behind this investigation was elaborated upon in Section 2.2. In short, multiword units which occur often in language are assumed to be more entrenched in the mental lexicon, and because of this, they should be easier to retrieve as wholes than multiword sequences that do not occur often. Whether frequency indeed does influence the storage and retrieval of multiword sequences as wholes should become visible in children's segmentation and processing.

Prepositional locative and transitive noun+verb-combinations have largely been investigated in an explorative way (see Section 2.2). Both types of multiword constructions have been argued to be among the basic constructional units of language, and some evidence has been produced in studies with adults. In this study, we investigate whether the finding that adults use these units in their language processing and segmentation could be confirmed for young children. The inclusion of prepositional locative and transitive noun+verb-combinations as stimulus items in the tasks allowed for an investigation of whether despite the strong effect of literacy, i.e., of knowing about word as units of language, highly entrenched multiword units tend to resist being broken up during segmentation and processing, i.e., whether they are treated as 'big words' (cf. Arnon, 2010; Dąbrowska, 2004; Tomasello, 2006).

To what extent the Dutch children kept the multiword expressions together can be gathered from Table 17 to 22 above. The multiword target unit responses given in these tables indicate how often the children

responded with these units. The figures tell us how often children provided the selected multiword targets like '*handen wassen*' as holes in the sentence segmentation task, how often they repeated these units in the last-part repetition task, and how often they repeated the clicks that had occurred within the units in the audio-recorded sentences in the click task (thus: '*handen X wassen*') in a position just before or just after the multiword targets – i.e., as '*handen wassen – click*', or '*click – handen wassen*'. For the mixed-words task, it was calculated how often children changed the multiword targets that had been presented in mixed order by the researcher to the regular word order in their repetitions – as this was supposed to show that children did not repeat the sentences word by word, but rather processed multiword units holistically. The results from the dictation task were excluded from the analyses for this section, as the selected multiword targets were not incorporated in that task.

In Table 17 to 22 no distinction was made between high and low frequent multiword targets. Frequency is assumed to influence the level of entrenchment. In order to test for a frequency effect, a Repeated Measures analysis was done, with the average percentages of the number of times that the pre-literate and literate children responded with the multiword targets after high and low frequency transitive and locative combinations as dependent factor, the low and high frequency (LF or HF) transitive noun+verb-combinations (NV) and locative prepositional combinations (PP) as dependent variable, and literacy as between factor. Age was again taken as a co-variable. Table 26 presents the results.

	Last-part rep. (N=74)		Sentence segm. (N=72)		Click (N=80)		Mixed words (N=82)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LF-PP	26.25	(20.47)	11.83	(16.58)	29.52	(25.71)	4.05	(11.76)
LF-NV	32.84	(23.16)	16.11	(18.67)	23.40	(21.94)	5.57	(9.56)
Total LF	29.46	(18.59)	13.79	(14.96)	26.10	(20.68)	4.75	(9.98)
HF-PP	23.54	(24.06)	10.64	(20.01)	33,81	(29.07)	9.08	(17.58)
HF-NV	36.24	(19.93)	13.34	(14.75)	19.89	(20.30)	3.09	(8.16)
Total HF	29.87	(19.34)	11.89	(15.34)	25.91	(22.57)	6.22	(11.95)
Total PP	24.97	(20.66)	11.20	(16.26)	30.35	(25.62)	7.29	(16.39)
Total NV	34.63	(18.96)	14.82	(15.10)	21.41	(18.01)	4.34	(8.86)

Table 26: Hypothetical multiword responses by frequency and constructional type per task in %, M (and SD)

Table 26 shows that there were in general big differences in the average percentages of multiword target unit responses after high and low transitive and locative multiword combinations in the different tasks, varying from 3.09% in the mixed-words task, to 36.25% in the last-part repetition task. Within tasks, the variation in percentages was much smaller, and patterns were recognizable.

As for the last-part repetition and sentence segmentation tasks, for answers after the same constructional type (thus only after noun+verb-combinations, NV, or only after prepositional phrases, PP), the differences in responses for low and high frequency targets are small. In general, the differences that exist seem to be between constructional types, thus between transitive noun+verb-combinations and locative prepositional phrases more than between high frequency (HF) and low frequency (LF). Moreover, where there is a difference between high and low frequency target items, the high frequency target is not always the one with a higher number of multiword target unit responses. This seems to go against the hypothesis that higher frequency multiword targets are more often processed as wholes.

Table 27 shows the results of the Repeated Measures analyses for each of the tasks, with frequency and constructional type as within-subject factor, multiword target units as dependent variable and age as covariate.

	Age			Frequency			Constructional type		
	F	P	η_P^2	F	P	η_P^2	F	p	η_P^2
Last-part rep. (N=74)	8.75**	.004	.11	3.47	.07	.05	.16	.69	.002
Sentence segm. task (N=72)	.22	.64	.003	2.56	.11	.04	11.87**	.001	.16
Click task (N=80)	9.08**	.004	.11	.054	.82	.001	.26	.61	.004
Mixed words (N=82)	.02	.90	.000	1.21	.27	.015	.002	.97	.000

Table 27: RM-anayses with frequency and constructional type as within-subject factor, age as covariate and hypothetical units as dependent variable

This table shows a significant main effect of age for the last-part repetition task and for the click task, with a medium effect size for both tasks. There was no significant main effect of frequency in any of the tasks, but there was

a significant main effect of constructional type in the sentence segmentation task, which shows that NV-targets more often led to multiword hypothetical unit responses than locative prepositional targets. Two two-way significant interactions were found. The first was a significant interaction of type by age in the sentence segmentation task, $F=17.46$, $p=.000$, $\eta_P^2=.20$. The second was a significant interaction of frequency by type in the mixed-words task, $F=13.08$, $p=.001$, $\eta_P^2=.14$. The interaction in the sentence segmentation task probably derived from the fact that there were more older children who provided multiword hypothetical target units as answers for the transitive targets than for the locatives, while the opposite was true for the younger children. The second interaction, of frequency by type in the mixed-words task, resulted from the fact that more multiword target unit responses were given for the transitive low frequency targets than for the locatives, whereas the opposite was found for the high frequency targets: for these, more multiword hypothetical units were provided for the locatives than for the transitives.

The fact that there was no main effect of frequency suggests that this factor did not play a major role in children's segmentation and processing of the target units that were included in the tasks. Constructional type also seemed to have only a limited impact.

7.2.4 The impact of literacy, frequency, constructional type and task combined

In the analyses so far (Section 7.2.1 to 7.2.3), the impact of the different factors was investigated for each of the tasks separately. In a Manova Repeated Measures analysis, it was investigated to what extent the factors literacy, frequency and constructional type influenced the single-word responses in all tasks combined, leaving out only the dictation task (see Chapter 6) and the eye-tracking task (see Section 7.4). This led to an analysis with literacy as between subject factor, frequency, constructional type and task as within subject factor, and age as covariate (N=60), and the single-word responses as dependent variable.

	F(1,58)	p	$\eta_P{}^2$
Age	.96	.33	.02
Frequency	2.43	.13	.04
Type	.98	.33	.02
Task	1.73	.18	.03
Literacy	27.50***	.000	.32
Frequency by age	1.74	.19	.03
Frequency by literacy	1.36	.25	.02
Type by age	.05	.82	.001
Type by literacy	.83	.37	.01
Task by age	1.25	.29	.02
Task by literacy	.68	.51	.01
Frequency by type	.002	.96	.000

Table 28: Manova RM analysis with literacy as between subject factor, frequency, type and task as within subject factor, age as covariate and single-word responses as dependent factor

The conclusion from this analysis, as presented in Table 28, shows that there was a significant main effect of literacy, but no other significant main effect was found for any of the other factors. There were no interactions either.

7.2.5   Conclusions on the impact of literacy, frequency, constructional type and task in Dutch

The analyses of single-word responses per task and for all tasks combined, suggested a major role of literacy. Literate children provided far more single-word responses than pre-literate children in all of the tasks. No significant influence from literacy was found in the last-part repetition task, but this may have resulted from the fact that literate children provided many other types of answers in addition to single-word responses and that this may be related to the structure of the task itself.

With respect to frequency and constructional type, the influence was not so clear: no influence of frequency could be found for children's multiword target unit responses, and the effect of constructional type seemed restricted to the sentence segmentation task.

It can therefore be concluded that the effect of literacy was much larger than any effects of frequency or constructional type. It is unclear whether this calls for a reassessment of the link between frequency and

entrenchment, as the frequency data computed for this current study were based on Google, and this may be problematic when used as an approximation of child language. A further discussion of this aspect of the methodology will be supplied in Chapter 8.

In relation to the pilot studies, presented in Chapter 4, it should be noted that the outcomes of the pilot version of some tasks differed from the results presented in this chapter – not in terms of significance levels, but in terms of single-word responses: in the pilot versions of the last-part repetition and sentence segmentation tasks, literate children provided far more single-word responses than in the main study. The reason for this difference may lie in the choice of target items: The pause was more often inserted after combinations of meaningful open class words, the multiword target units, in the main study than in the pilot sessions in the last-part repetition task. This may have led to fewer single-word responses. Likewise, the selection of the meaningful multiword target units may have triggered more multiword responses in the sentence segmentation task in a similar way (see also Section 2.1.2).

7.3　Results from the Turkish tasks

Before going into the results of the Turkish data, the reliability of the single-word responses given in response to multiword stimulus items was investigated (Cronbach's alpha). Accordingly, the reliability for the last-part repetition task and the sentence segmentation task was calculated over 50 target items. The reliability of both tasks appears to be high: for the last-part repetition task, Cronbach's α=.99, and for the sentence segmentation task α=.98.

7.3.1　General results from the Turkish tasks

It was first calculated how often the Turkish children provided the various response types. The same categories were used as for the Dutch data. These were the following (see Section 5.3.2 and 5.3.3):

(a)　Sub-word unit responses (e.g., '… *el – te – miz – le – mek* …', 'wash hands') hyphens indicating segment boundaries made).

(b)　Single-word responses (e.g., '… − *bahçenin* − *içinde* − ….', 'in the garden').

(c) Multiword responses that corresponded to the transitive and locative constructions that were target units (e.g., a segment consisting of the complete unit '*bahçenin içinde*' in the sentence segmentation task).

(d) Other multiword responses, including ones that included the target units (e.g., the repetition of the sentence '*Eve gelinçe el temizlemek*', 'Once she gets home, she washes her hands', in the last-part repetition task).

(e) Anticipations – in the last-part repetition task only.

(f) Other responses, such as deictic reactions or non-verbal responses (see Chapter 5).

Table 29 presents the average percentages of these response types (following the method described in Chapter 5).

	Grade	Last-part repetition	SD	Sentence segmentation	SD
Sub-word responses	K2	2.41	(10.78)	12.26	(24.49)
	E2	.00	(.00)	3.32	(8.28)
	Total	1.07	(7.19)	7.30	(17.79)
Word responses	K2	6.69	(10.54)	39.57	(19.97)
	E2	23.44	(31.82)	74.54	(26.03)
	Total	16.00	(25.90)	58.99	(29.16)
Hypothetical unit responses	K2	10.90	(13.36)	34.43	(17.60)
	E2	24.54	(19.58)	17.08	(19.39)
	Total	18.48	(18.25)	24.79	(20.37)
Other multiword responses	K2	71.35	(29.03)	8.39	(10.05)
	E2	51.07	(32.07)	2.29	(4.19)
	Total	60.08	(32.07)	5.00	(7.91)
Anticipation	K2	.11	(.49)	NA	NA
	E2	.00	(.00)		
	Total	.05	(.32)		
Other	K2	8.54	(14.26)	5.34	(5.35)
	E2	.95	(1.72)	2.77	(5.39)
	Total	4.32	(10.20)	3.92	(5.46)

Table 29: Responses per type and task for the Turkish children, per grade in %, M (and SD) (N=45)

As for the data from the Dutch children, an answer of one type automatically excludes the other categories. The percentages for the different categories in Table 29 are thus to some extent dependent upon each other.

A further remark that should be made here is that the percentages mentioned under the 'words' category will be slightly biased, since there were five single-word locative targets included in each task (i.e., nouns marked with locative case, such as '*bahçe-de*', 'garden – LOC'). If a child repeated or segmented this as one part, it was categorized as a 'single word'. In the Dutch versions of the tasks, there was no such bias, since all targets included there consisted of multiword sequences (see also Section 5.2.3).

Table 29 shows there was variation between the tasks: in the last-part repetition task, multiword units were the most frequent response type (on average 60.08%), whereas in the sentence segmentation task, multiword targets were often broken up into single words (58.99%). Answers that corresponded to target multiword units (24.79%) or other multiword sequences (5.00%) were less common in this task. Sub-word units were not provided often in either task, and responses that had to be categorized as 'other' were not common either (0-12.26%).

Apart from differences per task, there were also differences in the responses per grade. In the repetition task, both pre-literate (K1) and literate (E2) children provided 'other multiword answers' more frequently than any other category, mostly multiword sequences that were bigger than the target units, but the literate children answered with the target units more often than the pre-literates. These literate children also provided more single words than the pre-literate peer: respectively 23.44% versus 6.69% of the responses consisted of single words in the repetition task.

In the sentence segmentation task, the pre-literate children in K2 provided an almost equal number of single-word and hypothetical unit responses (39.57% and 34.43%), whereas the literate children showed a clear tendency for single-word responses in this task: almost 75% of the literate children's answers in the sentence segmentation task consisted of single-word units.

The Turkish data thus showed a difference in response types per task, and a difference in answers between pre-literate and literate children in both tasks. That literate children provided more single words in the sentence segmentation task than in the repetition task, may again be related to the relative brevity of the stimulus sentences (see Section 5.3.2). More importantly, the fact that literate children produced more single-word units

than pre-literate children suggests an effect of the literacy education these children had obtained.

In order to further investigate to what extent literacy affected the Turkish children's responses, the single-word responses were further analysed in T-tests. The results will be presented below.

7.3.2 The impact of literacy on each task in Turkish

Like with the Dutch children, it was expected that the literate Turkish children would provide more single-word reactions than their pre-literate peers. To check whether the literacy effect suggested on the basis of the findings above could be confirmed statistically, T-tests were carried out on the basis of the single-word reactions given by the children in the two tasks. In these analyses, the average percentages of single-word responses in the tasks were taken as within-subjects variables and literacy as a between-subjects factor. For the Turkish children, who were only selected from two grades, age was not taken as a co-variable, as age corresponded perfectly to whether children were literate or not. Table 30 presents the average percentages of single-word reactions for the Turkish pre-literate and literate children (N=45) in the last-part repetition task and sentence segmentation task.

	Grade	Last-part repetition	SD	Sentence segmentation	SD
Word	K2 (N=20)	6.69	(10.54)	39.57	(19.97)
	E2 (N=25)	23.44	(31.82)	74.54	(26.03)
	Total	16.00	(25.90)	58.99	(29.16)

Table 30: Single-word responses per task and group in %, M (and SD) (N=45)

The average number of single-word segmentations was much higher among the literate children in E2, than among the pre-literate children in K2. From the T-tests, it derived that this difference was significant for both tasks, $t_{last-part}(43)=-2.47$, p=.02; $t_{sentence\ segm}(43)=-4.95$, p=.000. This suggested indeed an important influence of literacy on the segmentations made by the Turkish children.

7.3.3 Entrenchment effect Turkish data

In order to test to what extent the Turkish multiword units were left as wholes, and to what extent this depended on the constructional type and on

the frequency of occurrence of these combinations, an analysis was made of the number of times that the Turkish children segmented multiword targets as units. In addition, it was investigated whether especially the literate children treated the single-word locative targets such as '*bahçede*' ('in the garden') as single units more often than their multiword equivalents ('*bahçenin içinde*'), as this could show to what extent children's segmentation is affected by literacy, or 'graphic relativity' (cf. Bugarski, 1993; see also Chapter 2).

In the Repeated Measures analysis done to answer this question, constructional types (transitive or locative multiword combination) and frequency were taken as within-subject factors, and the percentage of target unit responses as the dependent variable (see Table 31).

	Last-part repetition		Sentence segmentation	
	Mean	SD	Mean	SD
LF-PP	23.52	(25.44)	30.42	(28.23)
LF-NV	34.36	(29.35)	37.26	(31.45)
Total LF	29.20	(25.54)	33.77	(28.32)
HF-PP	16.46	(22.65)	21.48	(24.71)
HF-NV	24.35	(31.15)	35.37	(28.96)
Total HF	20.56	(25.51)	28.22	(23.45)
Total PP	20.11	(22.29)	26.00	(24.77)
Total NV	29.48	(28.96)	36.12	(28.24)

Table 31: Hypothetical unit responses by frequency and constructional type in %, M (and SD) (N=45)

Table 31 shows that children had fewer target unit responses after high frequency multiword targets than after low frequency targets. In addition, they provided more target unit responses, in both tasks, for transitive noun+verb-combinations than for locative multiword stimulus items.

These differences also showed substantial variation for the pre-literate and literate children taken together. It is only for the transitive noun+verb-combinations in the sentence segmentation task that the differences between high and low frequent targets are not so striking: 35.37% of the high frequent multiword targets were left as units and 37.26% of the low frequent ones.

A Repeated Measures analysis, with frequency and constructional type as within subject variable and the hypothetical unit responses as dependent

variable, showed both a significant main effect of frequency and of constructional type in both tasks, $F_{last-part\ freq}(1,44)=20.78$, p=.000, $\eta_P^2=.32$, $F_{last-part\ constr}(1,44)=16.68$, p=.00, $\eta_P^2=.28$; $F_{sent\ seg\ freq}(1,44)=6.86$, p=.01, $\eta_P^2=.14$, $F_{sent\ seg\ constr}(1,44)=16.25$, p=.000, $\eta_P^2=.27$. It is striking that in the last-part repetition task, there were more target unit responses for the low frequency targets than for the high frequency ones. This was in contrast, of course, to the hypothesis that high frequent units, supposed to be more easily stored and produced as wholes because of the higher entrenchment of such units, would elicit responses that would preserve their unit status. A possible explanation may be that more frequently occurring multiword sequences may be better known, and they may be more easily decomposed, as children are more aware of how they are built up out of constituent parts than for less frequently occurring sequences.

As for the effect of constructional type, it appeared that in both tasks children provided more target unit responses for the transitive NV-combinations than after for the locative expressions. Children often provided more than just the locative combination, including for example anticipations of what would follow the expression, e.g., a verb. This did not happen as often with the transitive combinations, which may seem more complete as meaningful units on their own.

The Turkish data also allowed for an analysis of the contrast between two very similar locative phrases, of which one appears as one word and the other as two words, to be referred to as the short and the long locative combinations (see Chapter 5). The expectation was that if the children were highly influenced by their knowledge of writing conventions, they would provide single-word responses not just for the short form but also for the long form. If they were instead affected more by the semantic integrity of the locatives, they would probably respond with the whole target unit, e.g., '*bahçenin içinde*'. The long form refers to one concept after all, or more precisely, one location. Table 32 shows the relevant data.

	Answer type	Last-part repetition task		Sentence segmentation task	
		Mean	**SD**	**Mean**	**SD**
Short form locative	Subword	1.11	(7.45)	9.16	(21.56)
	Single word	31.40	(32.68)	84.94	(23.42)
	More than unit	64.55	(33.45)	4.44	(9.36)
	Anticipation	0.25	(1.66)	–	–
	Other	2.47	(6.24)	1.13	(3.77)
Long form locative	Subword	0.67	(4.47)	7.94	(20.72)
	Single word	13.03	(28.99)	57.20	(34.22)
	Complete unit	21.07	(24.83)	27.08	(27.66)
	More than unit	58.05	(33.42)	4.51	(9.90)
	Anticipation	–	–	–	–
	Other	7.18	(17.01)	3.27	(7.76)

Table 32: Response types for short and long form of locatives in Turkish in the last-part repetition and sentence segmentation task in %, M (and SD) (N=45)

Table 32 shows that for the last-part repetition task, the children's preferred response consisted of more than just the locative constructions, in whatever form. Single-word responses occurred in 31.40% of the cases for the short form, and only in 13.03% of the cases for the long form.

In the sentence segmentation task, children provided more single-word responses both for the short (84.94%) and the long forms (57.20%). For the long form, however, the complete target unit was provided as one segment fairly often, in 27.08% of the cases.

In order to be able to confirm a literacy effect statistically, a Repeated Measures analysis was carried out, with the percentage of single-word responses for the shorter and longer locative forms and the percentage of complete unit responses for the longer form as dependent variables and literacy as between subject variable. The division of these responses by literacy are presented in Table 33.

Answer type after short / long form	Pre-literate or literate	Last-part repetition task		Sentence segmentation task	
		Mean	SD	Mean	SD
Word after short form	Pre-literate (N=20)	18.77	(26.70)	73.51	(26.40)
	Literate (N=25)	41.51	(33.97)	94.09	(16.06)
	Total (N=45)	31.40	(32.68)	84.94	(23.42)
Word after long form	Pre-literate (N=20)	3.50	(15.65)	35.54	(23.78)
	Literate (N=25)	20.66	(34.80)	74.53	(31.54)
	Total (N=45)	13.03	(28.99)	57.20	(34.22)
Complete unit after long form	Pre-literate (N=20)	13.92	(22.79)	39.31	(25.95)
	Literate (N=25)	26.79	(25.34)	17.30	(25.41)
	Total (N=45)	21.07	(24.83)	27.08	(27.66)

Table 33: Single-word responses for short and long locatives and complete unit responses for long form in the tasks, by literacy in %, M (and SD)

As Table 33 shows, the responses of pre-literate and literate children varied to a large extent for all response types, and the percentages of response types also varied highly. The Repeated Measures analysis showed that these differences were significant for the last-part repetition task: there was a main effect of response type, $F(2,86)=5.91$, $p=.009$, $\eta_P^2=.12$, and a main effect of literacy, $F(1,43)=9.36$, $p=.004$, $\eta_P^2=.18$. There was no significant interaction ($p>.05$), suggesting that pre-literate and literate children showed the same trends per answer category, but to smaller or larger degrees.

In the last-part repetition task, the literate children provided more single-word responses for the short and the long form, and more complete target units for the long form than the pre-literate children. Pre-literate children often gave longer responses. The expected literacy effect was confirmed for this task, although the fact that literate children also often provided the target unit for the longer form shows the literacy effect is limited.

In the sentence segmentation task, too, there were obvious differences between the pre-literate children's single-word and complete unit responses. The literate children provided more single-word responses, both with the short and the long locative stimulus items, whereas the pre-literates more often provided the whole construction when presented with the longer form. A Repeated Measures analysis with the types of response as dependent variable and literacy as between subjects factor showed main effects of response type and literacy, $F_{response\ types}(1,43)=50.34$, $p=.000$, $\eta_P^2=.53$,

$F_{literacy}(1,43)=9.42$, p=.004, $\eta_P^2=.18$, and a significant interaction of response type by literacy, $F(1,43)=15.99$, p=.000, $\eta_P^2=.27$. This interaction indicates that the difference between pre-literate and literate children was, as expected, much larger for the longer form than for the short form.

These results from the sentence segmentation task suggest an even clearer literacy effect than that of the last-part repetition task. It is not just that literate children provided fewer multiword target unit responses, a T-test also revealed this difference to be significant, t(43)=2.86, p=.007. The results overall were taken to confirm the hypothesis that literacy does affect the way in which children segment units in language.

7.3.4 The effects of literacy, frequency, constructional type and task combined in the Turkish data

The results above all represent results per task. In another Repeated Measures analysis, it was examined to what extent the factors literacy, frequency and constructional type influenced the single-word responses in the two tasks combined. This led to an analysis with literacy as between subject factor, frequency, constructional type and task as within subject factors, and the single-word responses as dependent variable (N=45). In Table 34, the primary results of this analysis are presented.

	F	P	η_P^2
Frequency	3.27	.08	.07
Type	77.27***	.000	.64
Task	81.04***	.000	.65
Literacy	18.07***	.000	.30
Frequency by literacy	1.49	.23	.03
Type by literacy	.71	.40	.02
Task by literacy	6.62*	.01	.13
Frequency by type	5.05*	.03	.11
Frequency by task	5.68*	.02	.12
Task by type	18.51***	.000	.30

Table 34: RM analysis with frequency, constructional type and task as within subject factor, literacy as between subject factor and single-word responses as dependent variable

Main effects were found of constructional type, task, and literacy (p=.000), and interactions of task by literacy, frequency by constructional type, frequency by task and task by constructional type. The interaction of task by literacy resulted from the fact that in the sentence segmentation task, far more literate than pre-literate children provided single-word responses, which was not the case in the last-part repetition task to the same extent. The interaction of frequency by constructional type was a result of the larger difference in the number of single-word responses after transitives and locatives in the sentence segmentation task than in the last-part repetition task. The interaction between frequency by task resulted from a relatively higher number of single-word responses after high frequency targets than after low frequency targets in the sentence segmentation task than in the last-part repetition task; and the interaction of task by constructional type resulted from a relatively higher number of single-word responses after locative targets than after transitives in the sentence segmentation task than in the last-part repetition task.

7.3.5   Conclusions Turkish data

The analyses of the Turkish data presented above show that, similar to the Dutch data, literacy seems to have a huge impact on the number of single-word responses children gave. Moreover, the Turkish data suggest that the frequency of occurrence of the multiword target units that were included in the tasks, as well as the constructional types and the nature of the tasks, affected the responses made. The frequency data need serious qualification, though: contrary to expectation, in the last-part segmentation task, children provided more multiword responses with low frequency targets than with high frequency targets. Tellingly, in the analysis of all factors combined, there was no main effect of frequency at all.

The analyses of responses given after short and long form locative constructions, consisting of either one or two words in writing, also suggested a literacy effect, especially in the sentence segmentation task. In the last-part repetition task, this effect was less obvious because literate children also provided the long locative construction more often as a single unit than the pre-literate children, who, in general, provided often longer responses: They often added a verb to the locative unit. This may be related to the fact that the locative constructions typically function grammatically as adjuncts, linked semantically to a verb.

The fact that literate children provided more single-word responses after the long, and also after the short, locative form in Turkish than the

pre-literates, confirms the hypothesis that literacy affects children's segmentations. That the literate children – as the pre-literate children – provided more single-word reactions to the short form locative construction than to the longer form may result from the fact that the short form consisted of a single word referring to a location. It may, however, also be related to the fact that the short form locative constructions had a more general meaning than the long forms, they could refer to all kinds of positions (e.g., 'on', 'at', 'in', 'over', etc.) and therefore, it could be expected that the short form locatives occur more frequently in Turkish than the longer form locatives, which refer to one specific location.

Bearing in mind that the Turkish data are only based on two, relatively offline, tasks, we can draw the conclusion that the results partly confirm the literacy effect that was found in the Dutch data. However, the nature of the tasks also affects the results, as became clear from the comparison between tasks as presented in Section 7.3.4.

Before going into a final discussion of how the results from the Dutch and Turkish children should be compared, the most online task conducted in this study will be discussed separately, in Section 7.4. This task, an eye-tracking task, was performed by Dutch children only.

7.4   Results from the eye-tracking task

For the eye-tracking data, only obtained from Dutch children, it was first of all analysed whether children's first fixations after hearing the first part of the audio input fell faster on the target pictures that represented frequently occurring noun+verb-combinations than on the counter target pictures, as they listened to phrases represented by the target or counter target pictures, and to what extent literacy affected any differences between children (see Chapter 5). The expectation was that children would look faster at high frequency targets than at counter targets that occur less often in everyday language. To check this hypothesis, a Univariate analysis was carried out, with the number of first fixations on the target or counter target as dependent variable (see Section 5.3.6), and the audio input (target or counter target, condition 1 and 2) and literacy (pre-literate or literate) as independent variables. Table 35 presents the results.

Audio-input	Picture	Average fixation number			
		Pre-literates (N=105)		Literates (N=135)	
		Mean	SD	Mean	SD
Target pronounced (condition 1)	Target	1.99	(0.19)	2.72	(0.17)
	Counter target	1.96	(0.19)	2.36	(0.17)
Counter target pronounced (condition 2)	Target	2.27	(0.19)	2.61	(0.17)
	Counter target	2.08	(0.19)	2.39	(0.17)

Table 35: Average first fixation number on target or counter target picture in relation to audio input (target or counter target) for pre-literates and literates

In general, Table 35 seems to show that pre-literate children looked faster at the frequent target and less frequent counter target pictures than the literate children in condition 1. For all children, fixation was faster on the counter target pictures in condition 1 than on the target pictures.

In condition 2, the children seemed to react slower than in condition 1, but both the pre-literate and the literate children looked faster at the pictures representing what they heard, the counter target, than at the target pictures.

The Univariate analysis showed that, in contrast to our expectation, there was no significant main effect from the pictures on the first fixations, $F(1,232)=.52$, $p>.05$, $\eta_P^2=.002$. There was also no main effect of the audio input (condition 1 or 2) that children received, $F(1,232)=2.45$, $p>.05$, $\eta_P^2=.01$. There was, however, a main effect from literacy, $F(1,232)=7.00$, $p=.01$, $\eta_P^2=0.3$. The pre-literate children fixated sooner on the target or counter target pictures than the literate children, which may indicate that they took a shorter 'route' to these pictures than their literate peers. There were no significant interactions of the three independent variables found.

We will now focus on whether the gaze routes followed the hypothesized direction. The question was whether the pre-literate and literate children took different routes in their fixations. The assumption beforehand was that literate children would more often decompose the multiword target that they heard into its constituent words – which were depicted as well – than the pre-literate children, who could be expected to rely on complete meaningful units without interference from the form of the unit in writing. We investigated how often the pre-literate and literate children showed one of the following two patterns in their eye-movement paths upon obtaining the audio input after segment 5 (see Section 5.3.6):

(1)   From distractor 1 (noun picture) to distractor 2 (verb picture) to the
      target (noun+verb picture).

(2)   From distractor 1 (noun picture) to the target or counter target
      (noun+verb picture)

Table 36 presents the number of times (in percentages) that route 1 and 2
occurred for the pre-literate and literate children.

	Route 1	Route 2
Pre-literates (N=105)	–	6.29
Literates (N=135)	0.59	9.31
Total	0.36	8.13

Table 36: Number of times (%) that children's eye movements showed route 1 or 2

As Table 36 shows, route 1 occurred only 0.36% of the time of all 1,672
eye-movement paths in the data. The pre-literate children did not take route
1 at all; the literate children had done so in 0.59% of their eye-movement
paths. Route 2 occurred in 8.13% of all eye-movement paths of the children
in total: 6.29% of the pre-literate children's eye-movement paths, and 9.31%
of the literate children's paths. These percentages may seem very small, but
note that they are very specific, route 1 even more so than route 2. In all
other cases, children's eyes showed different routes, for instance not starting
with a fixation on the picture that represented the noun of the multiword
unit. These are, however, not of interest for the research questions posed
here and they will therefore not be analysed further.

   Since route 1 did not often occur in the data at all, and the pre-literate
children never showed route 1 in the obtained data at all, it was not
analysed any further.

   An in-depth analysis of route 2, however, using a chi-square test, showed
with a statistical trend (χ^2 (1)=3.43, p=.06) that route 2 occurred more often
among all children for screens with which the target unit occurred in the
audio input (thus in condition 1) than for screens with which they heard the
counter target (condition 2). As the counter targets had lower frequencies of
occurrence than the target units, this finding suggests that for route 2, the
frequency of occurrence of the multiword units that were selected as targets
or counter targets did matter: The children's eye movements went more
often from 'hands' to 'wash hands' straight away than from 'hands' to the
less frequent 'fold hands' ('pray'). The fact that 'fold hands' occurs less often
in general in Dutch seems to have affected the results.

With respect to the literacy hypothesis, the results showed that 30.15% of all eye-movement paths in route 2 came from the pre-literate children (41 of their eye-movement paths could be assigned to route 2), and 69.85% (95 of their eye-movement paths) from literate children. Route 2 thus occurred more often among the literate than among the pre-literate children, and a chi-square analysis of this showed that this difference was significant (χ^2 (1)=4.48, p=.03).

This may indicate that, while the children did not seem to break up the multiword units into their constituent words, as would have been represented by route 1, literate children did seem to be affected by their knowledge that phrases consist of individual words. In contrast to pre-literate children, they fixated more often on the picture representing the first word of the target or counter target that they heard. This suggests that they did decompose the multiword target units. Interestingly, their eyes went from the first word immediately to the complete unit – for instance immediately from '*handen*' ('hands') to '*handen wassen*' ('wash hands').

In sum, the eye-tracking task performed by the Dutch children showed an effect from literacy on the way in which children processed the depicted noun+verb-combinations, and it showed that frequency probably affected processing as well, especially for literate children.

7.5   Comparing the Dutch and Turkish results in relation to the hypotheses

In Section 7.1 to 7.4, the data from the Dutch and Turkish tasks were analysed separately. In this section, the combined impact of literacy and typology will be assessed by comparing the Turkish and Dutch single-word responses in the two tasks that were conducted in both languages.

7.5.1 The impact of literacy and typological background (comparing Turkish and Dutch data)

To test whether children's typological background had an effect on their responses in the tasks, i.e., whether the typological background of the language the children happened to speak natively affected the kinds of segmentations they made, the number of single-word responses in the two tasks that were conducted both in Turkish and in Dutch – the sentence segmentation task and the last-part repetition task – were compared. It was

expected that Dutch children would have smaller units in their responses than Turkish children, as Dutch is an isolating language and Turkish an agglutinative language (see Chapter 2). Related to the impact of literacy – for which clear evidence was found in the task results given above – it was furthermore expected that especially literate children's responses would differ in Dutch and Turkish, while pre-literate children, due to their lack of knowledge about the writing conventions of their respective languages, would provide more similar units, often larger than single words.

In Table 37, an overview is given of the percentages of single-word reactions in the two tasks by background and literacy.

		Last-part repetition task		Sentence segmentation task	
		Mean	SD	Mean	SD
Pre-literate	Dutch (N=37)	22.01	(16.82)	19.18	(8.86)
	Turkish (N=16)	4.44	(6.60)	41.56	(18.72)
	Total (N=53)	16.71	(16.58)	25.93	(16.22)
Literate	Dutch (N=35)	38.59	(24.84)	44.57	(20.13)
	Turkish (N=25)	23.44	(31.82)	74.54	(26.03)
	Total (N=60)	32.28	(28.71)	57.06	(27.04)
Total	Dutch (N=72)	30.07	(22.55)	31.52	(19.93)
	Turkish (N=41)	16.03	(26.68)	61.67	(28.34)
	Total (N=113)	24.97	(24.95)	42.46	(27.40)

Table 37: Single-word responses by typological background and literacy per task in %, M (and SD)

Table 37 shows that, in general, children provided fewer single-word responses in the last-part repetition task than in the sentence segmentation task, and that overall, pre-literate children had fewer single-word responses than their literate peers. There seemed to be a substantial difference as well between the Dutch and Turkish children. Whether these differences were significant or not was tested in a Repeated Measures analysis, with task as within subject factor, literacy and background as between subject factors, the single-word responses as dependent variable and age as covariate. The results from this analysis are presented in Table 38.

	F	p	η_P^2
Task	2.87	.09	.03
Age	.21	.65	.002
Literacy	15.21***	.000	.12
Typological background	2.23	.14	.02
Task by age	.58	.45	.005
Task by literacy	3.53	.07	.03
Task by typological background	62.80***	.000	.37
Background by literacy	.63	.43	.006

Table 38: RM analysis with task as within subject factor, literacy and typological background as between, single-word responses as dependent variable and age as covariate

Table 38 shows that there was no main effect of age or task, but there was one of literacy. In line with our hypothesis, literate children provided significantly more single-word answers than pre-literate children. There was no main effect of typological background, suggesting that the language that children spoke did not affect the number of single-word responses they provided.

There was a significant interaction of task by typological background. This significance was caused by the fact that in the last-part repetition task, there were fewer single-word reactions by Turkish children than by Dutch children, whereas the opposite was true for the sentence segmentation task, in which the Turkish children provided more single-word reactions than the Dutch children. A possible explanation for this will be discussed in Section 8.1.

7.6   Interim conclusion

Overall, what can be concluded from these results is that typological characteristics of the children's native language did not affect the results much, while literacy significantly affected the degree to which children made use of single-word segmentations and decompositions. Among the Dutch children, the literacy effect was significant for all tasks except the last-part repetition and eye-tracking tasks, even when age was taken as a co-variable. In the last-part repetition task, the differences between pre-literate and literate children's single-word responses were not significant, and in the

eye-tracking task, the results were not clear with respect to the influence of literacy: a first analysis showed that pre-literate children were quicker to focus on the picture of the target or the counter target than the literate children, but the analysis of the routes by which children came to their fixations did not suggest a major influence of literacy. In fact, the total number of times that children's eyes crossed the depictions of all constituent parts of the multiword targets that were included in the task (thus both '*handen*', 'hands' and '*wassen*', 'wash' in the case of '*handen wassen*', 'wash hands') was so low that no significant difference could be detected. In Turkish, however, there was a literacy effect, also for the last-part repetition task, and the comparison between the responses to short and long form locatives confirmed a literacy effect as well.

In addition to the literacy effect, task effects were also found for Turkish. To what extent this can be related to the more offline or more online nature of the tasks will be discussed in Chapter 8, but for now, it can be noted that the materials prepared for the development of the tasks, and the way in which they were offered to the children, seemed to affect children's behaviour.

Frequency, assumed to be related to the entrenchment of multiword units in the mental lexicon of participants, did not have an obvious impact on the data. Higher frequency of the (target) units did not lead to significantly more multiword responses in any of the tasks, in Turkish or in Dutch, except the eye-tracking task (which was only done with Dutch participants), i.e., in the most online task that was conducted. However, this is only true for the routes that children took to come to the target or counter target; not for the speed with which they reached that goal.

The syntactic constructional type of the multiword targets did not yield clear-cut differences either. The children were given transitive noun+verb-combinations, in which both words have a lexical meaning, and locative prepositional phrases, in which only the noun varies in meaning. In Turkish, the locative combinations were significantly more often kept as wholes than the multiword transitive combinations, but in Dutch, such an effect was only found in the sentence segmentation task. It is unclear, unfortunately, whether this is a real effect or whether it might have to do with low reliability of the frequency data underlying the division into high and low frequency targets. This will be further discussed in Chapter 8.

Discussion and conclusions

This thesis has focused on the way in which monolingual Dutch, monolingual Turkish and, to a more limited extent, also bilingual Turkish-Dutch children segment language into units and how they process these units. On the basis of previous studies of metalinguistic development and of multiword units, several factors were expected to determine the results. These included whether the children were literate or not, the typological profile of the language they speak, and the frequency of co-occurring word sequences. Frequently occurring multiword combinations may be entrenched as complete units in the mental lexicon and processed as such.

In this chapter, the conclusions that derived from the analyses of the experimental tasks described in Chapter 4, 6 and 7 will be evaluated and put in a broader perspective: what are the implications of the findings, and how should we interpret them? In addition, suggestions for future research will be made.

Before we turn to this broader perspective, though, it is imperative to look once more at the findings, and see how they correspond with the more offline or more online nature of the tasks, as outlined in Chapter 3. After a discussion of the research questions in Section 8.1, Section 8.2 will discuss to what extent the tasks differ from each other, and how this relates to the proposed continuum. Finally, Section 8.3 contains some further remarks, including ideas for follow-up studies.

8.1   Reflection on general conclusions

In Chapter 2, the main research question and its sub-questions were presented. The main research question was:

What is the influence of literacy and typological background, and of the frequency of occurrence of specific multiword sequences, on the way in which children segment and on how they process spoken language?

The sub-questions posed were:

(1) How do offline and online tasks differ from each other?

(2) How can offline and online tasks be applied to investigate language segmentation and processing units among young, pre-literate children?

(3) What is the effect of typological background (being a speaker of Turkish or Dutch) on the units that children segment and process in language? And what is the effect of being bilingual in Turkish and Dutch?

(4) What is the effect of literacy on the units that children segment and process in language?

(5) Does the entrenchment of multiword sequences, as measured by frequency of occurrence of such sequences, affect the segmentation and processing of such sequences in young children?

(6) Is there an interaction between the effects of literacy, typology and the frequency of multiword sequences on the way in which children segment and process language?

This section provides the answers to each sub-question, followed by a general conclusion to answer the main question. Questions 1 and 2 will be discussed combined, as these two questions are closely connected to each other.

(1) How do offline and online tasks differ from each other?

(2) How can offline and online tasks be applied to investigate language segmentation and processing units among young, pre-literate children?

The question of how offline and online tasks differ was extensively discussed in Chapter 3. In previous studies, offline and online tasks have mainly been taken to differ in whether or not participants can or have to apply conscious knowledge, and whether or not they are granted time to think about their responses. In Chapter 3, this dichotomy between the types of tasks was questioned, and, based on previous literature on offline and online tasks and on research into children's metalinguistic development, a proposal was made for a continuum, ranging from more offline to more

online tasks. After this theoretical approach, the tasks developed for the current study were in Section 5.4 placed on the continuum. This placement will be discussed more into detail in Section 8.2.

So far, previous studies into language segmentation – which would involve more 'offline' tasks – by pre-literate and literate children and by adults have mainly focused on the development of metalinguistic awareness (see Goswami, 2009). Most studies conducted in this area raised the question whether children and illiterate adults segment specific units in language, such as words or phonemes (cf. Fox & Routh, 1975; Gombert, 1992; Holden & MacGinitie, 1972; Homer, 2000; Homer & Olson, 1999; Karpova, 1966, orig. 1955; Kolinsky, Cary & Morais, 1987; Kurvers, 2002; Kurvers & Uri, 2006; Lazo, Pumfrey & Peers, 1997; Melzer & Herse, 1969; Morais et al., 1986; Morais & Kolinsky, 1995; Morris, 1993; Olson, 1994; Ramachandra & Karanth, 2007; Ravid & Tolchinsky, 2002; Roberts, 1992; Tunmer, Bowey & Grieve, 1983). So far, the question what units people segment when they are not specifically asked to focus on particular types of units has not often been raised. It was addressed in Karmiloff-Smith et al. (1996), who asked their participants to repeat the last 'thing' in a last-part repetition task, in addition to a task in which they asked children to repeat 'the last word'. They showed that asking for the metalinguistic concept of 'words' led to different responses than when they asked for 'things'. Also Kurvers (2002) and Veldhuis (2011) asked participants to segment sentences bit-by-bit ('*in stukjes*', in Dutch). Avoiding metalinguistic terms in the instruction would allow researchers to investigate the preferred segmentation units without possibly biasing them.

In the current study, the use of specific metalinguistic terms such as 'words' was avoided in the language segmentation tasks. These tasks included a 'last-part repetition task', based on Karmiloff-Smith et al. (1996) and Kurvers and Uri (2006), a 'sentence segmentation task', as previously conducted in slightly different forms by Karpova (1966, orig. 1955), Holden and MacGinitie (1972), Ehri (1975), Berthoud-Papandropoulou (1978), Tunmer, Bowey and Grieve (1983), Bialystok (1986), Durgunoğlu and Öney (1999), and Kurvers (2002), and a 'dictation task'. This latter task had been applied before to investigate children's abilities to segment connected speech into words, in the form of a poem or a song or sentences given in advance (Chaney, 1989; Chau, 1997). In contrast to these predecessors, however, this time no text was given to the children; instead, they were asked to come up with their own story to avoid influence from well-known rhythmic patterns – particularly an issue with songs and poems. For the sentence segmentation and last-part repetition tasks, selected language

materials were prepared. Controlling for the linguistic stimulus material makes it possible to compare segmentations made by the children when provided with the same multiword sequences.

Compared to the language segmentation tasks, it was harder to find processing tasks, which are more online in nature, that could be performed by young pre-literate children and that were suitable for this study. Most experimental tasks that focus on language processing have involved reading or writing. In the end, we did find several tasks that looked promising, such as a click task, a self-paced listening task, a mixed-words task, a picture naming task, a production task and an eye-tracking task. As these tasks had generally not been used to investigate with what units people process language, adjustments had to be made, and the tasks we developed had to be tested extensively in pilot sessions. Eventually, the processing tasks that were included in this study were a click task, a mixed-words task and an eye-tracking task.

(3) What is the effect of typological background (being a speaker of Turkish or Dutch) on the units that children segment and process in language? And what is the effect of being bilingual in Turkish and Dutch?

Before any analyses were made with respect to the possible influence from typological background, literacy (see question 4 below), frequency of multi-word target items (see question 5 below) and interactions of these factors, it was investigated whether there were any correlations with sociolinguistic factors, with intelligence scores, or with school results in language proficiency tests. There turned out to be very few significant correlations (see Section 7.1).

In Chapter 4, 6 and 7, the influence of being bilingual in Turkish and Dutch and of typological background on the children's responses was analysed. The hypothesis was that the typology of a language, whether it is synthetic or agglutinative, or analytic or isolating, can be expected to influence the units that exist in people's mental lexicon. Turkish is an agglutinative language, in which most words are multi-morphemic (often combining one or more function morphemes with a content morpheme), and Dutch is a more isolating language, in which function morphemes are usually separate words. Testing children with a monolingual Turkish and a monolingual Dutch background enabled us to investigate the role of typology on children's language segmentation and processing units.

The initial idea was to include bilingual children. Studies of bilingualism have yielded ambiguous results with respect to the effects of bilingualism: Some argue that being bilingual is an advantage for the development of metalinguistic awareness, whereas others have claimed the opposite (cf. Bialystok, 1986; Bialystok, Craik & Luk, 2008; Francis, 1999). Comparing bilingual children's responses in each language to those from monolingual counterparts, would allow us to see whether there were any differences that could be ascribed to bilingualism.

The results from the pilot sessions, in which bilingual Turkish-Dutch children participated, could not confirm an influence from being bilingual (see Chapter 4): the responses from the bilingual Turkish-Dutch and the monolingual Dutch children were very similar. However, the bilingual children who participated in the pilot sessions were highly dominant in Dutch, as was confirmed by language tests that were administered (see Section 4.5.2). This will have affected the results. Since the primary school children who had participated in the pilot sessions generally appeared to be typical representatives of the bilingual community in the Netherlands, it was decided to not search further for more balanced bilingual subjects, and only work with monolingual Dutch and Turkish children.

The possible effect of typological background, which would be shown through different responses provided by the Dutch and the Turkish monolingual children on the sentence segmentation and last-part repetition tasks, was analysed in Section 7.3. Single-word responses were compared in the two languages, as word boundary-marking differs between them, Dutch being an analytical language and Turkish an agglutinative one. It turned out that there was no main effect of typological background.

There was, however, a significant interaction of task by background. This interaction was caused by the fact that the Turkish children's number of single-word responses was much lower in the last-part repetition task than that of the Dutch children, whereas the opposite was true for the sentence segmentation task.

Possibly, the fact that two different Turkish words were used to describe 'parts' or 'bits'–influenced the children's segmentations. The words used were '*bölüm*' and '*parca*'. These words can both be translated as 'part', but as can be argued about Karmiloff-Smith et al.'s (1996) choice for using the word 'thing' (instead of 'word', in their task), the connotation of these words may have affected the results. 'Thing' may also refer to the last mentioned proposition – especially for children who are often asked to repeat messages

or rules by their moms or by their school teachers. In Turkish, the word '*bölüm*' may refer to a fragment, a section or a chapter, whereas the word '*parca*' is more like a component or an item. It is not possible at this point to investigate the subjects' associations with these words, so we cannot know whether they influenced the results.

Secondly, the length of the sentences given in the task may have affected the length of multiword responses given by the children in the tasks. As discussed in Section 7.2.1, it may well be that the children provided shorter answers in the sentence segmentation task than in the last-part repetition task because of the input they obtained: the sentences in the sentence segmentation task were much shorter than those in the stories developed for the last-part repetition task. As items were also not connected to each other in the sentence segmentation task and the stimulus sentences were very simple and short, the set-up of the tasks may well have influenced the children's responses. In this case, what looks like it might be an effect of typology is really just an artefact of the methodological design.

A final methodological explanation may lie in the examples provided as part of the instructions. Several example sentences and possible responses were given for all tasks before the children were asked to perform the tasks (see Section 5.3 and Appendix 4 to 7). To avoid biasing one specific kind of segmentation, for instance along word boundaries, these examples contained both single-word segmentations and multiword ones. This way we hoped to impress on the children that they could freely come up with the answers they thought best. In the sentence segmentation task, stimulus sentences were short, so children were given single-word sample responses as well as short multiword sequences (of two or three words). In the last-part repetition task, in contrast, longer multiword sequences were also provided as possible responses, as the sentences given as examples to this task were longer. This holds especially for the Dutch versions of the task: for the Dutch children, the example sentences were also related to each other, as in the experimental material, and the first and second of the three example sentences were even repeated in their sentential context in the introduction to the task. This way of providing examples may have affected the children's responses. As children are generally known to imitate adults (cf. Tomasello, 2006), the (Turkish) children may (even more than the Dutch children) have tried to imitate the short example responses that they were given. Accordingly, their responses in the sentence segmentation task were shorter than in the last-part repetition task.

(4) What is the effect of literacy on the units that children segment and process in language?

In line with studies on the development of metalinguistic awareness (Coulmas, 1989; Homer, 2000; Morais, 1978; Kurvers, 2002; Olson, 1994, 1996; Ramachandra & Karanth, 2007), the literacy hypothesis posed in this study was that literate children would show more single-word units in their responses in the tasks, especially in the tasks that were focused on language segmentation, than their pre-literate peers, who had not yet been taught about words on paper. With respect to the tasks that focused on language processing, the hypothesis was not so strong, as previous work on language processing had so far not often looked for literacy effects (see Section 2.2). As shown in Chapter 6 and 7, the results showed a strong effect of literacy for both the Dutch and the Turkish children's single-word segmentations and processing, in all tasks.

From the qualitative analysis of the dictation task, in which children provided a large variety of different segments as units, it was concluded that pre-literate children produced significantly fewer single-word segmentations than their literate peers. Our finding that sub-word units occurred often as segments in the data from younger children may be related to the children's experience with dictation at school. It is possible that the task in which children were explicitly requested to dictate the sentences they wanted to say to the researcher, with reference to her writing them down, appeared to the children to be rather similar to the dictations that they often do in Dutch classrooms, from grade 1 on. In such exercises, the teacher provides the children with words and parts of words in the lower grades, before they are provided with longer sentences in higher grades. The small units produced by the younger children might reflect experience with this particular practice rather than what segments the children work with in practice when producing speech.

The quantitative analyses of the data obtained from the other tasks, both in Dutch and in Turkish, also showed a clear influence of literacy. In all tasks except the Dutch last-part repetition task, there was a main effect of literacy on children's single-word responses. The exception posed by the Dutch last-part repetition task may be related to the task itself: both in the Dutch and in the Turkish version of the task, the literate children did not clearly prefer a single-word segmentation strategy, as they did in the other tasks. Instead, they also produced a lot of multiword segmentations, including the target units inserted into the task as experimental items. In general, the literate children thus seemed to prefer responses that were longer than

single words in this task, rather than the single-word responses that they were expected to give on the basis of the literacy hypothesis. As discussed in Chapter 7, the last-part repetition task involved a meaningful, and for the children rather exciting, story, that was followed with interest. The pauses that were made by the researcher when reading the story may have been a source of irritation for the children, or as an indication for them – following what teachers in class do – to summarize past events or to anticipate upcoming events in the story. This may have contributed to the longer responses that children provided in the task.

A literacy effect was also found in the eye-tracking task. While listening to sentences that contained selected frequent and less frequent multiword sequences, children's viewing paths suggested that literates showed more segmental processing of the constituent words within multiword target units than the pre-literate children. This would be in line with previous studies, for instance Arnon (2010), who suggested that children start out with bigger chunks, and only come to segment such chunks into smaller parts as they get older. Possibly, this trend towards increasing decomposition of larger chunks is not merely related to age, but also to literacy, as it stimulates children to work with individual words.

The findings suggest a major role for literacy. That the effect sizes, partial eta squared, ranged from .12 to .29, showing medium to large effects, further supports this conclusion.

(5) Does the entrenchment of multiword sequences, as measured by frequency of occurrence of such sequences, affect the segmentation and processing of such sequences in young children?

The hypothesis with respect to frequency was mainly based on results from studies on multiword units in the tradition of Cognitive Linguistics. Most of these studies worked with adult subjects (cf. Barlow & Kemmer, 2000; Taylor, 2002; see also Section 2.2). Researchers have generally concluded that a higher frequency of occurrence of multiword sequences leads to better entrenchment in the mental lexicon, which, in turn, leads them to be processed to holistically.

In the current study, it was examined whether a similar story could be told about the way in which young pre-literate and literate children process units in language. In addition, it was investigated to what extent the frequency of occurrence of particular multiword sequences that were expected to be processed as wholes, would also be segmented holistically.

In order to find an answer to these questions, the same multiword sequences – consisting of either transitive noun+verb-combinations or of locative prepositional (in Dutch) and postpositional (in Turkish) phrases – were included in all but the dictation and eye-tracking tasks. These multiword sequences were checked for their frequency of occurrence in Google, which was used as a reference corpus. They were then included in a bin of 'high frequency items' or in a bin of 'low frequency items', and this was taken as an indication for their probable entrenchment levels: the results would then indicate to what extent we could find positive or negative evidence for the hypothesis that there would be a frequency effect.

As described in Chapter 6 and 7, the effect of frequency was not as clear as that of literacy. Several reasons for this can be considered.

First of all, as mentioned in Section 6.7, the frequency data as obtained from Google may not correspond well to children's language use, and therefore not provide good information on children's mental lexicon. A recent study in which the adult corpus SONAR and the child language corpus BasiLex were compared, showed considerable differences in the frequencies of occurrence of individual words (Tellings et al., submitted). Even though children's language development has been claimed to be largely dependent upon adults' language use, the language to which children are exposed in their daily lives differs from the mostly written corpus that is constituted by Google, which mostly features adult language use. This difference may be the reason why no large impact of frequency was found, even though the rather fine-grained measure of EMI- and MI-values was used to rank the frequencies of the multiword units used in the tasks, and even though these were compared in *Woordwerken* (Schrooten & Vermeer, 1994), to check whether children could reasonably be expected to be familiar with the language used in our experiments.

Perhaps, if a comparison had been possible between the stimulus items and a child language corpus, the results would have differed. Since such a corpus was, and is, not available in Dutch, the possibilities to investigate the effects of frequency of occurrence on the entrenchment of multiword sequences in children's mental lexicons are fairly limited. The option of creating a limited reference corpus, as Bannard and Matthews (2008), seemed not attractive, as young children's knowledge of language – including the multiword sequences occurring in speech – varies appreciably per person, and a small corpus would therefore soon become unreliable.

Secondly, it may be the case that it is not (merely) frequency which affects the entrenchment of multiword units in the mental lexicon. Other

factors may be at stake, such as diversity in language use, discourse function, salience and complexity (cf. Brandt et al., 2011). In ongoing research, Verhagen (see Verhagen, 2011) investigates whether factors such as individual differences in people's language use, pre-emption or salience affects the entrenchment of multiword sequences. Her experimental and focus group data show that indeed there is considerable individual variation in the extent to which subjects judge specific multiword sequences to be lexical units. In that sense, looking only at frequency as the basis for the analysis of the extent to which multiword sequences are segmented and processed as wholes may not necessarily suffice (also see Bannard & Matthews' (2008) claims about individual variation among children).

Third, the reason why there was no clear frequency effect may be related to the strong literacy-effects found: literacy may have overruled any other factors, due to which any effects of frequency were not very visible anymore. A more detailed analysis of how answers compared within more homogenous groups, for instance only pre-literate children or only literate children from a specific grade, could provide information on this matter.

Another issue, mentioned in Section 7.4, is that the relation between frequency of occurrence of a multiword sequence and its complete entrenchment works differently for adults than for young children. As the latter are still learning the language, a higher frequency of a multiword unit may not just lead to its storage, but also to a better knowledge or recognition of the constituent parts. Recall that the hypothesis was developed on the basis of previous studies with adults. For children a sequence such as '*handen wassen*' ('wash hands') may trigger segmentation and processing of the single words 'wash' and 'hands' more easily.

The effect of constructional type, whether the selected multiword sequences were transitive or locative combinations, was also investigated. The hypothesis behind taking these two types of combinations as target items was that they have both been demonstrated to function as processing units (cf. Doğruöz & Backus, 2009; Goldberg, 2006; Langacker, 2008; Tomasello, 2003, and Section 2.2.1). Whether the two types of constructions would be dealt with in a similar way by the children was, however, an open question. We were curious to see whether their different constituent parts (with the locative consisting of two closed class words and one from the open class, and the transitive construction consisting of only open class words, at least in Dutch) and the differences between the two languages (see Section 2.3) would lead to differences in responses.

Our analyses, as discussed in Chapter 7, showed that for Dutch there was only a main effect of constructional type in the sentence segmentation task, but that for Turkish constructional type was found to lead to a main effect in both the tasks conducted in this language. The Dutch finding resulted from the fact that children provided more target unit responses for the transitive than for the locative combinations in the sentence segmentation task. This was also the case in the last-part repetition task, but as the difference between the two constructional types was smaller in this task, no significant levels were reached. In the Turkish tasks, too, children significantly more often provided transitive than locative combinations as wholes. This was possibly related to the fact that the Turkish children in general provided a larger variety of responses after the locative combinations (which thus led to a lower mean of target unit responses, as the answer categories were dependent on each other, see Chapter 5).

As for the effect of constructional type, it could thus only be cautiously concluded that the Turkish children reacted differently to transitive combinations with the schematic template [N V] than to locative combinations with the structure [N+GEN POST+GEN.LOC]. For the Dutch children, this effect was only visible in one task, and therefore, no far-reaching conclusions could be drawn here. The reason for this limited effect of constructional type in Dutch may however also be related to the expressions that were selected as target items: they all contained clear meaningful constituent words, and they consisted of three words at most. A study in which different multiword sequences, with different schematic templates (for instance with more open slots), could provide further insights into the differences between specific constructional types on the way in which they are processed as wholes.

(6)   Is there an interaction between the effects of literacy, typology and the frequency of multiword sequences on the way in which children segment and process language?

The analyses of the influence of literacy, typological background and the frequency of the multiword target units in the tasks conducted in Dutch showed that there were no significant interactions of these factors on the answers that children provided in the tasks. There were significant interactions for the number of multiword target unit responses in the sentence segmentation and mixed-words tasks, but these involved interactions of constructional type by age (in the sentence segmentation task) and of frequency by constructional type (in the mixed-words task). The

factors that established the main focus in this study were not found to interact in any of the tasks conducted in Dutch.

In Turkish, significant interactions of the three factors were not found either. It appeared from the analyses of the single-word responses the children provided in the two tasks they performed that there were four significant interactions, but these involved only tasks by literacy, frequency by tasks, tasks by constructional type and frequency by constructional type. Significant interactions in which literacy and frequency played a role, were nowhere found.

The analysis of the influence of typological background, thus the comparison in responses provided in the tasks in Turkish and Dutch, did not show significant interactions with literacy or frequency either. The only interaction found in this comparison comprised of an interaction of tasks by typological background, while the difference in answers provided in the different tasks was in general found to be substantial as well.

Overall, however, literacy seemed to be the most affective factor, as a main effect of literacy was often found: there were more single-word responses in all tasks than multiword responses.

Answer to the main research question
All in all, this study showed an effect of literacy in all of the tasks both on children's language segmentation units and on the units they employed during language processing. There is, on the other hand, no unambiguous evidence confirming any impact of typological background and of the frequency of occurrence of particular multiword sequences. Further research would be needed to explore these issues (see Section 8.3).

8.2   Task effects and the placement of tasks on the proposed continuum

In Chapter 3, a proposal was advanced to distinguish experimental tasks on a continuum going from offline to online, as the simple dichotomy that is generally used seems not to suffice. By means of a set of ten criteria (see Section 3.2.2), the tasks conducted for the present study were placed on this continuum. The provisional placement of tasks was as follows (Figure 10):

OFFLINE ←———|————|————|——————————|——→ ONLINE
          Sent. Segm.   Last-part rep./Dict.   Mixed words/Click   Eye tracking

Figure 10: Overview of tasks along the offline-online continuum

The sentence segmentation task was assumed to be the most offline in nature, and the eye-tracking task the most online. The other tasks all had to be located somewhere around the middle of the continuum, which can be regarded as support for the idea that a two-way distinction between purely offline and purely online tasks is hard to maintain.

With respect to this placement of tasks on the continuum, it should be noted that their location could change if the ten criteria could be weighted. Still, it is not likely that this would drastically change the general degree of online- or offline-ness of the individual tasks.

One criterion that should also be addressed in the future when placing tasks on the continuum involves the role of the researcher in the interpretation of data obtained from the tasks. The criteria developed in Chapter 3 only addressed the set-up of tasks and requirements made of participants. The role of the researcher, e.g., what he/she needs to do with the data, and to what extent he/she needs to interpret data, was ignored. Possibly, however, a task in which the researcher is more free to interpret the data – as was for instance the case with the click task in this study – could be regarded as more offline in nature than tasks in which there is little space for discussion about what the data tell us.

The criteria that were considered in developing the continuum need to be reassessed as well. For example, while the instructions given to participants were considered, e.g., whether or not participants are aware of the (specific) purpose of a task, there are other aspects of the instruction that may well play a role. It makes a difference whether or not the instruction is very explicit, or what kinds of examples are given in the instruction. As mentioned in Section 8.1 above, children often tend to imitate adults, and it might be that a task with many examples leaves less space for children for their own thoughts or input, reducing the degree to which the task is online. Not providing any examples would have the opposite effect.

The same holds for how the set-up of a task figures in assigning a task its place on the continuum. It makes a difference whether the stimulus material consists of single, unrelated sentences or looks more like natural language use, for instance in the form of a story or conversation. The results sometimes suggested that the set-up of the tasks seemed to have affected the findings. The occurrence of short sentences in the sentence segmentation task seemed to lead to short responses, while the way in which children were engrossed in the story in the last-part repetition task seems to have led to more anticipations and longer responses. Factors like these influence children's attention to the task or their involvement in the meaning of what is said, and this may show in their responses.

Finally, the relation between a task and experiences that participants know from their own lives may be considered. The dictation task made explicit reference to writing and dictating, skills that children practice at school on a daily basis. Possibly, this reference, the children's 'world knowledge' about these two skills, and their related ideas on what written language looks like, affected their performance in the tasks, in which they wanted to provide correct responses. In a similar vein, the story in the last-part repetition task, with its target items occurring in a context, may have represented a relatively natural setting for the young children who participated in the tasks – they often listen to stories read out aloud to them, and are then usually focused on the message or the meaning of the story – as opposed to the unrelated sequence of short sentences that occurred in the sentence segmentation task. Such experience with the form of a task will influence the way in which children respond, as they are likely to provide answers that are similar to the ones they are usually supposed to give, for instance in class. Accordingly, the children will focus more consciously on the task and the responses they provide. This will make the task more offline than expected in advance.

In general, however, it can be concluded that the dichotomy that is generally assumed between offline and online tasks obscures too many details. A continuum is better suited to the diversity of factors that are involved in determining the degree to which a task can be considered either 'offline' or 'online'. Further investigation of the factors that determine the placement of a particular task on the continuum is, however, still needed.

8.3   Further remarks

I hope to have shown that this study into children's language segmentation and processing units has made use of rather innovative methodology, in which various factors were analysed at the same time, and for which new experiments had to be developed. The biggest challenge was that half of all participants were pre-literate. A battery of tasks was developed, and in all of them except the dictation task specific multiword sequences were included as test items to see whether children processed and segmented these as wholes, or in smaller constituent parts.

The selection of the multiword targets that were included, consisted of transitive and locative constructions, such as 'wash hands' and 'in the bath'. Their familiarity among the children who participated in the tasks was

checked via *Woordwerken* (Schrooten & Vermeer, 1994), and their frequency of occurrence was analysed using Google. The drawback of using this corpus was that it did not reflect language used towards and by children well. For a future study, using a child language corpus that is large enough to be representative of child language in general, language which is said to vary highly across individuals, would be a prerequisite. Possibly a corpus like Childes, but larger than it currently is for the Dutch data, would form a better basis for frequency assessments of child language. As such a corpus, with a larger size, does so far not consist for Dutch, however, it may be necessary to focus on children with a different language background.

Testing children who are used to a different writing convention or orthography may be interesting as well. As mentioned in Chapter 2 in relation to previous studies, it can be expected that children who are used to languages in which word boundaries are not marked in writing, such as Chinese or Japanese, show different behaviour concerning word boundary marking (cf. Bassetti, 2005; Chau, 1997; Homer, 2000; Hoosain, 1992; Veldhuis, Li & Kurvers, 2010; Veldhuis, 2011). Testing such children could not only provide more evidence about the psychological reality of words (cf. Olson, 1994), studying them could also show whether there are general or even universal units, e.g., units that get entrenched on the basis of frequency of occurrence, that are segmented and processed most easily. If universal patterns are not found, or only appear in the more offline tasks, then testing children with different writing systems and conventions in word marking could lead to further insights into the way in which graphic relativity (cf. Bugarski, 1993) affects language processing. Thus far, effects from graphic relativity have only been investigated on language segmentation; to see to what extent knowledge of specific writing systems affect people's language processing units as well is unknown.

Another issue that should be addressed is the effect of schooling. In order to interpret our results, the schools' curriculums were examined, but in a future study, this issue should be addressed in more detail. It is likely that the kind of literacy education that children – or adults – receive affects their metalinguistic notions about what constitutes the units of language.

That said, the set-up of the current study is not the only type that is possible. For the current study, mostly quantitative data were collected, from experiments that were prepared in advance, though item selection was done on the basis of natural corpus information. Collecting more information from observations of natural speech, which would involve more 'online' information than can be collected in experimental studies, may lead

to further insights into the issues central in this study. Likewise, the ever more popular neurolinguistic techniques, such as fMRI's and ERP-scans could be used to elicit information on the units children use (cf. Sekerina, Fernández & Clashen, 2008). For the current study, using these techniques was not an option.

The way in which target items were selected may also be reassessed. For the current study, it was decided to take two constructional types and analyse in a top-down approach whether children segmented those constructions or not (see Chapter 5). Doing a similar study with more constructional types is needed to see whether the findings from this study are representative for the processing of constructions in general.

Another matter that may be reconsidered in the future regards the participants in a study like the present one. In the current study, children from different schools and age groups were compared (see Chapter 5). A more extensive longitudinal study could possibly shed light on whether there is a specific sequence of units that children mostly come up with at different stages in their (language and literacy) development.

Moreover, in the main study only children aged four to nine were tested. These children had all acquired the language they were tested in as their first language. Arnon (2010) showed that adults tend to base linguistic decisions upon (metalinguistic) rules, so it would be interesting to test whether adults, who have been exposed to such rules for much longer than the oldest children who were tested in the current study, segment and process units in language differently. Does it matter that adults tend to apply more analytic learning strategies, i.e., do these learning strategies relate to language segmentation and processing? Does this differ between literate to illiterate adults?

Finally, a question that could be addressed in the future is to what extent bilingualism affects the units employed in language segmentation and processing. The pilot sessions with bilingual children showed that testing bilingual Turkish-Dutch children involved unexpected problems as the large majority of the Turkish-Dutch primary school children appeared to be highly dominant in Dutch, and only able to speak Turkish to a limited extent. As, however, it has been argued that bilinguals process language differently from monolinguals and that their metalinguistic development differs from monolinguals (cf. Bialystok, 1986a; Francis, 1999), the issue is still interesting.

References

Abrams, K. & Bever, T.G. (1969). Syntactic structure modifies attention during speech perception and recognition. *Quarterly Journal of Experimental Psychology, 21*, 280-290.

Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Altmann, G.T.M. (2004). Language-mediated eye movements in the absence of a visual world: The 'blank screen paradigm'. *Cognition, 93*. B79-B87.

Anderson, R.C., Goldberg, S.R. & Hidde, J.L. (1976). Meaningful processing of sentences. In H. Singer & R.B. Ruddell (Eds.), *Theoretical models and processing of reading*. Newark: International Reading Association.

Arnon, I. (2010). *Starting big: The role of multiword phrases in language learning and use*. Dissertation, re-distributed by Standford University.

Arnon, I. & Snider, N. (2010). More than words: Frequency effects for multiword phrases. *Journal of Memory and Language, 62*, 67-82.

Audacity [Computer program]. Retrieved from:*http://audacity.sourceforge.net/*

Bannard, C. & Lieven, E. (2009). Repetition and reuse in child language learning. In R. Corrigan, E. Moravcsik, H. Ouali & K. Wheatley (Eds.), *Formulaic language. Volume II: Acquisition, loss, psychological reality, functional explanations.* Amsterdam: John Benjamins.

Bannard, C. & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science, 19* (3), 241-248.

Barlow, M. & Kemmer, S. (2000). *Usage based models of language.* Stanford: Center for the Study of Language and Information Publications.

Bassetti, B. (2005). Effects of writing systems on second language awareness: Word awareness in English learners of Chinese as a foreign language. In V. Cook & B. Bassetti (Eds.), *Second language writing systems*. Clevedon: Multilingual Matters.

Bassetti, B. (2007). Bilingualism, biliteracy and metalinguistic awareness: Word awareness in English and Japanese users of Chinese as a second Language. *Birkbeck Studies in Applied Linguistics*, 2, 1-21.

Bassetti, B. (2009). Effects of adding interword spacing on Chinese reading: A comparison of Chinese native readers and English readers of Chinese as a second language. *Applied Psycholinguistics*, 30, 757-775.

Benveniste (1974) in Gombert (1992). *Metalinguistic development*. New York: Harvester Wheatsheaf.

Berthoud-Papandropoulou, I. (1978). An experimental study of children's ideas about language. In I. Sinclair et al. (Eds.), *The child's conception of language.* Berlin: Springer-Verlag.

Bialystok, E. (1986a). Factors in the growth of linguistic awareness. *Child Development*, 57, 498-510.

Bialystok, E. (1986b). Children's concept of word. *Journal of Psycholinguistic Research*, 15, 13-32.

Bialystok, E. (2007). Acquisition of literacy in bilingual children: A framework for research. *Language Learning*, 57, 45-77.

Bialystok, E., Craik, F. & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 34, 859-873.

Blumenfield, H.K. & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye tracking. *Language and Cognitive Processes*, 22 (5), 633-660.

Boersma, P. & Weenink, D. (2012). *Praat: Doing phonetics by computer* [Computer program]. Retrieved from: *http://www.praat.org/*

Bradley, L. & Bryant, P. (1983). Categorizing sounds and learning to read: A causal connection. *Nature*, 301, 419-421.

Brandt, S., Verhagen, A., Lieven, E. & Tomasello, M. (2011). German children's productivity with simple transitive and complement-clause constructions: Testing the effects of frequency and diversity. *Cognitive Linguistics*, 22 (2), 325-357.

Brown, P.F., Della Pietra, V.J., De Souza, P.V., Lai, J.C. & Mercer, R.L. (1992). Class-based n-gram models of natural language. *Computational Linguistics*, 18 (4), 467-479.

Bugarski, R. (1993). Graphic relativity and linguistic constructs. In R.J. Scholes (Ed.), *Literacy and language analysis.* Hillsdale: Lawrence Erlbaum.

Bybee, J. (2002). Phonological evidence for exemplar storage of multiword sequences. *SSLA*, 24, 215-221.

Carreiras, M., Seghier, M., Baquero, S., Estévez, A., Lozano, A., Devlin, J.T. & Price, C.J. (2009). An anatomical signature for literacy. *Nature*, *461*, 983-U245.

Chaney, C. (1989). I pledge a legiance tothe flag: Three studies in word segmentation. *Applied Psycholinguistics*, *10*, 261-281.

Chaney, C. (1992). Language development, metalinguistic skills, and print awareness in 3-year-old children. *Applied Psycholinguistics*, *13*, 485-514.

Chapin, P.G., Smith, T.S. & Abrahamson, A.A. (1972). Two factors in perceptual speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, *11*, 164-173.

Chau, L.Y. (1997). *The development of the 'word unit' concept by Cantonese-speaking children*. BA-thesis, University of Hong Kong.

Chomsky, N. (1957). *Syntactic structures.* Berlin: Walter de Gruyter GmbH & Co.

Chomsky, N. (1964) in Sekerina, I.A., Fernández, E.M. & Clashen, H. (2008). *Developmental psycholinguistics: On-line methods in children's language processing*. Amsterdam: John Benjamins.

Christophe, A., Dupoux, E., Bertoncini, J. & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America*, *95* (3), 1570-1580.

Church, K.W. & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics*, *16* (1), 22-29.

Clahsen, H. (2008). Behavorial methods for investigating morphological and syntactic processing in children. In I.A. Sekerina et al. (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing*. Amsterdam: Benjamins.

Clark, E.V. (1978). Awareness of language: Some evidence from what children say and do. In I. Sinclair et al. (Eds.), *The child's conception of language*. Berlin: Springer-Verlag.

Cohen, L. & Mehler, J. (1996). Click monitoring revisited: An on-line study of sentence comprehension. *Memory & Cognition*, *24* (1), 94-102.

Conklin, K. & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, *29*, 72-89.

Cook, V. (2011). Relating language and cognition: The speaker of one language. In V. Cook & B. Bassetti (Eds.), *Language and bilingual cognition.* New York: Psychology Press.

Cooper, R. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech

perception, memory, and language processing. *Cognitive Psychology*, *6*, 84-107.

Correa, J. & Dockrell, J.E. (2007). Unconventional word segmentation in Brazilian children's early text production. *Journal of Reading and Writing*, *20*, 815-831.

Coulmas, F. (1989). Units of speech and units of writing. In F. Coulmas (Ed.), *The writing systems of the world*. Oxford: Basil Blackwell.

Cozijn, R. (2006). Het gebruik van oogbewegingen in leesonderzoek. *Tijdschrift voor Taalbeheersing*, *28*, 220–232.

Cozijn, R., Commandeur, E., Vonk, W. & Noordman, L.G.M. (2011). The time course of the use of implicit causality information in the processing of pronouns: A visual world paradigm study. *Journal of Memory and Language*, *64*, 381-403.

Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Cutler, A., Mehler, J., Norris, D. & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, *24*, 381-410.

Cutler, A., Kearns, R., Norris, D. & Scott, D. (1993). Problems with click detection: Insights from cross-linguistic comparisons. *Speech Communication*, *13*, 401-410.

Dąbrowska, E. (2004). Rules or schemas? Evidence from Polish. *Language and Cognitive Processes*, *19*, 225-271.

Daniels, P. (2009) Grammatalogy. In D. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy*. Cambridge: Cambridge University Press.

Dellatolas, G., Braga, L.W., Souza Ldo, N., Filho, G.N., Queiroz, E. & Deloche, G. (2003). Cognitive consequences of early phase of literacy. *Journal of the International Neuropsychological Society*, *9*, 771-782.

Delogu, C., Conte, S. & Sementina, C. (1998). Cognitive factors in the evaluation of synthetic speech. *Speech Communication*, *24*, 153-168.

Doğruöz, A.S. & Backus, A. (2009). Innovative constructions in Dutch Turkish: An assessment of on-going contact-induces change. *Bilingualism: Language and Cognition*, *12* (1), 41-64.

Doherty, M.J. & Perner, J. (1998). Metalinguistic awareness and theory of mind: Just two words for the same thing? *Cognitive Development*, *13*, 279-305.

Durgunoğlu, A.Y. & Öney, B. (1999). A cross-linguistic comparison of phonological awareness and word recognition. *Reading & Writing*, *11*, 281-299.

Ehri, L.C. (1975). Word consciousness in readers and prereaders. *Journal of Educational Psychology*, *67*, 204-212.

Ehrismann, M. (2009). *Eén enkele eenheid? De wijze waarop gefixeerde expressies in het lexicon zijn opgeslagen*. Unpublished MA-thesis, Tilburg University, the Netherlands.

Ellis, N.C. & Cadierno, T. (2009). Constructing a second language: Introduction to the special section. *Annual Review of Cognitive Linguistics*, *7*, 111- 139.

Evans, M.A. & Saint-Aubin, J. (2005). What are children looking at during shared storybook reading: Evidence from eye movement monitoring. *Psychological Science*, *16*, 913-920.

Fernald, A., Zangl, R., Portillo, A.L. & Marchman, V.A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I.A. Sekerina et al. (Eds), *Developmental psycholinguistics: On-line methods in children's language processing*. Amsterdam: John Benjamins.

Fernández, E.M. & Smith Cairns, H. (2011). *Fundamentals of psycholinguistics.* Oxford: Wiley-Blackwell.

Field, J. (2004). *Psycholinguistics: the key concepts.* London: Routledge.

Fodor, J.A. & Bever, T.G. (1965). The psychological reality of linguistic elements. *Journal of Verbal Learning and Verbal Behavior*, *4*, 414-420.

Fox, B. & Routh, D.K. (1975). Analyzing spoken language into words, syllables and phonemes: A developmental study. *Journal of Psycholinguistic Research*, *4*, 331-342.

Francis, N. (1999). Bilingualism, writing, and metalinguistic awareness: Oral-literate interactions between first and second languages. *Applied Psycholinguistics, 20,* 533-561.

Fukuda, K. (1983). Click monitoring and the perceptual segmentation of speech sequences. *Psychologia*, *26*, 214-222.

Geva, E. (2006). Learning to read in a second language: Research, implications and recommendations for services. In R. Tremblay, R. Barr & R. Peters (Eds.), *Encyclopedia on early childhood development.* Montreal, CQ: Centre of Excellence for Early Childhood Development.

Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Gombert, J.E. (1992). *Metalinguistic development*. New York: Harvester Wheatsheaf.

Goswami, U. (2009). The basic processes in reading: Insights from neuroscience. In D.R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy*. Cambridge: Cambridge University Press.

Green, D.W. (1977). The immediate processing of sentences. *Quarterly Journal of Experimental Psychology*, *29*, 135-146.

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

Gürel, A. (1999). Decomposition: To what extent? The case of Turkish. *Brain and Language*, *68*, 218-224.

Harris, R. (2009). Speech and writing. In D.R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy.* Cambridge: Cambridge University Press.

Havelock, E. (1986). *The muse learns to write: Reflections on orality and literacy from antiquity to the present.* New Haven: Yale University Press.

Hirsh-Pasek, K., Gleitman, L.R. & Gleitman, H. (1978). What did the brain say to the mind? A study of the detection and report of ambiguity by young children. In I. Sinclair et al. (Eds.), *The child's conception of language*. Berlin: Springer-Verlag.

Holden, M.H. & MacGinitie, W.H. (1972). Children's conception of word-boundaries in speech and print. *Journal of Educational Psychology*, *63*, 551-557.

Homer, B.D. (2000). *Literacy and metalinguistic awareness: A cross-cultural study*. Dissertation, National Library Canada.

Homer, B.D. (2009). Literacy and metalinguistic development. In D.R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy*. Cambridge: Cambridge University Press.

Homer, B.D. & Olson, D.R. (1999). The role of literacy in children's concept of word. *Written Language and Literacy*, *2*, 113-140.

Hoosain, R. (1992). Psychological reality of the word in Chinese. *Advances in Psychology*, *90*, 111-130.

Houston, D.M., Santelmann, L.M. & Jusczyk, P.W. (2004). English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes*, *19* (1), 97-136.

Huizenga, H. (2010). *Aanvankelijk en technisch lezen.* Groningen: Noordhoff.

Huizenga, H. & Robbe, R. (2013). *Basiskennis taalonderwijs.* Groningen: Noordhoff.

Hsia, S. (1992). Developmental knowledge of inter- and intraword boundaries: Evidence from American and Mandarin Chinese speaking beginning readers. *Applied Psycholinguistics*, *13*, 341-372.

Hyönä, J., Niemi, P. & Underwood, G. (1989). Reading long words embedded in sentences: Informativeness of word halves affects eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 68-71.

Inagaki, K., Hatano, G. & Otake, T. (2000). The effect of *kana* literacy on the speech segmentation unit used by Japanese young children. *Journal of Experimental Child Psychology*, *75*, 70-91.

Juffermans, K. & Veldhuis, D. (2012). Comparing spelling and segmentation practices in three versions of a Mandinka text. *Tilburg Papers in Culture Studies*, 34. Tilburg: Tilburg University.

Juel, C. (2009). Beginning reading. In D.R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy*. Cambridge: Cambridge University Press.

Jusczyk, P.W. & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1-23.

Kamide, Y., Altmann, G.T.M. & Haywood, S.L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*, 133-156.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science.* Cambridge, MA: MIT Press.

Karmiloff-Smith, A., Grant, J., Sims, K., Jones, M.C. & Cuckle, P. (1996). Rethinking metalinguistic awareness: Representing and accessing knowledge about what counts as a word. *Cognition*, *58*, 197-219.

Karpova, S.N. (1966, orig. 1955). Osoznanie slovesnogo sostava rechi rebenkom doshkol'nogo vozrasta (The preschoolers' realisation of the lexical structure of speech). *Voposy Psikhol*, *4*, 32-55. Summary in: F. Smith & G.A. Miller (1966), *The genesis of language: a psycholinguistic approach*. Cambridge, MA: MIT Press.

Keller, F. & Lapata, M. (2003). Using the web to obtain frequencies for unseen bigrams. *Computational Linguistics*, *29* (3), 459-484.

Kemper, S. & Liu, C-J. (2007). Eye movements of young and older adults during reading. *Psychology and Aging*, *22*, 84-93.

Kolinsky, R., Cary, L. & Morais, J. (1987). Awareness of words as phonological entities: The role of literacy. *Applied Psycholinguistics*, *8*, 223-232.

Kosmidis, M.H., Tsapkini, K. & Folia, V. (2006). Lexical processing in illiteracy: Effect of literacy or education? *Cortex*, *42*, 1021-1027.

Kraak, A. (2006). *Homo loquens en homo scribens: Over natuur en cultuur bij taal.* Amsterdam: Amsterdam University Press.

Kurvers, J. (2002). *Met ongeletterde ogen. Kennis van taal en schrift van analfabeten.* Amsterdam: Aksant Academic Publishers.

Kurvers, J. & Uri, H. (2006). Metalexical awareness: Development, methodology or written language? A cross-linguistic comparison. *Journal of Psycholinguistic Research*, *35*, 353-367.

Ladefoged, P. & Broadbent, D.E. (1960). Perception of sequence in auditory events. *Quarterly Journal of Experimental Psychology*, *12* (3), 162-170.

Laghzaoui, M. (2011). *Emergent academic language at home and at school: A longitudinal study of 3- to 6-year-old Moroccan Berber children in the Netherlands.* Oisterwijk: Uitgeverij BOXPress.

Langacker, R.W. (2002). *Concept, image and symbol: The cognitive basis of grammar.* Berlin: Mouton de Gruyter.

Langacker, R.W. (2008). *Cognitive grammar: A basic introduction.* Oxford: Oxford University Press.

Lazo, M.G., Pumfrey, P.D. & Peers, I. (1997). Metalinguistic awareness, reading, and spelling: Roots and branches of literacy. *Journal of Research in Reading, 20,* 85-114.

Lehtonen, M. & Laine, M. (2003). How word frequency affects morphological processing in monolinguals and bilinguals. *Bilingualism: Language and Cognition, 6,* 213–225.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

Levinson, S.C. (1996). Language and Space. *Annual Review of Anthropology, 25,* 353-382.

Lin, T-J., Anderson, R.C., Ku, Y-M., Christianson, K. & Packard, J.L. (2011). Chinese children's concept of word. *Writing Systems Research, 3* (1), 41-57.

Litjens, P. (2001). Verhalen vertellen. In A. Mottart (Ed.), *Retoriek en praktijk van het schoolvak Nederlands 2000*. Gent: Academia Press.

Litjens, P., Schripsema, H., Damhuis, R. & Moggré, R. (2001). Kinderen vertellen verhalen. *De wereld van het jonge kind, 03,* 218-221.

Männel, C. & Friederici, A.D. (2008). Event-related brain potentials as a window to children's language processing. In I.A. Sekerina et al. (Eds), *Developmental psycholinguistics: On-line methods in children's language processing*. Amsterdam: John Benjamins.

Marinis, T. (2003). Psycholinguistic techniques in second language acquisition research. *Second Language Research, 19,* 144-161.

Marshall, J.C. & Morton, J. (1978). On the mechanics of Emma. In I. Sinclair et al. (Eds.), *The child's conception of language*. Berlin: Springer-Verlag.

McLuhan, M. (1964). *Understanding media: The extensions of man*. New York: McGraw-Hill.

Melzer, N.S. & Herse, R. (1969). The boundaries of written words as seen by first graders. *Journal of Literacy Research, 1* (3), 3-14.

Meroni, L. & Crain, S. (2003). On not being led down the kindergarten path. *Proceedings of the 27th Boston University Conference on Language Development* (pp. 531-544). Somerville, MA: Cascadilla Press.

Mitchell, D.C. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras & Ch. Clifton Jr. (Eds.), *The on-line*

study of sentence comprehension: Eyetracking, ERPs and beyond. New York: Psychology Press.

Morais, J. (1978). Spatial constraints on attention to speech. In J. Requin (Ed.), *Attention and Performance VII* (pp. 245-260). Hillsdale, NJ: Erlbaum.

Morais, J., Cary, L., Alegria, J. & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, *7*, 323-331.

Morais, J., Bertelson, P., Cary, L. & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition*, *24*, 45-64.

Morais, J., Alegria, J. & Content, A. (1987). The relationships between segmental analysis and alphabetic literacy: An interactive view. *Current psychology of cognition*, *7*, 415-438.

Morais, J. & Kolinsky, R. (1995). The consequences of phonemic awareness. In B. de Gelder & J. Morais (Eds.), *Speech and reading. A comparative approach.* London: Erlbaum, Taylor & Francis.

Morris, D. (1993). The relationship between children's concept of word in text and phoneme awareness in learning to read: A longitudinal study. *Research in the Teaching of English*, *27*, 133-154.

Mos, M. (2010). *Complex lexical items.* Utrecht: LOT.

Narain, G. & Verhoeven, L. (1994). Ontwikkeling van tweetaligheid bij allochtone kleuters. In R. Appel, L. Verhoeven & K. Jaspaert (Eds.), *Studies in Meertaligheid*, *5.* Tilburg: Tilburg University Press.

Nederlandse Taalunie (2004). *Corpus Gesproken Nederlands (CGN)* [Corpus of Spoken Dutch]. Retrieved from: *http://lands.let.ru.nl/cgn/home.htm*

Nederlands Jeugd Instituut (NJI) (2013, April 16). *Verdeling allochtone jeugd naar herkomst (2011)*. Retrieved from: *http://www.nji.nl/nl/Kennis/ Databanken/Cijfers-over-Jeugd-en-Opvoeding/Cijfers-over-Migrantenjeugd? view=content&refid=&recordnr=10000320&voor= &toon=tabel&tabelnr=2*

Nicoladis, E. & Genesee, F. (1996). Word awareness in second language learners and bilingual children. *Language Awareness*, *5*, 80-90.

Olson, D.R. (1994). *The world on paper. The conceptual and cognitive implications of writing and reading.* Cambridge: Cambridge University Press.

Olson, D.R. (1996). Towards a psychology of literacy: On the relations between speech and writing. *Cognition*, *60*, 83-104.

Ostrosky-Solis, F., Arellano Garcia, M. & Perez, P. (2004). Can learning to read and write change the brain organization? An electrophysiological study. *International journal of psychology*, *39*, 27-35.

Papandropoulou, I. & Sinclair, H. (1974). What is a word? Experimental study of children's ideas on grammar. *Human development*, *17*, 241-258.

Pattamadilok, C., Knierim, I., Kawabata Duncan, K.J. & Devlin, J.T. (2010). How does learning to read affect speech perception? *Journal of Neuroscience*, *30* (25), 8435-8444.

Pellicer, A. (2004). Segmentation in the writing of Mayan language statements by indigenous children with primary schooling. In T. Nunes & P. Bryant (Eds.), *Handbook of children's literacy*. Dordrecht: Kluwer Academic Publishers.

Peng, D.L., Orchard, L.N. & Stern, J.A. (1983). Evaluation of eye movement variables of Chinese and American readers. Information abstraction by Chinese and American readers. *Pavlovian Journal of Biological Sciences*, *18*, 94-102.

Petersson, K.M., Ingvar, M. & Reis, A. (2009). Language and literacy from a cognitive neuroscience perspective. In D.R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy*. Cambridge: Cambridge University Press.

Piaget, J. (1926, orig. 1923). *The language and thought of the child.* Guildford: Biddles Ltd.

Piaget, J. (1976). *The grasp of consciousness: Action and concept in the young child*. Cambridge, MA: Harvard University Press.

Pinker, S. (1994). *The language instinct: How the mind creates language.* New York: William Morrow and Company, Inc.

Pynte, J. & Kennedy, A. (2006). An influence over eye movements in reading exerted from beyond the level of the word: Evidence from reading in English and French. *European Journal of Cognitive Psychology*, *16*, 178-202.

Ramachandra, V. & Karanth, P. (2007). The role of literacy in the conceptualization of words: Data from Kannada-speaking children and non-literate adults. *Reading and Writing*, *20*, 173-199.

Raven, J.C. (1962). *Coloured Progressive Matrices: Sets A, A$_B$, B.* London: H.K. Lewis & Co., Ltd.

Ravid D. & Hora, A. (2009). From implicit to explicit language knowledge and intervention: Introduction to the Special Issue on intervention and metalanguage. *First Language*, *29*, 5-14.

Ravid, D. & Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *Journal of Child Language*, *29*, 417-447.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422.

Rayner, K. & Pollatsek, A. (1989). *The psychology of reading.* Hillsdale: Lawrence Erlbaum Associates, Inc.

Rayner, K., Li, X., Williams, C.C., Cave, K.R. & Wella, A.D. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Research*, *47*, 2714-2726.

Reis, A. & Castro-Caldas, A. (1997). Illiteracy: A bias for cognitive development. *Journal of the International Neuropsychological Society*, *3* (5), 444-450.

Roberts, B. (1992). The evolution of the young child's concept of word as a unit of spoken and written language. *Reading Research Quarterly*, *27* (2), 125-138.

Sapir, E. (1929). The status of linguistics as a science. *Language*, *5* (4), 207-214.

Sapir, E. (1970, orig. 1920). *Language: An introduction to the study of speech*. London: Hart-Davis.

Scheurwater, J. (2010). *Eye-tracking in kindergarten: Pilot study on the validity of an eye-tracking task and its use as an instrument to obtain information about the units in which language is processed online.* Unpublished Bachelor's thesis: Tilburg University, the Netherlands.

Schild, U., Roder, B. & Friedrich, C.K. (2011). Learning to read shapes the activation of neural lexical representations in the speech recognition pathway. *Developmental Cognitive Neuroscience*, *1*, 163-174.

Schrooten, W. & Vermeer, A. (1994). Woorden in het basisonderwijs: 15.000 woorden aangeboden aan leerlingen. *Studies in Meertaligheid*, *6*. Tilburg: Tilburg University Press.

Schwartz, M., Leikin, M. & Share, D.L. (2005). Bi-literate bilingualism versus mono-literate bilingualism: A longitudinal study of reading acquisition in Hebrew (L2) among Russian-speaking (L1) children. *Written Language and Literacy*, *8*, 179-207.

Sekerina, I.A., Fernández, E.M. & Clashen, H. (2008). *Developmental psycholinguistics: On-line methods in children's language processing*. Amsterdam: John Benjamins.

Sinclair, I., Jarvella, R. & Levelt, W.J.M. (1978). *The child's conception of language*. Berlin: Springer-Verlag.

Siyanova-Chanturia, A., Conklin, K. & Van Heuven, W. (2011). Seeing a phrase 'time and again' matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Language, Memory, and Cognition*, *37*, 776-784.

Slobin, D.I. (1978). A case study in early language awareness. In I. Sinclair et al. (Eds.), *The child's conception of language*. Berlin: Springer-Verlag.

Smiskova-Gustafsson, H. (2013). *Chunks in L2 development: A usage-based perspective.* Groningen: Wöhrmann print service WPS.

Smith, M.A. (2012). How to do a Raven matrices test. In *How to do a Raven matrices test/IQ Mindware.* Retrieved January 24, 2014, from: *http://www.iqmindware.com/iq-mindware/how-to-do-a-raven-matrices-test/*

Snedeker, J. & Thothathiri, M. (2008). What lurks beneath: Syntactic priming during language comprehension in preschoolers (and adults). In I.A. Sekerina et al. (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing*. Amsterdam: John Benjamins.

Soveri, A., Lehtonen, M. & Laine, M. (2007). Word frequency and morphological processing in Finnish revisited. *The Mental Lexicon*, *2*, 359-385.

Stefanowitsch, A. & Gries, Th.S. (2003). Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, *8* (2), 209-243.

Sulzby, E. & Teale, W. (1991), Emergent literacy. In R. Barr et al. (Eds.), *Handbook of reading research*, Vol. 2. New York: Longman.

Swiney, D. & Cutler, A. (1979). The access and processing of idiomatic expressions. *JVLVB*, *18*, 523-534.

Taylor, J.R. (2002). *Cognitive grammar*. Oxford: Oxford University Press.

Tellings, A., Hulsbosch, M., Vermeer, A. & Van den Bosch, A. (submitted). BasiLex: An 11.5 million words corpus of Dutch texts written for children. *CLIN-Journal*, *4*.

Tolchinsky, L. (2004), Childhood conceptions of literacy. In T. Nunes & P. Bryant (Eds.), *Handbook of children's literacy*. Dordrecht: Kluwer Academic Publishers.

Tolchinsky, L. & Teberosky, A. (1998). The development of word segmentation and writing in two scripts. *Cognitive Development*, *13*, 1-24.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, *74*, 209-253.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2006). Acquiring linguistic constructions. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology*. New York: Wiley.

Trueswell, J.C. (2008). Using eye movements as a developmental measure within psycholinguistics. In I.A. Sekerina et al. (Eds), *Developmental psycholinguistics: On-line methods in children's language processing*. Amsterdam: John Benjamins.

Trueswell, J.C., Sekerina, I., Hill, N.M. & Logrip, M.L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*, 89-134.

Tunmer, W.E., Bowey, J.A. & Grieve, R. (1983). The development of young children's awareness of the word as a unit of spoken language. *Journal of Psycholinguistic Research*, *6*, 567-594.

Uri, H. (2001). Bilingualism, monolingualism, and metalinguistic awareness. In I.A. Hvenekilde & J. Nortier (Eds.), *Meetings at the crossroad. Studies of multilingualism and multiculturalism in Oslo and Utrecht*. Oslo: Novus.

Van Dyken, J.R. & Kutsch Lojenga, C. (1993). Word boundaries: Key factors in orthography development. In R.L. Hartell (Ed.), *Alphabets of Africa*. Dakar: UNESCO and SIL.

Van de Guchte, C. & Vermeer, A. (2003). Een passende woordkeus: Het kiezen van woorden voor woordenschatlessen. *Toegepaste Taalwetenschap in Artikelen*, *69*, 9-23.

Van Kleeck, N. (1982). The emergence of linguistic awareness: A cognitive framework. *Merrill-Palmer Quarterly*, *28*, 237-265.

Van Kleeck, N. (1984). Assessment and intervention: Does 'meta' matter? In G. Wallach & K. Butler (Eds.), *Language learning disabilities in school-age children*. Baltimore: Williams & Wilkins.

Velan, H. & Frost, R. (2007). Cambridge University versus Hebrew University: The impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review*, *14* (5), 913-918.

Veldhuis, D. (2011). *The relativity of 'graphic relativity' among students and native-speakers of the multigraphic language Japanese.* Unpublished Mphil-thesis, Leiden University, the Netherlands.

Veldhuis, D. (2012). The effect of literacy on the offline segmentation and online processing of high and low frequent chunks of language. In N. de Jong, K. Juffermans, M. Keijzer & L. Rasier (Eds.), *Papers of the Anela 2012 applied linguistics conference*. Delft: Eburon.

Veldhuis, D., Li, J. & Kurvers, J. (2010). Segmentatie van gesproken Nederlands en Chinees. De invloed van geletterdheid en schriftsoort. *Toegepaste Taalwetenschap in Artikelen*, *83*, 79-89.

Veldhuis, D. & Backus, A. (2012). More than frequency: Multiword combinations in pre-literate children's language segmentation. In S. Kersten, C. Ludwig, D. Meer & B. Rüschoff (Eds.), *Language learning and language use: Applied linguistics approaches*. Duisburg: Universitätsverlag Rhein-Ruhr.

Veldhuis, D. & Kurvers, J. (2012). Offline segmentation and online language processing units: The influence of literacy. *Written Language & Literacy*, *15* (2), 165-184.

Veldhuis, D., Vermeer, A. & Yan, D. (2012). Hu meertalig? De bruikbaarheid van een sneltoets tweetaligheid voor kinderen. *Pedagogische Studiën*, *3*, 144-158.

Verhagen, V. (2011). *Individual differences in entrenchment of multiword units: Evidence from a Magnitude Estimation task*. Unpublished MA-thesis, Tilburg University, the Netherlands.

Vitu, F. & O'Regan, J.K. (1995). A challenge to current theories of eye movements in reading. In J.M. Findlay, R. Walker & R.W. Kentridge (Eds.), *Eye movement research: Mechanisms, processes and applications*. Amsterdam: Elsevier Science Publishers.

Weber, A. & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*, 1-25.

Weighall, A.R. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Child Psychology*, *99*, 75–95.

White, S.J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, *34* (1), 205-223.

Whorf, B. L. (2012). *Language, thought, and reality: selected writings of Benjamin Lee Whorf* (J. B. Carroll, S. C. Levinson, & P. Lee, Eds.). Cambridge, MA: MIT Press.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Wray, A. (2008). *Formulaic language: Pushing the boundaries.* Oxford: Oxford University Press.

Yao, Y. (2011). Interword spacing effects on reading Mandarin Chinese as a second language. *Writing Systems Research*, *3* (1), 23-40.

Zhang, W., Yoshida, T., Tang, X. & Ho, T.-B. (2009). Improving effectiveness of mutual information for substantial multiword expression extraction. *Expert Systems with Applications*, *36*, 10919-10930.

Schools that participated in the project

Pilots in Dutch, Turkish and Dutch, or Turkish only, in the Netherlands at the following schools:

– BS Christoffel, Tilburg
– BS De Bogaard, Ravenstein
– BS De Vlashof, Tilburg
– BS Wandelbos, Tilburg
– Jenaplanschool 't Hoge Land, Epe
– PCBO Da Costaschool, Rotterdam
– OBS Combinatie 70, Rotterdam
– OBS De Boog, Rotterdam
– OBS De Sleutel, Tilburg
– OBS Pantarijn, Rotterdam

Final tests conducted at:

– Dokter Bosschool, Utrecht, the Netherlands
– Two schools in Turkey: one in Istanbul and one near Istanbul

Target items from the pilot:
One word in Turkish, several words in Dutch

Last-part repetition task: pause after locative prepositional combinations: in total 28 breaks in and after hypothesized units and filler items

Item number	Target in Dutch	Translation in English
18	in bad / bed	*in bath / bed*
20	onder de kist / krant	*under the box / newspaper*
23	op de stoel / slee	*on the chair / luge*
27	op de wc / weg	*on the toilet / road*

Sentence segmentation task: in total 17 sentences that included locative prepositional, possessive and transitive combinations and fillers in two versions (in table below distinguished by a slash forward)

Item number	Target in Dutch	Translation in English
2	mijn spiegel / mijn schouder	*my mirror / my shoulder*
4	mijn adres / mijn adem	*my address / my breath*
6	in het bos / in de buik	*in the forest / in the belly*
7	mijn jurk / jas	*my dress / coat*
9	hun kapper / koffer	*their hairdresser / suitcase*
10	in de lucht / in het licht	*in the air / light*
13	mijn tas / tand	*my bag / tooth*
14	in de koffie / keuken	*in the coffee / kitchen*
17	op de mat / muur	*on the mat / wall*

Click task: in total 13 sentences with clicks inserted in and right after hypothesized units. Clicks occurred within locative prepositional and possessive combinations. // marking the click.

Item number	Target in Dutch	Translation in English
5	mijn // droom	*my // dream*
7	in // de kamer - // de kelder	*in // the room - // the cellar*
8	hun // spiegel	*their // mirror*
10	in // de kerk - // de kast	*in // the church - // the cupboard*
13	onder // de caravan - // de kast	*under // the caravan - // the cupboard*

SPL: in total 11 sentences with breaks after various single and multiword sequences. Breaks after target locative prepositional and possessive combinations in two versions, represented by a slash forward in the table below.

Item number	Target in Dutch	Translation in English
2	*mijn schouder / spiegel*	*my shoulder / mirror*
9	*mijn adem / adres*	*my breath / address*
5	*mijn jas / jurk*	*my bag / dress*
7	*mijn tand / tas*	*my tooth / bag*
11	*hun koffer / konijn*	*their suitcase / rabbit*

Manual adjustments of fixations

The following steps were taken in assigning, or unassigning, fixations manually in the program Fixation (Cozijn, 2006).

- First fixation(s, sometimes up to children's third fixation) were unassigned if they fell at or around the middle of the screen. Fixations in those places were an effect of the cross children had seen on the screens presented to them right before the target screens occurred.

- Fixations that fell outside the screen were also unassigned, unless they were short and in between two fixations in the same quadrant. In the latter case, they were regarded as blinks and assigned to the quadrant of the fixation before and after.

- Fixations were assigned to another quadrant if they were relatively short in duration and the next fixation was in the other quadrant. In such a case, the fixation could be regarded as a 'stop while eyes were jumping' from one quadrant to the next.

- Fixations at the end of a series of looks at a screen were unassigned if these fell on or around the middle (and it could not be stated for sure where the child was gazing at).

- Fixations, as presented by the program, were moved all together (and assigned to a different position) if all fixations occurring on the screens suggested that they had to be moved.

- If it could not be determined for a particular fixation halfway down a session of fixations on a particular screen (for instance because such a fixation appeared in the middle of two quadrants, without being a blink), the data obtained for this screen for this child was removed from further analyses.

– For blinks where the eyes only moved quickly in a short fixation to another quadrant, that fixation was assigned to the quadrant where fixations occurred before and after the blink.

– The data of screens for which children had fewer than five fixations were removed from analyses. In such cases, children were probably gazing. In such cases, the eye-tracking data would not lead to reliable analyses.

If four or more of the ten target screens that a child had seen were removed from further analyses for one of the reasons mentioned above, then all the data from this child were removed from the analyses.

Selected items and MI/EMI values

Google frequency data and MI/EMI values of Dutch targets

Low MI NV-combinations	Translation in English	MI
tafel dekken	*set the table*	**-0.01**
handen wassen	*wash hands*	**-0.04**
hond aaien	*pet a dog*	**-0.13**
muur verven	*paint a wall*	**-0.67**
bellen blazen	*blow bubbles*	**-1.68**
kast opruimen	*clean a cupboard*	**-1.73**
mop vertellen	*tell a joke*	**-1.94**
verstoppertje doen	*do hide and seek*	**-3.69**
zandkasteel maken	*build a castle of sand*	**-6.09**
das omslaan	*wrap a scarf around*	**-12.74**

High MI NV-combinations	Translation in English	MI
veters strikken	*tie laces*	**5.80**
neus snuiten	*blow your nose*	**4.94**
tanden poetsen	*brush teeth*	**4.06**
verstoppertje spelen	*play hide and seek*	**2.94**
boterham smeren	*prepare a sandwich*	**2.87**
touwtje springen	*jump rope*	**0**
bladzijde omslaan	*turn a page*	**2.32**
liedje zingen	*sing a song*	**1.87**
das omdoen	*do a scarf around*	**1.46**
taart bakken	*bake a cake*	**1.12**

Low EMI PP-combinations	Translation in English	EMI
in de thee	*in the tea*	**-2.88**
in de koffer	*in the suitcase*	**-3.00**
in de koffie	*in the coffee*	**-3.70**
in het bad	*in the bath*	**-4.39**
onder de kast	*under the cupboard*	**-4.48**
op de fluit	*on the flute*	**-5.47**
in de kom	*in the bowl*	**-7.03**
op de skelter	*on the skelter*	**-7.96**
onder de tv	*under the TV*	**-8.57**
op de step	*on the step*	**-5.85**

High EMI PP-combinations	Translation in English	EMI
in de tuin	*in the garden*	**2.15**
onder de kraan	*under the tap*	**0.16**
in het bos	*in the forest*	**-0.55**
in de kamer	*in the room*	**-0.80**
op de fiets	*on the bike*	**-1.41**
in de kist	*in the box*	**-1.68**
in de keuken	*in the kitchen*	**-2.05**
op de schommel	*on the swing*	**-2.32**
onder de tafel	*under the table*	**-2.67**
op de stoel	*on the chair*	**-2.23**

Dutch Google size estimated at 631,250,000 (September 2011)

Google frequency data and MI values of Turkish targets

Low MI NV-combinations	Translation in English	MI
el temizlemek	*wash hands*	**-4.23**
ayak silmek	*wipe feet*	**-0.44**
saati söylemek	*tell the time*	**-0.08**
sayfa değiştirmek	*flip the page*	**0.39**
köpek okşamak	*pet a dog*	**0.83**
dal kirmak	*break branches*	**2.59**
kahve yapmak	*make coffee*	**4.17**
meyve yemek	*eat fruits*	**5.08**
dil çikarmak	*stick out tongue*	**5.15**
sofra kurmak	*set the table*	**5.81**

High MI NV-combinations	Translation in English	MI
balık tutmak	*catch a fish*	**9.87**
lastik şişirmek	*pump tyres*	**10.00**
üçgen çizmek	*draw triangle*	**10.52**
odun_kesmek	*cut firewood*	**10.82**
mum_yakmak	*light a candle*	**11.69**
saç_taramak	*comb hair*	**11.71**
diş_fırçalamak	*brush teeth*	**11.76**
kravat_takmak	*put on(wear) a tie*	**12.07**
bağcık_bağlamak	*tie shoelaces*	**12.15**
sigara_içmek	*smoke (cigarette)*	**12.73**

Low EMI PP-combinations	Translation in English	MI
elmanın altında	*under the apple*	**-0.15**
gazetenin altında	*under the newspaper*	**0.90**
tereyağının üzerinde	*on the butter*	**1.23**
bahçenin üzerinde	*over the garden*	**2.43**
yangının içinde	*in the fire*	**2.68**
evin yanında	*next to the house*	**2.82**
paranın üzerinde	*on banknotes*	**2.94**
mutfağın içinde	*in the kitchen*	**3.10**
bardağın üzerinde	*above the glass*	**3.46**
okulun içinde	*in the school*	**4.89**

High MI PP-combinations	Translation in English	MI
kızağın üzerinde	*on the sledge*	**7.53**
valizin içinde	*in the suitcase*	**7.97**
ormanın içinde	*in the forest*	**7.98**
halının üzerinde	*on the carpet*	**8.13**
koltuğun üzerinde	*on the couch*	**8.13**
sandalyenin üzerinde	*on the chair*	**8.29**
masanın üzerinde	*on the table*	**8.83**
sepetin içinde	*in the basket*	**8.99**
kutunun içinde	*in the box*	**10.06**
taburenin üzerinde	*on the stool*	**10.25**

Single-word locatives	Translation in English – in the context of tasks
bahçede	*over the garden*
koltukta	*on the couch*
masada	*on the table*
mutfakta	*in the kitchen*
sandalyede	*on the chair*

Sentences occurring in the sentence segmentation task, Dutch and Turkish

Example sentences in Dutch

No.	Sentence in Dutch	Translation in English
1	Ik vind groene appels erg lekker.	*I like green apples very much.*
2	Nienke heeft een rode trui.	*Nienke has a red sweater.*
3	Anne heeft een gele auto.	*Anne has a yellow car.*

Sentences in Dutch

No.	Sentence in Dutch	Translation in English
1	Sander vindt veters strikken heel makkelijk.	*Sander finds tying his laces very easy.*
2	Kun je die doos in de kamer zetten?	*Can you put the box in the living?*
3	Je moet een das omslaan in de winter.	*You have to wear a scarf in winter.*
4	Nina moet haar voeten vegen. [filler]	*Nina has to wipe her feet. [filler]*
5	Mama is in de keuken aan het koken.	*Mom is cooking in the kitchen.*
6	Je moet je tanden poetsen voor je gaat slapen.	*You have to brush your teeth before bed.*
7	Pim zit op de skelter van zijn broer.	*Pim is on his brother's skelter.*
8	Ik kan heel mooi tekenen! [filler]	*You draw very nice! [filler]*
9	Jakob wil verstoppertje doen in de pauze.	*Jakob wants to play hide and seek during the break.*
10	Zullen we in de tuin gaan spelen?	*Shall we play in the garden?*
11	Dit is een boek over beren en olifanten. [filler]	*This is a book on bears and elephants. [filler]*
12	Zal ik een boterham smeren straks?	*Shall I make a sandwich later on?*
13	Wij hebben in het bos gespeeld.	*We have played in the forest.*
14	Zullen we op de computer spelen? [filler]	*Shall we play on the computer? [filler]*
15	Pien vindt touwtje springen erg leuk.	*Pien thinks skip the rope is a very fun game.*
16	De bal is onder de kast gerold.	*The ball rolled under the cupboard.*
17	Sanne moet de kast opruimen van de juf.	*The teacher told Sanne to tidy the cupboard.*
18	Heb jij een rode fiets? [filler]	*Have you got a red bike? [filler]*

No.	Sentence in Dutch	Translation in English
19	Ik zie de Donald Duck onder de tv liggen.	*I see a Donald Duck-comic under the TV.*
20	Maaike wil piano of gitaar spelen. [filler]	*Maaike wants to play piano or guitar. [filler]*
21	Zal ik je een mop vertellen over 3 Belgen?	*Shall I tell a joke about 3 Belgians?*
22	Wie heeft er op de fluit geblazen?	*Who blew the whistle?*
23	Ik heb een oranje jas. [filler]	*I have an orange coat. [filler]*
24	De kleren in de kist zijn erg oud.	*The clothes in the box are very old.*
25	Sam zit op voetbal. [filler]	*Sam plays football. [filler]*
26	Tijn wil verstoppertje spelen in de pauze.	*Tijn wants to play hide and seek during the break.*
27	Fleur gaat een zandkasteel maken in de zandbak.	*Fleur is going to build a sandcastle in the sandbox.*
28	Joris vindt hutten bouwen leuk. [filler]	*Joris likes building tree houses. [filler]*
29	Je moet je neus snuiten van mama.	*Mom tells you to wipe your nose.*
30	Menno vindt de tafel dekken stom.	*Menno thinks setting the table is stupid.*
31	Zal ik je beer in de koffer stoppen?	*Shall I put the bear in the suitcase?*
32	Oma gaat een taart bakken morgen.	*Grandma is going to bake a cake tomorrow.*
33	Wil jij een film kijken vanavond? [filler]	*Do you want to watch a film tonight? [filler]*
34	Opa wil in de koffie drie scheppen suiker!	*Grandpa likes three spoons of sugar in his coffee!*
35	Vind jij een hond aaien niet eng?	*Don't you think patting dogs is scary?*
36	Ik kan heel hard rennen! [filler]	*I can run very fast! [filler]*
37	Doe je wat voer in de kom van de vissen?	*Can you put some fishfood in the fishbowl?*
38	Papa gaat een muur verven thuis.	*Father is going to paint a wall at home.*
39	Iris gaat bij haar tante en oom logeren. [filler]	*Iris is going to sleep over at her aunt and uncle's.*
40	Ik zag Bram op de schommel zitten.	*I saw Bram sitting on the swing.*
41	Je moet een das omdoen in de winter.	*You have to wear a scarf in winter.*
42	Zal ik de hond uitlaten straks? [filler]	*Shall I walk the dog later on?*
43	Ik vind bellen blazen erg leuk.	*I like making bubbles very much.*
44	Ik heb mijn handen onder de kraan gewassen.	*I have washed my hands under the tap.*
45	Floor kijkt een film. [filler]	*Floor is watching a film. [filler]*
46	De hond ligt onder de tafel te slapen.	*The dog is sleeping under the table.*
47	Zie je die grote wolk in de lucht? [filler]	*Do you see that large cloud in the sky?*
48	Zal ik de bladzijde omslaan en verder lezen?	*Shall I turn the page and read on?*
49	Ben jij op de fiets gekomen?	*Did you come by bicycle?*
50	Wij gaan een liedje zingen met de juf.	*We are going to sing a song with the teacher.*
51	Els koopt bananen en mandarijnen. [filler]	*Els buys bananas and tangerines. [filler]*
52	Sara is op de step naar school gekomen.	*Sara came to school on a step.*
53	Wil jij suiker in de thee hebben?	*Do you want sugar in the tea?*
54	Ik vind vuurwerk afsteken eng. [filler]	*I think lighting fireworks is scary.*
55	Het water in het bad is koud.	*The water in the tub is cold.*
56	Sem zet zijn fiets in het rek. [filler]	*Sem puts his bicycle in the bicycle-stand.*
57	Ik ben op de stoel geklommen.	*I climbed on a chair.*
58	Je moet je handen wassen voor het eten.	*You have to wash your hands for supper.*

Example of sentences used in the Turkish version of the task

No.	Sentence in Dutch
1	X is sitting on a red chair. [19]
2	Mehmet went to school.
3	My mother did the washes.
4	My aunt made spinach/pilav dish. (some of the children had a specific wish for a particular dish – S.D.)
5	Bugs Bunny went his home yesterday. (this was a popular animation among the children)
6	X and X played hide-and-seek.

No.	Sentence in Turkish	Translation in English
1	Baltayla odun kesmek babamın işidir.	*It's my father duty to cut wood with the axe.*
2	Annem bağcık bağlamak konusunda bana yardım etti.	*My mom helped me to tie shoelaces (with tying shoelaces)*
3	Çoraplar valizin içinde kalmış.	*The socks were left in the suitcase.*
4	Cem'in saçı tereyağının üzerinde duruyordu.	*Cem's hair was on the butter.*
5	Çilekler sepette ezilmişler.	*The strawberries got squashed in the basket.*
6	Erkeklere her gün kravat takmak zor gelir.	*For men it's difficult to put on a tie everyday.*
7	Top oynarken ormanın içinde sincap gördük.	*When we were playing football we saw a squirrel in the forest.*
8	Öğretmenler okulda toplantı yaptılar.	*The teachers had a meeting in the school.*
9	Sokakta yürürken dil çıkarmak çok ayıptır.	*It's very rude to stick out your tongue when you're walking on the street.*
10	Öğrencilerin okulun içinde koşması yasak.	*It is forbidden for students to run in(side) the school.*
11	Beş yaşında piyano çalmaya başladım.	*I started playing the piano when I was 5 years old.*
12	Ateş yakmak için dal kırmak çok kolay.	*It is very easy to break branches to light a fire.*
13	Kışın kızağın üzerinde kaydık.	*In winter we slid down on the sledge.*
14	Gazete okurken hızlıca sayfa değiştirmek çok zordur.	*It is very difficult to flip the page fast when you're reading the newspaper.*
15	Babam "evin yanında bağırmayın" dedi.	*My dad said 'do not shout next to the house'.*
16	Çocukken üçgen çizmek kolay değildir.	*It is not easy to draw a triangle when you're young.*
17	Yumurtalar sepetin içinde kırılmış.	*The egg had broken in the basket.*
18	Her gün meyve yemek sağlık için çok faydalıdır.	*It is very healthy to eat fruits every day.*
19	Ayakkabılarını kutunun içinde temizledi.	*She cleaned her shoes in(side) the box.*

[19] S.D.: dr. Seza Dogruöz, the researcher who was responsible for the part of the Building Blocks Project focused on the typological influence on children's perception of basic building blocks in language. She prepared and conducted, after the item-selection was made, the tasks (including examples) in Turkish, together with Derya Demircay, a PhD-student at Tilburg University.

No.	Sentence in Turkish	Translation in English
20	Sokaktaki adam bana saati söylemek için durmuş.	*The man on the street stopped to tell me the time.*
21	En sevdiğim şarkıcı konser verdi.	*My favorite singer gave a concert.*
22	Sabah bahçenin üzerinde bir uçak uçuyordu.	*In the morning a plane was flying over the garden.*
23	Babam arabadan lastik şişirmek için indi.	*My dad got out of the car to pump the tyre.*
24	Cüzdandaki parada Atatürk resmi var.	*There is a picture of Atatürk on the banknote in the wallet.*
25	Okulda çayla simit satılıyor.	*Tea and simit is sold in school.*
26	Ünlü oyuncu yaşama veda etti.	*The famous actress passed away.*
27	Sinek bardağın üzerinde uçuyordu.	*The fly was flying above the glass.*
28	Annem bardakta su kalmamış dedi.	*My mother said 'there is no water left in the glass'.*
29	Kek için sekiz yumurta bir bardak süt lazım.	*We need 8 eggs, and 1 glass of milk for the cake.*
30	Çocuklar kızakta çok eğleniyorlar.	*The children are having a lot of fun on the sledge.*
31	Doktorlar sigara içmek zararlıdır diyor.	*Doctors say that it is unhealthy to smoke (cigarettes).*
32	Ninem koltukta dinleniyor.	*My granny is resting on the couch.*
33	Okuldayken hep sandalyede otururuz.	*We always sit on chairs in school.*
34	Evde sofra kurmak annemin işidir.	*At home it is my mother's duty to set the table.*
35	Havlarken köpek okşamak cesaret istiyor.	*It requires courage to pet the dog when it's barking.*
36	Kadınla çocuğu yangının içinde kalmış.	*The woman and her baby were left with the fire.*
37	Banyodan sonra saç taramak daha zordur.	*It's more difficult to comb your hair after a shower.*
38	Yağmur yağınca dışarı çıkamadık.	*We couldn't go out when it was raining.*
39	Çocuklar elmanın altında kurt buldular.	*The children found a worm under the apple.*
40	Fatma çok koşarsa yorulur.	*Fatma gets too tired if she runs too much.*
41	Annem paranın üzerinde çok kir olur dedi.	*My mother said there is a lot of dirt on banknotes.*
42	Kediler çok yaramaz olur.	*Cats are very naughty.*
43	Adam taburenin üzerinde oturuyordu.	*The man was sitting on the stool.*
44	Eve girmeden önce ayak silmek lazım.	*You need to wipe your feet before entering the house.*
45	Kardeşim mutfağın içinde fare yakaladı.	*My brother caught a mouse in the kitchen.*
46	Yastığı sandalyenin üzerinde bırakmışsın.	*You left the pillow on the chair.*
47	Okulda masada hep tebeşir olur.	*At school there is always chalk on the table.*
48	Annem el temizlemek çok iyi bir şey dedi.	*My mother said that washing hands is a very good thing.*
49	Adamlar bahçede yemek yiyordu.	*The men were eating in the garden.*
50	Basketbol oynamayı çok seviyorum.	*I love playing basketball.*
51	Eve gelince kahve yapmak insanı rahatlatır.	*When you come home it calms you down to make coffee.*

52	Pastaya mum yakmak için kibrit aldık.	*We bought matches to light the candles on the cake.*
53	Dedem gazetenin altında gözlüklerini unutmuş.	*My grandpa forgot his glasses under the newspaper.*
54	Büyükbabam balık tutmak için deniz kenarına gitti.	*My grandfather went to the seaside to catch fish.*
55	Ankaranın havası buz gibi olur.	*The weather gets freezing in Ankara.*
56	Küçükken koltuğun üzerinde zıplardım.	*When I was young I used to jump on the couch.*
57	Eve geldiğimde mutfakta üç kişi vardı.	*When I got home there were 3 people in the kitchen.*
58	Kedi halının üzerinde mışıl mışıl uyuyordu.	*The cat was fast asleep on the carpet.*
59	Kitaplar masanın üzerinde duruyordu.	*The books were on the table.*
60	Yatmadan önce diş fırçalamak önemlidir.	*It is important to brush your teeth before going to bed.*

Stories prepared for the last-part repetition task, in Dutch and Turkish

Stories in Dutch

Example sentences:
(// marks pauses; bold script indicates example provided by researcher; regular script indicates possible answers as provided by the researcher)

1) **Er was eens een klein meisje**//… [meisje/ klein meisje]
2) **Er was eens een klein meisje dat Roodkapje heette. Ze woonde in een groot bos**//… [bos/ een groot bos]
3) **Ze woonde in een groot bos. Op een dag ging ze naar haar oma**//… [oma/ haar oma/ ging ze naar haar oma]

Story 1:
// marks pauses; bold script indicates target; {marks point where backtracking started. English translation below the Dutch version.

Vandaag is Sanne jarig. {Ze wordt 1. *al*// zes jaar. Vanochtend {mag Sanne eerst 2. **in het bad** van haar moeder. Daarna {gaat ze 3. **op de step**// naar school. Het is koud. Daarom {moet Sanne wel een 4. **das omdoen**// van haar moeder.

{Op school mag Sanne 5. **op de stoel**// gaan staan. {De kinderen van de klas gaan een 6. **liedje zingen**// voor Sanne. Dan mag ze trakteren. Ook mag {ze kiezen wat ze in 7. *de*// pauze wil doen. {Sanne wil het liefste 8. **op de schommel**// spelen. {Dat vindt ze net als 9. **bellen blazen**// erg leuk.

Als Sanne weer thuis is {mag ze haar moeder helpen met 10. **een taart bakken**// voor 's avonds. {Sanne moet meel en melk 11. **in de kom**// doen. Maar {dat is heel moeilijk, en 12. *daarom*// morst Sanne op de grond en tegen de muur. {Het lijkt wel op 13. **muur verven**// met deeg! {Er is ook deeg 14. **onder de tafel**// terechtgekomen. {Het is een 15. *vieze*// bende in de keuken. Zelfs 16. **in de thee**// zitten kruimels.

Snel {ruimt Sannes moeder 17. *alles*// op. {Ze wast de kom 18. **onder de kraan**// af.
{Sanne moet haar 19. **handen wassen**// voor ze een nieuwe taart gaan bakken.

{'s Avonds komen opa 20. *en*// oma. {Ze willen 21. **op de fiets**// komen, en ze blijven
ook eten.

{Sanne gaat snel 22. **de tafel dekken**// voordat opa en oma er zijn. Eerst moet {het
speelgoed 23. *van*// tafel. {Ze stopt het 24. **in de koffer**// die mama vorig jaar gekocht
heeft. Nu is de tafel weer netjes. Sanne zet de thee, koffie en de taart vast klaar.

Als opa en oma er zijn, krijgen ze eerst koffie. {Opa wil 25. **in de koffie** graag drie
schepjes suiker. Dat vindt hij lekker.

Dan krijgt Sanne haar cadeautjes. Eerst krijgt ze van oma een mooi boekje. {Dat gaat
ze 26. *meteen*// bekijken. {Ze wil net de 27. **bladzijde omslaan**// als opa met zijn
cadeautje binnenkomt. Dat cadeautje is een hondje! {Een hele 28. *lieve*// hond met
zwarte en witte vlekjes. {Sanne vraagt of zij de 29. **hond aaien**// mag. {Dit is wel het
30. *mooiste*// verjaardagscadeau dat Sanne had kunnen verzinnen….

*Today, Sanne has her birthday. She's six years old already. This morning, Sanne is allowed by
her mother to bathe in the tub. Afterwards she goes to school on her step. It's cold. Therefore
her mother insists that she wears a scarf.*

*At school, Sanne may climb on her chair. The children of her form will sing her a song. Then,
she's allowed to hand out treats. She can also choose what to do during intermission. Sanne
prefers to play on the swing. She likes that, as she likes making bubbles, very much.*

*When she has returned home, Sanne may help her mother backing a cake for that evening.
Sanne has to put flour and milk in a bowl. But that is difficult and that is why she spills on
the ground and against the wall. It looks like painting the walls with dough! Dough also
ended up under the table. It is a dirty mess in the kitchen. There are even crumbs in the tea.*

*Quickly, Sanne's mother tidies everything. She washes up the bowl under the tap. Sanne has
to wash her hands before baking another cake.*

*In the evening grandpa and grandma visit. The want to come by bicycle and stay over for
supper. Sanne quickly sets the table, before grandpa and grandma arrive. First, the toys have
to be taken off the table. She puts these in the suitcase mama bought last year. Now the table is
tidy again. Sanne arranges the tea, coffee and the cake.*

*When grandpa and grandma arrive, they get coffee first. Grandpa takes his coffee with three
spoons of sugar. That's how he likes it.*

*Then, Sanne receives het presents. First, she gets a nice booklet from grandma. She
immediately starts to examine it. She is about to turn a page, when grandpa enters with his
present. That present is a puppy! A very adorable dog with black and white spots. Sanne asks
if she can pet the dog. This is the best birthday-present Sanne could have imagined…..*

Story 2:

Niek logeert bij zijn nichtje Maaike. Dat vindt hij erg leuk. {Bij Maaike mogen ze altijd tot 1. *heel*// laat opblijven. Ook {kunnen ze fijn 2. **in de tuin**// spelen. Maaike speelt ook fluit. {Niek vindt het leuk als hij ook eventjes 3. **op de fluit**// mag blazen.

Maar {eerst moeten Niek 4. *en*// Maaike naar school. Omdat Niek er is {mogen ze 5. **een boterham smeren**// zelf doen. {Dat moet 6. **in de keuken**// op het aanrecht. Dan {moeten ze hun **7. tanden poetsen**// van de moeder van Maaike. En daarna moeten ze hun schoenen aantrekken. {Niek kan 8. **veters strikken**// ook al helemaal zelf.

Het is koud. Daarom {moeten ze wel 9. **een das omslaan**// van Maaikes moeder. Dan {kunnen ze 10. *eindelijk*// naar school.

Op school moeten ze de klas opruimen. Niek en Maaike {moeten 11. **een kast opruimen**// van de juf. {De spullen die kapot zijn moeten ze 12. **in de kist**// stoppen, die de juf heeft klaargezet.

Op de kast staat de tv en {in de kast 13. *liggen*// boeken en spelletjes. {Ook 14. **onder de kast**// liggen papieren. En {Niek en Maaike vinden 15. **onder de tv**// een stripboek. {Opruimen is 16. *een*// flinke klus. Er is ook veel stof {en 17. *daarom*// moet Niek niezen. "Je moet {je 18. **neus snuiten**// Niek", zegt juf. Dat helpt.

Na school {wil Niek graag 19. **verstoppertje doen**// met Maaike. Maar Maaike heeft daar geen zin in. {Zij wil liever 20. **een zandkasteel maken** in de zandbak. "{Wil je dan 21. **op de skelter**// rijden?", vraagt Niek. Maar {daar heeft Maaike 22. *ook*// al geen zin in. {Zij vindt 23. **touwtje springen**// veel leuker. {Het is 24. *moeilijk*// hoor, om te bedenken wat je samen wil doen.

Dan heeft Niek een idee! {Hij wil een 25. **mop vertellen**// aan Maaikes vader. Maaikes {vader zit 26. **in de kamer**// te lezen. Snel rent {Niek naar hem toe 27. *en*// vertelt de mop. Maaikes vader moet hard lachen.

"Wat gaan jullie nu doen?", vraagt Maaikes vader. "{Willen jullie nog even 28. **verstoppertje spelen**// voor het eten? Als jullie willen, {kunnen we dat wel 29. **in het bos**// gaan doen?" Dat willen Maaike en Niek wel. {In het bos spelen is 30. *natuurlijk*// hartstikke leuk!

Niek stays over at his cousin Maaike. He likes that. At Maaike's they're always allowed to stay up very late. They can also play in the garden. Maaike also plays the flute. Niek likes it when he is allowed to play on the flute as well.
But first, Niek and Maaike have to go to school. Because Niek is staying over, they may prepare their own sandwiches. They have to do this in the kitchen, at the counter. Then,

*Maaike's mother makes them clean their teeth. Afterwards, they have to put on their shoes.
Niek can tie his laces all by himself.*

*It's cold. That is why Maaike´s mother makes them wear a scarf. Then, finally, they can go to
school.*

*At school, they have to tidy the classroom. The teacher tells Niek and Maaike to clean up a
cupboard. They have to put the broken things in a box, that the teacher provided.*

*On top of the cupboard is a television and in the cupboard are books and games. Under the
cupboard there are some papers as well. Niek and Maaike discover a comic-book under the
cupboard as well. Tidying is quite an endeavor. It's also very dusty, so Niek has to sneeze.
"You have to blow your nose, Niek", the teacher says. That's good advice!*

*After school, Niek wants to play hide and seek with Maaike. But Maaike doesn't want to. She
would rather build a sandcastle in the sandbox. "Do you want to drive the go-card then?",
Niek asks. But Maaike isn't interested in that either. She prefers to play skip the rope. It's
hard to think up a common plan!*

*Then, Niek has an idea! He wants to tell Maaike's father a joke. Maaike's father is reading in
the living room. Quickly, Niek joins him and tells the joke. Maaike's father laughs loudly.*

*"What are you going to do now?", Maaike´s father asks. "Do you want to play hide and seek
before supper? If you like, we can go play it in the woods?" Maaike and Niek do want to.
Playing in the woods is of course great fun!*

Stories in Turkish

Example sentences:

*(// marks pauses; bold script indicates example provided by researcher; regular script indicates
possible answers as provided by the researcher)*

1) **X ate Spinach last night**//…
2) **The favorite team**//… ["galatasaray" or "team" or "favorite team" or some other
   answer]
3) **X went for a picnic yesterday**//…
4) **The mother of X has made X dish last night**// …
5) **X has a Barbie doll**//…

Story 1:

*(// marks pauses after targets; (X) marks pauses after fillers; {marks point where backtracking
started; bold script indicates targets)*

Defne ile Cem aynı okula gidiyorlardı. Her gün okuldan çıkınca {önce **bahçede**//
oynarlardı. {Sonra da **mutfakta**// bir şeyler yerlerdi. Bir gün {okuldan eve gelirken
köpek okşamak// için durdular. {Eve gelince **el temizlemek**// için banyoya gittiler.
Anneleri de {bu arada **sofra kurmak**// için mutfağa gitmişti. {Banyoda **taburenin
üzerinde**// bir elma buldular. {Dikkatle bakınca **elmanın altında**// kıpırdayan bir şey

gördüler. Bunun küçük bir kurt olduğunu görünce **çok** (X) korktular. Koşa koşa mutfağa geldiler. {Anneleri **sandalyenin üzerinde**// oturmuş soğan doğruyordu. {Onlara **koltukta**// oturmalarını söyledi. {Kendisine **kahve yapmak**// için fincan çıkardı. {Daha sonra da **sigara içmek**// için balkona çıktı. Dışarda yağmur yağıyordu. Balkondan içeri girince {terliklerin halının üzerinde// ayak izi bıraktığını fark ettiler. {Defne **koltuğun üzerinde**// Cemle oynuyordu. Ayak izlerini görünce {annesine **ayak silmek**// çok önemli dedi. Akşam **olunca** (X) Cem eve gitti. Defnenin babası geldi. {Ahmet Amca **balık tutmak**// için deniz kenarına gitmişti. {Ailece **masada**// yemek yerken televizyona baktılar. Yemekte balık vardı. {Defne **sandalyede**// rahat oturamadı. O sırada elektrikler gitti. {Defnenin annesi **mum yakmak**// istedi. Yemekten önce Ahmet Amca {gazeteyi **masanın üzerinde**// unutmuştu. Defnenin eli muma çarptı **ve** (X) gazete birden alev aldı. {Defne **yangının içinde**// kaldı. Babası alevlere **su** (X) atarak hemen yangını söndürdü. {Bu arada **gazetenin altında**// para olduğunu fark etti. {Ne yazık ki **paranın üzerinde**// delikler oluşmuştu. Artık kullanılamaz **hale** (X) gelmişti. {Bu arada **mutfağın içinde**// yoğun duman vardı. Annesi mutfağı havalandırdı. Babası sakinleşmek için dışarı çıktı. Defne de bu **zor** (X) günün ardından erken yatmak istedi. {Annesi **diş fırçalamak**// çok önemli diye hatırlattı. {Defne **saç taramak**// da önemli dedi. Annesi de güldü ve saçlarını okşadı.

Defne and Cem were going to the same school. When they got out of school first they would play in the garden. Then they would eat something in the kitchen. One day when they were coming home from school they stopped to pet a dog. When they arrived at home they went to the bathroom to clean (wash) their hands. In the meanwhile their (her) mother had gone into the kitchen to set the table. They found an apple on the stool in the bathroom. When they looked closely they saw something moving under the apple. They were really scared when they saw that it was a little worm. They came running to the kitchen. Their (her) mother was sitting on the chair, peeling onions. She told them to sit on the couch. She took a cup to make coffee for herself. Then she went out to the balcony to smoke (a cigarette). It was raining outside. When she came back inside from the balcony they noticed that her slippers made a footmark on the carpet. Defne was playing with Cem on the couch. When she saw the footmarks she said that it is important to wipe your feet. When it was evening, Cem went home. Defne's father came home. Uncle Ahmet had been to the sea side to catch fish. As they were eating at the table as a family they looked at the television. They had fish for dinner. Defne could not sit comfortable in the chair. At that moment, the electricity went out. Defne's mother wanted to light a candle. Before dinner, Uncle Ahmet had forgotten the newspaper on the table. Defne's hand hit the candle and the newspaper suddenly caught fire. Defne stayed within the fire. Her father extinguished the fire immediately by throwing water on it. During this he noticed that there was some money under the newspaper. Unfortunately there were holes in the banknote. It had become unusable. In the meanwhile, there was thick smoke in the kitchen. Her mother aired the kitchen. Her father went out to calm down. Defne wanted to go to bed early after an eventful day. Her mother reminded her that it is very important to brush your teeth. Defne added that it is also very important to comb your hair. Her mother smiled and caressed her hair.

Story 2:

Aylin ile Barış çok iyi iki arkadaştı. {İkisi de **okulun içinde**// çok sıkıldılar. Çünkü {herkes **üçgen çizmek**// için uğraşıyordu. Aylin de Barış **da** (X) çoktan ödevlerini bitirmişti. Dışarda kar yağıyordu. Paltolarını giyip bahçeye çıktılar. {O sırada **bahçenin üzerinde**// kuşlar uçuyordu. {Çok karınları acıktı ama **okulda**// yiyecek satılmıyordu. Bu yüzden **gizlice** (X) karşıdaki bakkala gittiler. Bakkal tereyağlı tost yapıyordu {ama **tereyağının üzerinde**// bir parmak toz vardı. Tosttan vazgeçtiler. {Onun yerine **bardakta**// limonata gördüler. Fakat {dikkatli bakınca **bardağın üzerinde**// ölü karıncalar gördüler ve mideleri bulandı. {O sırada raftaki **kutunun içinde**// bir yüzük gördüler. {Hemen **bağcık bağlamak**// için eğilirmiş gibi yapıp yüzüğü ceplerine koydular. {Bakkaldan çıkınca bir **evin yanında**// durdular. Aylin yüzüğü takar takmaz {kendilerini bir **ormanın içinde**// buldular. {Karşıdan da kızağın üzerinde// bir dede geliyordu. {Dikkatli bakınca **kızakta**// bir valiz gördüler. {Tabi ki **valizin içinde**// ne olduğunu merak ettiler. Kızak yakınlaşınca dede valizden bir sepet çıkardı. {Çocuklar "dede, **sepetin içinde**// ne var?" diye sordular. Dede "**sepette**// elma var" dedi. {Sonra da gülümseyerek "**meyve yemek**// çok önemlidir ve elmayı her ısırışınızda bir dilek dileyebilirsiniz" dedi. {İkisi de çok sevindi. Barış {"ben büyükler gibi **kravat takmak**// istiyorum" dedi. {Aylin de saate bakmayı bilmediğini ama birisi sorunca **saati söylemek**// istediğini belirtti. Dede bu dileklerin olması için {Barışa **odun kesmek**// Aylin'e de yemek yapmak görevini verdi. İkisi de bu şartlara çok şaşırdı ama **hemen** (X) kabul ettiler. Görevlerinin başına gittiler. {Barış **dal kırmak**// için dolaşmaya başladı. Aylin de yemek **için** (X) malzeme aramaya başladı. O sırada yerde eski bir gazete gördü. {Resimlerine bakarken **sayfa değiştirmek**// istedi. {Bir sayfada **dil çıkarmak**// çok ayıptır yazıyordu. Bir sayfada da para remi vardı {ama **parada**// bir gariplik vardı. Çünkü üstünde hayvanlar oynuyordu. Aylin çok şaşırdı ama yemek için malzeme aramaya devam etti. Ormanda giderken bir bisiklet buldu **ama** (X) lastikleri inmişti. O sırada aklına bir fikir geldi {ve **lastik şişirmek**// için pompa aradı. Bisikletin **yanındaki** (X) çantada bir pompa buldu. Lastikleri şişirdi ve Barış'ı bulmaya gitti.

Aylin and Barış were really close friends. They both got very bored in school. Because everyone was trying to draw triangles. Both Aylin and Barış had already finished their assignments. It was snowing outside. They put on their coats and went outside. At that time birds were flying over the garden. They were really hungry but they didn't sell any food at school. So they secretly went to the market across. The market had toasted sandwiches with butter but there was an inch of dust on the butter. They gave up the idea of a toasted sandwich. Instead they saw lemonade in a glass. However, upon closer inspection they saw dead ants on the glass and they were disgusted. At that moment they saw a ring in the boxes on the shelf. Immediately they bent down pretending to tie their shoelaces and put the ring in their pocket. When they got out of the market they stood next to a house. As soon as Aylin put the ring on they found themselves in a forest. Towards them came an old man on a sledge. When they looked carefully they saw a suitcase in (on) the sledge. Naturally they wondered

what was in the suitcase. As the sledge came near them the old man took out a basket from the suitcase. The children asked 'grandpa what is in(side) the basket?'. The old man said 'there are apples in the basket'. And he added 'it is very important to eat fruits and every time you bite into the apple you can make a wish'. Both children were very happy. Barış said 'I want to put on (wear) a tie like grown-ups do'. Aylin said she did not know how to tell the time but wished to tell the time when someone asked. The old man told Barış to cut firewood and Aylin to cook food. Both of them were really surprised by these requirements but they accepted right away. They set to work. Barış started to walk around to break branches. Aylin started to look for ingredients to cook. Then she saw an old newspaper on the ground. While looking at the pictures she wanted to flip the page. On one page it wrote that it's rude to stick out (your) tongue. One another page there was a picture of a banknote but there was something strange on the banknote. Because there were pictures of animals dancing on it. Aylin was very surprised but she continued to look for ingredients to cook. When she was walking in the woods she found a bike but the tyres were flat. Then she thought of something (came up with an idea) and looked for a pump to pump the tyres. She found a pump in the bag next to the bike. She pumped the tyres and went to look for Barış.

Sentences occurring in the click task (Dutch only)

Example sentences:
(// marks place where click was inserted)

1) **Ik vind groene appels**// **erg lekker** (*I really like green apples*)
2) **Eet jij altijd met**// **vork en mes?** (*Do you always eat with fork and knive?*)
3) **Anne heeft een**// **gele auto** (*Anne has a yellow car*)

Sentences click task

No.	Sentence in Dutch	Translation in English
1	Ik vind veters// strikken heel makkelijk.	*I think tying laces is really easy.*
2	Kun je die doos in// de kamer zetten?	*Could you put that box in the living room?*
3	Je moet een das// omslaan als het koud is.	*You have to wear a scarf when it's cold.*
4	Wil jij een film kijken// vanavond? [filler]	*Do you want to watch a film tonight?*
5	Mama is in de// keuken aan het koken.	*Mama is cooking in the kitchen.*
6	Je moet je tanden// poetsen voor je gaat slapen.	*You have to clean your teeth before going to bed.*
7	Pim zit op// de skelter van zijn broer.	*Pim sits on his brother's go-card.*
8	Ik kan heel// hard rennen! [filler]	*I can run real fast!*
9	Zullen we verstoppertje// doen in de pauze?	*Shall we play hide and seek during intermission?*
10	Zullen we in// de tuin gaan spelen?	*Shall we play in the garden?*
11	Dit is een boek// over beren en olifanten. [filler]	*This is a book on bears and elephants*
12	Zal ik je boterham// smeren straks?	*Shall I prepare your sandwich later on?*
13	Wij hebben in het// bos gespeeld.	*We have played in the forest.*
14	Zal ik de hond uitlaten// straks? [filler]	*Shall I walk the dog later on?*
15	Ik vind touwtje// springen erg leuk.	*I really like skip the rope.*
16	De bal is onder// de kast gerold.	*The ball ended up under the cupboard.*
17	Sanne moet een kast// opruimen van juf.	*The teacher made Sanne tidy up the cupboard.*
18	Zie je// die grote wolk in de lucht? [filler]	*Do you see that large cloud in the sky?*
19	In de kast onder de// tv ligt mijn speelgoed.	*In the cupboard under the television are my toys*

No.	Sentence in Dutch	Translation in English
20	Maaike// wil piano of gitaar spelen. [filler]	*Maaike wants to play the piano or the guitar.*
21	Zal ik je een mop// vertellen over 3 Belgen?	*Shall I tell a joke about three Belgians?*
22	Wie heeft er op// de fluit op de geblazen?	*Who blew the whistle?*
23	Ik heb// een oranje jas. [filler]	*I have an orange coat.*
24	De kleren in de// kist zijn erg oud.	*The clothes in the box are very old.*
25	Sem zet zijn// fiets in het rek. [filler]	*Sem puts his bicycle in the bicycle-stand.*
26	Tijn wil verstoppertje// spelen in de pauze.	*Tijn wants to play hide and seek.*
27	Fleur wil een zandkasteel// maken in de zandbak.	*Fleur wants to build a castle of sand in the play ground.*
28	Ik vind vuurwerk afsteken// eng. [filler]	*I think setting fireworks on fire is scary.*
29	Je moet je neus// snuiten van mama.	*Mummy says that you have to blow your nose.*
30	Mijn zus moet de tafel// dekken van mama.	*Mummy makes my sister set the table.*
31	Zal ik je beer in de// koffer stoppen?	*Shall I put your bear in the suitcase?*
32	Oma gaat een taart// bakken als ik kom.	*Granny is going to bake a cake when I get there.*
33	Wil jij een film kijken// vanavond? [filler]	*Do you want to watch a film tonight?*
34	Papa is in// de koffie aan het roeren.	*Daddy is stirring (in) his coffee.*
35	Vind jij een hond// aaien niet eng?	*Don't you think petting a dog is scary?*
36	Ik kan kan heel// hard rennen! [filler]	*I can run really fast!*
37	Doe je wat voer in de// kom van de vis?	*Can you put some food in the fishbowl?*
38	Papa gaat een muur// verven thuis.	*Daddy is going to paint a wall at home.*
39	Iris gaat bij haar tante en oom// logeren. [filler]	*Iris is going to spend the night at her aunt and uncle's.*
40	Ik zag Bram op// de schommel zitten.	*I saw Bram sitting on the swing.*
41	Je moet een das// omdoen in de winter.	*You have to wrap a scarf around in winter.*
42	Zal ik de hond uitlaten// straks? [filler]	*Shall I walk the dog later?*
43	Ik vind bellen// blazen erg leuk.	*I like making bubbles.*
44	Ik heb mijn handen onder de// kraan gewassen.	*I have washed my hands under the tap.*
45	Floor kijkt// een film. [filler]	*Floor is watching a film.*
46	De hond ligt onder// de tafel te slapen.	*The dog is sleeping under the table.*
47	Zie je// die grote wolk in de lucht? [filler]	*Do you see that big cloud in the air?*
48	Zal ik de bladzijde// omslaan en doorlezen?	*Shall I turn the page and read on?*
49	Ben jij op de// fiets gekomen?	*Did you come by bike?*
50	Wij gaan een liedje// zingen met de juf.	*We are going to sing a song with the teacher.*
51	Els koopt// bananen en mandarijnen. [filler]	*Els buys bananas and clementines.*
52	Sara is op// de step naar school gekomen.	*Sara has come to school by step.*
53	Wil jij suiker in de// thee hebben?	*Do you want sugar in your tea?*
54	Ik vind vuurwerk afsteken// eng. [filler]	*I think setting fireworks on fire is scary.*
55	Het water in// het bad is koud.	*The water in the bath is cold.*
56	Sem zet zijn// fiets in het rek. [filler]	*Sem puts his bike in the rack.*
57	Ik mocht op de// stoel gaan staan.	*I was allowed to stand on the chair.*
58	Je moet je handen// wassen voor het eten.	*You have to wash your hands before dinner.*

Sentences occurring in the mixed-words task (Dutch only)

Example sentences:
1) **Juf vertelt verhaal een** (*The teacher tells story a*)
2) **Oma koopt appels en peren** (*Granny buys apples and pears*)
3) **Madelief heeft trui een roze** (*Madelief has sweater a pink*)

Sentences mixed-words task

| No.	Sentence in Dutch	Translation in English
1	Mijn zus moet de dekken tafel vanavond.	*My sister has to the set table tonight.*
2	Zit je beer koffer de in?	*Is your bear suitcase the in?*
3	Oma gaat een bakken taart.	*Granny is going to a bake cake.*
4	Nina kan goed hinkelen. [filler]	*Nina can jump well.*
5	Ik wil graag suiker koffie de in.	*I would like sugar coffee the in.*
6	Vind jij een aaien hond niet eng?	*Don't you think petting a dog is scary?*
7	Opa rookt een sigaar. [filler]	*Grandpa is smoking a cigar.*
8	Mama doet pap kom de in.	*Mummy does breakfast bowl the in.*
9	Papa gaat een verven muur thuis.	*Daddy is going to a paint wall at home.*
10	Iris gaat bij haar tante en oom logeren. [filler]	*Iris is going to spend the night at her aunt and uncle's.*
11	Bram zit schommel op de.	*Bram is swing on the.*
12	Je moet een omdoen das.	*You have to a wrap around scarf.*
13	Sanne heeft een blauwe rugzak. [filler]	*Sanne has a blue backpack.*
14	Ik vind blazen bellen erg leuk.	*I think blow bubbles is great fun.*
15	Ik houd mijn handen kraan onder de.	*I hold my hands tap under the.*
16	Floor kijkt een film. [filler]	*Floor is watching a film.*
17	De hond ligt tafel de onder.	*The dog is table the under.*
18	Joris eet een pannenkoek. [filler]	*Joris is eating a pancake.*
19	Zal ik de omslaan bladzijde en doorlezen?	*Shall I the turn page and read on?*
20	Ben jij fiets op de gekomen?	*Have you come bike on the?*
21	Wij gaan een zingen liedje.	*We are going to a sing song.*
22	Els koopt bananen en mandarijnen. [filler]	*Els buys bananas and clementines.*
23	Sara gaat step op de.	*Sara goes step on the.*
24	Wil jij suiker thee in de hebben?	*Do you want sugar tea in the?*

| No.	Sentence in Dutch	Translation in English
25	Ik ga naar opa en oma. [filler]	*I am going to granny and grandpa.*
26	Het water bad het in is koud.	*The water bath the in is cold.*
27	Heb jij een rode fiets? [filler]	*Do you have a red bike?*
28	Ik stond stoel op de.	*I was standing chair on the.*
29	Voor het eten moet je je wassen handen.	*You have to your wash hands before dinner.*
30	Kun jij al zelf strikken veters?	*Can you already tie laces?*
31	Zet die doos maar kamer de in.	*Put the box room the in.*
32	Je moet een omslaan das als het koud is.	*You have to a wrap around scarf when it's cold.*
33	Els gaat naar school. [filler]	*Els goes to school.*
34	Mama is keuken in de.	*Mummy is kitchen in the.*
35	Voor het slapen moet je je poetsen tanden.	*Before sleeping, you have to your brush teeth.*
36	Pim zit skelter de op.	*Pim is cart the on.*
37	Fie lust graag appels en tomaten. [filler]	*Fie likes apples and tomatoes.*
38	Zullen we zo doen verstoppertje?	*Shall we play hide and seek?*
39	Zullen we tuin in de spelen?	*Shall we play garden in the?*
40	Dit is een boek over beren en olifanten. [filler]	*This is a book about bears and elephants.*
41	Zal ik je smeren boterham straks?	*Shall I your prepare sandwich later?*
42	Ik speel graag bos het in.	*I like playing forest the in.*
43	Puck voert haar vis. [filler]	*Puck is feeding her fish.*
44	Ik vind springen touwtje erg leuk.	*I like jump rope very much.*
45	Er loopt een spin kast onder de.	*There is a spider cupboard under the.*
46	Sanne moet van juf een opruimen kast.	*Sanne has to a cupboard tidy (from the teacher).*
47	Tom heeft een blauwe telefoon. [filler]	*Tom has a blue cellphone.*
48	Ligt er iets tv onder de?	*Is there something TV under the?*
49	Maaike wil piano of gitaar spelen. [filler]	*Maaike wants to play piano or guitar.*
50	Zal ik je een vertellen mop?	*Shall I a tell joke?*
51	Wie heeft er fluit op de geblazen?	*Who has blown whistle on the?*
52	Ik heb een oranje jas. [filler]	*I have an orange coat.*
53	De verkleedkleren liggen kist de in.	*The play clothes are cupboard the in.*
54	Ik hou van computeren. [filler]	*I like gaming.*
55	Tijn wil spelen verstoppertje.	*Tijn wants play hide and seek.*
56	Fleur wil een maken zandkasteel.	*Fleur wants to a make sandcastle.*
57	Mieke heeft paarse schoenen. [filler]	*Mieke has purple shoes.*
58	Je moet je snuiten neus van mama.	*You have to your blow nose from mummy.*

Eye-tracking materials

Targetscreen pictures
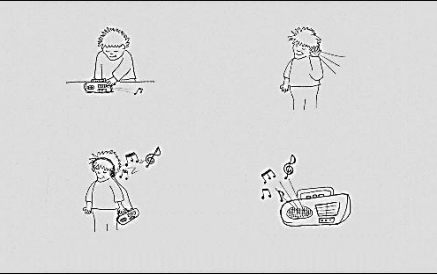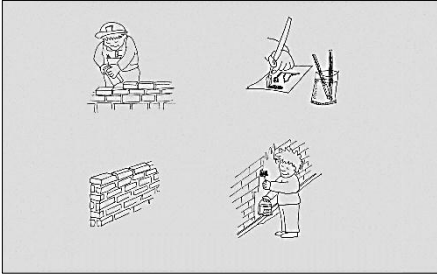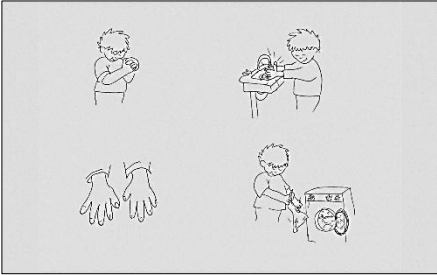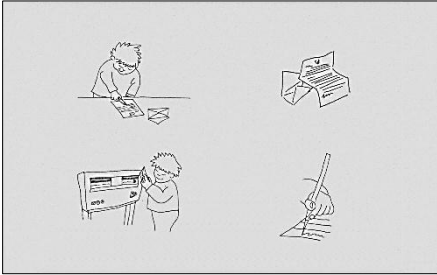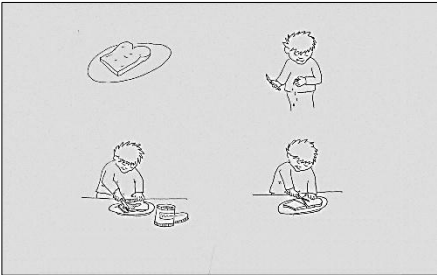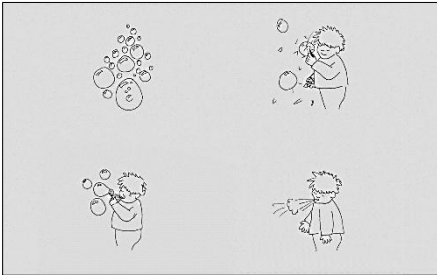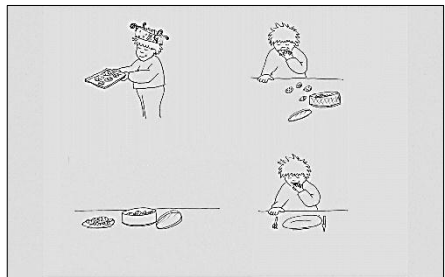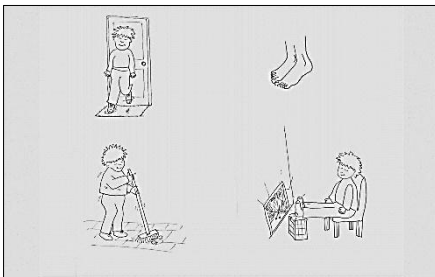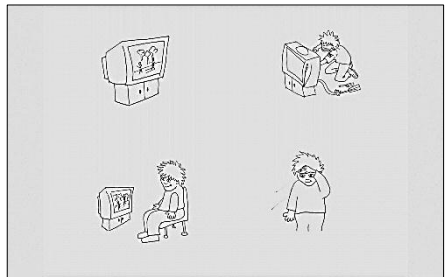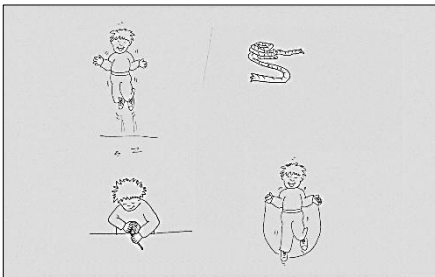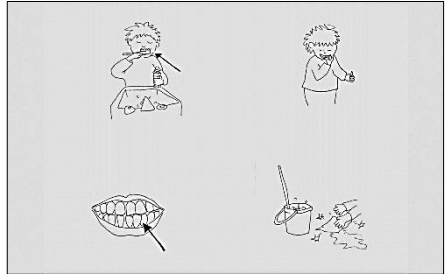
Example and filler pictures
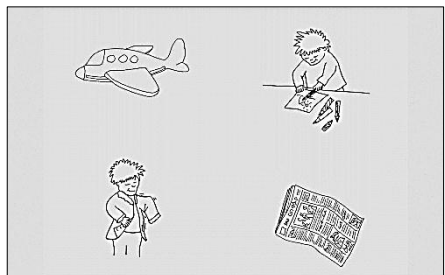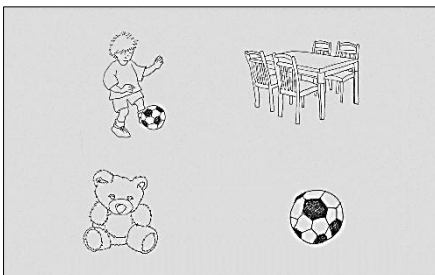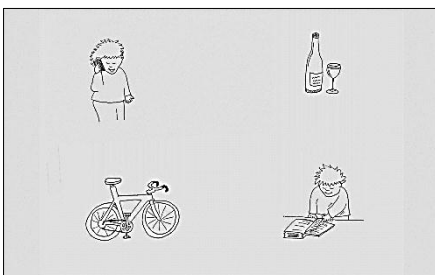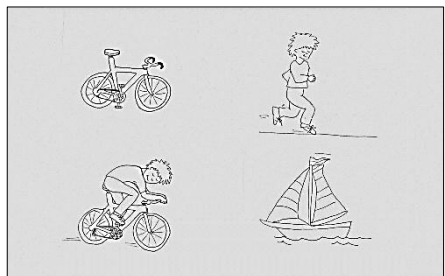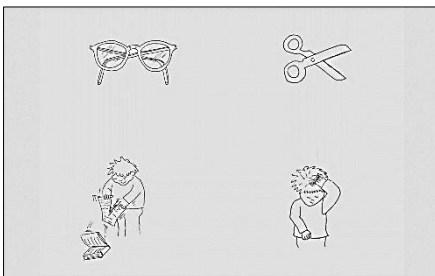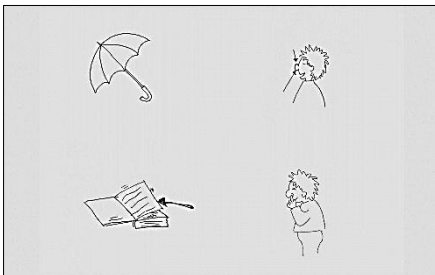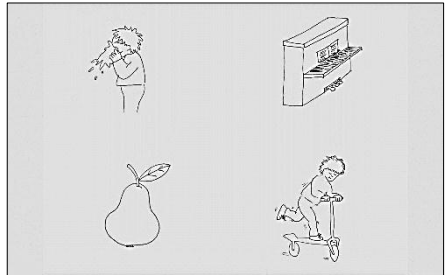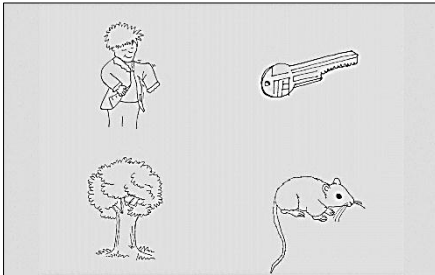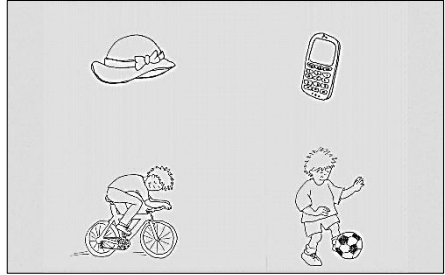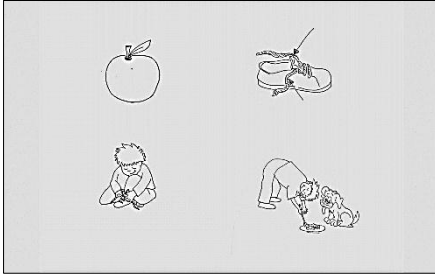
Target items eye-tracking task

Version 1	Version 2	Translation in English	Translation in English
bellen kapotmaken	bellen blazen	*break bubbles*	*blow bubbles*
boterham snijden	boterham smeren	*cut a sandwich*	*spread a sandwich*
brief schrijven	brief posten	*write a letter*	*post a letter*
handen vouwen	handen wassen	*fold hands (prey)*	*wash hands*
muur verven	muur bouwen	*paint a wall*	*build a wall*
muziek luisteren	muziek aandoen	*listen to music*	*put on music*
taart eten	taart bakken	*eat a cake*	*bake a cake*
tanden poetsen	tand trekken	*brush teeth*	*pull out teeth*
touw oprollen	touw springen	*roll a rope*	*jump rope*
tv kijken	tv repareren	*watch TV*	*repair a TV*
voeten warmen	voeten vegen	*warm feet*	*clean feet*
koekjes eten	koekjes uitdelen	*eat biscuits*	*share biscuits*

Children's multiword segments (per grade) in the dictation tasks

Grade	Multiword segments	Translation in English	Raw frequency	Assigned to category
K1	*en met*	and with	70,700,000	En X
	en ik	and I	67,000,000	En X
	de hele tijd	all the time	35,200,000	DET(+ADJ)+N
	in het weekend	in the weekend	33,200,000	PP
	in de trein	in the train	32,500,000	PP
	in het park	in the park	23,700,000	PP
	in de boot	in the boat	17,500,000	PP
	op vakantie	on holidays	14,400,000	PP
	en toen	and then	12,700,000	En X
	en toen	and then	12,700,000	En X
	en toen	and then	12,700,000	En X
	naar het zwembad	to the swimming pool	10,100,000	PP
	hadden we	had we	8,490,000	SUBJ+V
	naar een hotel	to a hotel	7,480,000	PP
	gingen we	went we	3,470,000	SUBJ+V
	gingen we	went we	3,470,000	SUBJ+V
	gingen we	went we	3,470,000	SUBJ+V
	we gingen	we went	3,450,000	SUBJ+V
	gisteren heb ik	yesterday have I	2,330,000	SUBJ+V
	de schuur	the shed	2,210,000	DET(+ADJ)+N
	ik moest eerst	I must first	1,230,000	SUBJ+V
	heel leuk	great fun	1,230,000	Other
	met de tram	by tram	886,000	PP
	het deed	it did	880,000	SUBJ+V
	wij gingen	we went	576,000	SUBJ+V
	wij gingen	we went	576,000	SUBJ+V
	mocht ik	may I	521,000	SUBJ+V
	in een bolletje	in a ball	429,000	PP
	m'n knieën	with knees	241,000	DET(+ADJ)+N
	allemaal leuke dingen gedaan	all nice things done	192,000	Other
	ging d'r	went there	189,000	Other

Grade	Multiword segments	Translation in English	Raw frequency	Assigned to category
	ik mocht 't alleen	I was allowed it alone	10 (NB: *Ik mocht HET alleen*: 188,000)	Sent
	ik ging een keertje	I went a time	186,000	SUBJ+V
	spelletjes doen	play games	182,000	DET+OBJ+V
	een rondleiding krijgen	get a tour	160,000	DET+OBJ+V
	gisteren ging ik	yesterday I went	137,000	SUBJ+V
	de snoepjes	the sweets	122,000	DET(+ADJ)+N
	pannenkoeken eten	eat pancakes	112,000	DET+OBJ+V
	en toen met de trein	and then by train	46,800	En X
	naar Schiermonnikoog	to Schiermonnikoog	44,600	PP
	naar het strand toe	to the beach (to)	38,300	PP
	het lammetje	the lamb	31,200	DET(+ADJ)+N
	het lammetje	the lamb	31,200	DET(+ADJ)+N
	gewone kleren	normal clothes	30,200	DET(+ADJ)+N
	gaan we pizza eten	we go eat pizza	27,800	Sent
	ging zwemmen	went swimming	26,000	Pred
	m'n handen omhoog	with my hands raised	23,600	Other
	en toen staan	and then standing	21,700	En X
	een oranje bloem	an orange flower	21,300	DET(+ADJ)+N
	ging ik thuis eten	I went eating at home	19,400	Sent
	toen gevallen	then fallen	16,300	Other
	en drinkt melk	and drinks milk	13,500	En X
	een schat zoeken	search a treasure	13,300	DET+OBJ_V
	had ik Melissa	I had Melissa	13,000	Sent
	eet gras	eat grass	10,900	Pred
	zwempak aan	swimming suit on	9,800	Other
	gingen we pannenkoeken eten	we went eating pancakes	8,980	Sent
	ik ging naar Texel	I went to Texel	6,630	Sent
	mama slapen	mummy sleep	4,960	DET+OBJ+V
	verstoppertje doen	play hide and seek	3,590	DET+OBJ+V
	op m'n zusje passen	look after my sister	2,220	DET+OBJ+V
	een lammetje in de klas	a lamb in the classroom	1,510	DET(+ADJ)+N **with PP**
	ik mocht 't alleen	I was allowed it alone	10 (NB: *Ik mocht HET alleen*: 188,000)	Sent
	had een blauwe vlek	had a blue stain	9	Pred
	Spiderman pak gekregen	got Spiderman suit	9	DET+OBJ+V
	ik ging naar Turkije op vakantie	I went to Turkey on holidays	2	Sent
	had ik vissen gevangen	I had caught fish	1	Pred
	ik had deze pet gekocht	I had bought this cap	1	Sent
	heb de hangmat opgehangen	have put the hammock	0	Pred

Grade	Multiword segments	Translation in English	Raw frequency	Assigned to category
	met papa	with daddy		
	ik ging met m'n zwempak zwemmen	I went swimming with my swimming suit	0	Sent
	omdat 'ie dan z'n moeder terug kan vinden	because he can ten find his mother back	0	Sent
	kon gisteren ook dragen	could bear yesterday too	0	Pred
	ik ging toen op het strand spelen	I went then playing on the beach	0	Sent
	aan oma Volkel	to granny Volkel	0	PP
K2	*en dan*	and then	52,900,000	En X
	moet je	must you	50,400,000	SUBJ+V
	ik vind	I find	36,900,000	SUBJ+V
	en daar	and there	29,700,000	En X
	maak je	make you	22,300,000	SUBJ+V
	als je	if you	17,600,000	Other
	mag je	may you	13,100,000	SUBJ+V
	ik wil ook	I want too	13,000,000	SUBJ+V
	vandaag heb ik	today have I	4,890,000	SUBJ+V
	ik ging	I went	3,400,000	SUBJ+V
	ik ging	I went	3,400,000	SUBJ+V
	ik ging	I went	3,400,000	SUBJ+V
	m'n moeder	my mother	2,350,000	DET(+ADJ)+N
	mijn broer	my brother	2,140,000	DET(+ADJ)+N
	het was leuk	it was fun	2,000,000	Sent
	mijn zus	my sister	1,820,000	DET(+ADJ)+N
	m'n zus	my sister	624,000	DET(+ADJ)+N
	een deken	a cover	434,000	DET(+ADJ)+N
	en ikzelf	and I self	422,000	En X
	en m'n vader	and my father	240,000	En X
	in m'n haar	in my hair	178,000	PP
	ook op hockey	on hockey too	93,800	PP
	een armband gemaakt	made a bracelet	72,900	DET+OBJ+V
	tekening maken	make a drawing	70,600	DET+OBJ+V
	nog eventjes spelen	play for a while	25,500	DET+OBJ+V
	had ik gezwommen	had I swum	22,600	Sent
	had ik gezwommen	had I swum	22,600	Sent
	op turnen	doing gymnastics (lit, on gymnastics)	19,800	PP
	hut bouwen	build a tree house	14,100	DET+OBJ+V
	naar de bos	to the forest	11,900	PP
	die is groen	that is green	7,180	Sent
	en m'n andere zus	and my other sister	7,090	En X
	naar de hockey kijken	watch the hockey	1,320	DET+OBJ+V
	heb ik touwtje gesprongen	have I jumped rope	411	Sent
	heeft een zwembad gewonnen	has won a pool	7	Pred
	m'n parfum lekker	my perfume nice	2	DET(+ADJ)+N

Grade	Multiword segments	Translation in English	Raw frequency	Assigned to category
	ik heb parfum spuit	I have perfume spray	0	Sent
	zag er blauw met wit uit	looked blue with white	0	Pred
E1	*opa en oma*	granddad and granny	17,500,000	DET(+ADJ)+N
	opa en oma	granddad and granny	17,500,000	DET(+ADJ)+N
	1 april	1 April	12,500,000	Other
	en toen	and then	9,810,000	En X
	er waren	there were	6,170,000	SUBJ+V
	en rijden	and drive	994,000	En X
	soort huis	kind of house	598,000	DET(+ADJ)+N
	op een bunker	on a bunker	171,000	PP
	vriendje geweest	friend been	11,900	DET+OBJ+V
	waarin je kan verdwalen	in which you can get lost	8,550	Sent
	met het blauwe lintje	with the blue ribbon	5,140	PP
	negen puppy's	nine puppies	1,040 (NB: negen puppies: 1,300; 9 puppy's: 2,190; 9 puppies: 2,580)	DET(+ADJ)+N
	heleboel beesten	numerous animals	1,400	DET(+ADJ)+N
	en de muren waren zacht	and the walls were soft	1	Other
	gingen we in een bunker zitten	went we sitting in a bunker	0	Sent
	en stenen en auto's gespeeld	and stones and cars played	0	En X
E2	*met een*	with a	360,000,000	PP
	en de	and the	277,000,000	En X
	en de	and the	277,000,000	En X
	naar de	to the	195,000,000	PP
	naar de	to the	195,000,000	PP
	één dag	one day	26,600,000 (NB: 1 dag: 30,200,000)	DET(+ADJ)+N
	op de fiets	by bike	50,000,000	PP
	waar je	where you	44,100,000	Other
	heel veel	many many	35,900,000	Other
	in zijn hand	in his hand	15,700,000	PP
	op vakantie	on holidays	14,300,000	PP
	en toen	and then	9,810,000	En X
	en toen	and then	9,810,000	En X
	als ik	if I	4,390,000	Other
	drie keer	three times	3,980,000 (NB: *3 keer*:	DET(+ADJ)+N

Grade	Multiword segments	Translation in English	Raw frequency	Assigned to category
			4,290,000)	
	daar zijn	there are	6,560,000	Sent
	voor je mond	for your mouth	5,260,000	PP
	ging ik	went I	3,950,000	SUBJ+V
	en tv kijken	and watch TV	2,530,000	En X
	één dag	one day	2,390,000	DET(+ADJ)+N
	en ik haalde	and I got	2,260,000	En X
	ik kan	I can	1,940,000	SUBJ+V
	een vriendje	a friend	723,000	DET(+ADJ)+N
	De Vliegende Hollander	The Vliegende Hollander	718,000	DET(+ADJ)+N
	met een camper	with a camper	700,000	PP
	de tas	the bag	696,000	DET(+ADJ)+N
	in zijn gedachten	in his thoughts	642,000	PP
	in kan	can in	601,000	Other
	en een looping	and a looping	525,000	En X
	in de straaljager	in a jet	454,000	PP
	zei die	said he	289,000	SUBJ+V
	pake en beppe	granddad and granny (Frisian)	281,000	DET(+ADJ)+N
	en de diabolo	and the diabolo	262,000	En X
	speelde hij ook mee	he played with (me) too	225,000	SUBJ+V
	in de Python	in the Python	203,000	PP
	een grap uit	a joke out	158,000	DET(+ADJ)+N
	pannenkoeken eten	eat pancakes	116,000	DET+OBJ+V
	en de bobsleebaan	and the bobsleigh track	86,300	En X
	een kapje	a cap	65,800	DET(+ADJ)+N
	een 1 april grap uit	a April the first joke out	45,100	DET(+ADJ)+N
	de hoepel	a hoolahoop	40,400	DET(+ADJ)+N
	elfjes en trollen	fairies and trolls	38,900	DET(+ADJ)+N
	de paardentram	the horsetram	35,600	DET(+ADJ)+N
	snoepjes op	sweets on	32,500	Other
	met een vriendinnetje spelen	play with a friend	32,400	DET+OBJ+V
	frietjes eten	eat fries	29,800	DET+OBJ+V
	en Joris en de Draak	and Joris and the Dragon	13,300	En X
	Efteling gegaan	gone Efteling	8,480	DET+OBJ+V
	het schommelschip	the swing-ship	3,960	DET(+ADJ)+N
	streetdance geweest	streetdance been	608	DET+OBJ+V
	de slakkenbaan	the snail route	69	DET(+ADJ)+N
	spugen als je misselijk bent	vomit if you are sick	4	Other
	feestjes van mijn neefjes	party of my nephews	2	DET(+ADJ)+N
	en ook nog de Droomvlucht	and also still the Droomvlucht	1	En X
	doppel maak	make a double	1	DET+OBJ+V
	wat ie zou fantaseren	what he would fantasise	1	Sent

Correlation scores on single-word responses by background factors

(1) **Correlation scores on single-word responses by sociolinguistic factors, school tests and Raven's intelligence test for Dutch children in K1 and K2**

		Raven	*Kleutertaal* Language test	TV frequency	Reading at home frequency
% single-word responses dictation	Pearson corr.	.117	.032	-.120	.214
	Sig. (2-tailed)	.547	.863	.513	.247
	N	29	31	32	31
% single-word sentence segm.	Pearson corr.	-.047	.165	.122	.044
	Sig. (2-tailed)	.793	.343	.480	.800
	N	33	35	36	34
% single-word last-part repetition	Pearson corr.	.079	.191	.048	.145
	Sig. (2-tailed)	.675	.265	.776	.399
	N	34	36	37	36
% single-word responses click	Pearson corr.	.359*	.150	.069	-.002
	Sig. (2-tailed)	.044	.369	.669	.989
	N	32	38	41	40

(2) **Correlation scores on single-word responses by sociolinguistic factors, school tests and Raven's intelligence test for Dutch children in E1 and E2**

		Raven	CITO Woordenschat Language test	TV frequency	Reading at home frequency
% single-word responses dictation	Pearson corr.	-.222	.193	-.136	.209
	Sig. (2-tailed)	.221	.274	.445	.235
	N	32	34	34	34
% single-word sentence segm.	Pearson corr.	-.406*	-.290	-.210	-.207
	Sig. (2-tailed)	.016	.096	.226	.234
	N	35	34	35	35
% single-word last-part repetition	Pearson corr.	.134	.077	.077	-.052
	Sig. (2-tailed)	.437	.661	.657	.765
	N	36	35	36	36
% single-word responses click	Pearson corr.	.473**	.357*	.007	.094
	Sig. (2-tailed)	.003	.028	.965	.571
	N	37	38	39	39

(3) **Correlation scores on single-word responses by TV frequency and reading frequency at home for all Turkish children**

		TV frequency	Reading at home frequency
% single-word sentence segmentation task	Pearson corr.	-.082	.206
	Sig. (2-tailed)	.615	.203
	N	40	40
% single-word last-part repetition task	Pearson corr.	.157	.150
	Sig. (2-tailed)	.334	.356
	N	40	40

List of abbreviations

ADJ	adjective
ADV	adverb
DET	determiner
GEN	genitive
GEN.LOC.	general locative
HF	high frequency
LF	low frequency
LOC	locative
N	noun
NV	noun+verb-combination
OBJ	object
POST	postposition
PP	prepositional phrase
Pred	predicate
PREP	preposition
Sent	sentence
SUBJ	subject
V	verb
Vpres.	verb in present tense

Nederlandse samenvatting

De hoofdvraag in dit proefschrift is wat de basiseenheden van taal (de 'building blocks of language') voor kinderen zijn wanneer zij taal horen of produceren en hoe de basiseenheden in taal verschillen voor kinderen met een verschillende taalachtergrond – Nederlands, Turks, of Turks-Nederlands – en met een verschillende kennis van (geschreven) taal. Hierbij is tevens de vraag gesteld in hoeverre veelvoorkomende woordcombinaties, zoals 'handen wassen', als de basiseenheden van taal beschouwd zouden kunnen worden, in plaats van losse woorden die doorgaans als de basiseenheden van taal worden gezien.

Om een antwoord te vinden op deze vragen zijn voor dit onderzoek zes verschillende experimenten uitgevoerd. Van deze experimenten waren er enkele meer gericht op de manier waarop kinderen taal bewust in stukjes hakken, en probeerden enkele inzicht te verkrijgen in de manier waarop kinderen taal onbewust segmenteren als zij taal horen of produceren. Deze taken, die in Hoofdstuk 4, 6 en 7 besproken worden, vallen onder respectievelijk 'meer offline' en 'meer online' taken – een categorisatie die in dit proefschrift ook besproken zal worden (zie Hoofdstuk 3).

De deelnemers aan de taken waren eentalige Nederlandse, eentalige Turkse en (in 'pilotsessies') tweetalige Turks-Nederlandse kinderen in de leeftijd van 3-10 jaar. De reden om het onderzoek op deze groepen te richten was ten eerste dat kinderen in deze leeftijdscategorie leren lezen en schrijven. Dit gaf ons de kans het effect van geletterdheid te bestuderen op de manier waarop kinderen taal segmenteren en eenheden in taal verwerken. Daarnaast zou het testen van eentalige Nederlandse, eentalige Turkse en tweetalige Turks-Nederlandse kinderen ons de mogelijkheid bieden om de invloed van een bepaalde taal (agglutinerend of isolerend) en een- of tweetaligheid te onderzoeken.

Dat geletterdheid een effect zou hebben op de eenheden die kinderen in taal herkennen en verwerken, was te verwachten op basis van eerder onder-

zoek naar met name metalinguïstisch bewustzijn – het bewustzijn van 'taal als taal' en van structuren in talen. In lijn met deze onderzoeken (zie bv. Fox & Routh, 1975; Gombert, 1992; Holden & MacGinitie, 1972; Homer, 2000; Homer & Olson, 1999; Karpova, 1966, orig. 1955; Kolinsky, Cary & Morais, 1987; Kurvers, 2002; Kurvers & Uri, 2006; Lazo, Pumfrey & Peers, 1997; Melzer & Herse, 1969; Morais et al., 1986; Morais & Kolinsky, 1995; Morris, 1993; Olson, 1994; Ramachandra & Karanth, 2007; Ravid & Tolchinsky, 2002; Roberts, 1992; Tunmer, Bowey & Grieve, 1983) was onze hypothese dat geletterde kinderen taal vaker langs woordgrenzen segmenteren en verwerken dan ongeletterde kinderen, die nog geen instructie op school hebben gehad over wat woorden zijn en voor wie daarom woorden nog niet per se de basiseenheden van taal zouden hoeven zijn (zie Hoofdstuk 2).

Dat de taalachtergrond van kinderen de eenheden die zij in taal zouden herkennen en verwerken zou beïnvloeden, was te verwachten op basis van de structuren van de talen die in dit onderzoek meegenomen werden. Het Turks is een agglutinerende taal, waarin woorden vaak meerdere morfemen bevatten – meer dan in het Nederlands. Nederlands is daarentegen een isolerende taal, waarin functiewoorden en (andere) betekenisvolle woorden doorgaans gescheiden worden. Zo bestaat het zinsdeel 'op de tafel' in het Nederlands uit drie woorden; in het Turks uit één, namelijk '*masada*' (*masa* = tafel; *-da* = algemene plaatsaanduiding). Door te kijken hoe Nederlandse en Turkse kinderen in hun moedertaal zinnen in stukjes segmenteren of verwerken, hoopten wij te kunnen bepalen of en wat de invloed van een bepaalde taal – of ook het kennen van een bepaalde geschreven taal – is op de eenheden die kinderen herkennen en verwerken in taal.

Dat daarnaast ook het al dan niet tweetalig zijn van kinderen hun bewustzijn beïnvloedt van eenheden in taal, en mogelijk ook de eenheden die zij in meer online taken verwerken, werd ook op basis van eerder onderzoek verwacht. Onderzoek naar tweetaligheid heeft aangetoond dat kinderen die op jonge leeftijd twee talen leren spreken, eerder metalinguïstische kennis verwerven (Bialystok, 1986b; Bialystok, Craig & Luk, 2008; Francis, 1999). Daar het bewustzijn van eenheden in taal ook onder het metalinguïstisch bewustzijn valt, werd in dit onderzoek, in een exploratieve pilotstudie, bekeken of tweetalige Turks-Nederlandse kinderen inderdaad een groter bewustzijn van woorden in gesproken taal lieten zien. In dit proefschrift is hiervoor geen bewijs gevonden, maar dit heeft mogelijk te maken met de grote dominantie in het Nederlands van de tweetalige Turks-Nederlandse kinderen in dit onderzoek: de kinderen waren simpelweg niet 'gebalanceerd tweetalig', en daarmee niet per se representatief voor de beoogde doelgroep in deze studie.

De uitkomsten van dit onderzoek geven meer inzichten in de basiseenheden die Nederlandse en Turkse kinderen in taal herkennen. Zolang niet met zekerheid gesteld kan worden dat woorden psychologische realiteit hebben (het tegenovergestelde is vaak ook beweerd, zie bijvoorbeeld Olson, 1994, 1996), kan de vraag gesteld worden of het aanleren van deze eenheden de taalverwerving bij kinderen niet eerder verstoort dan stimuleert (zie Hoofdstuk 2). In de Cognitieve Linguïstiek is ook wel gesteld dat de eenheden die mensen in taal verwerken niet uit kleine stukjes of uit woorden bestaan, maar dat er in taal grotere eenheden te herkennen zijn die mensen verwerken. Deze multiwoord-eenheden, waaraan doorgaans wordt gerefereerd als '(multiwoord) constructies', '(lexicale) chunks', of 'complex lexical items' (zie Hoofdstuk 2), zouden even goed als de basiseenheden van taal beschouwd kunnen worden door kinderen, vooral als zij vaak voorkomen in ons taalgebruik (zie Hoofdstuk 2).

In **Hoofdstuk 1** van dit proefschrift wordt deze achtergrond ook beschreven – zij het uitgebreider dan hierboven. In dit hoofdstuk wordt daarnaast ook een uitgebreidere schets en uitleg gegeven van welke onderwerpen in welke hoofdstukken van dit proefschrift aan bod komen.

Hoofdstuk 2 bevat het theoretisch kader van dit proefschrift.

Onderzoek naar de 'meer offline' ontwikkeling van metalinguïstisch bewustzijn bij kinderen (het bewustzijn van structuren van taal) heeft een relatie aangetoond tussen woordbewustzijn en geletterdheid: kinderen, en volwassenen, worden zich doorgaans pas bewust van het feit dat alfabetische talen uit woorden bestaan als zij leren lezen en schrijven, en die woord-eenheden tegenkomen. Op grond hiervan zou verwacht kunnen worden dat voor geletterden woorden ook als de basiseenheden van taal kunnen fungeren in meer onbewuste taalverwerking.

In 'meer online' onderzoek binnen de Cognitieve Linguïstiek, dat meer gericht is op onbewuste processen, is echter al vaak beweerd dat de eenheden die mensen verwerken in taal niet uit enkele woorden bestaan, maar dat er multiwoord-eenheden zijn – combinaties van woorden die vaak samen voorkomen – die mensen verwerken als zij taal produceren of horen. Volgens de Cognitieve Linguïstiek zou een combinatie van bijvoorbeeld 'handen' met 'wassen', in 'handen wassen', als één eenheid verwerkt kunnen worden. Minder frequent voorkomende sequenties, zoals bijvoorbeeld 'handen vouwen', zouden echter niet als één eenheid gezien of verwerkt hoeven te worden.

Naast deze algemene theorieën over eenheden in taal, wordt in Hoofdstuk 2 ook omschreven hoe talen met verschillende structuren op verschillende wijze gesegmenteerd en verwerkt zouden worden, en hoe de resultaten van 'offline' taken, die gericht zijn op bewuste oordelen van mensen over taal, zich verhouden tot resultaten die gevonden zijn in 'online' taken, die gericht zijn op onbewuste mentale processen. Ook wordt in dit hoofdstuk kort omschreven hoe bovenstaande bevindingen hebben bijgedragen aan (methodologische) keuzes die voor dit onderzoek zijn gemaakt, en welke aspecten hierbij in overweging zijn genomen.

Hoofdstuk 3 is een verdieping op het verschil tussen offline en online taken. In de literatuur zoals die tot op heden ingaat op beide soorten taken, wordt het verschil in offline en online doorgaans gekenmerkt door (eerder genoemde) verschillen in bewustzijn/onbewustzijn en door de tijdsdruk die er in taken is. Er wordt meestal gesteld dat taken waarin géén tijdsdruk is (en participanten dus rustig hun antwoorden kunnen overwegen) tot de offline taken behoren, terwijl taken met tijdsdruk normaliter als online worden beschouwd (zie Fernández & Smith Cairns, 2011; Sekerina, Fernandéz & Clashen, 2008).

In dit hoofdstuk wordt deze dichotomie in de indeling van taken in twijfel getrokken, en wordt een continuüm van taken voorgesteld – van taken die meer of minder offline of online zijn. Reden hiervoor is dat er, behalve het bewustzijn en de tijdsdruk, veel meer factoren in literatuur over offline en online taken genoemd worden. Deze factoren omvatten bijvoorbeeld de methode van aanbieding van de taak (op schrift of auditief), het al dan niet bekend zijn van het specifieke doel van een taak, de taal die gebruikt wordt in een taak (natuurlijk taalgebruik of aangepast), en dat wat participanten in een taak moeten doen (taal manipuleren, definiëren of enkel verwerken). Factoren als deze worden doorgaans echter niet meegerekend in de categorisatie van taken. Om een compleet beeld te vormen en taken goed te kunnen categoriseren, zouden ze echter wel in ogenschouw genomen dienen te worden. In dit onderzoek wordt daarom een lijst gegeven van tien punten, waarvan bekeken zou moeten worden of een taak ze heeft (in dat geval zou het punt een '1' moeten krijgen), of niet (dit levert een '0' op). De lijst van punten, zoals gegeven in Figuur 1 in dit onderzoek, ziet er als volgt uit:

	Offline kenmerk
1	Bewust van een taaltaak
2	Voorbereid materiaal
3	Onnatuurlijk taalgebruik in de taak
4	Talige input is afgerond op het moment dat participant aan de beurt is
5	Geen tijdslimiet
6	Beoordeling gevraagd
7	Manipulatie gevraagd
8	Talige verwerking gevraagd
9	Specifiek doel van de taak bekend
10	Methode van aanbieden van de taak is lezen of schrijven

Figuur 1: Omschrijving van tien offline kenmerken

Wanneer deze lijst van punten naast verschillende taken wordt gelegd, blijkt wel dat een continuüm van meer offline naar meer online taken waarschijnlijker is dan een dichotomie: taken verschillen immers in het aantal punten dat zij in totaal voor deze offline kenmerken krijgen.

In **Hoofdstuk 4** komen de opzet en resultaten van verschillende vooronderzoeken (de pilotstudies) aan bod. Deze pilotstudies zijn uitgevoerd voor de opzet van de taken die in dit proefschrift worden behandeld. De reden om een heel hoofdstuk aan deze pilots te wijden is tweeledig; enerzijds leidden de pilotstudies tot verscheidene methodologische beslissingen die uitgebreide argumentatie verdienen; anderzijds hebben de pilotstudies, die met eentalige Nederlandse kinderen en met tweetalige Turks-Nederlandse kinderen werden uitgevoerd, tot inzichten op het gebied van een- en tweetaligheid geleid die interessant zijn, maar die in dit proefschrift verder niet meer aan bod zullen komen.

De methodologische beslissingen die op basis van de pilotstudies zijn genomen, omvatten allereerst besluiten met betrekking tot de taken: welke taken in de hoofdstudie uitgevoerd zouden worden, en welke niet. Twee taken die in pilotsessies zijn uitgevoerd, een productietaak en een 'self-paced listening' taak, zijn in de uiteindelijke testsessies niet meer meegenomen, omdat deze – hoewel ze tot interessante en betrouwbare resultaten zouden kunnen leiden – niet valide of betrouwbaar werden bevonden onder de groep kinderen (ongeletterd en geletterd) die werd getest. Ook een taak waarin kinderen in de pilotsessies moesten aanwijzen welke wijze van

spelling paste bij plaatjes van samengestelde zelfstandige naamwoorden en locatieven, is in de uiteindelijke testsessie niet uitgevoerd. Deze taak was toch meer gericht op het (metalinguïstische) idee van kinderen over geprinte tekst dan op de onderzoeksvragen die in dit onderzoek gesteld waren.

Andere methodologische besluiten die op grond van de pilots zijn genomen, zijn het uitsluiten van kinderen uit groep 5 in de uiteindelijke testsessies (de resultaten van deze kinderen weken heel weinig af van die van kinderen uit groep 4, en de taken leken voor kinderen in groep 5 te eenvoudig en te saai om betrouwbaar te worden bevonden), en het niet verder testen van tweetalige Turks-Nederlandse kinderen, omdat deze vaak zo dominant in het Nederlands bleken dat men zich kan afvragen in hoeverre de kinderen nog echt als 'tweetalig' bestempeld kunnen worden.

De verschillen die naar voren kwamen in antwoorden in de taken die de eentalige Nederlandse en tweetalige Turks-Nederlandse kinderen in de pilots hebben uitgevoerd, moeten ook in dit licht worden bekeken. De verschillen die opdoemden (en die waren er zeker, de tweetalige kinderen scoorden veelal lager op woordsegmentatie in de taken dan hun Nederlandstalige klasgenootjes) zijn wellicht niet enkel te verklaren vanuit de tweetaligheid van de Turks-Nederlandse kinderen, maar veeleer vanuit het feit dat deze kinderen in het algemeen een achterstand in het Nederlands leken te hebben – zeker in vergelijking met hun eentalige Nederlandse klasgenoten. In die zin zijn de data van de tweetalige kinderen dus wel interessant, maar zij geven niet per definitie een antwoord op de vraag hoe tweetaligheid het idee van basiseenheden in taal beïnvloedt.

Hoofdstuk 5 beschrijft de methodologie van het onderzoek. In dit hoofdstuk wordt uitgebreid omschreven welke taken in de uiteindelijke testsessies met de kinderen zijn uitgevoerd: een dicteetaak (dictation), een zins-segmentatietaak (sentence segmentation), een 'herhaal-het-laatste-stukje'-taak (last-part repetition), een klikjestaak (click task), een gemixte-woorden-taak (mixed-words task) en een oogbewegingstaak (eye-tracking task). In dit hoofdstuk wordt omschreven hoe de taken waren opgezet, hoeveel kinderen aan de taken deelnamen, welke procedures werden uitgevoerd, en hoe antwoorden gecategoriseerd werden voor de interpretatie ervan.

In Hoofdstuk 5 wordt tevens uitgebreid ingegaan op de selectie van items voor de taken: in alle taken, behalve in de dicteetaak, kwamen bepaalde multiwoord-eenheden voor, die gebruikt werden om te bepalen of kinderen taal in woorden, multiwoord-eenheden of andere eenheden segmenteerden en verwerkten. Deze items, die bestonden uit locatieve constructies, zoals 'in de boot' en transitieve werkwoordsvormen zoals

'handen wassen' of 'een taart bakken', werden gekozen op basis van hun relatieve frequentie van voorkomen in een taalcorpus. Voor het meten van deze frequentie werden de zogenaamde 'Mutual Information value' of 'Enhanced Mutual Information value' gebruikt (Church & Hanks, 1990; Zhang et al., 2009). Deze maten geven weer hoe vaak twee of meer woorden achter elkaar voorkomen in verhouding tot het totaal aantal keer dat zij in een bepaald corpus (in dit onderzoek Google) voorkomen.

Daarnaast wordt in dit hoofdstuk beschreven in welke mate de taken als 'meer offline' of 'meer online' beschouwd kunnen worden, en worden zij geplaatst op het continuüm dat in Hoofdstuk 3 werd voorgesteld. Volgens de tien kenmerken die voor deze plaatsing van taken bekeken zouden moeten worden, en die ook uitgebreid in Hoofdstuk 3 zijn besproken, zou de verdeling van taken op het continuüm er als volgt uitzien:

OFFLINE ⟵———————┼————┼————————┼————————————→ ONLINE

Sent. Segm.   Last-part rep./Dict.   Mixed words/Click   Eye tracking

Figuur 2: Overzicht van taken op het voorgestelde offline-online continuüm

Deze verdeling van taken op het continuüm was in Hoofdstuk 5 puur theoretisch; in de conclusies in Hoofdstuk 8 wordt verder ingegaan op hoe de resultaten in de taken zich tot deze verdeling verhouden.

In **Hoofdstuk 6** is een kwalitatieve analyse gemaakt van de antwoorden in de dicteetaak. In deze taak werd kinderen gevraagd een (zelf verzonnen) verhaaltje 'stukje voor stukje' te vertellen. Dit zou informatie moeten geven over de manier waarop kinderen zelf gesproken taal segmenteren.

In Hoofdstuk 6 wordt weergegeven wat voor soort stukken kinderen in hun verhalen segmenteerden, en hoe deze stukken zich verhouden tot in de literatuur (m.n. in de Cognitieve Linguïstiek) eerder genoemde 'chunks' of 'multiword units', die hypothetisch de eenheden zijn die mensen in taal verwerken. Daarnaast wordt in dit hoofdstuk geanalyseerd in hoeverre de eenheden die kinderen zelf in hun verhalen segmenteerden, corresponderen met veelvoorkomende eenheden. Dit bleek – in vergelijking tot het corpus dat gebruikt werd voor dit onderzoek, namelijk Google – weinig te zijn: de multiwoord-eenheden die kinderen in hun dictees noemden, waren niet alleen hoogfrequente multiwoord-combinaties in Google; de kinderen gaven ook multiwoord-eenheden in hun dictees die maar één keer, of zelfs helemaal niet, voorkwamen in Google.

Hoofdstuk 7 behandelt de resultaten van de kinderen op de taken die uiteindelijk in het Nederlands en in het Turks werden uitgevoerd.

Allereerst worden in dit hoofdstuk resultaten besproken van achtergrondfactoren, zoals hoe vaak kinderen thuis boeken lazen of voorgelezen werden, of hoe vaak zij televisie keken, schoolresultaten op het gebied van taal, en de resultaten van een intelligentietest (Raven, 1962). Deze werden, nog voor een verdere analyse werd uitgevoerd, vergeleken met het aantal keer dat kinderen in de taken taal langs woordgrenzen gesegmenteerd hadden, of verwerkt leken te hebben. De resultaten van deze analyses toonden minieme effecten: alleen in de klikjestaak werden effecten van de scores op de intelligentietest gevonden. Dit suggereert dat deze factoren niet doorslaggevend zijn voor de manier waarop kinderen taal langs woordgrenzen segmenteren of lijken te verwerken.

In lijn met deze resultaten worden in Hoofdstuk 7 vervolgens de resultaten van de Nederlandse kinderen in de taken – met uitzondering van de oogbewegingstaak – besproken. Hieruit blijkt dat geletterdheid de significante factor is die bepaalt of kinderen taal langs woordgrenzen segmenteren of verwerken als zij taal produceren of horen. De frequentie van het voorkomen van de multiwoord-eenheden kon worden getest aan de hand van target items die in alle taken behalve in de dicteetaak voorkwamen, en die onderverdeeld waren in items met een hoge en items met een lage frequentie. Uit de analyse van antwoorden gegeven na deze items, moest echter worden opgemaakt dat behalve voor de oogbewegingstaak, de relatieve frequentie van het voorkomen van multiwoord-eenheden zoals 'handen wassen' en 'handen vouwen' geen grote invloed had: er werd geen hoofdeffect van frequentie gevonden op het aantal keer dat kinderen de multiwoord-eenheden (de hoogfrequente, in vergelijking tot de laagfrequente) als geheel als antwoord gaven.

In de oogbewegingstaak, die in Hoofdstuk 7 apart werd besproken omdat de antwoorden van deze taak niet direct te vergelijken waren met die uit de andere taken, werd ook onderzocht of er een invloed van de frequentie van voorkomen van multiwoord-eenheden en geletterdheid bij de kinderen zichtbaar was op de eenheden die zij verwerken in taal. Uit de analyses van de snelheid waarmee de ogen van de kinderen bewogen naar bepaalde plaatjes die de multiwoord-eenheden weergaven, en het kijkpad dat kinderen aflegden om bij targetplaatjes te komen, suggereerden een invloed van zowel frequentie als van geletterdheid.

Uit de data van de Turkse kinderen – die bestond uit resultaten op twee segmentatietaken waarvan ook een Nederlandse versie was – kon eveneens geconcludeerd worden dat de geletterde kinderen veel vaker langs woord-

grenzen segmenteerden dan de ongeletterde kinderen. Het effect van geletterdheid leek hier dus ook doorslaggevend op het aantal woord-segmentaties dat de kinderen gaven. Net als bij de Nederlandse kinderen toonden de resultaten van de Turkse kinderen geen invloed van de frequentie van voorkomen van multiwoord-targets op de segmentatie: hoogfrequente multiwoord-targets werden ook in het Turks niet vaker als geheel genoemd in de antwoorden dan laagfrequente targets.

Een vergelijking van de Turkse en Nederlandse data toonde voorts dat de taalachtergrond van kinderen op zichzelf geen significante invloed had op het aantal woordsegmentaties dat zij maakten in de segmentatietaken (de zins-segmentatietaak en de 'herhaal het laatste stukje'-taak). Het effect van frequentie en target-type (transitief en locatief) was ook niet eenduidig in beide talen. Wel bleek dat in beide talen het type taak een effect had op het aantal woordsegmentaties dat kinderen als antwoord gaven. In Hoofdstuk 8 wordt verder ingegaan op hoe dit samenhangt met de 'meer offline' of 'meer online' basis van de taken.

In **Hoofdstuk 8** wordt allereerst antwoord gegeven op de deelvragen die in Hoofdstuk 1 van dit proefschrift gesteld werden. Hierna worden de resultaten van de taken, en de conclusies die daaruit voortkwamen en al in Hoofdstuk 7 benoemd waren, nogmaals besproken, maar dan in het kader van het voorgestelde continuüm. Dat na een analyse van de taken en de factoren die van invloed zijn op het bepalen van het online of offline zijn, veel taken rond het midden van het continuüm lagen, wordt hierbij als doorslaggevend aangevoerd voor het feit dat alleen een dichotomie tussen offline en online taken niet houdbaar is. Het voorgestelde continuüm lijkt passender, hoewel een nadere analyse van onder andere de weging van de verschillende factoren die van invloed zijn bij het bepalen van de mate waarin een taak offline of online genoemd kan worden, nog nodig zou zijn in vervolgonderzoek.

Naast deze belangrijke conclusie aangaande de methodologische vraag die in dit proefschrift in Hoofdstuk 3 uitgebreid aan de orde is gekomen, worden in Hoofdstuk 8 uiteraard ook de resultaten die gevonden zijn in de experimentele taken zoals uitgevoerd voor dit onderzoek in een breder perspectief geplaatst. Zo wordt er gesteld dat het nader bestuderen van het geletterdheidsonderwijs van volwassenen en kinderen nodig zou zijn om te kunnen bepalen wat het effect daarvan is op de eenheden die zij in taal herkennen; wordt er besproken dat andere onderzoeksmethoden, bijvoorbeeld meer natuurlijke observaties of juist het toepassen van meer geavanceerde online technieken als ERP's of fMRI's nog meer informatie

zouden kunnen verschaffen over de taalverwerkingseenheden van kinderen; en wordt de keuze voor bepaalde items nogmaals bediscussieerd – en hoe een keuze voor andere, of meer soorten items, zou kunnen bijdragen aan inzichten in de representativiteit van de bevindingen in dit proefschrift voor andere multiwoord-eenheden in taal.

Ook wordt de mogelijkheid besproken om soortgelijk onderzoek nog uit te voeren in andere talen dan het Turks en het Nederlands, die beide gebruik maken van een alfabetisch schrift waarin woordgrenzen door middel van spaties worden gemarkeerd. Soortgelijk onderzoek met kinderen die gewend zijn aan talen waarin woorden op schrift niet duidelijk gemarkeerd worden, zoals in het Chinees en Japans (zie Hoofdstuk 2), zou niet alleen tot meer inzicht in de 'psychologische realiteit' van woorden kunnen leiden (cf. Olson, 1994), het zou ook kunnen tonen of er algemene of universele eenheden zijn die, bijvoorbeeld op basis van de frequentie van voorkomen, het makkelijkst worden herkend en verwerkt. Dit zou duidelijk kunnen maken of de 'grafische relativiteit' (cf. Bugarski, 1993), die in meer offline studies is gevonden, ook van invloed is op taalverwerking.

Ook de noodzakelijkheid om (ongeletterde) volwassenen te testen, en mogelijk ook een longitudinale studie uit te voeren, wordt benoemd. Volwassenen maken gewoonlijk gebruik van meer analytische strategieën (cf. Arnon, 2010), waarin ze vaker regels toepassen dan kinderen. De vragen die hierbij gesteld kunnen worden, zijn in hoeverre deze strategieën van invloed zijn op taalsegmentatie en verwerking, en of er verschillen te zien zijn tussen geletterde en ongeletterde volwassenen. Een longitudinale studie zou daarnaast kunnen bijdragen aan inzichten in een mogelijke relatie tussen de (taal)ontwikkeling van kinderen en bepaalde eenheden die in meer of mindere mate als basiseenheden worden beschouwd.

En een laatste vraag, die na afronding van het huidige onderzoek onbeantwoord is gebleven, maar waarnaar zeker meer onderzoek verricht zou moeten worden, betreft het effect van 'gebalanceerde tweetaligheid' op de eenheden die kinderen herkennen en verwerken in taal. In dit onderzoek is het niet mogelijk geweest om 'gebalanceerde' tweetaligen te testen, daar de tweetalige Turks-Nederlandse kinderen die aan dit onderzoek meededen (al op jonge leeftijd) dominant in het Nederlands bleken, en het Turks slechts in beperkte mate bleken te beheersen. Omdat eerder onderzoek echter heeft getoond dat tweetaligen taal niet op dezelfde manier hoeven te verwerken als eentalige kinderen, en dat hun metalinguïstische bewustzijn ook anders zou worden ontwikkeld (cf. Bialystok, 1986a; Francis, 1999), blijft het interessant te bekijken of, en hoe, tweetaligheid de segmentatie- en verwerkingseenheden in taal beïnvloedt.

Tilburg Dissertations in Culture Studies

This list includes the doctoral dissertations that through their authors and/or supervisors are related to the Department of Culture Studies at the Tilburg University School of Humanities. The dissertations cover the broad field of contemporary sociocultural change in domains such as language and communication, performing arts, social and spiritual ritualization, media and politics.

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 attitude-meting*. 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 Code-switching 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.

18   Sinan Çankaya. *Buiten veiliger dan binnen: in- en uitsluiting van etnische minderheden binnen de politieorganisatie*. Supervisors: Ruben Gowricharn and Frank Bovenkerk, 24 October 2011.

19 Femke Nijland. *Mirroring Interaction. An Exploratory Study into Student Interaction in Independent Working*. Supervisors: Sjaak Kroon, Sanneke Bolhuis, Piet-Hein van de Ven and Olav Severijnen, 20 December 2011.

20 Youssef Boutachekourt. *Exploring Cultural Diversity. Concurrentievoordelen uit multiculturele strategieën*. Supervisors: Ruben Gowricharn and Slawek Magala, 14 March 2012.

21 Jef Van der Aa. *Ethnographic Monitoring. Language, Narrative and Voice in a Carribbean Classroom*. Supervisors: Jan Blommaert and Sjaak Kroon, 8 June 2012.

22 Özel Bağcı. *Acculturation Orientations of Turkish Immigrants in Germany*. Supervisors: Guus Extra and Kutlay Yağmur, 3 October 2012.

23 Arnold Pannenborg. *Big Men Playing Football. Money, Politics and Foul Play in the African Game*. Supervisor: Wouter van Beek, 12 October 2012.

24 Ico Maly, *N-VA. Analyse van een politieke ideologie*. Supervisors: Jan Blommaert and Sjaak Kroon, 23 October 2012.

25 Daniela Stoica. *Dutch and Romanian Muslim Women Converts: Inward and Outward Transformations, New Knowledge Perspectives and Community Rooted Narratives*. Supervisors: Enikö Vincze and Jan Jaap de Ruiter, 30 October 2012.

26 Mary Scott. *A Chronicle of Learning: Voicing the Text*. Supervisors: Jan Blommaert, Sjaak Kroon and Jef Van der Aa, 27 May 2013.

27 Stasja Koot. *Dwelling in Tourism. Power and Myth Amongst Bushmen in Southern Africa*. Supervisor: Wouter van Beek, 23 October 2013.

28 Miranda Vroon-van Vugt. *Dead Man Walking in Endor. Narrative Mental Spaces and Conceptual Blending in 1 Samuel 28*. Supervisor: Ellen van Wolde, 19 December 2013.

29 Sarali Gintsburg. *Formulaicity in Jbala Poetry*. Supervisors: Ad Backus, Sjaak Kroon and Jan Jaap de Ruiter, 11 February 2014.

30 Pascal Touoyem. *Dynamiques de l'ethnicité en Afrique. Éléments pour une théorie de l'État multinational*. Supervisors: Wouter van Beek and Wim van Binsbergen, 18 February 2014.

31 Behrooz Moradi Kakesh. *Het islamitisch fundamentalisme als tegenbeweging. Iran als case study*. Supervisors: Herman Beck and Wouter van Beek, 6 June 2014.

32   Elina Westinen. *The Discursive Construction of Authenticity: Resources, Scales and Polycentricity in Finnish Hip Hop Culture*. Supervisors: Sirpa Leppänen and Jan Blommaert, 15 June 2014.

33   Alice Leri. *Who is Turkish American? Investigating Contemporary Discourses on Turkish Americanness*. Supervisors: Odile Heynders and Piia Varis, 9 September 2014.

34   Jaswina Elahi. *Etnische websites, behoeften en netwerken. Over het gebruik van internet door jongeren*. Supervisors: Ruben Gowricharn and Sjaak Kroon, 10 September 2014.

35   Bert Danckaert. *Simple Present*. Supervisors: Jan Blommaert and Odile Heynders, 29 October 2014.

36   Fie Velghe. *'This is almost like writing': Mobile phones, learning and literacy in a South African township*. Supervisors: Jan Blommaert, Sjaak Kroon and Piia Varis, 3 December 2014.

37   Nico de Vos. *Lichamelijke verbondenheid in beweging. Een filosofisch onderzoek naar intercorporaliteit in de hedendaagse danskunst*. Supervisors: Odile Heynders and Frans van Peperstraten, 16 December 2014.

38   Danielle Boon. *Adult literacy education in a multilingual context: Teaching, learning and using written language in Timor-Leste*. Supervisors: Sjaak Kroon and Jeanne Kurvers, 17 December 2014.

39   Liesbeth Hoeven. *Een boek om in te wonen. De verhaalcultuur na Auschwitz*. Supervisors: Erik Borgman and Maaike de Haardt, 21 January 2015.

40   Laurie Faro. *Postponed monuments in the Netherlands: Manifestation, context, and meaning*. Supervisors: Paul Post and Rien van Uden, 28 January 2015.

41   Snezana Stupar. *Immigrants regulate emotions in the same way as majority members in the Netherlands*. Supervisors: Fons van de Vijver and Johnny Fontaine, 30 January 2015.

42   Jia He. *The general response style from a cross-cultural perspective*. Supervisors: Fons van de Vijver and Alejandra del Carmen Dominguez Espinosa, 4 February 2015.

43   Dorina Veldhuis. *Effects of literacy, typology and frequency on children's language segmentation and processing units*. Supervisors: Ad Backus, Jeanne Kurvers and Anne Vermeer, 1 April 2015.