Accessing word meaning: Semantic word knowledge and reading comprehension in Dutch monolingual and bilingual fifth-graders
Cremer, M.

Citation for published version (APA):
Word knowledge is one of the key elements in reading comprehension and by extension in school success. At the same time, it is not quite clear which components of lexical knowledge play a role in reading. Is it enough to recognize the words we read? Do we need an in-depth understanding of their meaning? Is it also important how fast or easily we access meaning? The studies in this thesis investigated to what extent differences between children in word knowledge and in underlying lexical-semantic processes are predictive of differences in reading comprehension.

In three quantitative studies, lexical-semantic skills and reading comprehension are investigated for Dutch monolingual and bilingual minority children. Previous research indicates disadvantages for bilingual minority children in both vocabulary and reading.

The analyses in this thesis show delays for bilingual minority children in reading comprehension, semantic word knowledge and speed of accessing semantic and lexical information in the face of comparable word decoding. For both groups, semantic word knowledge is a relevant predictor of reading comprehension. Semantic classification speed but not priming makes a small additional contribution to reading comprehension. A final model shows that the differences between language groups are not fully mediated by semantic word knowledge and speed, leaving room for factors other than lexical-semantic differences. The findings suggest that sufficient semantic knowledge of word meaning and to a small extent fast access to meaning facilitate reading comprehension.

This book is of interest to scholars in the field of psycholinguistics, second language acquisition and applied linguistics.
Accessing word meaning
Semantic word knowledge and reading comprehension
in Dutch monolingual and bilingual fifth-graders
Accessing word meaning
Semantic word knowledge and reading comprehension
in Dutch monolingual and bilingual fifth-graders

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 21 juni 2013, te 14.00 uur

door

Marjolein Cremer

geboren te Zuidlaren
Promotiecommissie

Promotor: Prof. dr. J.H. Hulstijn

Co-promotor: Dr. J.J.M. Schoonen

Overige Leden: Prof. dr. A.E. Baker
               Prof. dr. C.L.J. de Bot
               Prof. dr. P.F. de Jong
               Prof. dr. N. Schmitt
               Dr. A.R. Vermeer

Faculteit der Geesteswetenschappen
Acknowledgements

The thesis that is before you is the report of a research project that could not have been accomplished without the help and support of many. I would like to name a few of them here.

First of all, I want to thank my supervisor, Rob Schoonen, for his advice, his critical questions, valuable suggestions, stimulating discussions, detailed comments, hands-on experience, and for always having an open door. I could not have wished for a better teacher.

I warmly thank Jan Hulstijn, my promotor and captain of our research group, for following all developments closely, for his advice, support, and constructive comments.

My sincere thanks go to my test assistants, Meike de Beer, Daphne Dingshoff, Jolien Janson, Jorien Dirksen and Iske Bakker, and to the school boards, teachers and all the pupils of the participating schools for their assistance in the carrying out of this research project.

I thank all my colleagues at the Amsterdam Center for Language and Communication - in particular, my mates, the CASLA research group, Gerdien Kerssies and Elly ten Berge, Judith Rispens, Dirk-Jan Vet, and Bart de Boer for their valuable advice, and Margriet Heim for correcting the Dutch summary of this thesis - for making this a fantastic journey.

Amsterdam, maart 2013
# Table of contents

1 General introduction 1

1.1 Vocabulary and school success 1

1.2 Semantic word knowledge 2

1.2.1 Access to semantic word knowledge 4

1.2.2 Semantic priming 6

1.3 Individual differences between learners 7

1.3.1 Bilingual minority children 8

1.4 The organisation of this thesis 10

2 Word knowledge and reading: theories and empirical insights 13

2.1 What’s in a word? 13

2.1.1 Semantic and associative relations between words 16

2.1.2 The development of semantic word knowledge 17

2.2 Word associations as a reflection of semantic networks 21

2.3 The accessibility of word knowledge 25

2.4 The role of word knowledge in reading comprehension 30

2.4.1 Word knowledge and reading in monolingual and bilingual children 37

2.5 Open issues and focus of the present work 45

3 Do word associations assess word knowledge? A comparison of monolingual and bilingual, child and adult word associations 49

3.1 Background 50

3.1.1 Word knowledge: lexical and semantic network 51

3.1.2 From context dependent to meaning based 53

3.1.3 Assessing word associations: categorization and elicitation 55
4 The role of accessibility of semantic word knowledge in monolingual and bilingual fifth-grade reading

4.1 Background
   4.1.1 Multidimensionality of word knowledge
   4.1.2 Semantic versus contextual word knowledge
   4.1.3 Word knowledge and reading
   4.1.4 Availability and accessibility of semantic word knowledge

4.2 Method
   4.2.1 Participants
   4.2.2 Materials
   4.2.3 Procedure
   4.2.4 Analyses

4.3 Results
   4.3.1 Descriptives
   4.3.2 Monolingual and bilingual children
   4.3.3 Relationships with reading comprehension
   4.3.4 Contributions to reading comprehension

4.4 Discussion and conclusion
Chapter 1

General introduction

1.1 Vocabulary and school success

For children to be successful at school, extensive vocabulary knowledge is needed. Word knowledge is an important predictor of academic achievement and it plays a central role in cognitive development, especially in relation to literacy and learning (Cunningham & Stanovich, 1997; Snow, Burns, & Griffin, 1998). For children the acquisition of word knowledge is a continuous process that involves learning word labels, meanings and uses. This happens in parallel with their conceptual development: as they learn about the world, they learn about words. Although all children seem to learn words quickly, research shows that there are large differences between children in lexical knowledge. Disadvantages in word knowledge are found in particular for bilingual children from an immigrant or minority background whose home language is different from the language of society (Cito, 2002). Although under favourable circumstances bilingual children can become highly proficient in both of their languages, minority children seem to have difficulty coping with the language of schooling. Differences between children in vocabulary knowledge have often been linked to differences between children in reading comprehension (cf. Mancilla-Martinez & Lesaux, 2010; Nakamoto, Lindsey & Manis, 2008; Proctor, Carlo, August & Snow, 2005), but it is not clear which components of word knowledge and lexical processing underlie this relationship.

Children’s lexical development involves more than acquiring a large vocabulary. Besides knowing many individual words, language learners need to
know how words are used, which associations go with words and they need to know about the relationships between words. In addition, for fluent reading learners need to be able to access this knowledge fluently and recognize and understand words quickly, as readers construct the meaning of a text on the basis of their understanding of the words in the text. A piece of text such as *Sam will also bring a hammer. He never forgets his tools* will be difficult to process without an understanding of how the words *hammer* and *tools* relate to each other. As school language and school texts become increasingly complex and abstract, learners’ knowledge of word meaning and of relationships between words becomes ever more important. Understanding how word knowledge feeds into the reading comprehension process is important not only from a theoretical perspective. It may also contribute to the knowledge needed to design better tailored reading programs to target reading delays.

1.2 Semantic word knowledge

Researchers have traditionally distinguished between two aspects of vocabulary knowledge: breadth (size) and depth (Anderson & Freebody, 1985). Breadth refers to how many words are known; depth refers to how well those words are known. The concept of depth can be defined comprehensively as including a whole range of dimensions such as spelling, pronunciation, grammatical features, and collocations (Richards, 1976); it can also more specifically refer to semantic word knowledge in terms of knowledge of context-independent word meaning and abstract, semantic relations (Schoonen & Verhallen, 2008; Verhallen, 1994). Although breadth and depth are closely related (Meara & Wolter, 2004), they are constructs that can be measured separately (Qian, 1999, 2002; Schoonen & Verhallen, 1998). The importance of depth or richness of word knowledge may be that, as it increases, words can be used more flexibly, and their meaning can be readily appreciated and accessed within multiple contexts (Anderson & Freebody, 1981; Beck, McKeown & Kucan, 2002).

Several studies have investigated vocabulary depth in terms of knowledge of semantic relations between words. Knowledge of word meaning develops from
personal, contextual meanings into more abstract, semantic meanings as learners repeatedly encounter words in language use (Nelson, 2007). As such they build up a semantic network of word relations or associations. Verhallen (1994; Verhallen & Schoonen, 1993) found evidence for differences in children’s semantic word knowledge. She compared 9-11-year-old Turkish-Dutch and monolingual Dutch children in their assignment of meaning to simple Dutch words. One important difference between the Turkish-Dutch and the Dutch children was that the Turkish-Dutch children received less Dutch input at home. On the basis of structured interviews involving a definition task, she found striking differences in the kind of knowledge monolingual and bilingual children had of familiar words. The bilingual children consistently attributed fewer meaning aspects to simple Dutch words and their knowledge was more context-specific such that they would associate a nose with context-specific knowledge such as dripping or handkerchief whereas monolingual children mentioned more semantically related meaning aspects such as smelling or body part. These differences in associations suggest differences between children in semantic networks. In a follow-up study, using a standardised paper-and-pencil word association test, similar differences in semantic word knowledge emerged and these were found to predict learners’ reading comprehension (Schoonen & Verhallen, 1998). Qian (1999) investigated vocabulary depth with young adult learners of English and used a word association format measuring knowledge of meaning and collocations of presented stimulus words as opposed to more contextual knowledge. He found that semantic differences between learners significantly predicted their reading performance beyond the contribution of vocabulary breadth. These studies show that semantic weakness is associated with poor reading comprehension (Qian, 1999; Schoonen & Verhallen, 1998).

Since the studies discussed above used non-timed measures of vocabulary knowledge, they do not reveal much about the role of processing components of word knowledge. Hence, it remains unclear whether good comprehenders are simply better at offline tasks such as paper-and-pencil tests or whether they actually have different semantic networks and process meaning more efficiently, resulting in better comprehension. Differences in semantic processing between learners have
indeed been found (Schreuder & Flores d'Arcais, 1989; Nation & Snowling, 1999, 2004) and such differences are taken to reflect differences in underlying semantic representations. For example, Nation and Snowling (2004) found that knowledge of related words in terms of semantic fluency and synonym judgement contributed unique variance to individual differences in reading comprehension. Considering such processing differences, a possible explanation for the relationship between semantic word knowledge and reading comprehension may lie in how word meaning is processed. It may be that when reading words, less-proficient comprehenders activate less well-developed concepts and other kinds of related words than proficient comprehenders, due to differences in their semantic word knowledge.

1.2.1 Access to semantic word knowledge

Processing aspects of language use have recently become more important in language learning studies (cf. Segalowitz, 2010). In addition to the roles of breadth and depth in reading comprehension, several researchers consider fluent processing as a dimension of word knowledge that may be supportive of reading. Fluency refers to the ease and speed with which learners can access and use words, as opposed to simply recognising the words and knowing about how to use them, which is what more traditional vocabulary tests measure (Daller, Milton & Treffers-Daller, 2007).

Beck and her colleagues (2002) suggested that fluency of access to word meanings may be a key factor in explaining differences in outcomes of vocabulary training studies. Vocabulary training that effects the fluency with which word meanings are accessed has an impact on reading comprehension, but vocabulary instruction that does not lead to sufficient fluency of access often does not generalize to reading comprehension (cf. Jenkins, Pany & Schreck, 1978).

What then does lexical fluency entail? There are many basic questions regarding fluency that are largely unanswered. The terms fluency and automaticity are used interchangeably in the literature. Based on work in reading theory and cognitive sciences (Logan, 1988; Perfetti, 1985; Stanovich, 1990), Wolf, Miller and Donnelli (2000) define automaticity as “a continuum in which processes are
considered automatic when they are fast, obligatory and autonomous and require only limited use of cognitive resources” (p. 377). They use the term automaticity in relation to underlying component processes and fluency to refer to fast word identification and comprehension outcomes. In line with this we could say that fluent reading comprehension is preceded by fluent word identification, which in turn may result from automatic activation of word meaning.

It has been suggested that word knowledge and access to that knowledge in terms of automatic sub lexical processes are interdependent and are a prerequisite for reading fluency. Initially, in his verbal efficiency theory, Perfetti (1985) posited that word identification, the rapid retrieval of a word’s phonology and meaning, was a limiting factor in comprehension. In the more recent lexical quality hypothesis (Perfetti & Hart, 2001, 2002) the knowledge component is also emphasised: apart from speed it is important that the reader has the ability to retrieve the meanings that are needed in a given context. In other words, the quality of the reader’s lexical representations needs to be sufficient. Perfetti and Hart propose that the quality of lexical representations drives fast processing and efficient word identification so that processing resources can be devoted to higher order comprehension (Perfetti & Hart, 2001). The accessibility of word knowledge builds on learners’ underlying semantic representations and the interconnectedness of lexical-semantic information. Wolf and colleagues posit that as word meanings become well established through frequent encounters, they become more quickly accessible (Wolf, Miller, & Donnelly, 2000). In this thesis we will use the term accessibility to refer to both speed of access to and activation of word meaning.

A recent study with 203 third-grade readers investigated whether breadth, depth and fluency of lexical knowledge were distinguishable (Tannenbaum, Torgesen & Wagner, 2006). The researchers defined fluency as the rate at which learners access the meaning of a word. For this they administered two tasks: the Word Use Fluency subtest of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002) and an experimenter-developed semantic category fluency test. The word use fluency test measured the number of target words correctly used in sentences during a one-minute testing period; in the
semantic category fluency test the children named as many items as possible from eight categories such as farm animals and fruits. Depth was measured using two definition tests, involving multiple meanings and the naming of attributes. Tannenbaum and colleagues found that the three dimensions of word knowledge were not completely distinguishable from each other using the tests employed. A two-factor model of breadth and depth/fluency provided the best fit to the data. They found that both breadth and depth accounted for unique variance to reading comprehension, although the contribution of depth/fluency was small. Over half of the variance in reading comprehension that was explained by the vocabulary measures was variance that the two vocabulary factors had in common. The researchers suggest that depth and fluency are influenced by similar types of experiences with words and that speed of access to word meaning improves as word meaning is reinforced and flexibility of use is obtained.

1.2.2 Semantic priming

A more psycholinguistic way of investigating the processing of word meaning is through semantic priming. Semantic priming refers to the unconscious speeding up of word recognition when a target word is preceded by a semantically related word. Priming effects reflect the automatic co-activation of related words. A study by Nation and Snowling (1999) with 10-11-year-old children compared proficient and less-proficient readers in the degree of automatic activation of semantically related words. They measured semantic priming effects in a continuous, auditory lexical decision task. On average the two groups differed significantly in semantic processing. Although both groups of comprehenders showed priming for function-related words (e.g., broom - floor), there were differences in priming for abstractly related words (category co-ordinates such as airplane - train). For category coordinates, less-proficient comprehenders only showed priming if the semantic category pairs were also commonly associated. The researchers conclude that in the absence of such explicit co-occurrence of words in language use less-proficient comprehenders are less sensitive to abstract semantic relations. Such findings suggest that children’s reading comprehension problems may be associated with less
effective semantic processing, or reduced accessibility of semantic knowledge. What remains unclear here is whether individual differences between children in the accessibility of word meaning – for example as reflected in semantic priming – make a unique contribution to explaining variance in their reading comprehension scores.

1.3 Individual differences between learners

Much research that tries to find out how language works focuses on the fine-grained cognitive processes underlying language use in an average language user, e.g., how do we store words in memory, how do we process acoustic signals. At the same time there is research, partly grounded in educational practice, that focuses on individual differences between language users, e.g., why are some learners more successful than others, what type of instruction suits which learners (for an early example of this tradition, see Cronbach 1957). This individual differences research systematically compares groups of language users that differ in age, proficiency, language background, to name but a few variables. As Roberts and Meyer (2012) point out, the two ‘paradigms’ cannot do without one another. Studying individual differences is important because any psychological theory of language should be able to predict and explain differences between language users. Relating differences in variables such as age or intelligence to differences in learning outcomes is a good way of studying how several variables together affect a target behavior. In this vein, Andringa, Olsthoorn, van Beuningen, Schoonen and Hulstijn (2012) compared variation in native and non-native listening comprehension. Their results showed that for both groups linguistic knowledge differences explained variation in listening comprehension, but for the native speakers, processing speed also contributed substantially, whereas for the non-native speakers there was a significant contribution of reasoning ability (IQ). Such comparisons are important for pinpointing whether the contributions of general cognitive components to using language depend on the speakers’ linguistic proficiency. This thesis focuses on understanding the relation between subcomponents of lexical knowledge and reading proficiency. To this end it relates individual differences at the level of
lexical knowledge and processing to individual differences at the higher order level of reading comprehension. In that, this thesis combines an individual-differences approach and a process-oriented approach.

1.3.1 Bilingual minority children
Bilingual children from a minority background form a population of special interest for an investigation of the relationship between word knowledge and reading. Delays in word knowledge and reading comprehension have consistently been reported for this population. Bilingual minority children often also come from lower socio-economic backgrounds and it is often hard to distinguish the two factors, let alone control for these in experimental research.

Differences between children have been reported for vocabulary breadth as well as depth. First-grade, bilingual Hispanic children in the US have been reported to have poorer receptive knowledge of L2 words than their (English) first language (L1) speaking age mates (Umbel, Pearson, Fernandez, & Oller, 1992). In the Netherlands, children from non-Western immigrant communities on average lag behind their Dutch peers at school in language subjects as well as mathematics, as measured with national curriculum tests (Central Bureau voor de Statistiek, 2008). Regarding language skills, minority children lag behind not only in vocabulary breadth, but also in knowledge of word meaning. National surveys in grades 1 to 3 show that children with a home language other than Dutch show considerable delays for reading and listening comprehension as well as knowledge of word meaning and meaning relations, but not in spelling and sentence building (Cito, 2002). Cross-sectional data collected from fourth-grade Spanish-speaking and English-only children from four schools in the US corroborate that bilingual children not only know fewer English vocabulary words but also that their knowledge of word meaning is poorer in comparison to their monolingual age mates (August, Carlo, Lively, Lippman, McLaughlin & Snow, 1999). Importantly, these differences in in-depth knowledge of word meaning do not always show on the surface.

Parallel to delays in word knowledge significant delays in reading performance of bilingual children have been signalled even though these children
have experienced the same education as monolingual children (August, Carlo, Dressler & Snow, 2005). Smits and Aarnoutse (1997) found that, throughout primary education, bilingual children do not fall behind in decoding skills such as spelling and word reading, but show poorer performance in tasks such as reading comprehension and vocabulary.

Research consistently shows that delays in reading comprehension and vocabulary are difficult to overcome (Biemiller, 2005). A Dutch study (Aarnoutse, Van Leeuwe, Voeten, & Oud, 2001) reported a two-year delay between proficient and less-proficient reading comprehenders that persisted throughout primary school. A Canadian study (Farnia and Geva, 2007) showed that after years of schooling ethnic minority children do not catch up with their monolingual peers. It is often difficult for teachers to target these delays since it remains unclear which aspects of vocabulary knowledge are problematic and how these contribute to reading comprehension. A Dutch study by Verhallen, Schoonen and Appel (2001) showed that so-called deep word knowledge is trainable. Eight-to-ten-year old children who were relatively weak at knowledge of semantic word relations followed a short training programme (12 hours) focused on three word domains (animals, occupations, vehicles) and significantly improved their word knowledge in comparison to a control group. The improvements did not however generalise to other domains and to reading comprehension.

The reported delays of bilingual minority children for knowledge of word meaning seem to run parallel to their poorer reading comprehension skills. This raises the question to what extent individual differences in lexical (semantic) knowledge and processing can explain individual differences in reading comprehension. To understand the relationship between these differences, and to gain an insight into the comprehension problems of these learners, this thesis investigates to what extent word associations, semantic word knowledge and speed of access to that knowledge are different in monolingual and bilingual children, in the Dutch context. Researchers need to understand the precise nature of the problems of children who struggle with reading comprehension in order to design
effective educational interventions and provide literacy education that fits the needs of both monolingual and bilingual children.

### 1.4 The organisation of this thesis

The goals of this thesis are to provide an insight into the extent to which underlying processes help explain the relationship between semantic word knowledge and reading comprehension, as well as to pinpoint the lexical-semantic differences between monolingual and bilingual minority children since for the latter a notorious delay in reading performance is reported. Since research suggests a role for semantic word knowledge and the accessibility of semantic word knowledge in reading comprehension, these are components that deserve further investigation. If fast access to word meaning supports comprehension, differences in accessibility may be expected to help explain differences in reading performance. We distinguish two important aspects of semantic word knowledge: *availability*, the knowledge itself, and *accessibility*, the speed with which that knowledge is activated. This information may help better explain the differences between proficient and less-proficient comprehenders and between monolingual and bilingual children. We extend the fluency dimension to unconscious activation of semantic word knowledge. The research on lexical-semantic processing in an educational context is still scarce. This thesis contributes to filling that gap.

Chapter 2 provides an overview of the literature relevant to the relationship between word knowledge and reading and presents the research questions of this thesis in more detail. Chapter 3 reports on a study investigating qualitative differences between monolingual and bilingual children and adults in word knowledge as reflected by their word associations to target words. Chapter 4 presents a study examining learners’ speed of access to semantic word knowledge as assessed with a speeded categorization task. The contributions of semantic word knowledge and speed of access to reading comprehension are investigated. Chapter 5 presents a study addressing speed of access by using measures of lexical decision and semantic classification, and investigates differences between children in semantic priming. The contributions of speed of access and semantic priming to
reading comprehension are investigated. Together, the empirical studies in Chapters 3 to 5 assess the relation between components of word knowledge and processing and reading comprehension. Finally, Chapter 6 summarises and discusses the findings and presents the overall conclusions of this thesis.
Chapter 2

Word knowledge and reading: theories and empirical insights

This chapter presents and discusses the theoretical and empirical background to this thesis and presents the research questions addressed in the studies presented in Chapters 3 to 5. In section 2.1 word knowledge as a multidimensional construct is discussed; section 2.1.1 explains the difference between semantic and associative relations between words, followed by a discussion of the development of semantic word knowledge in section 2.1.2. In section 2.2 findings from word association studies are reviewed. Section 2.3 discusses how semantic word knowledge is accessed and processed; section 2.4 moves from the fine-grained processes of word processing to the higher-level process of reading. The relevance of components of word knowledge for reading comprehension is discussed. In section 2.4.1 bilingual minority children are discussed as a group of special interest. Finally, section 2.5 summarises remaining issues and presents the research questions that are addressed in this thesis.

2.1 What’s in a word?

Word knowledge has for a long time been regarded as a one-dimensional, single construct, assessed by means of a traditional vocabulary size test. Viewing words as
separate, countable items allows for estimates of vocabulary size: eight-year-olds acquiring English as their L1 are estimated to have a vocabulary of around 6,000 words (Biemiller, 2005) with a figure of around 45,000 words for an average high school senior in the US (Nagy & Anderson, 1984). Research conducted by Vermeer (2001) among large numbers of elementary school children in the Netherlands, suggests that 12-year old native speakers of Dutch already have a receptive vocabulary of 16,000 words. However, learning words is more than acquiring an item yes or no: there is growing evidence that lexical knowledge is a multi-dimensional construct, or actually that it comprises multiple constructs.

The complexity of lexical knowledge was already discussed by Cronbach (1942) and is nicely summed up by Nagy and Scott (2000) who state that word knowledge is gradient, multidimensional, multiple-layered, interrelated and heterogeneous. The dichotomy between vocabulary breadth and depth (Anderson & Freebody, 1985; Verhallen & Schoonen, 1998) or quantity and quality of word knowledge was already mentioned in section 1.2. Breadth refers to the number of words a language learner is familiar with, while depth relates to how well words are known. Put differently, while breadth refers to how many words have meaning for the individual, depth refers to how much meaning the words have for the individual. Depth of word knowledge can be defined broadly, ranging from knowledge of a word’s spoken and written form, its meaning, the words it collocates with, the grammatical patterns it occurs in, to the derivatives that can be made from it, and so forth (Henriksen, 1999; Nation, 1990; Nation, 2001; Richards, 1976). Importantly, depth is a gradual concept. At the most basic level, a word can be recognised but not well defined. As it is learned more in depth, the word can be defined in greater detail. Finally, relations can be made between the word and other words, several meanings of a word can be learned and the word can be used in different contexts (Tannenbaum, Torgesen, & Wagner, 2006). Several word-knowledge continuum frameworks have been proposed to illustrate the notion of degree of word knowledge, ranging from partial to full-fledged representations (see e.g., Henriksen, 1999; Schwanenflugel, Stahl & McFalls, 1997; Wesche & Paribakht, 1996; Wolter, 2001). An example of how knowledge of depth can be assessed is Wesche &
Paribakht’s (1996) Vocabulary Knowledge Scale instrument which measures depth of word knowledge using a five-point scale ranging from complete unfamiliarity, through recognition of the word form and some idea of its meaning, to the ability to use it with grammatical and semantic accuracy in a sentence.

Depth of word knowledge can also be more narrowly defined in terms of core meanings and semantic relations (Schoonen & Verhallen, 2008; Verhallen, 1994). Semantic depth has been measured by having test takers generate definitions or sentences containing target words, or by assessing a test taker’s ability to link words to – and distinguish them from – related words. An example of the latter approach is Read’s word-associates format (1993). Read used a multiple-choice test format which provides a practical way of assessing how well particular words are known. In the task, university students are presented with a stimulus word (e.g., *team*) and have to identify related words from a set of eight words. The relations are meaning-based (*group*), collocational (*sport*) or analytic/definitional (*together*). The researcher found differences between students in how well familiar words were known. Qian (1999) tested young adult Chinese and Korean ESL learners in Canada, using a similar format to Read’s. He found that the learners’ depth scores were correlated to vocabulary breadth and to reading comprehension. Another example of qualitative word knowledge testing is a study by Schmitt and Meara (1997) who tested Japanese students’ word association knowledge and verbal suffix knowledge of English and found those to be related to vocabulary size and general language proficiency.

The multidimensional conception of word knowledge is well-illustrated by the ‘lexical space’ formulated by Daller, Milton and Treffers-Daller (2007). In this model, word knowledge consists of vocabulary breadth, depth of word knowledge, and fluency (accessibility). As Figure 2.1 below shows, breadth, depth and fluency may develop independently and reach different levels for different words and different learners. Moreover, the three dimensions may have interdependent thresholds for growth, such that depth may grow only when a certain level of vocabulary breadth has been acquired, or that fluency may quickly improve as word knowledge deepens.
Although the dimensions of breadth and depth of word knowledge are inherently related (Meara & Wolter, 2004), several studies have shown that they are distinct constructs. Despite the substantial correlation between depth and breadth, measures of depth make a unique contribution to explaining variance in reading performance, beyond vocabulary size (Ouellette, 2006; Qian, 1999, 2002; Schoonen & Verhallen, 1998).

### 2.1.1 Semantic and associative relations between words

Theoretically, depth of word knowledge is distinct from breadth because depth of word knowledge reflects the degree to which concepts and contexts are linked together. This interrelatedness of words in the mental lexicon is expressed in the widely used network metaphor (Aitchison, 2003). The denser the network a word is in, the deeper or greater the knowledge of that word (Nagy & Herman, 1987). Lexical networks illustrate the intricate link between breadth and depth. When a new word is learned, this influences the whole system as a result of the item’s connectedness with other items (Meara & Wolter, 2004).
Connections between words are formed, strengthened or weakened as a function of their co-occurrence in language use (cake - candles) or due to overlap in meaning or semantic features (dog - bear). The former are commonly referred to as associative relations, the latter as semantic relations. Association strength of two words varies depending on how frequently they co-occur or how common it is for people to associate them; semantic relatedness varies depending on the degree of semantic overlap between two words. Since associative relations are based on real-world co-occurrence they are generally context-bound and sometimes subjective. Yet, for many words, association norms can be established (e.g., black – white, table - chair). Whereas associative relations may emerge between entire words, semantically related words can be said to be connected in terms of certain features. For instance, the words dog and bear overlap in the sense that they are both hairy, belong to the category animals, and have four legs. The words dog and cat share those features too and in addition both belong to the category pets. The more features two words share, the stronger they are related. At the same time, not all meanings can be described through features. What are the features of red? Can that word have meaning without reference to other colours or to objects of that colour? Lyons (1963: 59) posits that the meaning of a word is not locally contained in the lexical unit itself, but rather is determined by its relations with other words. This thesis focuses on depth as indicated by knowledge of (abstract) semantic relations between words.

2.1.2 The development of semantic word knowledge
Connections in the mental lexicon develop from idiosyncratic or context-bound associations into more abstract, semantic word knowledge such as category coordinate or superordinate relations, so that the lexicon contains both associative and semantic relations. In adults, knowledge of basic meanings and facts is assumed to have a hierarchical taxonomic structure (Medin, Ross, & Markman, 2005). Nelson (1996) reported that hierarchical category organisation is not well established until the early school years. Children’s word knowledge is built up continuously on the basis of words they encounter in particular contexts (dog – leash
From their initial, context-bound knowledge, children are assumed to gradually abstract and refine category information (e.g., dog – animal) (Nelson, 1977). As their understanding of the world increases, children’s personal meanings develop into conceptual ‘core’ meanings (Nelson, 1974) or more conventional and ‘shared meanings’ (Nelson, 2007). As Mandler (1983) noted, taxonomic categories, such as subordinates and superordinates, are based on shared meaning and, unlike contextual or thematic categories, are atemporal, nonspatial, and hierarchical in format. Petrey (1977) showed that in free word association tasks, young children are more likely to respond to cue words with contextually related words (example from a medical context: examine – needle) than with abstract, semantic associations (examine – look (at), check). She describes this development as a shift in children’s word knowledge from episodic to semantic: “[y]oung children associate primarily to the stimulus word's perceived contexts, older subjects to its abstract semantic content” (p. 57). As such, children’s conceptual and linguistic development is intertwined (Gopnik & Meltzoff, 1992). Understanding of variation in (the development of) semantic relations between words in the mental lexicon is not only of theoretical importance, it has also been linked to reading performance (Nation & Snowling, 1999; Qian, 1999; Schoonen & Verhallen, 1998). Below, I will discuss a number of studies that have investigated the development of abstract, semantic word knowledge.

It has been shown that young children’s categorizations are initially idiosyncratic and that their category boundaries become reliably less fuzzy with age. Alexander and Enns (1988) asked 3-, 4-, 5-, and 24-year-old participants to categorise a continuum of puppets. There was more agreement among five-year-old children than among three-year-olds, and the five-year-olds were more consistent across different categorisation tasks (free sorting, selection from a set of distractors and naming). When subjects were asked to explain their categorisations, younger children gave primarily idiosyncratic and unclassified responses (“they are all friends”; “he wants to be [an X]”), whereas older children referred mainly to specific visual properties of the puppets.
In a speeded category verification task, Jerger and Damian (2005) tested the recognition of exemplar pictures that were more or less related to the category of clothing by 4-14-year-olds and adults. Whereas adults were equally accurate in classifying typical (*pants*) and a-typical category objects (*glove*), children were more accurate with typical category objects. Children were also significantly less accurate than adults in classifying out-of-category related items (*necklace*). The authors found more pronounced age-related improvement in accuracy for atypical objects than for typical objects and for related out-of-category objects than for unrelated out-of-category objects, which may ‘reflect children’s increasing specification of the properties characterizing a category’ (Jerger & Damian, 2005: 67). Similar results were found in an experiment with printed words with adults, indicating that typicality and relatedness effects did not reflect merely picture-related processes.

In category generation tasks, young children are better at generating words that fit a certain context or script than words that fit abstract, taxonomic categories. They seem to prefer experience-based categories, which are also referred to as thematic or slot-filler categories. For example, *cereal* is a slot-filler item in the child’s category or script of *[foods I eat for breakfast]*. Nelson and Nelson (1990) used a category-generation task in which children name items within different categories. Kindergartners generated comparable numbers of items in the contextually constrained slot-filler condition (e.g., name foods that you eat for breakfast) and in the taxonomic condition (e.g., name foods), whereas second-graders (age 8 years) generated many more items in the taxonomic condition than in the slot-filler condition. Given these findings, Nelson and Nelson (1990) emphasized the role of development and experience in the shift from a slot-filler to a taxonomic strategy as the preferred strategy for categorizing vocabulary between ages five and eight. This shift in salience from thematic to taxonomic organisation seems to hold up across languages and industrialized and schooled societies. For example, Yu and Nelson (1993) compared monolingual Korean-speaking five- and eight-year-olds’ performance on a category-generation task. Similar to English-speaking children, young Korean-speaking children produced comparable numbers of items in the slot-filler and taxonomic conditions, whereas older children produced more items in the
taxonomic condition. Peña, Bedore and Zlatic-Giunta (2002) found similar qualitative changes in category generation for bilingual children. Although younger Spanish-English bilingual children ($M$ age = 5;1) generated approximately equal numbers of items in both conditions, older Spanish-English bilingual children ($M$ age = 6;5) were beginning to demonstrate a taxonomic bias.

Both context-bound and taxonomic relations between words are basic organizational principles of the lexicon. Although thematic relations have frequently been found to be the basis of children's but not adults' classification, Lin and Murphy (2001) found that when thematic relations are meaningful and salient, they have significant influence on adults' category construction (sorting), inductive reasoning, and verification of category membership. After a series of ten experiments, the authors conclude that concepts function closely with knowledge of scenes and events and that this knowledge has a role in adults' conceptual representations. Because thematic associations show a strong real world contiguity, they may sometimes co-activate each other more strongly than taxonomic relations in lexical tasks. For example, when asking people to come up with the first word that comes to mind the association between cake and candles may be stronger than the logical semantic relation between cake and bread.

The studies discussed above show that classifying on the basis of abstract, taxonomic relations becomes increasingly prevalent across tasks up to the preteen years (for an elaborate discussion, see Bjorklund, 2005). Education is said to promote knowledge of taxonomic or hierarchical categories (Nelson & Nelson, 1990). An early study in a non-industrialised society showed that farmers who received little education were more likely than secondary school students to use context-bound relations rather than taxonomic ones to categorise entities (Scribner, 1974). Taxonomic knowledge becomes especially important as children start using more abstract reasoning, for example in reading comprehension. By the end of elementary school (around the age of 11) children may be expected to have both a firmly rooted context-bound knowledge of words and a well-developed context-independent, semantic knowledge of words.
2.2 Word associations as a reflection of semantic networks

Many researchers have investigated differences between children in knowledge of semantic relations between words by using word associations. In the free word association task language users are presented with a set of stimulus words one-by-one and are asked to produce the first word that comes to mind. Reported associations are assumed to reflect connections between words in their respondent’s lexicon. As such, they may reflect the semantic representation of words (De Groot, 1989). Word associations are assumed to spring both from conceptually related representations and from representations that commonly co-occur in language use. Word association studies have compared different groups of learners such as younger versus older or native versus nonnative speakers. We might expect adult and native speaker networks to be denser and more highly organised than similar networks generated by children or L2 learners. However, consistent differences in word associations between learner groups have been hard to pin down, as the following discussion of association studies will show.

In the 1960s and 1970s, word association behaviour was used to study the developmental organisation of the L1 lexicon (e.g., Entwisle, 1966; Ervin, 1961; Nelson, 1977). Some L1 association research had shown that between 6-8 years of age a shift from syntagmatic responses (cold - outside) to paradigmatic responses (cold - hot) occurs (Cronin, 2002; Entwisle, 1966). Syntagmatic responses usually belong to a different grammatical class and occur with the stimulus word in a syntactic sequence; paradigmatic responses belong to the same grammatical class as the stimulus word. Syntagmatic responses have been said to be given before many semantic features have been acquired, whereas paradigmatic responses have been taken as indicative of a more developed semantic system (see e.g., Clark, 1970).

What is often overlooked in the literature on this topic is that Entwisle (1966) also found that so-called late syntagmatic responses are typical of adults, although these seem different from children’s syntagmatic responses.

The syntagmatic-paradigmatic distinction has often been criticised for being a distinction based merely on word class (Fitzpatrick, 2006; Namei, 2004; Petrey, 1977), a researcher-imposed distinction. Petrey (1977) has shown that syntagmatic
responses could spring from either semantic or episodic storage. For example, when presented with the word *add*, children gave the episodic responses *flour* (syntagmatic) and *cook* (paradigmatic); these would be classified differently in S-P terminology even though they are related in the same way to *add*: they come from the same situation or script. Petrey proposed distinguishing between episodic and semantic responses. Recent studies have used semantic-based classification of word association responses, distinguishing responses that are more or less meaning-related to the stimulus word (Namei, 2004; Fitzpatrick, 2006). For example, Fitzpatrick (2006) distinguished a meaning-based, position-based, form-based, and an erratic category with 17 subcategories.

L2 word association studies have shown that L2 learners have quite distinctive associations from native speakers (Riegel, Riegel, & Meyer, 1967; Riegel & Zivian, 1972). Some studies have claimed that the L1 mental lexicon is organized mainly on a semantic basis, while the organization of the L2 mental lexicon in the early stages of development is phonologically based, with phonological responses indicating less profound lexical knowledge (Meara, 1983). However, Nissen and Henriksen (2006) found what they called “a surprising majority” of syntagmatic responses in the L1 word association test they administered to 25 Danish 17-18-year-olds (p. 389), while they had expected a higher proportion of semantic responses. Namei (2004) found that phonological links occur also at advanced levels of proficiency, both in the L1 and L2, depending on degree of word knowledge: words that are barely known may elicit phonological associations, whereas well-known words are connected to other words mainly on a semantic basis (see also Wolter, 2001). Unclear is whether differences between L1 and L2 children and adults have to do with cognitive or linguistic development.

Namei (2004) found evidence for an age-related semantic shift in word associations. She compared free word associations of 100 Persian-Swedish bilingual students (age 6-22) to a Swedish and a Persian L1 control group. All bilinguals were socialized in Persian before they were exposed to Swedish to any meaningful degree. About half of the bilinguals were introduced to Swedish before age three. The latest age of onset for Swedish was age 15. The majority of the students rated
their proficiency to be higher in Swedish than Persian. Namei operationalised paradigmatic responses as semantically related to the stimulus word (*bitter - taste*), and syntagmatic responses as associations that form a phrase with the stimulus word (*nice – dog; table - cloth*); she also distinguished form-based responses. Semantic responses increased between the ages of six and ten in all groups, even though bilinguals are exposed to much less input in their languages, especially their mother tongue. The Swedish and Persian L1 and L2 groups showed roughly the same developmental response pattern with more semantic, as well as more abstract, low frequency syntagmatic (referred to by Namei as late syntagmatic) responses (*deep – hole; moon – distant*) with increasing age. Namei describes this age-related, cross-linguistic increase in semantic responses as “a function of cognitive development during the same period, which indicates that the development of logical and semantic relationships takes place as children grow older and gradually come to a better understanding of implicit relationships such as contrast, synonymy, hyponymy and meronymy, in addition to gaining knowledge about syntactic relations” (p. 381). At the same time, overall, L1 groups gave significantly more semantic responses and more abstract syntagmatic responses than the bilingual group in each of their languages. Also, these responses appeared at a younger age for the L1 speakers than for the bilinguals and for the bilinguals these appeared earlier in Swedish than in Persian. The bilinguals lagged behind the respective comparison groups three school years in the development of late syntagmatic associations in each language. Namei attributes these dissimilarities to the lower amount of language input for the bilinguals, especially in the mother tongue (Persian). Possibly, the cognitive development and corresponding semantic shift happen somewhat later in the bilinguals resulting in a lexical delay.

In the development of their semantic networks, bilinguals can be influenced by the levels of proficiency they have in their second language as well as the age at which they acquired the language and the way they learned and use it. Some L2 studies used the association task as a test of native(like)ness or proficiency in the L2, comparing L2 learners’ responses with those of native speakers (Kruse, Pankhurst, & Sharwood Smith, 1987; Schmitt, 1998; Sökmen, 1993). It has proved difficult to
establish stereotypical responses and to use word associations as a measure of proficiency (e.g., Wolter, 2002). When confined to first responses, reliability is generally high for words with smaller sets of associates and stronger primaries. Also, highly frequent words tend to elicit predictable words, especially among native speakers. But this homogeneity tends to disappear when words are less frequent (Fitzpatrick, 2007).

Li, Zhang and Wang (2011) tested bilingual students’ awareness of thematic and taxonomic relations in relation to their proficiency. A word association and forced-choice categorisation task showed equal awareness of taxonomic relations in L1 and in L2. Interestingly, however, students were more aware of thematic than taxonomic relations in L1, but less aware of thematic than taxonomic relations in L2. In a similar vein, Fitzpatrick and Izura (2011) found very few collocational associations (sheep – black, through the idiom black sheep) in the L2 compared with the L1 but a higher proportion of semantic associations (sheep – animal) in the L2 compared with the L1. These findings might suggest that the lexical-semantic organisation of the L2 is more guided by conceptual constraints and less by the conventional use of the language. Additionally, association responses in the L1 are generally faster than in the L2 (Fitzpatrick & Izura 2011; Namei 2004).

The empirical studies reviewed suggest that lexical acquisition and storage are likely to be affected by different factors for monolinguals and bilinguals and for a speaker’s L1 and L2. In (late) L2 learners the cognitive development that boosts semantic word knowledge is already complete and their developing L2 lexicon builds on the already well-formed L1 lexicon. Moreover, the setting and the way in which the L1 and L2 are learned affect their development. For child bilinguals, on the other hand, cognitive and lexical development are intertwined and may well be expressed in both the first and second language. Studies into associative behaviour have produced data that are largely inconclusive. The idea of semantic development is well-established, but an increase in semantic association responses has not been shown unanimously. It is still unclear to what extent word associations of younger and older L1 and L2 learners differ systematically (Fitzpatrick & Izura, 2011; Namei, 2004; Nissen & Henriksen, 2006). At the same time, the value of word
association responses as a reflection of semantic capacities remains unclear. Do word associations reflect semantic relations, lexical co-occurrence or associations between words recently used by the speaker? The association process may simply reflect some active bit of knowledge, rather than the quality or extent of a person’s knowledge. Semantic responses signal semantic knowledge, but when speakers do not produce (many) semantic responses this may not necessarily mean that they do not have much semantic knowledge.

2.3 The accessibility of word knowledge

Measures of vocabulary depth are typically concerned with declarative knowledge, which learners can consciously access and report in a vocabulary test, rather than the more implicit procedural knowledge that underlies word recognition, proficient listening comprehension or fluent conversational speech (Read, 2004). The ease with which language users use their vocabulary skills is fundamental to fluent speaking, listening, reading and writing. In fluent conversation we retrieve two or three words per second from a lexicon that contains tens of thousands of items (Levelt, Roelofs, & Meyer, 1999). Written words can be recognized in around 200 ms (Sereno & Rayner, 2003). In order to understand a word, we need to process the form or sound of that word, recognize the word itself and activate its meaning as well as related meanings. In his Verbal Efficiency theory Perfetti (1985) postulated that meaning access and related processes need to be fluent to save up time and cognitive resources to devote to the higher order interpretation processes such as reading. Several studies have shown differences between learners in the speed and automaticity with which they can access their word knowledge.

Word naming tasks show that words can be read out loud without accessing the lexicon. A direct mapping between spelling and sound, at least in regular words, allows for the pronunciation to be “read off” the word, without necessarily identifying that word in the lexicon. This is shown by the fact that participants can read aloud both words and non-words: the orthographical and phonological processes involved do not require access to the lexicon, although of course such processes themselves are based on regularities in the lexicon. At the same time,
naming times have been found to be shorter for words than for nonwords, and for high frequency words shorter than for low frequency words, indicating that a lexical search procedure is also involved (Forster & Chambers, 1973). Exception words, with irregular spelling patterns, are generally assumed to be named with lexical support (Nation & Cocksey, 2009).

For word recognition, language users need to access their lexicon. In lexical decision tasks speed of access is measured, that is, the time needed to identify a stimulus word as an existing word or not. For this they need to find a match in their lexicon. The meaning of the word need not be activated. It has been argued that lexical decision does not require semantic processing (Shelton & Martin, 1992). Other research has shown that lexical decision responses are influenced by meaning integration processes, since lexical decision performance is inhibited by incongruous sentence contexts and facilitated by congruous sentence contexts (De Groot, 1985).

For word meaning to be activated, tasks require processing at the semantic level, such as category decision tasks (e.g., is the stimulus animate, does it denote an entity bigger than the screen you are looking at, is it fast or slow). For a semantic decision, in addition to recognising the word, its semantic network must be activated. In a rich semantic network, there are many relations and interconnectedness is high. For example, the word *teeth* may readily activate semantically related words such as *bite, mouth, jaw, eat*. It may also activate contextual associations such as *tooth ache, false, dentist, grandma*. The density of the lexical-semantic network surrounding a word influences how fast it is accessed or recognised. Tasks requiring semantic processing generally take longer than mere lexical tasks (Bueno & Frenck-Mestre, 2008; De Groot, 1990). The time course of the different processing levels is still subject of debate. Although semantic processing is probably completed after lexical identification, there is some indication that basic semantic information (e.g., is something good or bad; strong or weak) is involved in the recognition process from the start (Wurm, Vakoch, & Seaman, 2004).

Several studies have reported semantic feature effects on the speed of word recognition and semantic processing in young adults. Pexman and colleagues have
reported faster lexical decision and naming responses (Pexman, Lupker, & Hino, 2002) and faster semantic categorisation and self-paced reading times (Pexman, Holyk, & Monfils, 2003) for words for which participants could list many semantic features, as opposed to words for which participants could list fewer features. Experiments by Buchanan, Westbury and Burgess (2001) indicate that a large semantic neighbourhood (the number of words that could appear in similar lexical contexts) facilitates lexical decision and naming performance. However, in their study the lexical decision task seemed more reliant on semantic processing than naming. In addition, Pexman, Hargreaves, Siakaluk, Bodner and Pope (2008) showed that number of semantic neighbours, number of semantic features and contextual dispersion (the content areas in which words have been experienced) were related to lexical decision and semantic categorization performance. Higher values on each of the three semantic richness measures were correlated to faster responses in both tasks. Moreover, the three measures were significant predictors of responses in at least one task (lexical decision and/or semantic categorisation). The three measures were only modestly correlated, suggesting they do not all tap the same construct. These findings indicate that richness of semantic word knowledge influences word recognition and semantic processing.

Data from a timed association study suggest that degree of relatedness speeds up processing. Fitzpatrick and Izura (2011) found that adults’ free association responses were faster when stimulus and response were related in more than one way, for example by form and by meaning (postman – postbox) or by meaning and by collocation (spider – web) as opposed to only meaning related words. Responses that co-occur in the language and have a meaning relation with the cue word (meaning and collocation) were produced significantly faster than any other response type. In their association experiment, Fitzpatrick and Izura further distinguished between two types of semantic responses: responses whose meaning is equivalent (sofa – couch, prince – king) and responses whose meaning is related but not equivalent to the cue word (party – celebrate). The latter type of response took significantly longer to produce than the former. These results suggest that words with a direct semantic connection to the cue word or connected by more than one
route (form and meaning, meaning and collocation) are quickly and strongly activated.

Differences have been found between adults and children in semantic word knowledge and categorisation speed. The category verification study by Jerger and Damian (2005) discussed earlier showed not only that younger participants were less accurate at categorising objects as clothing but also that they were slower. Across the board, participants were faster at categorising typical objects (pants) and unrelated out-of-category objects (soup) than atypical (glove) and related out-of-category objects (necklace). However, younger participants were slower at responding to atypical category members and to related out-of-category objects. This suggests that for younger participants semantic category knowledge is less well developed, resulting in slower responses.

The works discussed above concern tasks that require conscious processing (categorise, list features, etc.). As such they do not directly tell us about language users’ underlying semantic representations. Priming studies can provide us with such information. Priming occurs when a target word (e.g. nurse) is recognised faster because of the preceding presentation of a related item, the prime (e.g. doctor) (Meyer & Schvaneveldt, 1971). The assumption is that “activation of the prime causes the target to be activated faster. Conversely, if a target word is activated faster (primed), you can be sure that the priming word must have been activated” (Altman, 1997: 71-72). The temporary facilitation observed in priming tasks is commonly attributed to the automatic spread of activation among semantically related elements in a memory network (Anderson, 1983; McNamara, 1992). Alternative explanations have also been put forward (Masson, 1995; Plaut & Booth, 2000; Ratcliff & McKoon, 1988). Some evidence suggests that semantic priming may not always be as short-lived as previously thought (Becker, Moscovitch, Behrmann & Joordens, 1997; Joordens & Becker, 1997). In a study with undergraduate students, Becker and colleagues demonstrate a semantic priming effect spanning many intervening items and lasting much longer than a few seconds (Becker, Moscovitch, Behrmann & Joordens, 1997). In a continuous animacy decision task, students were presented with a list of 15 prime words followed by a
list of 30 target words, with a two-minute pause in between. There was one prime per target. All words were preceded by an on-screen fixation cross which lasted one second. The set up resulted in an average lag of 21.5 items between a given prime and the corresponding word in the target list. On average, students were 35 ms faster at making semantic decisions for primed words than for unprimed words. Regardless of the exact nature of the underlying mechanisms or the duration of priming effects, semantic priming appears to reflect cognitive processes that depend on the interconnectedness or embedding of lexical knowledge in memory.

Words that are not only semantically related but also commonly associated have been found to show a significant increase in the size of the priming effect: the so-called associative boost (Lucas, 2000; Moss, Ostrin, Tyler, & Marslen-Wilson, 1995). To avoid this additional effect, several researchers have tried to dissociate semantic and associative effects by devising non-associated prime-target pairs (Shelton & Martin, 1992). Some researchers have claimed there is no evidence of priming based purely on association (Lucas, 2000) and that there is always some degree of featural overlap. McDonough & Trofimovich (2009) posit that differences between learners in semantic priming may provide a window into differences in lexical-semantic organisation.

Several researchers have connected semantic priming benefits to children’s developing semantic networks throughout childhood (Nakamura, Ohta, Okita, Ozaki, & Matsushima, 2006; Nation & Snowling, 1999). Perraudin and Mounoud (2009) compared children and adults’ instrumental and categorical knowledge to investigate the development of conceptual organisation. In a priming paradigm, they tested the automatic activation of instrumental and categorical relations, using a naming task and a categorical decision task. The results showed that on both types of task, adults and 9-year-old children showed instrumental and categorical priming effects. However, five-year-old children showed mainly instrumental priming effects while categorical effects remained marginal. Furthermore, the magnitude of the instrumental priming effects decreased with age. These processing differences suggest that young children’s word knowledge, in terms of semantic category knowledge and abstract semantic relations between words, is not yet fully
developed.

The studies reviewed above indicate that rich semantic word knowledge is needed for fluent word recognition and processing. Fast word recognition in adults seems to depend on the quality of underlying word representations. Semantic processing skills appear to be still in development in young children as they are less sensitive to semantic category information. It is unclear to what extent individual differences between children with regard to speed of word recognition and sensitivity to semantic priming are related to differences in language proficiency, in particular reading comprehension.

2.4 The role of word knowledge in reading comprehension

The relevance of lexical knowledge and skills for reading comprehension seems self-evident (cf. Mancilla-Martinez & Lesaux, 2010; Nakamoto, Lindsey, & Manis, 2008; Proctor et al., 2005). Words carry meaning and hence are fundamental for understanding text. Yet, the exact nature of the relationship between word knowledge and reading is not altogether clear. The sheer quantity of words known (breadth) is strongly predictive of reading comprehension, yet little is understood about how quality of word knowledge (depth) affects comprehension. To what extent do processing components of word knowledge affect reading? Is it sufficient that semantic word knowledge can be accessed consciously, or can readers only benefit from semantic word knowledge when it is activated automatically?

According to the simple view of reading (Gough & Tunmer, 1986) reading comprehension is the product of word decoding and general language comprehension. Vocabulary knowledge is subsumed under language comprehension. Word decoding could be viewed as a prerequisite for reading: words need to be decoded, before they can be comprehended. It has been stressed that the automatization of word decoding skill contributes to fluent reading levels (Perfetti, 1992; Stanovich, 2000). Decoding skill becomes less correlated to reading comprehension as children progress through elementary school (Sticht & James, 1984; Verhoeven, 1990; Verhoeven & Van Leeuwe, 2008), in contrast to vocabulary knowledge.
The importance of vocabulary size for reading comprehension has been well established in L1 research (Anderson & Freebody, 1981; Anderson & Freebody, 1983), L2 research (Laufer, 1992; 1996), and L1 and L2 research (Van Gelderen, Schoonen, De Glopper, Hulstijn, Simis, Snellings & Stevenson, 2004). Laufer (1992; 1996) showed correlations of up to .75 between vocabulary size and reading comprehension in English as a second language for first-year university students in Israel. There are indications that the relation between word knowledge and reading comprehension in L1 grows with age (Snow, 2002; Van Gelderen, Schoonen, Stoel, De Glopper & Hulstijn, 2007). Torgesen, Wagner, Rashotte, Burgess and Hecht (1997) found that vocabulary was a significant factor in the prediction of reading comprehension: second-grade vocabulary explained 24% of the variance in fourth-grade reading, whereas third-grade vocabulary explained 43% of the variance in fifth-grade reading comprehension. De Jong and Van der Leij (2002) found additional contributions of vocabulary size and listening comprehension to third-grade reading comprehension, after word decoding and first-grade reading comprehension had been taken into account. They state that the additional effect, after the autoregressive effect was controlled for, might reflect the outcome of a process of reciprocal causation since vocabulary and listening comprehension do not develop independently from reading comprehension. The reciprocity of the relationship is intuitively plausible: vocabulary enables comprehension, while comprehending what you read allows for meanings of words in the text to be inferred. It has indeed been shown that skilled (adult) readers are better than less skilled readers at learning new word meanings (Perfetti, Wlotko, & Hart, 2005).

The quality of word knowledge has also been shown to affect reading comprehension. In a longitudinal study using structural equation modelling, Verhoeven & Van Leeuwe (2008) found a substantial effect on reading comprehension of scores on a (receptive) word meaning test in which children had to identify the correct meaning from four possible meanings. Throughout the grades, knowledge of word meanings was related reciprocally to reading comprehension. This is in line with earlier research (Muter, Hulme, Snowling, & Stevenson, 2004; Snow, 2002). In a study with young adult readers, Braze, Tabor, Shankweiler and
Mencl (2007) showed that orally assessed word knowledge (comprising receptive vocabulary knowledge and expressive knowledge of word meaning) captured unique variance in reading comprehension after decoding skill and listening comprehension were taken into account. The contribution of word knowledge overlapped considerably with the contributions of decoding and listening comprehension, but was not wholly contained within them.

Ouellette (2006) investigated the unique contributions of vocabulary breadth and depth to reading performance. He assessed depth with a definition and a synonym task. Results showed that semantic depth made a significant contribution to reading comprehension even when vocabulary breadth and decoding skills were controlled. In addition, the variance in reading comprehension accounted for by receptive and expressive vocabulary breadth was shared in all hierarchical regression models tested, suggesting that the inclusion of only one of these measures of breadth is sufficient in the analysis of reading comprehension. Overall, vocabulary breadth and depth together explained 28.5% of the variance in reading comprehension. In a study with nine-year-old children, Nation and Snowling found that poor comprehenders were weaker on tasks of both receptive and expressive vocabulary (recognize synonyms, match figurative usage to meaning, define words, provide multiple meanings) than controls matched for decoding ability (Nation & Snowling, 1998). In line with this are outcomes from studies using a multiple-choice, word association format, based on Read (1993) which showed that success at identifying abstract meaning aspects of stimulus words was a unique contributor in the prediction of reading comprehension levels, in addition to the prediction afforded by vocabulary size scores (adults: Qian, 1999; children: Schoonen & Verhallen, 1998). Facility with decontextualized language has further been shown to be related to children’s reading ability.

There is strong evidence that breadth and depth of word knowledge overlap. Vermeer (2001) found correlations between breadth and depth to be so high that he concluded there is no conceptual distinction. Several studies use composite measures to create stronger predictors of reading (cf. Braze et al., 2007). Qian (2002) assessed the contributions of three vocabulary measures to reading
comprehension: a vocabulary depth measure (a word associates format measuring knowledge of synonymy, polysemy and collocation), a breadth measure (Vocabulary Levels test, Nation, 1983) and a TOEFL vocabulary (breadth) measure (measuring knowledge of synonyms). Qian reported that the three measures were similarly useful in predicting performance on a TOEFL reading comprehension test. When breadth and depth were entered in succession, vocabulary depth predicted 13% of variance in reading beyond the Vocabulary Levels measure and 14% of variance beyond the TOEFL measure.

Tannenbaum, Torgesen and Wagner (2006) stress the importance of considering fluency in the study of word knowledge. They examined the roles of three dimensions of word knowledge, breadth, depth and fluency, in third-grade reading comprehension. For breadth, receptive vocabulary and knowledge of word definitions were measured. Depth was assessed using a multiple meanings and an attributes test both of which involve word definitions. Fluency was measured using the Word Use Fluency subtest (Good & Kaminski, 2002) and a semantic category fluency task. Structural equation modeling revealed that a 2-factor model of breadth and depth/fluency provided the best fit to the data indicating that the three dimensions of word knowledge were not completely distinguishable. In the model, breadth was a significant predictor of reading comprehension, but the depth/fluency factor was not. A confirmatory factor analysis did reveal strong and significant correlations between reading comprehension and the breadth and depth/fluency factors respectively (.79 and .71). When combining standardized scores for the breadth measures and standardized scores for the four depth/fluency measures into a composite breadth and a composite depth/fluency score, Tannenbaum and colleagues did find a significant contribution of depth/fluency to reading comprehension in a regression analysis. The magnitude of the variance unique to depth/fluency was small (2%). The study also showed that over half of the variance in reading comprehension that was explained by the vocabulary measures was variance that the two vocabulary factors had in common. Tannenbaum and colleagues posit that depth of knowledge about meaning is highly associated with fluency of accessing a word’s meaning. These results show that the exact roles of
depth and accessibility of semantic word knowledge in reading comprehension are not quite clear.

Other studies have compared the semantic processing speed - or fluency - of differently skilled reader groups. Children with comprehension weaknesses were found to be slower and less accurate at deciding whether two spoken words are synonyms, especially for low-imageability items (Nation & Snowling, 1998a). At the same time, these children’s rhyme judgements did not differ significantly. These results indicate semantic but not phonological deficits. Weak comprehenders also produced fewer words in a speeded semantic category generation task, but not in a rhyme fluency task. A third experiment showed that poor comprehenders were slower at reading words with irregular spelling patterns and low-frequency words. These findings indicate that poor comprehenders have problems reading words that are typically read with support from semantics. In line with this are findings of oral vocabulary weaknesses in eight to ten year olds (Ricketts, Nation, & Bishop, 2007). Ricketts and colleagues found that children with poor reading comprehension were less able to verbally define words and read fewer exception words correctly. In spite of these outcomes it remains unclear to what extent individual differences in the accessibility of semantic knowledge help predict differences in reading comprehension.

Limitations in word knowledge have been suggested to be causally related to reading comprehension failure (Cromley & Azevedo, 2007). Vocabulary training studies reflect the importance of fast access to semantic word knowledge for reading comprehension. An experiment by Beck and colleagues (2002) revealed that instruction that involves multiple repetitions helps improve the speed of accessing the word’s meaning. According to Beck and colleagues good instruction helps children develop knowledge of the core concept of the word and how the word is used in different contexts to develop flexible knowledge about a word that contributes to reading comprehension. In studies that provided this rich vocabulary instruction, learners showed gains in both word knowledge and comprehension of text containing the words taught (Beck, McCaslin, & McKeown, 1980; Beck & McKeown, 1983; McKeown, Beck, Omanson, & Perfetti, 1983). Vocabulary that
does not produce sufficient fluency of access was found to not generalize to reading comprehension (cf. Jenkins, Pany, & Schreck, 1978).

The interrelationships among semantic word knowledge, lexical processing speed and reading comprehension are well captured in Perfetti’s verbal efficiency theory (1985) and in the lexical quality hypothesis (Perfetti and Hart, 2001, 2002). It proposes that the quality of lexical representations influences the ease with which those representations can be accessed. Quality here refers to the degree of orthographic, phonological and syntactic-semantic information and their integration. It is claimed that fast access to (semantic) word knowledge supports efficient word recognition as well as comprehension. High-quality word representations allow for automatic word processing, which enables children to use their mental resources for the comprehension of text, allowing them to use reading as a tool to acquire new concepts and information (Perfetti, 1998; Samuels & Flor, 1997). This again reflects the recurrent nature of the interaction: word knowledge allows for comprehension, comprehension allows for reading practice, reading practice strengthens word knowledge, and so forth.

There are indications that not only the speed of access but also the automatic activation of semantic word knowledge – as reflected by semantic priming – benefits reading comprehension. Nation and Snowling (1999) tested 11-year-old children and found that, although both good and poor comprehenders showed priming for function-related words (e.g. *broom – floor; shampoo – hair*), poor comprehenders only showed priming for words related through semantic categories (e.g. *cat – dog; airplane – train*) if these pairs were also strongly associated. Betjemann and Keenan (2008) also found evidence of semantic processing weaknesses in poor comprehenders. They assessed priming in children with reading disability and in age-matched controls (mean age 11;5), in visual and auditory lexical decision tasks. In the visual task, children with reading disability were found to have deficits in semantic (*ship – boat*), phonological/graphemic (*goat – boat*), and combined (*float – boat*) priming. The same pattern of semantic priming deficits also occurred in auditory lexical decisions, suggesting that the semantic deficits are a general semantic problem and not confined to the visual modality.
Moreover, the children with reading disability did not show greater priming in the combined condition than in the phonological/graphemic condition alone, suggesting that they are not getting any additional benefit from the semantic relatedness in the combined primes. The poor comprehenders also showed less priming than reading-age matched controls, suggesting that their priming deficits are not simply due to a lower reading level but are due to the reading disability in particular. These semantic deficits may contribute to both word reading and comprehension problems seen in children with reading disability.

Although most studies compare groups of proficient and less proficient readers on measures of semantic knowledge, Larkin, Woltz, Reynolds and Clark (1996) related individual differences in semantic priming to differences in reading comprehension. They administered measures of repetition priming and semantic priming to a sample of 60 sixth-graders. The children had to decide whether the two words in a word pair were synonyms or were unrelated. Priming could occur between consecutive pairs with up to three intervening pairs. There could be a like or positive match (e.g., big huge as the prime followed after zero, one, or two intervening items by large giant as the target) or a different or negative match (e.g., child city as the prime item followed after zero, one, or two intervening items by kid town as the target). The researchers report an average semantic priming effect of 75 ms across conditions, with more facilitation for positive compared to negative match items, which approached significance. There was no systematic effect of lag on priming. The semantic priming measure was found to account for 26% of the variance in reading comprehension. The repetition priming measure did not correlate significantly with reading. Larkin and colleagues conclude that individual differences in the spread of activation during verbal processing may underlie differences in some aspects of reading ability.

Weekes, Hamilton, Oakhill and Holliday (2008) had 32 children aged from nine to eleven study spoken words that were semantic associates (e.g., bed, rest, and awake) or phonological associates (e.g., pole, bowl, and hole) followed by free recall and a recognition test containing non-studied critical words (e.g., sleep and roll). Results showed reduced recall and recognition of critical words in the semantic
condition but not in the phonological condition for poor comprehenders, which shows that the tendency to infer themes from studied words in the semantic task is reduced in children with comprehension difficulties. Weekes and colleagues conclude that poor comprehenders are less sensitive to abstract semantic associations between words because of reduced gist memory. Gist traces represent interpretations of concepts (meanings, relations, and patterns) that are retrieved as a result of connecting the meaning across events. In this way, the gist is remembered, not the exact information (Kintsch & Yarbrough, 1982). Gist processes are generated automatically during text reading as a result of activation spreading across associative connections between words (Reyna & Kiernan, 1994).

In sum, although the importance of word knowledge for reading comprehension has been shown (cf. Mancilla-Martinez & Lesaux, 2010; Nakamoto et al., 2008; Proctor, Carlo, August, & Snow, 2005), the precise nature of the relationship and the components of word knowledge involved are not well understood. Empirical studies investigated participants from grade 1 up to adulthood and employed different operationalisations of depth ranging from knowledge of definitions and abstract category knowledge to combined depth/fluency measures. The distinguishability of breadth, depth and fluency is still unclear and a unique contribution of semantic word knowledge to reading comprehension has not been shown unanimously (see Vermeer 2001). There is research that suggests that also the automatic activation of semantic word knowledge – as reflected by semantic priming – explains variance in reading scores (Larkin et al., 1996). Differences between good and poor comprehenders in automatic activation of semantic word knowledge may suggest a delay in word knowledge development for poor readers. However, most studies concern group comparisons with reading-impaired children and controls. Research is needed to establish to what extent individual differences in semantic priming contribute to individual differences between children in reading.

2.4.1 Word knowledge and reading in monolingual and bilingual children

The link between vocabulary development and reading comprehension and by
extension academic achievement has been shown for both L1 (Snow, Porche, Tabors, & Harris, 2007) and L2 learners (Droop & Verhoeven, 2003; Proctor et al., 2005; Schoonen & Verhallen, 1998). However, in some learning contexts young L2 learners may not reach sufficient levels of word knowledge and reading to be able to achieve academically. One such group are bilingual learners from a minority background. Numerous studies have shown that bilingual children who are exposed to less input at home in the society’s main language fall behind their monolingual peers in language proficiency. Differences are well-established for vocabulary breadth (Appel & Vermeer, 1998; August, Carlo, Dressler, & Snow, 2005; Verhoeven & Vermeer, 1996). For example, Umbel and colleagues found that first-grade, bilingual Hispanic children in Miami in the US have poorer receptive knowledge of second language (L2) English words (PPVT-R) than their (English) L1-speaking peers (Umbel, Pearson, Fernandez, & Oller, 1992). Children speaking both English and Spanish at home (ESH) outperformed children speaking only Spanish at home (OSH) with the ESH group scoring low to average compared to monolingual norms.

In addition, significant delays in reading comprehension have been reported for bilingual minority children, despite the fact that these children have received the same education as monolingual children. In a research review, August and colleagues report that children speaking English as a second language who experience slow vocabulary development are less able to comprehend text at grade level than their English-only peers (August, Carlo, Dressler & Snow, 2005). Large and persistent gaps between the reading comprehension of language-minority and English-only children are mentioned as a result (see also Proctor et al., 2005). Droop and Verhoeven (2003) compared low SES Dutch third and fourth graders to the skills of low SES minority third and fourth graders from a Turkish or Moroccan background living in the Netherlands. The minority children were found to lag behind the Dutch children in reading comprehension and oral language proficiency. Furthermore, the oral Dutch skills of the minority children played a more prominent role in the explanation of their reading comprehension skills than the oral-language skills of the Dutch children, however. Delays in vocabulary and reading
comprehension have been shown to persist throughout elementary school (Aarnoutse, Van Leeuwe, Voeten, & Oud, 2001; Biemiller, 2005). Farnia and Geva (2007) show that even after years of schooling ethnic minority children do not catch up with their Canadian monolingual peers.

There is ample research that shows that bilingual children do not fail in decoding skills such as spelling and word reading, but lag behind in comprehension skills such as reading comprehension and vocabulary (Droop & Verhoeven, 2003; Smits & Aarnoutse, 1997; Verhoeven & Van Leeuwe, 2012). In a study with Spanish-speaking English language learners in fourth grade, Proctor and colleagues found that decoding played a less predictive role in reading comprehension than tests of oral vocabulary knowledge and listening comprehension. They conclude that given adequate L2 decoding ability, L2 vocabulary knowledge is crucial for improved English reading comprehension outcomes for English language learners (Proctor et al., 2005). Nakamoto, Lindsey and Manis (2007) conducted a longitudinal study with Hispanic English language learners and found that the children’s reading comprehension but not their word decoding began to fall behind the native English-speaking sample, starting in grade three (~nine years old). Similarly, in a large-scale study in the Netherlands on word knowledge, reading and writing literacy, Verhoeven en Vermeer (2006) found major delays in reading skills of ethnic minority groups in comparison to their native peers, while decoding differences between children were small. When Verhoeven and Van Leeuwe (2012) compared the reading comprehension performance, word decoding and listening comprehension skills of Dutch L1 and L2 learners (N= 1,293 and 394 respectively) throughout the primary grades, they found that the levels of word decoding were more or less equal in the two groups, whereas the L2 learners stayed behind their L1 peers in listening as well as reading comprehension. Linear structural models were computed separately for the early (1-2), intermediate (3-4) and upper grades (5-6). At each grade level the model was the same for the L1 and L2 learners. With progression of grade, the impact of word decoding on reading comprehension decreased, while the impact of listening comprehension showed an increase to the same extent in the two groups of learners. Indeed, it has been suggested that bottom-
up skills such as decoding largely drive the reading process in younger children (seven - nine years old) (Shankweiler, Lundquist, Katz, Stuebing, Fletcher, Brady, Fowler, Dreyer, Marchione, Shaywitz, 1999). So, despite delays for L2 learners in listening and reading comprehension, the relationships between decoding, listening and reading comprehension can be modeled in the same way for L1 and L2 learners.

Although bilingual children’s delays regarding reading comprehension and vocabulary breadth are now well established, differences between children in vocabulary depth are less well investigated. Vermeer (2004) compared Dutch monolingual and bilingual minority children in second grade (seven years old) on measures of lexical richness, definition skill and vocabulary breadth. He reported significant delays for bilingual children for all three measures. Lexical richness was operationalized as word variation and word frequency in the children’s spontaneous speech. Correlations between both lexical richness and definition skill ($r = 0.71$) and vocabulary breadth ($r = 0.50$) were substantial. In another study, Vermeer (2001) notes the danger that verbal definition tasks measure not just vocabulary, but also definition or abstract description skill. Cross-sectional data from the US, collected from fourth-grade Spanish-speaking and English-only children from four schools, also point to differences between L1 and L2 children (Carlo, August, McLaughlin, Snow, Dressler, Lippman, Lively & White, 2004). Two tasks examined children’s understanding of word meaning in production and comprehension: a polysemy task and a word association task (Schoonen & Verhallen, 2008). The results showed that English language learners not only know fewer English words but that their in-depth knowledge of word meaning is lacking as well.

Two studies explicitly looked at differences in children’s meaning assignment to words. Beside differences in reading skill, Verhoeven and Vermeer (2006) found that Mediterranean minority children’s competence in meaning assignment to written words increasingly lags behind those of their native Dutch peers with a delay of two to three school years by grade four (~ten years old). Verhallen (1994) investigated children’s semantic word knowledge as reflected by meaning assignment to words in an extended definition task. She interviewed 40 monolingual Dutch and 40 bilingual Turkish children (ages nine and eleven) about
the meaning of simple Dutch words like book, nose and hair. The analyses of the children’s associations and definitions showed that the bilingual children consistently mentioned fewer meaning aspects and came up with more syntagmatic and idiosyncratic associations than their monolingual Dutch peers, who provided more paradigmatic and context-independent meanings. For example, for the question ‘What is a nose?’ a typical bilingual response would be ‘With a nose you can smell’, whereas a typical monolingual response would be ‘A nose is a body part you can smell and breathe with’. Such differences in depth have been called hidden differences, since they are not obvious on the surface, as all children seem to more or less ‘know’ such simple words. When the researchers used a standardised, multiple-choice word association format, differences in meaning assignment between the learner groups persisted (Schoonen & Verhallen, 2008). In the Word Association Test (WAT), Schoonen and Verhallen asked children to identify the three words related to a target word. Children were required to make a distinction between words that are closely semantically related to the stimulus word and words that are associated with the word in a more incidental, context-bound way. The task is designed to have children consciously analyze and select appropriate meaning relations (see Figure 2.2). Findings showed that Dutch bilingual minority children were significantly worse at identifying related words for a given target word. Schoonen and Verhallen posit that this may signal a delay in semantic organization of the lexicon of these bilinguals. In a regression analysis, it was shown that depth scores obtained from the WAT contributed to children’s reading comprehension scores (Schoonen & Verhallen, 1998). It is not exactly clear what this contribution of depth to reading comprehension means. The depth scores may reflect differences in children’s underlying representations of word meaning and thus relate to reading performance. However, it may also be that monolingual children outperform bilingual minority children on the WAT simply because they are better problem solvers, rather than that there are differences in the children’s underlying semantic representations.
In contrast with the above is the study by Vermeer (2001) mentioned in section 2.4 who compared the breadth and depth of word knowledge of monolingual and bilingual minority kindergartners (mean age 5;6). For breadth, he used a receptive picture vocabulary task and a word description task (scored on a three-point scale); for depth, he used a structured interview task similar to the one used by Verhallen (1994). Monolingual children obtained higher breadth scores and in the interviews they mentioned more characteristics; however, there were no differences in the types of associations given in the interviews. Vermeer found high correlations between the breadth and depth measures for both monolingual and bilingual children, which lead him to question the conceptual distinction between vocabulary breadth and depth. He states that “the denser the network around a word, the richer the set of connections around that word, the greater the number of words known, and the deeper the knowledge of that word” (p. 231). The absence in this study of an empirical difference between breadth and depth does not mean that a theoretical distinction does not exist. They are useful as separate constructs, when contrasting for example a speaker who knows a large number of words to a limited degree with

**Figure 2.2** Sample item from the WAT, translated from Dutch, with correct answers (Schoonen & Verhallen, 2008: 219)
a speaker with a small vocabulary but an indepth, interconnected knowledge of the words he or she knows. If breadth and depth could be trained separately, uneven profiles could result.

The relation between children’s word knowledge and reading comprehension may be different for monolingual and bilingual minority learners. Verhoeven (2000) compared native Dutch-speaking children and minority children in the first two grades of elementary school on tasks of vocabulary knowledge, word decoding, word spelling and reading comprehension. He found word decoding to be highly comparable for the two groups, but the minority children were found to be less efficient on spelling and reading comprehension than their monolingual Dutch peers. A structural model of reading comprehension showed that vocabulary knowledge had a stronger impact for monolingual than for bilingual children. This is in line with Droop and Verhoeven (2003) who found that the oral skills of bilingual children contributed more to reading comprehension than was the case for their monolingual peers. Verhoeven (2000) concludes that children learning to read in a second language should be helped to build their word knowledge and that reading instruction should be matched to this knowledge.

Proctor, Uccelli, Dalton and Snow (2009) suggest that the role of depth of vocabulary knowledge in reading comprehension may depend on children’s levels of oral comprehension. In a study focusing on deep vocabulary intervention, they had a group of 35 bilingual and monolingual fifth-grade children working on developing depth of knowledge of eight words, culminating in an activity in which the children produced captions for images related to each word. The researchers scored the captions using a four-point depth scale. Results indicated a significant additional effect of depth of word knowledge (3%) in predicting reading comprehension after decoding and oral comprehension, particularly for children with average to strong oral language skills (Proctor et al., 2009).

Some remarks regarding the language background of bilingual minority children are in place here. These learners differ in a number of respects from so-called ‘balanced’ bilinguals. They are learners with less input and less varied input than their monolingual peers: research shows that bilingual minority children have
delays in comparison to their monolingual peers when exposure to the second language is limited (Hancin-Bhatt & Nagy, 1994; Vermeer, 1992). Children’s vocabulary development in the first language has been strongly related to parents’ socioeconomic background and education and to amount and quality of language exposure (Weizman & Snow, 2001). Some researchers emphasize that sufficient exposure to both languages may help bilingual learners in catching up with their monolingual peers (Umbel & Oller, 1994; Vermeer, 2001). Because of this unfavourable situation, it has been suggested that bilingual children acquire abstract semantic word knowledge later (Schoonen & Verhallen, 2008: 231). In the Dutch case, lexical delays of bilingual children have been shown to hold not only for Dutch but also for the home language (see Verhallen, Özdemir, Yuksel, & Schoonen, 1999), also with respect to vocabulary size (Appel & Schaufeli, 1990). Bilingual minority children are often faced with the complex task of learning to read in a language they are not accustomed to speak before they enter primary education. Even if they do speak the home and the majority language, often one of the languages appears to be significantly weaker than the other (depending on the domain) (Meisel, 2007). Research has shown that problems with the spoken second language may have an impact on reading processes, especially in the domain of reading comprehension skills (Geva & Verhoeven, 2000; Verhoeven, 2000).

The literature reviewed indicates a clear gap between monolingual and bilingual minority children in word knowledge and reading proficiency. There is some evidence that vocabulary differences pertain to semantic depth and knowledge of semantic word relations and associations. Most studies that focus on vocabulary depth use off-line measures. Little is known about the extent to which bilingual minority children fall behind in processing components of word knowledge such as the accessibility of semantic word knowledge, in terms of speed of access and semantic priming. Research is needed to investigate whether semantic priming differences, as have been found for good and poor comprehenders (Nation & Snowling, 1999), exist between Dutch monolingual and bilingual minority children and whether these differences can explain the differences in reading performance between these groups. Moreover, the relation between semantic word knowledge,
semantic processing and reading comprehension may be different for Dutch monolingual and bilingual children (Droop & Verhoeven, 2003; Verhoeven, 2000, Proctor et al., 2009).

2.5 Open issues and focus of the present work

The literature reviewed shows that there are differences between children in their abstract, semantic knowledge of words and in the accessibility of that knowledge. Children’s word knowledge develops from initially context-dependent to more context-independent or abstract. Differences between children and adults in abstract knowledge of word meaning have been reported. One line of research investigates differences in word associations between monolingual and bilingual learners of different ages and proficiency levels. However, association studies fail to show consistent patterns. Research is needed to investigate to what extent word knowledge and word associations of monolingual and bilingual children differ systematically (Fitzpatrick & Izura, 2011; Namei, 2004; Nissen & Henriksen, 2006). Also, the value of word associations as a reflection of semantic word knowledge needs to be addressed.

One limitation of previous work examining individual differences in children’s semantic knowledge is that off-line tasks requiring conscious processing have mostly been used (see also Read, 2004). For example, in the study by Braze, Tabor, Shankweiler and Mencl (2007), expressive vocabulary was measured by asking children to define words. It is possible, therefore, that the poor comprehenders’ lower performance was due to difficulties verbalizing concise definitions for words, rather than a reflection of deficient understanding. Thus, outcomes from such tasks do not allow for claims about the status of children’s underlying semantic representations.

The research reviewed indicates a link between rich semantic word knowledge and fast access to lexical and semantic information. Adults’ knowledge of word meaning seems to be related to fast performance in lexical and semantic tasks (e.g., Pexman, Lupker & Hino, 2002). Semantic processing skills appear to be still in development in young children as they have been found to be less sensitive to
semantic category information (Perraudin & Mounoud, 2009). However, it is unclear to what extent there are individual processing differences between children in terms of speed of access to lexical and semantic information and sensitivity to semantic priming.

Previous studies show a relation between reading performance and both speed of access and the activation of semantic word knowledge – as reflected by semantic priming. Important here is that most studies concern group comparisons and do not relate individual differences in processing skills to individual difference in reading comprehension. It needs to be investigated to what extent individual differences in speed of access to lexical and semantic information and semantic priming can account for individual differences in reading comprehension.

The literature indicates a clear gap between monolingual and bilingual minority children in word knowledge and reading proficiency. Unclear is whether bilingual children fall behind in processing components of word knowledge, such as its accessibility in terms of speed of access to lexical and semantic information and in terms of semantic priming. Research needs to address whether semantic priming differences, as have been found for good and poor comprehenders (Nation & Snowling, 1999), exist between Dutch monolingual and bilingual minority children and whether these differences can explain the differences in reading performances between these groups.

In sum, the respective roles of semantic word knowledge, speed of access to lexical and semantic information and the activation of semantic information in reading comprehension deserve further investigation. Dutch bilingual minority children are included in this thesis because they are the children at risk with respect to (lack of) ‘deep’ and accessible word knowledge.

This review leads us to the following research questions that are addressed in the subsequent chapters:

1. Are there structural differences between Dutch monolingual and bilingual minority children and adults in the heterogeneity of their word association responses to simple words? (Chapter 3)
2. To what extent do age and language background play a role in the patterns of word associations of Dutch monolingual and bilingual minority children and adults? (Chapter 3)

3. Do Dutch monolingual and bilingual minority children differ in the availability of their semantic word knowledge? (Chapters 4 and 5)

4. Do Dutch monolingual and bilingual minority children differ in the speed with which they access lexical and semantic information about (individual) words? (Chapters 4 and 5)

5. Do Dutch monolingual and bilingual minority children show different priming effects for semantically related words? (Chapter 5)

6. Do availability of semantic word knowledge, speed of access and sensitivity to semantic priming explain individual differences in reading comprehension between Dutch monolingual and bilingual minority children? (Chapters 4 and 5)

The study reported on in Chapter 3 was set up as a preparatory study to generate association norms for the studies in Chapters 4 and 5. The data allow for an investigation of research questions 1 and 2. A comparison of word associations of monolingual and bilingual minority children from two distinct age groups pinpoints cognitive and lexical differences and addresses the role of word associations in the assessment of word knowledge. We hypothesize that deep vocabulary knowledge as reflected in semantic word knowledge is more prominent and hence more easily accessible in the mental lexicon of Dutch monolingual children than in the lexicon of Dutch bilingual minority children and that a preference for semantic word associations shows in monolingual children’s word association responses (research question 1). If differences in word associations between monolinguals and bilinguals hold for children and for adults, then they are most likely due to lexical differences (too little exposure to the words to come up with native-like associations); if differences ‘disappear’ with age, then they most likely spring from semantic, conceptual differences (research question 2).
Chapter 4 assesses children’ semantic word knowledge and speed of access (research questions 3 and 4) with an offline semantic word knowledge task and a speeded categorization task. We predict significant differences between monolingual and bilingual minority children in the availability of their semantic word knowledge as well as in speed of access to word knowledge. We expect these differences to contribute to differences between the children in reading comprehension (research question 6).

Chapter 5 investigates differences in children’ semantic word knowledge (research question 3) as well as in their speed of accessing lexical and semantic information (research question 4) on the basis of a lexical decision and a semantic classification task. Research question 5 is addressed through the semantic priming component in those tasks. We expect differences between monolingual and bilingual minority children in semantic word knowledge, in speed of access and in the activation of abstract, semantic word knowledge as reflected in semantic priming. On the basis of regression models of reading comprehension, Chapter 5 addresses research question 6. We expect differences in availability of semantic word knowledge, speed of access and sensitivity to semantic priming to contribute to explaining variance between the children in reading comprehension.

In sum, the empirical studies in Chapters 3 to 5 assess differences between Dutch monolingual and bilingual minority children with respect to word knowledge and reading, and they investigate the relation between word knowledge and reading comprehension. This research thus contributes to understanding some of the mechanisms underlying individual differences in reading comprehension between these children.
Chapter 3

Do word associations assess word knowledge? A comparison of monolingual and bilingual, child and adult word associations

Differences in word associations between Dutch monolingual and bilingual minority children can reflect differences in how well seemingly familiar words are known. In this chapter, we will compare monolingual and bilingual, child and adult free word associations. As briefly explained in section 2.5, comparing language and age groups in this task allows for an indication of cognitive and lexical differences in word association behaviour (research questions 1 and 2). The chapter starts with a brief introduction to the word association paradigm and relevant earlier studies (section 3.1). It introduces the idea of semantic networks (section 3.1.2) and the difference between contextual and semantic knowledge (section 3.1.2). Section 3.1.3 on categorization and elicitation procedures is followed by a discussion of the research methodology used (section 3.2). The results are presented in section 3.3 and discussed in section 3.4.

1 This chapter is an adapted version of Cremer, M., Dingshoff, D, De Beer, M. & Schoonen, R. (2011). Do word associations assess word knowledge? A comparison of L1 and L2, child and adult word associations. International Journal of Bilingualism. 15(2), 187-204. The study reported on was set up as a preparatory study to generate association norms for the studies in Chapters 4 and 5. Part of the research was carried out as course work by the second and third authors under the supervision of Prof. dr. Jan Hulstijn.
3.1 Background

Word knowledge develops from experiences with the world around us. Apart from learning labels for items or objects (a complicated process in itself), a language learner has to learn the exact extension of a word’s meaning (i.e. which other objects can be called the same name) and has to see in what way words fit together; in other words, he or she has to discover relations between words and thus build up a semantic network (Aitchison, 2003; Henriksen, 1999). Understanding relations between words means having a rich and a densely interconnected mental lexicon, which is considered an important feature of developing language proficiency. Word knowledge is often described in terms of a lexical and a semantic level (Kroll, Michael, Tokowicz, & Dufour, 2002; Potter, 1979; Snodgrass, 1984). The lexical level consists of the words in all of a speaker’s languages; the underlying concepts are stored at the abstract, semantic level (Kroll and Stewart, 1994). Knowledge of relations between words can be studied through a word association task, which asks people to respond with the first word that comes to mind for a number of stimulus words. Such associative behavior is presumed to reflect how words and concepts are structured and interrelated in the mind (Deese, 1965; Szalay & Deese, 1978). The word association paradigm has long been used in psychology to screen patients for possible idiosyncratic associations (Lukavsky, 2004; Mohr, Graves, Gianotti, Pizzagalli, & Brugger, 2001; Slechta, 2001). In L1 studies, it has been used to investigate children’s lexical and semantic development (Cronin, 2002). Schmitt (1998) claims that word associations reflect the strongest mental connections between words in the mind. They result from experience with the associated words. Somewhat more recently, word association tasks have been used to compare the organizational principles of the L1 and L2 lexicon (Fitzpatrick, 2006; Sheng, McGregor, & Marian, 2006; Zareva, 2007), and they can also be seen as reflecting lexical and conceptual processing (Van Hell & De Groot, 1998). Differences in types of association responses are sometimes taken as an indicator of language proficiency (Zareva, 2007), but proficiency level has not consistently been found to predict response patterns (Kruse, Pankhurst, & Sharwood Smith, 1987). So far, studies have failed to find consistent behavior patterns. Some have found more or
less stable association patterns for native speakers showing that responses become more meaning-based after late childhood. More varied responses have been found for nonnative speakers (Meara, 1983). It remains uncertain to what extent L1 or L2 speakers behave as homogeneous groups, or, how differently dispersed their responses are across different response categories.

In (young) children lexical and conceptual development are intertwined. As the first language is acquired, young children learn words (lexical labels for concepts in the world) and they learn concepts (what concepts in the world mean, how they relate to other concepts, where concept boundaries lie, and so on) (Verhallen, 1994). When we compare monolingual L1 and bilingual L2 development in a given language, it is not always clear whether differences found have a lexical and/or conceptual origin (Verhallen & Schoonen, 1993). Some of the L2 ‘problems’ may not occur in bilingual children’s L1, suggesting L2 language difficulties; others may show in their L1 too, suggesting broader, conceptual problems (Verhallen, Özdemir, Yuksel, & Schoonen, 1999; Verhallen & Schoonen, 1998).

In the present study, we investigate differences in word knowledge of child and adult first and second language speakers of Dutch; we compare their Dutch word associations. A combined analysis of child and adult association behavior allows us – to some extent – to tease apart language background effects (monolingual versus bilingual) and cognitive development effects (child versus adult). To prevent a possible influence from respondents with different L1 backgrounds, we focus on a ‘homogeneous’ bilingual group, that is, Turkish L1 speakers with Dutch as an L2.

### 3.1.1 Word knowledge: lexical and semantic network

Researchers have distinguished different dimensions of lexical knowledge, ranging from knowledge of a word’s form, position, semantic network, collocations and associated words, to polysemy, formal definitions and its receptive and productive use, to name but a few (Cronbach, 1942; Henriksen, 1999; Nation, 2001; Richards, 1976; Wesche & Paribakht, 1996). Depth of word knowledge can comprise all of the
aforementioned aspects of word knowledge. Read (2004) focuses on word meaning and suggests that we can understand depth as learners’ ability to distinguish semantically related words and, more generally, their knowledge of how individual words are linked to each other.

Most models of multilingual word processing distinguish between a lexical and a semantic level. In the Bilingual Interactive Activation Plus (BIA+) model (Dijkstra and Van Heuven 1998, 2002), lexical (and sublexical) levels involve orthographic and phonological information processing and a semantic level contains semantic/conceptual representations. Connections at and between the levels lead to dynamic activation spread between levels during language processing. The Revised Hierarchical Model (RHM), developed by Kroll and Stewart (1994), also contains a lexical and a conceptual store. Both lexical and semantic representations are shared across languages (Kroll et al. 2002); that is to say, a speaker’s languages ‘tap into’ one common lexical store and one common conceptual store. This means that words from a speaker’s languages are linked in terms of, for example, orthography and phonology but also in terms of meaning and conceptual features. These connections underlie word processing. Even though both conceptual and lexical links are active in semantic memory, the strength of such links differs as a function of (L2) fluency (Dijkstra, 2007). Whereas in the RHM L1 word use is conceptually mediated from the start, L2 learners acquire lexical links between L2 and L1 before they are able to conceptually mediate L2 word use. Thus, for a beginning learner an L2 word will most likely be translated into its L1 word and with that, activate its semantic/conceptual underpinnings.

Quality of word knowledge is often seen as a feature of the semantic network (Aitchison, 2003; Henriksen, 1999). Assuming that words get their meaning from their relations to other words, acquiring more meanings and related words leads to a deepening of word knowledge: the network becomes denser and highly interconnected. Henriksen states that understanding the relations among the items is a prerequisite for a more precise understanding of each individual item (Henriksen, 1999). Similarly, Meara and Wolter emphasize the interdependence of vocabulary size and depth and organization (2004). The idea of networks also grounds Read’s
Word Associates Format (1993) in which test takers have to find relations between a stimulus word and a given set of other words. This approach is also the focus of the multiple-choice, word association test format for children (WAT) developed by Schoonen and Verhallen (2008).

3.1.2 From context dependent to meaning based
Children build their lexicon from scratch by adding and reorganizing their lexical semantic network, a never-ending process. Links between words and concepts, from one or several languages, change as a function of salience, exposure to various uses, fluency and such. Words that co-occur in the language encountered will be stored together. Thus, associative relations between words are based on the likelihood of two words co-occurring in the language (Wettler, Rapp, & Sedlmeier, 2005). As language is learnt, word knowledge is gradually abstracted and refined from children’s functional, idiosyncratic knowledge (Nelson, 1977, 1982) into more meaning-based, semantic connections. As mentioned in section 2.1.2, Petrey (1977) describes this knowledge development as a dramatic shift from episodic to semantic. Using responses from Entwisle’s (1966) free word association data, she shows that “[e]arly associations ... depend not on semantic content but on episodic experience” (1966: 65). Very young children respond with contextually related words (example from a medical context: examine – needle) rather than with inherent semantic associations (examine – look (at), check), which become common after third grade (Petrey, 1977). Episodic kindergarten responses to dark, such as bed and sleep (both indirectly related to dark), disappear entirely by third grade, when light has become a common response. This illustrates the abstraction to more context-independent associations. This semantic development was also shown by Verhallen and Schoonen (1993), who set up an interview task and found that nine-year-old children mentioned significantly more context-dependent meaning aspects than eleven-year-olds, who mentioned relatively more abstract, context-independent meaning aspects.

Similar differences in quality of knowledge have been found for monolingual and bilingual children for, apparently, known words. This is remarkable, as all young children seem to acquire word knowledge rapidly, even
children who have been raised in another language than the language used at school. Verhallen (1994, Verhallen & Schoonen, 1993) has shown that children with Dutch as a second language not only have less semantic knowledge of familiar words compared to their monolingual peers, their semantic knowledge is also more context-dependent and more subjective. Most of these children who speak Dutch as an L2 have an immigration background and do not speak Dutch at home. This causes a mismatch between home environment and school setting in terms of language use. The lacking school achievements of these bilingual minority children may lie at the lexical (word) and/or at the semantic (conceptual) level. Because of this mismatch between home and school, children with Dutch as an L2 may have a delay in semantic development and abstraction in their L1 as well: their mental lexicon may not (yet) contain (strong) connections between semantically related words but rather between subjectively, context-dependently related words. This is a feature of the semantic level and, consequently, will also show in their L1. Verhallen and colleagues (Verhallen et al., 1999; Verhallen & Schoonen, 1998) indeed found that children’s semantic knowledge about (Turkish) L1 words was even more limited than their relatively limited semantic knowledge in L2.

Alternatively, if word knowledge displayed in association tasks is only explained by linguistic contiguities and associative learning (Wettler et al., 2005), Dutch bilingual minority learners may not have had enough opportunities for associative learning, because of less exposure.

Insights into the organization and development of language learners’ mental lexicons serve both assessment and research purposes. If we consider word association tasks as a kind of language test, we should bear in mind that association responses reveal subconscious connections in a constantly changing mental word base; connections between words and concepts are constantly being added, strengthened and weakened. Free word associations may therefore be an indicator of the language learners’ current states of mind, rather than a deliberate and reliable expression of their word knowledge.
3.1.3 Assessing word associations: categorization and elicitation

When using word associations to assess word knowledge, the elicitation procedure used, the selection of stimulus words, and the classification of responses (see Fitzpatrick, 2006) have to be approached carefully.

The different kinds of association given by participants can be categorized in different ways. Using a meaning-based approach, De Groot (1980) distinguishes objective versus subjective semantic relations. Objective semantic relations hold for all users of the language; examples include superordinates, subordinates and synonyms. Subjective semantic relations hold for individuals; an example could be \textit{vase-mother}. De Groot further distinguishes phonologically related responses, word extensions and idiomatic responses. Another, commonly used, way of categorizing association responses is in terms of the syntagmatic-paradigmatic distinction. This distinction is mainly based on word class. Paradigmatic associations are those in which stimulus word and response word belong to the same word class (e.g., noun–noun); syntagmatic associations have a word class other than the stimulus word (e.g., noun–verb). Often the substitutability of paradigmatic words is contrasted with syntagmatic words occurring together with stimulus words in discourse (Cronin, 2002). The paradigmatic-syntagmatic approach to categorizing associations has received criticism since the 1970s. Petrey (1977) labels the categories paradigmatic and syntagmatic inadequate for describing changes with age in word associations, “because [they] can designate nothing but syntactic properties” (1977: 69). She prefers the terms episodic and semantic (see section 2.1.2). When given the word \textit{add}, children gave the episodic responses \textit{flour} (syntagmatic) and \textit{cook} (paradigmatic); both are related in the same way to \textit{add} (they come from the same situation) but would be misclassified as opposites by S-P terminology. She adds that, “syntagmatic responses properly so-called are highly problematic because either semantic or episodic storage could produce them” (1977: 66). Nissen and Henriksen (2006) criticize the syntagmatic-paradigmatic shift as they find a surprisingly strong native speaker preference for syntagmatic responses. According to Fitzpatrick the syntagmatic-paradigmatic distinction imposes artificial constraints on the exploration of response types, as their contents are difficult to define. Many studies
do indeed add an ‘other’ category (sometimes included in the phonological) for difficult to classify responses, but this might be inherent to any classification scheme. In her categorization scheme Fitzpatrick uses four main categories: meaning-based, position-based, form-based and erratic associations. Each main category is divided into subcategories, detailing the type of relation between stimulus and response word. Fitzpatrick advises the use of subcategories, as only general main categories do not allow a very precise method of categorization (2006). In addition to categorizing the associations, it is insightful to know how dispersed the responses are across the association categories, in other words, do (groups of) speakers give many different associations or do they have preferences? Since L2 associations are often claimed to be less predictable than L1 associations, L1 and L2 speakers may show differently dispersed responses. L1 and L2 associations may for example cluster in different categories.

Different studies use different procedures of eliciting word associations and findings are diverse. Schmitt (1998) advocates eliciting multiple responses. He uses word association responses to determine whether a participant’s associative network is native-like. Although the assumption is that the participant responds with the strongest associate, Schmitt questions whether this will always be the case (1998: 391). He pleads for multiple responses to one stimulus word as this might better capture the richness of a subject’s associative network which can then be compared to a native speaker’s network. Others prefer the single-response procedure as it provides a reliable index of a word’s strongest associates (Nelson, McEvoy, & Dennis, 2000). Nelson and colleagues found that second responses added new but weak items to the set, and, when the primary associate was not produced as the first response, it tended not to be produced on the second. Furthermore, the first response may act as a stimulus for the second response thus eliciting a chain of responses and “such responses are not independent of the first” (2000: 891). Therefore, the present study adopts the single-response paradigm.

Finally, the kind of stimulus words used, influence the associations elicited. Many researchers found an influence of word class on response types (Cronin, 2002; Nissen & Henriksen, 2006). Nouns, for example, elicit paradigmatic responses more
often than verbs or adjectives do. There are inherent differences between word classes that need to be taken into account when comparing studies. Also word frequency affects the probability of response types (Cronin, 2002). Highly frequent words tend to elicit predictable words, especially among native speakers. This homogeneity of native speaker responses, compared to those of L2 speakers, tends to disappear when words are less frequent (Fitzpatrick, 2007). De Groot (1989) looked at imageability of words and found that concrete words elicit more homogeneous responses than abstract words. In our study, the list of stimulus words consisted mostly of nouns; words had varying frequencies but all occurred in a Dutch corpus of words used in primary school materials (Schrooten & Vermeer, 1994).

Differences in word knowledge between speakers might show in free word association tasks. However, the question is to what extent word association tasks can be considered assessment tools for quality of word knowledge. Can we assume a continuous development towards more and more (abstract) meaning-related word associations, preferably in one category, for example superordinates? We should bear in mind that word associations are not necessarily the result of abstract, semantic knowledge. Language use, such as amount or recentness of exposure to co-occurring words, contributes to the likelihood that one word triggers the other in a spontaneous word association task. In this chapter, we address structural differences in word associations at two levels. Firstly, we investigate possible differences in degree of dispersion of word associations across different types of associations between L1 and L2 children and L1 and L2 adult speakers of Dutch (research question 1). Secondly, we investigate differences for these speakers in distributional patterns of word associations across four categories, and within those main categories across subcategories, thereby assessing whether possible differences between L1 and L2 speakers hold for both children and adults (research question 2).

Differences in word associations between L1 children and L2 children might be caused by differences both at the lexical and at the semantic developmental level. These two levels can be investigated in the comparison of child L1 and L2 and adult L1 and L2 speakers. If differences between L1 and L2 groups hold for children
and for adults, they are most likely due to lexical differences (too little exposure to the words to come up with native like associations); if differences ‘disappear’ with age, they most likely reflect semantic, conceptual differences. Thus, comparison of L1 and L2 speakers from two distinct age groups could further reveal the role of word associations in the assessment of word knowledge.

3.2 Method
To answer the foregoing questions, the free word associations of over 400 children and a smaller number of adults were analyzed and compared. This study was originally set up to generate child word association norms. The adult associations were collected later by the second author as part of course work. For this reason, the number of adults is smaller than the number of children.

3.2.1 Participants
There were 422 children and 54 adults participating in the study. Children were tested in February 2007 as part of a larger study; adults were tested in an exploratory follow-up in March and April 2008. The original number of children tested was 534. Of the 534 original respondents, 112 were excluded from the current analyses as they spoke first languages other than Turkish (or Dutch). The 422 children were either monolingual speakers of Dutch (n = 389), or bilingual speakers with Dutch as a L2 and Turkish as the L1 (n = 33). Ages ranged from eight to thirteen. Children were pupils from seven primary schools in the Randstad, the western part of the Netherlands. Socio-economic background of schools varied. A third of the children were in grade six, the final year of primary school (eleven – twelve years old), a third in grade five (ten – eleven years old), and a third in grade four (nine – ten years old).

The adult participants spoke Dutch either as their L1 (n = 41) or L2 (n = 13). Again, second language speakers of Dutch spoke Turkish as their first language. The age of the adults ranged from 17 to 59. L2 adults were recruited by the second author through a Dutch learning institute located at the university, through Turkish student societies, and through personal connections. The adult sample included 32
women and 22 men. Level and amount of education among adults varied from lower level tertiary education to technical school to university education. Most L2 adults were born in the Netherlands; one man had lived there since a year.

3.2.2 Materials

A set of 118 stimulus words was used. Word selection was inspired by the Word Association Task (Schoonen & Verhallen, 2008) and determined by the frequency of occurrence in a word corpus based on reading materials for children in primary school (Schrooten & Vermeer, 1994). Words were 108 nouns, two verbs and eight adjectives of varying length. The words varied in frequency but all words occurred in the Schrooten and Vermeer corpus; high-frequent words from the corpus were selected as much as possible to ensure participants’ familiarity with the stimuli. The complete word list with the 118 randomized words was completed by all adults. Because of their shorter attention span, the children did not respond to as many as 118 words. Each child completed a 59-word list. The original list was split into two lists (I and II, see Appendix A) each of 59 words; both lists were randomized into five different orders, resulting in ten versions. This was done to prevent order effects and to prevent cheating during classroom administration. Both lists had been piloted with two children in the target age to assess the suitability of words and of the list length.

The list was preceded by a written instruction. At the end of the word list, a number of questions were included regarding participants’ background. These were somewhat different for children and adults. The children answered questions about the country in which they had been born, which language they had learnt first and which language(s) their mother and father, respectively, spoke with them at home. The adults were asked additional questions about their schooling and profession.

3.2.3 Procedure

All children were tested at school during a morning class with the paper-and-pencil version of the task. Adults were tested individually. Only Dutch was spoken. Participants were instructed to give one response only. To ensure that all child
participants understood the instruction a number of examples were discussed with
them in class. After completing the association task, the responses of a few children
from each class were investigated by means of a retrospective interview. For
practical reasons, this was not done for adults. All responses were entered into a
database where they were categorized manually.

Categorisation scheme
The coding scheme used to categorize the responses is adapted from the
classification systems used by Petrey (1977), De Groot (1980), and Fitzpatrick
(2006) as discussed in Section 2.1.2. In order to trace the prominence in the lexicon
of meaning aspects of word knowledge, our coding scheme has as its main
distinction responses that are directly (inherently) related to the meaning of the
stimulus word and responses that are only indirectly or subjectively related to the
meaning of the stimulus word. This contrasts with Fitzpatrick’s system, but is
similar to that of De Groot. Direct meaning-related words are related irrespective
of the context in which they appear. Thus, in our study, Fitzpatrick’s examples lecture–
university, stability–baby, and cultural–cathedral (2006: 133) would not be direct
meaning-related. In addition to the two meaning-based categories, we identify a
form-based and an ‘other’ category. The last two are not based on a meaning
relation. Form-based responses contain both responses based on orthography and on
phonology; it thus includes the traditional ‘clang’ (sound-based) category. Clang
associations are often characteristic of the associations of beginning learners (Meara,
1983). The ‘other’ category contains non-classifiable responses, repetitions and null
responses. The four main categories are subdivided into more specific subcategories
allowing more detailed analyses, as is shown in Table 3.1. Table 3.1 shows
examples of each category.
Table 3.1 Categorization scheme

<table>
<thead>
<tr>
<th>Category</th>
<th>Example</th>
</tr>
</thead>
</table>
| 1. Direct meaning-related | 1. coordinate (dog-cat)  
2. subordinate (car-porsch)  
3. superordinate (deer-animal)  
4. antonym (fake-real)  
5. partonomy 1 (part-whole) (teeth-mouth)  
6. partonym 2 (whole-part) (elephant-trunk)  
7. context-independent characteristic (sledge-snow)  
8. goal/target (knife-cut)  
9. synonym (simple-easy) |
| 2. Indirect meaning-related | 1. subjective association (motor bike-cool)  
2. composite word (apple-tree)  
3. context-dependent characteristic (strong-muscles) |
| 3. Form-based association | 1. change of affix (dog-dogs)  
2. similar form, other meaning (hash-harsh) |
| 4. Other | 1. non-classifiable (volcano-fish)  
2. repetition (ocean-ocean)  
3. no response |

Categorizing responses

In our scheme, collocations can belong to various categories depending on the collocation and the exact relationship between the two parts. A collocation such as bloem – plukken (cf. English pick – [a] flower) belongs to category 2.3. The stimulus-response pair neus – snuiten is a collocation (cf. English blow [your] nose) belonging to category 1.7. Functional associations such as goal-target pairs (category 1.8) are taken as aspects of abstract, semantic knowledge (a knife cuts, irrespective of the context). Not all associations were clear-cut for classification. In Dutch kaars-vet (candle-wax/greasy), it was unclear whether the response vet is
meant to be the second part of the composite word kaarsvet (candle wax, category 2.2) or whether it is a separate adjective referring to a candle as being greasy (category 1.7). Yet, the unlikely co-occurring of kaars and vet separately compared to the occurrence of the word kaarsvet decided for the composite word option.

Difficult cases show that relatedness in meaning has to be weighed for each individual word pair. Two words can be related on a continuum from clearly not related through indirectly related to clearly related (i.e., prototypical category members). One pair of context-independently related words can be more context-independently related than another (compare sledge-snow and ocean-fish).

To ensure reliable categorization, part of the collection of association responses was categorized twice by independent raters and interrater reliability was determined. The responses to 28 words (out of 118, i.e., 23.7%) were categorized twice. This was only done for the child data. The raters assigned 98.5% of responses to the same subcategory.

3.2.4 Analyses

All word associations were classified into one of the 17 categories (Table 3.1). As a first step in assessing respondents’ word associations, we addressed variation in kinds of associations given. Do L1 and L2 speakers prefer one or two types of associations or do they respond with all types of associations? For this, we computed Gini’s concentration index (Wickens, 1989: 130ff.) for each participant\. The lower the index, the less the associations are spread across the 17 categories. A participant who uniformly responds with the same type of association scores 0. Maximum spread across the categories is achieved when 3–4 associations occur in each category for children (59/17) or 6–7 for adults (118/17). In these cases the concentration index is maximum, that is .94. To check whether the two sets of stimulus words in the two lists for the children were similar enough to treat them as interchangeable, we assessed Gini’s concentration index in the adult data for the words of list I and list II separately. Remember that adults responded to all words,

---

2 Gini’s concentration index is defined as: $C_g = 1 - \Sigma \pi_j^2$, with \( \pi_j \) referring to the proportion of associations in category \( J \).
whereas the children responded to the words of one of the lists. It turned out that the correlation between the two lists was .81, which is a reassuring split half reliability.

Next, analyses involving log linear model fitting were used to analyze the distributional patterns of word associations across the categories for both L1 and L2 speakers, children and adults at once (see later). First, response patterns are studied at the level of the four main categories; then, the subcategories in each main category are analyzed.

In a log linear analysis, observed frequencies are described in terms of main effects and interaction effects (see Everitt, 1977; Wickens, 1989) in the most parsimonious way possible. Our unit of analysis is the word association. Word associations are aggregated across words and participants. Within a group (adults vs. children) all participants responded to the same number of words (118 and 59 respectively) and thus contributed equally to the group’s distributional pattern. The three-dimensional table of Category by Age by Language Background is analyzed in terms of effects of each of the dimensions and their interactions, as is done in an ANOVA. Main effects indicate that the levels of a dimension are not equally frequent. In our analyses this is a trivial finding: it indicates that there are more L1 associations than L2 associations, and more child associations than adult associations, and more associations in one category than in another. The first two are artifacts of the design, because we have more L1 than L2 participants, and more children than adults. The latter effect is not surprising either, because there is no theoretical reason to expect an equal distribution across the categories involved. The interesting effects are those in which categories interact with person variables (age and language background), which means that certain groups of participants responded with word associations more frequently in a certain category than other groups. A three-way interaction of Age, Language Background and Category would imply that different associative behavior between L1 and L2 speakers is mediated by age, meaning that differences change over time, or that age differences differ for L1

---

3 An alternative approach is to compute the percentages of responses or mean score for a certain category per person, and then to compare these ‘scores’ across groups (see Fitzpatrick, 2006). However, the comparisons to be made are interdependent: a high percentage of associations for one category implies low percentages for the other categories.
and L2 speakers.

The outcomes of the log linear model analyses for main categories and subcategories will be interpreted in two stages: first the search for the best data description will be described, which is a trade-off between fit and parsimony, and second the interpretation of the parameters for age, language background, and category in the selected model will be discussed. Because of the large number of word associations and $\chi^2$’s sensitivity to sample size, models will easily be rejected. To allow for a more comprehensive evaluation of models, descriptive measures will be reported, beside the $\chi^2$ test. A normed fit index (NFI) $\Delta$ indicates the relative increase in fit compared to a base model and the perfect fit of a saturated model (Bonett & Bentler, 1983). As a starting point, we took a base model including the main effects and the interaction between Age and Language Background, as these effects are inherent to the design of the study.

### 3.3 Results

A total of 31,270 association responses was categorised, of which the larger part comes from children (24,898), especially L1 children (22,951) (see Table 3.2).

<table>
<thead>
<tr>
<th>Age</th>
<th>Language Background</th>
<th>Main Categories</th>
<th>Total N (100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Direct Meaning-related</td>
<td>Indirect Meaning-related</td>
</tr>
<tr>
<td>Child</td>
<td>L1</td>
<td>47.9</td>
<td>45.3</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>43.1</td>
<td>40.2</td>
</tr>
<tr>
<td>Adult</td>
<td>L1</td>
<td>46.1</td>
<td>53.0</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>39.9</td>
<td>57.2</td>
</tr>
</tbody>
</table>

The general pattern is that most of the reported associations are, directly or indirectly, meaning-related (93.6%). Few associations are form-based or unclear (‘Other’, 6.3%); these are almost exclusively given by the L2 children. This indicates that respondents were familiar with most stimulus words, as we had
intended. The adults have a small preference for Indirect rather than Direct Meaning-related associations (mostly expressed by the L2 adults). L2 children give less meaning-related responses than L1 children, in both the direct and indirect category. The relations between types of associations, language background, and age will be analysed in more detail when we fit log linear models.

### 3.3.1 Degree of dispersion

Analyses of Gini’s concentration index show that participants’ word associations spread across the 17 response categories. All groups score high, that is, all groups give associations of all types (see Table 3.3).

<table>
<thead>
<tr>
<th>Age</th>
<th>Language Background</th>
<th>L1</th>
<th>L2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child</td>
<td>M</td>
<td>.81</td>
<td>.80</td>
</tr>
<tr>
<td></td>
<td>sd</td>
<td>(.056)</td>
<td>(.091)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>389</td>
<td>33</td>
</tr>
<tr>
<td>Adult</td>
<td>M</td>
<td>.78</td>
<td>.73</td>
</tr>
<tr>
<td></td>
<td>sd</td>
<td>(.064)</td>
<td>(.115)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>41</td>
<td>13</td>
</tr>
</tbody>
</table>

A two-way ANOVA with Age and Language Background as fixed factors reveals a significant effect of Age ($F = (1,472) = 21.2, p < .001, \eta_p = .043$) and of Language Background ($F = (1,472) = 6.7, p = .010, \eta_p = .014$): children’s responses are more dispersed than those of adults; L1 responses are more dispersed than L2 responses. The age effect is larger than the language background effect. The standard deviations show that the groups of L2 speakers differ slightly more among themselves in their dispersion than the L1 speakers.
3.3.2  Response patterns for main categories

Log linear analyses show in which respect response patterns are different for the groups. The fit of different models is summarized for main categories and subcategories (Tables 3.4 and 3.5 respectively).

Table 3.4  Model fit for main categories

<table>
<thead>
<tr>
<th>Models</th>
<th>χ²/df</th>
<th>p</th>
<th>Δ NFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. [A<em>L][C</em>L]</td>
<td>507.5/6</td>
<td>&lt;.001</td>
<td>.31</td>
</tr>
<tr>
<td>3. [A<em>L][C</em>A]</td>
<td>315.0/6</td>
<td>&lt;.001</td>
<td>.57</td>
</tr>
<tr>
<td>5. [C<em>A</em>L]</td>
<td>0/0</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

C=Category, A=Age, L=Language Background; [ ]=interactions including lower order interactions and main effects.

The model fit indexes in Table 3.4 show that statistically speaking a model without the three-way interaction (i.e., model 4) should be rejected (χ²(3) = 18.8, p < .001). However, when considering the NFI, we see that the model with just the two-way interactions with Category describes the data nearly as well as the saturated model (model 5). Dropping any of the two two-way interactions with Category causes a noticeable loss in model fit (i.e., model 2 and 3). A closer look at the relevant partial associations, between Category and Age, and Category and Language Background, shows that the Age differences are more prominent than the Language Background differences (χ²(3) = 631.3, p < .001 and χ²(3) = 233.4, p < .001 respectively. Not shown in the table). This is also indicated by the NFI: the Age by Category interaction (model 3) improves the base model fit with 57 per cent; the Language Background by Category interaction (model 2) brings about an improvement of 31 per cent.

When we zoom in on the two two-way interactions with Category, it shows that the main differences appear in the two meaning-related categories: adults’ word
associations are more meaning-related than children’s associations (whose responses also include some form-based or unclear (‘Other’) associations), and L1 speakers’ word associations are meaning-related more so than those of L2 speakers which is mostly caused by the L2 children.

3.3.3 Response patterns for subcategories

Similar analyses were performed for the distribution within the four main categories. However, we should bear in mind that these comparisons are relative to the number of associations within the main category (and not to a group’s grand total).

The nine direct meaning-related subcategories contain a total of 14,679 associations. Again, the three-way interaction is statistically significant ($\chi^2(8) = 21.0, p = .007$; see Table 3.5), but otherwise negligible. Two-way interactions also provide a very adequate description of the data. According to the partial associations, the major of the two two-way interactions with Category is Age by Category ($\chi^2(8) = 369.8, p < .001$), rather than Language by Category. The effects related to language background are minor. This is corroborated by the NFI indicating that the base model together with only the Category by Age interaction covers 83% of the possible improvement in fit. If we consider the parameter estimates, we see that the differences contributing most to the age effect are the relatively frequent mentioning by the children of superordinates, partonyms (part-whole related words) and functionally related words (goal-target), and their infrequent mentioning of subordinates and antonyms compared to the adults.

The analysis of the 14,610 Indirect Meaning-related associations shows a similar pattern (see Table 3.5). The three-way interaction is statistically significant; this effect is mainly caused by a relatively higher number of composite words of the L2 children. Of the two two-way interactions with Category, the Age by Category interaction is by far the largest, covering 86 % of the possible improvement from base model to saturated model. The children respond with context-dependent characteristics of stimulus words more frequently than the adults.

Relatively few word associations are Form-based, namely 234. Their distribution across the eight cells of Age by Language Background by two Form-
Based subcategories shows a significant, and according to the NFI substantial, three-way interaction effect ($\chi^2(1) = 5.1, p = .024$). However, we should be cautious interpreting this interaction because the (L2) adults had hardly any word associations in this category. For the L2 children 33 word associations were classified, all of which were Changes of affix. For the L1 children on average four-fifths of their form-based associations were Changes of affix. One fifth were clang associations. Going by the two-way interactions, Language Background seems to be more related to Category than Age is. However, because of the few adult responses, it remains uncertain whether this possible Language Background effect is mediated by Age, as the three-way interaction suggests.
Table 3.5 Model fit for subcategories

<table>
<thead>
<tr>
<th>Model (Direct Meaning-related)</th>
<th>$\chi^2$/df</th>
<th>p</th>
<th>Δ NFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. [A<em>L][C</em>L]</td>
<td>410.8/16</td>
<td>&lt;.001</td>
<td>.07</td>
</tr>
<tr>
<td>3. [A<em>L][C</em>A]</td>
<td>77.1/16</td>
<td>&lt;.001</td>
<td>.83</td>
</tr>
<tr>
<td>4. [A<em>L][C</em>A][C*L]</td>
<td>21.0/8</td>
<td>.007</td>
<td>.95</td>
</tr>
<tr>
<td>5. [C<em>A</em>L]</td>
<td>0/0</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model (Indirect Meaning-related)</th>
<th>$\chi^2$/df</th>
<th>p</th>
<th>Δ NFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. [A<em>L][C</em>L]</td>
<td>178.2/4</td>
<td>&lt;.001</td>
<td>.05</td>
</tr>
<tr>
<td>4. [A<em>L][C</em>A][C*L]</td>
<td>20.3/2</td>
<td>&lt;.001</td>
<td>.89</td>
</tr>
<tr>
<td>5. [C<em>A</em>L]</td>
<td>0/0</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model (Form-based)</th>
<th>$\chi^2$/df</th>
<th>p</th>
<th>Δ NFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. [A<em>L][C</em>A][C*L]</td>
<td>5.1/1</td>
<td>.024</td>
<td>.67</td>
</tr>
<tr>
<td>5. [C<em>A</em>L]</td>
<td>0/0</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model (Other)</th>
<th>$\chi^2$/df</th>
<th>p</th>
<th>Δ NFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. [A<em>L][C</em>A]</td>
<td>56.0/4</td>
<td>&lt;.001</td>
<td>.06</td>
</tr>
<tr>
<td>4. [A<em>L][C</em>A][C*L]</td>
<td>0.9/2</td>
<td>.623</td>
<td>.98</td>
</tr>
<tr>
<td>5. [C<em>A</em>L]</td>
<td>0/0</td>
<td>1.00</td>
<td>1</td>
</tr>
</tbody>
</table>

* Due to the relatively small number of observations in the Form-based category some cell frequencies were very low (0 or 1), causing some inconsistency in the Pearson $\chi^2$ for nested models. Therefore, the log-likelihood ratio $\chi^2$ was computed for this subcategory.
The three ‘Other’ subcategories included a total of 1747 word associations (only 5.6%). It is remarkable that the three-way interaction is clearly non-significant. The partial associations support what is indicated by the NFI: the Language Background by Category interaction has more effect than the Age by Category interaction; both are significant ($\chi^2(2) = 59.6, p < .001$ and $\chi^2(2) = 9.5, p = .008$). The relatively small number of associations of the adults should make us cautious again. The adults did not repeat any of the stimulus words whereas the children did. The non-responses were most frequent among the L2 participants, which might indicate that those stimulus words were not familiar enough to them.

In sum, the analyses show that all respondents give mostly meaning-related associations: L2 children less than L1 children and L2 adults clearly more Indirect Meaning-related associations (Table 3.2). Responses of the children and of the L1 speakers are more dispersed than of adults and of L2 speakers. The in-depth log linear analyses show that the distributional patterns of association responses differ among the groups. The age groups differ most. The adults give more meaning-related responses than the children: children give some Form-based (e.g., changes of affix) and Other responses while adults hardly do. The L1 speakers give more meaning-related responses than the L2 speakers: this is mostly caused by the L2 children’s relatively high number of Form-based and Other word associations. Within the Direct Meaning-related category, the adults tend to respond more with subordinates and antonyms, while the children prefer partonyms, functionally related words and superordinates. Within the Indirect Meaning-related category, children mention context-dependent characteristics more often than adults. The few Form-based associations were mostly given by the children; the few Other associations were mostly given by the L2 children.

3.4 Discussion and conclusion

In this study, we compared word associations of Dutch child and adult monolinguals and bilinguals with Turkish as their L1. We looked at differences between language and age groups with respect to degree of dispersion of word associations (research question 1) and distributional patterns of responses across response categories.
The general patterns of responses are the same for L1 and L2 children and adults: most associations are meaning-related. Differences between the groups of speakers are generally small. Adults’ responses are more meaning-related than children’s responses: this is mainly due to the relatively many ‘Other’ associations of the L2 children. This may be because of their possible unfamiliarity with some of the stimulus words. However, in the retrospective interviews children only incidentally reported not knowing a word. That may also contribute to the fact that the L2 children give less meaning-related responses than their L1 peers, both in the Direct and Indirect category. Interestingly, the adult respondents have a preference for Indirect Meaning-related responses. Because the adults do not mainly give associations of the most abstract, conceptually meaningful sort, we cannot consider Form-based, Indirect Meaning-related and Direct Meaning-related as a scale of word knowledge development (see Namei, 2004). However, this preference for less abstract associations need not surprise us if we consider that factors other than semantic, conceptual development (for example, language use) may cause certain associations, not necessarily Direct Meaning-related, to become more prominent in the mental lexicon and to be triggered in a spontaneous association task (see Nelson et al., 2000; Wettler et al., 2005, and also Nissen & Henriksen, 2006).

The effect of language background is relatively small when we look at response patterns within the two Meaning-related main categories. Age is the dominant factor, which may imply that conceptual development is more important, both in L1 and L2: adult L2 learners are most likely to resort to their semantic or conceptual knowledge developed in the L1. In other words, this semantic knowledge may not be very language specific or, at least, it is available for both L1 and L2 associations, provided that the L2 learners are familiar with the corresponding L1 word forms. Age as well as language background are relevant factors in comparing responses within the Form-based and Other categories. Most of these types of associations are given by children; in the Other category mostly by the bilingual children. The differences between monolingual and bilingual speakers may be related to familiarity with the stimulus words. If the stimulus word is unfamiliar to a
participant, he or she has no or not many semantic or conceptual links between that word and other words and may resort to morphologically related words (Form-based), repetitions or non-responses (Other). In those cases, the word association task behaves as a common vocabulary test, which runs counter to the usual purpose of word association tasks, namely to investigate what kind of knowledge respondents have.

The variation across categories differs significantly between L1 and L2 speakers, especially in the adult group, but the effect is rather small. Also, children’s responses are more spread out across categories than adults’ responses; adults hardly give form-based or unclear (Other) associations. As was recently pointed out by Fitzpatrick (2006), monolingual responses are not as homogeneous as is often assumed. From an assessment perspective, the idea of homogeneous responses is problematic. Using word associations in testing presumes some underlying scale ranging from 'poor' to 'good'. However, the operationalization of this presumed scale is difficult, if not impossible. Clang and other form-based associations are often considered indicators of poor development of the mental lexicon. (Direct) meaning-related associations are considered indicative of well-developed word knowledge. The discussion about homogeneity of responses usually refers to the homogeneity across (L1) speakers. In this way, homogeneity is a feature of a group rather than of an individual participant. In contrast, if we want to apply the concept of homogeneity to participants, a participant should stick to a few categories of abstract direct-meaning relations (as would be shown by a low Gini concentration index). But a less proficient language learner only responding with one type of form-based association will also score low on the concentration index. The question is whether this is what we mean by homogeneous responses: probably not. Another simple solution would be just to count up the number of ‘good’ associations, for example the number of superordinates. In this case, presented with the stimulus word black, the common association white (a coordinate) would not be scored as ‘good’, whereas color (a superordinate) would. The common association made by adult native speakers would thus remain unrewarded. In our view this illustrates that responses to word association tasks cannot be measured along a certain ruler of
increasing quality or meaning-relatedness, and as such, free word association tasks do not fit well in the paradigm of (language) proficiency testing (see Kruse et al., 1987). Responses are indicative of (recent) exposure rather than ‘good’ or ‘not so good’ word knowledge, which in itself can be interesting for research. To assess participants’ word knowledge, free word association tasks should be less ‘free’. If the instruction is: “what is the first word that comes to your mind when you see/hear black?” the participant cannot do anything but respond with the first word, be it white, purple or hair, although she may well know that color is the superordinate. If researchers want to tap into the latter kind of knowledge they had better use a more controlled procedure, such as asking for superordinates (e.g., in a definition task) or testing whether participants recognize related words from a limited set of alternatives (see formats developed by Meara and Wolter, 2004; Read, 1993; and Schoonen and Verhallen, 2008).

Limitations

We should bear in mind that this study has some limitations. Respondents should ideally be tested in both their L1 and their L2 to be able to make more conclusive claims about conceptual development and more profoundly explore parallel development in speakers’ L1 and L2. Furthermore, to really know whether differences disappear with age, a longitudinal design would be valuable.

Instead of having respondents write down their responses in a written association task, researchers in subsequent association studies would do well to record spoken responses (Van Hell & De Groot, 1998). A spoken association task, using a voice key for example, is likely to tap into relations between words in the mind in a more direct way, once the intervening writing process is taken out. Moreover, writing down responses could be an obstacle for L2 learners. Additionally, response latencies in an online design could give valuable information about the automaticity of (L2) word processing. Here, the Revised Hierarchical Model would predict longer response latencies for (beginning) L2 speakers, because for them, conceptual knowledge is accessed through the L1 lexicon rather than directly through L2.
The selection of stimulus words for the present study was determined by the goals of a larger research project. All words occurred in a word corpus of primary school language and did so with various frequencies. A precondition for valid word associations is that the words are, at least superficially, known to respondents. The somewhat larger percentages of Other responses for the L2 groups suggest that this was not always the case. An investigation of the Other responses shows that they were distributed rather randomly and were not concentrated around any words in particular. To further explore the relationship between L1 and L2 knowledge in the bilingual group, stricter word selection could be helpful (see Fitzpatrick, 2007). A post-hoc screening of the stimulus words displaying the largest differences in associative behavior between Dutch and Turkish participants did not point to specific word features that might be linked to semantic or translation differences. In future work, the inclusion of stimulus words with different meaning extensions in Dutch and Turkish could provide interesting insights into cross-linguistic association behavior.

Finally, the focus of this study is on word associations of monolingual Dutch and bilingual Turkish–Dutch speakers. We restricted our bilingual sample to speakers with a homogeneous L1 language background (Turkish). This restriction makes the bilingual group limited in size and in L1. Furthermore, the adult bilingual group is rather small, as it proved relatively difficult to find enough speakers in that group. This limits the generalizability of the results, especially those concerning the adult bilinguals. Replication with a larger group of bilingual adults could corroborate our findings.

Our results show that there are small differences between Dutch child and adult monolingual speakers and bilingual speakers with Turkish as their L1 with respect to the amount of dispersion of their word associations across categories: responses of children and of L1 speakers are more dispersed than of adults and of L2 speakers (research question 1). The issue of homogeneity of responses between and within subjects underscores the ‘problematic’ status of the free word association task as an assessment tool (for word knowledge). In the same vein, the results show that Direct Meaning-related associations are not necessarily ‘optimal’ or ‘proficient’
word associations, assuming that L1 adults are the most proficient group (see Nissen & Henriksen, 2006). The lack of such an ‘optimum’ further problematizes the role of word associations in assessment. In spite of these ‘problems’, free word association tasks can give insight into a language user’s current state of mind, that is, the strength of certain word relations, irrespective of the origin of the word associations. Our different findings for monolingual and bilingual children, and for monolingual and bilingual adults in distributional patterns across categories show the influence of bilingualism on the development of word associations (research question 1 and 2). At the same time, the prominent effect of age on meaning-based associations emphasizes the role of conceptual development in word association behaviour: the children did not show the adults’ preference for Indirect Meaning-related responses. Also, the adult data did not show as many Form-based responses as given by the children. With maturity, these shallow associations seem to ‘disappear’. Whether children’s word knowledge develops from episodically structured into conceptually structured or not, is an interesting question, but deserves an investigation using more pre-structured, receptive paradigms.
Chapter 4

The role of accessibility of semantic word knowledge in monolingual and bilingual fifth-grade reading

In this chapter we will investigate differences between monolingual and bilingual minority children in availability of semantic word knowledge (research question 3) and in speed of access to such knowledge (research question 4). We assess to what extent lexical-semantic differences can help explain differences between children in reading comprehension (research question 6). The chapter starts with a discussion of the multidimensionality of word knowledge and relevant previous research on vocabulary knowledge and reading. Sections 4.1.4 and 4.2 outline the present study and the methodology used. The results are presented in section 4.3 and discussed in section 4.4.

4.1 Background

Acquiring knowledge of words is an important aspect of language learning. Young children seem to acquire lexical knowledge very rapidly, and even children who have been raised in another language than the language used at school seem to expand their vocabulary rather quickly. Lexical knowledge is one of the key factors

5 This chapter is an adapted version of Cremer & Schoonen (2012). The role of accessibility of semantic word knowledge in monolingual and bilingual fifth-grade reading. Applied Psycholinguistics. Published online on 10 August 2012. http://dx.doi.org/10.1017/S0142716412000203.
in reading comprehension (Tannenbaum, Torgesen, & Wagner, 2006), which in turn is fundamental to school success (Carlo et al., 2004; Vermeer, 2001). Several studies have shown that bilingual children who receive less input at home in the society’s dominant language lag behind their monolingual peers in lexical knowledge. Differences have been reported for vocabulary size (Appel & Vermeer, 1998; August, Carlo, Dressler, & Snow, 2005; Verhoeven & Vermeer, 1993). For example, Umbel and colleagues found that first-grade, bilingual Hispanic children in the US have poorer receptive knowledge of second language words than their (English) L1-speaking peers (Umbel, Pearson, Fernandez, & Oller, 1992).

Striking differences have also been reported for the kind of knowledge monolingual and bilingual children have of seemingly familiar words. Verhallen (1994; Verhallen & Schoonen 1993) compared 9-11-year-old Turkish-Dutch and monolingual Dutch children. She investigated their semantic word knowledge as reflected by meaning assignment to words in an extended definition task. She found that the bilingual children consistently attribute fewer meaning aspects even to very simple Dutch words and that their knowledge is also more context-specific. Data from fourth-grade Spanish-speaking and English-only children from four schools in the US corroborate that English language learners not only know fewer English vocabulary words but that their knowledge of word meaning is lacking as well (August et al., 1999). Two tasks examined children’s understanding of the multiple meanings of words in both production and comprehension. The data showed a clear gap in scores to the advantage of the monolingual children.

Together with differences in lexical knowledge significant delays in reading comprehension skills of bilingual children are signalled despite their having received the same education as monolingual children. In a review of research August and colleagues report that children speaking English as a second language who experience slow vocabulary development are less able to comprehend text at grade level than their English-only peers (August, Carlo, Dressler & Snow, 2005). They mention large and persistent gaps between the reading comprehension of language-minority and English-only children as a result (see also Proctor, Carlo, August & Snow, 2005). In a longitudinal study with Hispanic English language learners
Nakamoto and colleagues found that the children’s reading comprehension but not their word decoding began to fall behind the native English-speaking sample, starting in the third grade (~ nine years old) (Nakamoto, Lindsey, & Manis, 2007). Smits and Aarnoutse (1997) also found that throughout primary education bilingual children do not fail in decoding skills such as spelling and word reading, but lag behind in comprehension skills such as reading comprehension and vocabulary. Word decoding and reading comprehension have indeed been shown to become less correlated as children progress through elementary school (Curtis, 1980; Sticht & James, 1984; Verhoeven, 1990), but at the same time early word reading is a good predictor of later reading comprehension (Mancilla-Martinez & Lesaux, 2010; Nakamoto, Lindsey & Manis, 2008). More recently, Droop and Verhoeven showed significant delays in reading comprehension but not decoding for a group of third and fourth grade minority children when compared to monolingual Dutch children (2003). Finally, in a large-scale study in the Netherlands on word knowledge, reading and writing literacy, Verhoeven en Vermeer (2006) found Mediterranean minority children’s competence in meaning assignment to written words increasingly lags behind those of their native Dutch peers with a delay of two to three school years by grade four (~ ten years old). While decoding differences between children were small, they found major delays in reading skills of ethnic minority groups in comparison to their native peers.

Delays in reading comprehension and vocabulary are extremely difficult to overcome. Aarnoutse and colleagues (Aarnoutse, Van Leeuwe, Voeten, & Oud, 2001) report a two-year delay between Dutch proficient and less-proficient reading comprehenders that persists throughout elementary school. In a study conducted in Canada, Farnia and Geva (2007) show that even after years of schooling ethnic minority children do not catch up with their monolingual age mates. In line with this is the finding that early gaps in word knowledge persist throughout elementary school (Biemiller, 2005). Although the connection between children’s lexical knowledge and their reading skill has repeatedly been shown (Nakamoto et al., 2008; Proctor et al., 2005), it is unclear exactly which aspects of lexical knowledge play a supportive role in reading comprehension.
In sum, differences in lexical knowledge between monolingual and bilingual children are not restricted to vocabulary size but also hold for knowledge of word meaning. Reported delays in lexical knowledge run parallel to delays in reading comprehension. This raises the question to what extent individual differences in lexical-semantic knowledge and processing can explain individual differences in reading comprehension. To understand the relationship between these differences, a closer look at the acquisition of lexical knowledge and some of the processes involved is in place.

4.1.1 Multidimensionality of word knowledge

Although lexical knowledge is most commonly thought of and assessed as the number of words known, or breadth of vocabulary, it is now increasingly clear that depth or richness of the representation of the individual words known is also a crucial dimension of variability (Ordóñez, Carlo, Snow, & McLaughlin, 2002; Wesche & Paribakht, 1996). As was pointed out in Chapter 1, depth of word knowledge comprises several dimensions referring to the levels that any lexical representation consists of, such as phonological, semantic, syntactic, and – in written language – orthographic information (Nagy & Scott, 2000; Nation, 1990; Read, 2004). All these aspects can be known to varying degrees. Although different frameworks distinguish different aspects of lexical knowledge or ability (Aitchison, 2003; Chapelle, 1994; Henriksen, 1999; Laufer, 1998; Nation, 1990), virtually all approaches concur in stressing the distinction between breadth and depth. In addition, Chapelle’s (1994) framework includes a processing component, needed to gain access to lexical knowledge. Similarly, Nagy and Scott define knowing a word as being able to “recognize it in connected speech or in print, to access its meaning, to pronounce it – and to be able to do these things within a fraction of a second” (2000: 273). In this thesis depth refers only to the semantic level, i.e. to knowledge of word meaning, and not, for example, to syntactic knowledge. Accessibility of semantic word knowledge is considered a separate component.

The recognition of a breadth and a depth dimension of lexical knowledge is seen in models of the mental lexicon. Following Levelt, Roelofs and Meyer (1999),
vocabulary storage involves lexical representations of the stored phonology or sound patterns of words within the lexicon, along with semantic representations of word meaning. Current (multilingual) models of lexical access such as the adapted Bilingual Interactive Activation (BIA+) model (Dijkstra & Van Heuven, 1998; 2002) and the Revised Hierarchical Model (RHM) (Kroll & Stewart, 1994) also distinguish a lexical (form) and a semantic (meaning) level. Meara and Wolter (2004) distinguish between vocabulary size and organisation in which organisation reflects the complexity of the lexical network. The models show that the mental lexicon is an active system: new links are perpetually being formed with implications for the organisation of the whole network. From this follows that the development of breadth and depth of word knowledge is interrelated, and that words do not have meanings in isolation.

4.1.2 Semantic versus contextual word knowledge

The acquisition of word meaning is an incremental process. Nelson (1977; Nelson, 1984) proposed that young children represent information in generalised event-based scripts (e.g. ‘getting dressed in the morning’ or ‘eating lunch’) maintaining the spatial and temporal relationships between objects. Initially, word knowledge is idiosyncratic and limited to the specific contexts in which words are encountered. From this script-based knowledge, Nelson hypothesizes that children abstract and refine category information. This happens through repeated encounters with concepts, words and related concepts and words. In this way children develop their personal and sometimes idiosyncratic meanings into more conventional and so-called ‘shared meanings’ of words (Nelson, 2007). Whereas very young children have been shown to respond with contextually related associations rather than with abstract, semantic associations, a child will normally have a more or less context-independent, abstract understanding of many words some time within the five-to-nine-year-old age range (Nelson, 1977; Petrey, 1977). Thus, by the end of elementary school, around the age of eleven, children may be expected to have both a firmly rooted contextual knowledge of words and a well-developed context-independent, semantic knowledge of words.
In this study we distinguish between contextual and semantic word knowledge, a distinction that is consistent with the opposition between episodic and semantic memory used by Petrey (originating from Tulving, 1972). The distinction is similar to De Groot’s (1980) distinction between subjective versus objective semantic relations; it diverges from the much used word-class based syntagmatic-paradigmatic distinction (e.g., Cronin, 2002). The term semantic word knowledge will refer to knowledge that is inherent to the word’s meaning and not limited to any specific context a word can occur in, e.g., knowing that a dog is an animal rather than knowing that it is cute. In that sense semantic refers to more abstract knowledge or a shared meaning of words (cf. Nelson, 2007). Because cute is not part of the core meaning of dog but is only indirectly or circumstantially related to its meaning in a certain context such knowledge will be called contextual. Verhallen (1994; Verhallen & Schoonen, 1993) showed that bilingual children associated qualitatively different meaning aspects with words that their monolingual peers. Bilinguals have less rich semantic representations of simple words: they showed more contextual word knowledge (nose – to wash your nose) and less semantic, hierarchical word knowledge (nose – body part, to blow your nose) and mentioned fewer defining characteristics of words. Of course there is no absolute but a gradual difference between semantic and contextual word knowledge. The distinction simplifies the actual continuum that exists as we go from more contextually to more semantically related words. The relation between robin and bird seems to have more semantic content than the relation between wing and bird which again is more semantic than the relation between tree and bird.

We store and retrieve both semantic and contextual information about words. When encountering the word dog the semantically related word animal and the contextually related word cute may both be needed and become activated, depending on the situation, the speaker’s intention, or the task at hand. Because of its context-independent nature, it may be semantic word knowledge that is particularly useful in reading comprehension.

As knowledge of word meaning becomes more abstract, words can be used more flexibly and their meaning can be more readily accessed within multiple
contexts (Anderson & Freebody, 1981; Beck, McKeown, & Kucan, 2002). In this thesis, accessibility is used as a neutral term for speed (reflecting ease) of accessing word meaning. Accessibility does not refer to the quality of the lexical network per se. Yet it is the interconnectedness of lexical-semantic information on which accessibility builds (Perfetti & Hart, 2002). Fast access to word knowledge, be it semantic or contextual, may well facilitate comprehension. It is unclear to what extent readers differ in the accessibility of their semantic word knowledge and whether that has its own effect on reading comprehension.

4.1.3 Word knowledge and reading

The unique role of semantic word knowledge in reading comprehension has been shown repeatedly. Qian (1999; 2002) used a depth-of-vocabulary-knowledge test measuring knowledge of synonymy, polysemy and collocations and tested English L2 university students. He found that scores on the depth test uniquely contributed to the prediction of reading comprehension levels. He even concluded that depth and breadth of lexical knowledge are equally important in predicting reading comprehension. Nation and Snowling (2004) found that performance of young children on semantic tasks predicted a substantial amount of variance (15.1%) in reading comprehension when entered into separate regression models following age, nonverbal intelligence, and phonological skills. Ouellette (2006) reported a unique contribution of depth (knowledge of synonyms) of 8% of explained variance beyond vocabulary size, age, nonverbal IQ, decoding and visual word recognition in fourth grade readers. Schoonen and Verhallen (1998) found a unique contribution to reading comprehension of children’s semantic, context-independent word knowledge (3-7%) as reflected in their performance on the Word Association Task.

Yet, the exact link between semantic word knowledge and reading comprehension is not understood in a clear way even now, as a result of its complexity. Research testifies to the recurrent nature of component interactions (e.g., Beck, Perfetti, & McKeown, 1982). Word knowledge allows comprehension, comprehension allows reading practice, reading practice strengthens word knowledge, and so forth. Ease of accessing semantic word knowledge is important
as semantic word knowledge may play a role in both the identification of words (at least in non-transparent orthographies) and in comprehension (Perfetti, Landi, & Oakhill, 2009). Perfetti posits that “[t]his dual role of word meanings places lexical semantics in a pivotal position between word identification and comprehension” (2009: 241). Longitudinal studies with young Spanish-speaking learners of English have shown that both word decoding fluency and vocabulary knowledge at a younger age are predictive of later reading comprehension proficiency (Mancilla-Martinez & Lesaux, 2010; Nakamoto et al., 2008; Proctor et al., 2005). Such mechanisms are reflected in intervention studies. Beck and colleagues (2002) suggested that speed of access to word meanings may be an important factor in explaining differences in outcomes from vocabulary training studies. They showed that instruction that involves multiple repetitions helps improve the speed of accessing the word’s meaning. Studies that provided this rich vocabulary instruction showed gains in both word knowledge and comprehension of text containing the words taught (Beck, McCaslin, & McKeown, 1980; Beck & McKeown, 1983; McKeown, Beck, Omanson, & Perfetti, 1983). Vocabulary instruction that does not produce sufficient fluency of access was found to not generalise to reading comprehension (cf. Jenkins, Pany, & Schreck, 1978).

Accessing or activating words is easier when they are related. In addition to being more or less semantically related words may be associated with each other. Apart from type of association (see Chapter 3), there is also strength of association. Association strength between two words springs mainly from their common co-occurrence in language use and is reflected in association responses. The greater the association strength between two words, the easier they activate one another (Fitzpatrick, 2007). Words may be weakly associated yet semantically similar (e.g., radish-beets) and they may also be highly associated yet semantically dissimilar (e.g., coat-rack) (examples from Thompson-Schill, Kurtz, & Gabrieli, 1998). In general, word recognition studies show more co-activation between related words (priming) when words are not only semantically but also associatively related (e.g., dog-cat) as compared to purely semantically related words (e.g., apple-lemon): the so-called associative boost (Abad, Noguera, & Ortells, 2003). Controlling for
association strength between stimuli allows one to measure sensitivity to purely semantic information, for example during reading.

There is mounting psycholinguistic evidence that children who are less proficient in comprehension have problems with semantic processing (Nation & Snowling, 1998a, 1999, 2004; Ricketts, Nation, & Bishop, 2007). In a study with 10-11-year-old children, Nation and Snowling (1999) found that both proficient and less-proficient comprehenders showed priming for function-related words (e.g., *broom-floor*), but for the category co-ordinates (e.g., *aeroplane-train*), less-proficient comprehenders only showed priming if the category pairs also shared high association strength. They conclude that in the absence of explicit co-occurrence less-proficient comprehenders are less sensitive to abstract semantic relations. Furthermore, less-proficient comprehenders were found to be slower to generate semantic category members (but not rhymes) than proficient comprehenders. Less-proficient comprehenders also scored lower on a synonym judgement task (Do *boat* and *ship* mean the same thing?), although not on a rhyme judgement task (Do *rose* and *nose* rhyme?) (Nation & Snowling, 1998a). Semantic variables such as concreteness also distinguish proficient from less-proficient readers. Using recall tasks, Nation, Adams, Bowyer-Crane and Snowling (1999) found that less-proficient comprehenders showed normal sensitivity to phonological manipulations (length and lexicality) but, consistent with their semantic weaknesses, their recall of abstract words was poor. Considering that word knowledge develops from contextual to more abstract, less-proficient comprehenders’ limited sensitivity to semantic relations may hint at a delay in word knowledge development, which in itself might be due to less exposure to (academic) language. The research reviewed suggests that children’s comprehension problems may be associated with reduced semantic word knowledge and less effective semantic processing or reduced accessibility of semantic knowledge.
4.1.4 Availability and accessibility of semantic word knowledge

In the study reported in this chapter, we investigate to what extent semantic word knowledge supports reading comprehension. We teased apart two important aspects of semantic word knowledge: availability, the knowledge itself, and accessibility, the speed with which that knowledge is activated (research questions 3 and 4). This information may help better explain the existing differences between proficient and less-proficient comprehenders, and between monolingual and bilingual children (research question 6).

If indeed both the availability and accessibility of semantic word knowledge play a role in reading comprehension, we expect measurements of both constructs to be related to measurements of reading comprehension. Second, if semantic word knowledge is more easily accessible in the lexicon of monolingual than of bilingual children, monolingual children will show a stronger inclination towards semantically related words (as opposed to contextually related words) than bilingual children. Moreover, monolingual children will display shorter response times to semantically related words than bilingual children. Third, if speed of access to semantic word knowledge plays a role in reading comprehension, individual differences in speed of access ought to help predict individual differences in reading comprehension. If accessibility of semantic meaning aspects is an additional dimension of word knowledge, it will predict individual differences in reading even after individual differences in the availability of semantic knowledge have been taken into account.

4.2 Method

To investigate differences between children in reading comprehension and in accessibility of semantic word knowledge, and the relationship between these constructs, several tests were administered to monolingual and bilingual children. Performance regarding the following constructs was measured: reading comprehension, availability of semantic word knowledge (both a paper-and-pencil task and a computer task), accessibility of semantic word knowledge (reaction times in a computer task), and word decoding. A word decoding test was included as a
control variable, because word decoding fluency influences performance on the accessibility measure (reaction times). Language background was assessed in a questionnaire in order to assign children to monolingual and bilingual groups.

4.2.1 Participants

Data were collected from a total of 156 children. None of the children had taken part in the study in Chapter 3. Of the original sample, 21 children were excluded: six due to dyslexia and fifteen due to missing data on (some of) the tests. This resulted in a final sample of 135 children (66 girls and 69 boys). The children were recruited from grade five (Dutch grade seven, the pre-final year of primary education) of nine primary schools in the western part of the Netherlands: eight schools were located in cities, one in a smaller town. Socio-economic status of school populations was intermediate to low. The age ranged from 10;7 to 12;9 years old, which means that a number of children (16%) are older than is usual for this grade. However, this percentage is in line with the distribution of age in national surveys (i.e. 20%), and there was no statistically significant age difference between the two language groups (F<1). The sample consisted of children who spoke Dutch as their only language (monolingual children) and children who spoke Dutch in addition to a minority language (bilingual children). This latter group consisted of children from families with an immigrant background who spoke a range of non-European languages as their first or second language, the two main languages spoken being Moroccan (Arabic and Berber) (N=24) and Turkish (N=16). The use of Dutch at home ranged from no Dutch spoken at all to mostly Dutch. Children who spoke European languages at home (e.g., German) in addition to Dutch had been excluded. In total, twelve children were born outside the Netherlands (mean age of arrival 3;3); three children arrived to the Netherlands after Kindergarten (ages 6, 7, and 8, respectively). All participants were free from diagnosed language disorders or speech impediments. For each participant, permission to take part was obtained through the school. Table 4.1 provides an overview.
Table 4.1 Description of the participants

<table>
<thead>
<tr>
<th>Language background</th>
<th>N</th>
<th>Gender (F / M)</th>
<th>Mean age (yr; m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolingual</td>
<td>65</td>
<td>30 / 35</td>
<td>11;5 (0;5)</td>
</tr>
<tr>
<td>Bilingual</td>
<td>70</td>
<td>36 / 34</td>
<td>11;6 (0;6)</td>
</tr>
<tr>
<td>Total</td>
<td>135</td>
<td>66 / 69</td>
<td>11;6 (0;6)</td>
</tr>
</tbody>
</table>

Note. F = female, M = male, yr = year, m=months, s=standard deviation

4.2.2 Materials

Reading Comprehension

Reading comprehension (RC) was measured using a standardised Dutch reading test for children in the final years of primary education (Dutch grades six to eight), ‘Begrijpend Lezen 678’ (Aarnoutse & Kapinga, 2006). The test has been normed nationally on a sample of 42 schools. Validity and reliability are reported as satisfactory (reliability in terms of Cronbach’s $\alpha=.83$). The test correlates well with other standardised tests of reading comprehension. It comprises an answer sheet and a booklet containing seven different reading passages, ranging in length from 122 words to 288 words. Each passage is accompanied by six or seven questions: three or four multiple-choice questions and two to four true/false statements, resulting in a total of 44 questions. Four questions are word-level questions; the other questions are above word level, including literal and inferential questions and overview questions of larger scope. In this sample, the distribution of reading comprehension scores did not deviate significantly from normality: skewness was -.088 (se .209)
and kurtosis \(-.608\) (se \(.414\)). Internal consistency reliability, Cronbach’s alpha, in this sample was \(.75\).

**Semantic word knowledge**

To test the availability of semantic knowledge of words the *Word Association Task* (WAT, Schoonen & Verhallen, 2008) was administered. This forced-choice task is based on Read’s format (1993) and has been developed for children from nine to twelve years old. In this paper-and-pencil task children have to connect a stimulus word (e.g., *vegetable*) with three out of six surrounding words (see Figure 2.2). The surrounding words are either semantically related to the stimulus word (e.g., *plant, lettuce, food*) or they are more contextually related or unrelated to the stimulus word (e.g., *plate* (i.e., object), *warm, strong*). Children are asked to identify the three semantically related words by drawing lines from the stimulus word to the three words that “really belong to it”. The distinction between semantic and contextual relations is a gradual one: children have to weigh and compare relations (*vegetable - plant* vs. *vegetable - strong*), reject too indirect i.e. contextual relations, and recognise content-based, semantic relations. Selecting the semantically related words is seen as an indication of well-developed semantic word knowledge (Schoonen & Verhallen, 2008). The WAT was administered in each class. As part of the instruction two examples were worked out and explained in class. Two versions of the task were created (with a different ordering of the items) and randomly assigned to the children. An item was scored correct only if all three semantic relations had been identified. A correct item received 1 point and an incorrect item received 0 points\(^6\). WAT scores were slightly skewed: \(-.579\) (se \(.209\)); kurtosis did not deviate significantly from normality: \(.341\) (se \(.414\)). Internal consistency reliability in this sample was \(.74\).

---

\(^6\) Schoonen and Verhallen (2008) explored different scoring procedures, including a graded scoring, and found that the various scoring procedures are highly intercorrelated.
Semantic word knowledge and its accessibility

To test the accessibility of semantic word knowledge a computerised semantic decision task, C-WAT, was designed. The task is inspired by Schoonen and Verhallen’s WAT (2008) and presents participants with a stimulus word followed by two instead of six words: one semantically-related word (the target), and one contextually-related word (the distracter). The participant is required to identify the semantically-related word as quickly as possible. Response times (RTs) reflect the speed of accessing the word’s meaning. RTs to targets are recorded from the onset of the target and distracter word on the screen. RTs to distracters are discarded. Similarly to the WAT, the C-WAT also records response accuracy (whether the semantic or contextual relation is chosen) reflecting availability of semantic word knowledge. Because speed of access to semantic word knowledge (RTs) is the primary focus of this task, response accuracy needs to be high in order to get enough valid RT data. Thus, the task was made relatively easy, compromising the accuracy measure.

As in the WAT, the distinction between targets and distracters is a gradual one. Targets are related to stimulus words either through hyponymy (flower-tulip), hypernymy (apple-fruit), synonymy (steal-rob) or meronymy (candle-wick), or they are a defining characteristic of the stimulus word (cigar-tobacco). Contextually-related words (distracters) are not directly related in meaning to stimulus words. However, for distracter words to be appealing they had to be associated to the stimulus words as well. To control for association strength, C-WAT stimulus words were selected from the stimuli of the word association task reported on in Chapter 3. The word association norms collected in that study guided the selecting of targets and distracters. The stimulus words chosen for the C-WAT all had strong associates that were not semantically related to it; these associates were the distracters. In addition, the association norms were used to ensure that the target words were not (strong) associates of the stimulus words. In all but four cases did targets not occur as association response. In these four cases, association strength between stimulus words and distracters was greater than association strength between stimulus words and targets. This procedure resulted in 69 items consisting of a stimulus word, a
semantically-related, non-associated target and an associated, contextually-related distracter. The test consisted of 69 test items preceded by six practice items.

For the C-WAT stimulus words, the high frequent stimuli from the association task were selected as much as possible to ensure children’s familiarity with them. Overall word length of targets and distracters was equal. Word class of targets and distracters was kept as balanced as possible. Item order was randomised; target position (left or right, see Figure 4.1) was balanced in the task and across participants. The C-WAT was piloted with several adults and children and was found to work satisfactorily.

Figure 4.1 C-WAT example item

The target fruit is inherently related to the meaning of the stimulus word apple, irrespective of context. The distracter healthy is contextually related to apple, but the two are strongly associated.

Item analyses showed that some items received accuracy scores not significantly different from chance, which had not been the case in the pilot. Because of this, the stimulus set was re-evaluated retrospectively by a larger group of eighteen adults. Fourteen items were identified as not unanimously agreed upon by the adults and were excluded. This is rather strict, as we wanted to be confident that the correct answers were clear. The resulting item set consisted of (69-14=) 55 items and was used for further analysis. The distributions of accuracy scores and mean RT scores (for 55 items) did not deviate significantly from normality: For
accuracy skewness was -.013 (se .209) and kurtosis -.779 (se .414); for mean RT scores skewness was .210 (se .209) and kurtosis -.297 (se .414). It is common practice to transform RT scores logarithmically or to take the inverse (Ratcliff, 1993) to avoid problems with the assumption of normal distributions. However, given the current distribution of our RT scores, we refrained from transforming the RT scores. Checks showed that the RT scores correlated between .89 and .99 with the transformed scores. In the final analyses these effects were only noticeable in the third decimal. Internal consistency reliability of the accuracy data was .52. This low reliability is largely due to the relative easiness of the test. The reliability as expressed by Cronbach’s α of the 55 RT scores was .93.

**Word decoding**

A standardised test of decoding skill was administered: the ‘Drie Minuten Test’ (Verhoeven, 1992), which has a good reliability and validity. Children were asked to read two word lists out loud. The first list consisted of 150 monosyllabic existing Dutch words (CCVCC); the second list consisted of 120 multisyllabic words: 60 two-, 30 three-, and 30 four-syllable words. There is a break between the two lists. For each list the score is the number of words read correctly in one minute. The two measurements in this study are strongly correlated (Pearson’s r = .829) and show a high split-half reliability of .905; therefore, the two list scores are combined (summed) to derive a single score for word decoding which is used in further analyses. The distribution of the observed word decoding scores deviates somewhat from normality: skewness is -.653 (se .209) and kurtosis 1.735 (se .414).

**Language background**

Participants’ language background was determined by administering a questionnaire on language use at home and at school outside the classroom and by conducting an interview with each participant and with the teachers. In the individual interviews

---

7 The test consists of three lists. We decided to omit the simple list (containing only VC, CV and CVC words) as it is very easy for 11-year-old children, and to have participants thus read two instead of three lists.
participants were asked about language use at home and outside school (i.e. which language is spoken to parents, siblings, other family members, and friends). Through the questionnaire background information was also obtained about such things as country and date of birth, and time of residence in the Netherlands.

4.2.3 Procedure

Children were tested in their schools, on a regular school day. The WAT and reading comprehension task were administered in class; C-WAT and Word Decoding were done individually. Testing in class started in the morning and lasted approximately 45 minutes. Individual sessions took approximately 25 minutes. Individual testing was done by one of two trained test assistants or the principle investigator. Instructions followed a strict protocol. Tasks were administered to all children in the same order. Testing started with the WAT. When all children had finished, the reading comprehension task was administered. Children who had finished the second task were asked to go with one of the test administrators to take the two individual tasks. The individual session started with the word decoding task, which lasted about three minutes. Next, the C-WAT was administered. C-WAT data were collected on two identical laptop computers; the task was run in E-prime Version 1.1 (Schneider, Eschman, & Zuccolotto, 2002). Before starting the C-WAT, the test assistants explained the task by means of five example items on paper boards. The assistants used the wording of the instruction as given in the test (see below) as well as various paraphrases. The importance of responding as quickly but accurately as possible was stressed. The task began with an on-screen instruction and six practice items. On-screen instruction was the following: “You will see a +, followed by a word. Below that word, two words will appear. Of the two words, choose the word that really belongs to the word at the top8”. A fixation cross appeared on the screen for 500 ms and was followed by a stimulus word. After 800 ms the target and the distracter appeared side by side under the stimulus word: the three words on the screen formed a triangle (see Figure 4.1). This triad stayed on the screen until a response was made by pressing either the m (for the word to the right) or the z (for

---

8 Note that all instructions and examples are translations of the original Dutch test.
the word to the left) key on the keyboard, which are located directly above the space bar, to the right and left respectively. If no response was detected after 10000ms, the screen went blank. The inter-item interval was 2000ms. Time settings were based on pilot runs with children. After the child had completed the C-WAT, the participant questionnaire was administered. As part of this questionnaire the researchers conducted a short interview about language dominance.

### 4.2.4 Analyses

Item analyses for the WAT and Reading Comprehension test did not result in changes to the data set. The word decoding test yields the number of words read; therefore, no item analysis was done. The C-WAT data set was examined for outliers and missing data (i.e. no response detected). We adopted a lenient criterion for outliers to prevent the removal of too many (valid) RTs. Outliers were defined as RTs that were 3 SD units above or below the item mean. Missing data constituted only 0.5% of the data. The percentage of incorrect responses was 19%. Outliers and RTs for incorrect trials were replaced by missings. Fifteen out of 150 participants were excluded from the analyses, because they had 30% or more missing RT data. In the RT data set missing data were imputed with R (Van Buuren & Oudshoorn, 1999). This was done five times. Means were calculated in the five imputed data sets and were averaged into an overall mean RT score for each participant.

After psychometric properties of the tests had been established, analysis of variance procedures were used to test for overall group differences. Correlations between measures were computed and hierarchical regression analyses were used to identify which of the measures best account for variance in reading comprehension. A path analysis confirms the final regression model as an adequate description of the data.

### 4.3 Results

#### 4.3.1 Descriptives

In Table 4.2 the means, standard deviations and reliabilities for the tests of reading comprehension, availability of semantic word knowledge (WAT and C-WAT),
speed of access to semantic word knowledge (C-WAT), and word decoding are presented. Performance in none of the tests reached ceiling. All tests were reliable, with the exception of the C-WAT accuracy measure. This measure had little variance as items that were too difficult had been removed from the task to ensure high accuracy performance and to allow for a reliable measure of reaction times (see Materials).

<table>
<thead>
<tr>
<th>Measures</th>
<th>Items</th>
<th>M</th>
<th>SD</th>
<th>Rel.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Comprehension</td>
<td>44</td>
<td>30.2</td>
<td>5.5</td>
<td>.75</td>
</tr>
<tr>
<td>WAT (Word knowledge)</td>
<td>30</td>
<td>20.1</td>
<td>4.5</td>
<td>.74</td>
</tr>
<tr>
<td>C-WAT Accuracy</td>
<td>55</td>
<td>47.3</td>
<td>3.5</td>
<td>.52</td>
</tr>
<tr>
<td>C-WAT RT</td>
<td>55</td>
<td>2362.1</td>
<td>499.2</td>
<td>.93</td>
</tr>
<tr>
<td>Word Decoding</td>
<td>270</td>
<td>169.5</td>
<td>26.0</td>
<td>.91</td>
</tr>
</tbody>
</table>

Note: Items = number of test items; M = mean; SD = standard deviation; Rel = reliability (Cronbach’s alpha, and for Word Decoding split-half reliability).

4.3.2 Monolingual and bilingual children

Analyses of variance (ANOVAs) revealed significant group differences for measures of reading comprehension and availability of semantic word knowledge (WAT). Monolingual children performed significantly better at both the reading comprehension task, F(1, 133) = 4.201, \( p = .042 \), effect size partial eta squared (\( \eta^2_p \)) is .031, and at the written availability task (WAT), F(1,133) = 16.608, \( p = .000 \), \( \eta^2_p \) = .111. Differences in availability as measured in the C-WAT (accuracy scores) were not statistically significant, F(1,133) = 2.907, \( p = .091 \), \( \eta^2_p \) = .021. Regarding the accessibility measure, Table 4.3 shows that monolingual children seem to have been faster at identifying semantically related words, but this difference is not statistically significant, F(1,133) = 0.649, \( p = .422 \), \( \eta^2_p \) = .005, probably due to the relatively large individual differences (see standard deviations). For word decoding,
scores on the combined measure showed that the bilingual group outperformed the monolingual group, but this difference is not statistically significant, \( F(1,133) = 1.841, p = .177, \eta_p^2 = .014 \). As was mentioned above, word decoding speed may affect the accessibility RT measure and hence word decoding differences could possibly mask accessibility differences. However, an ANCOVA of the accessibility scores with word decoding as a covariate still shows no significant difference between the two groups, \( F(1,132) = 1.777, p = .185, \eta_p^2 = .013 \). The effect size, \( \eta_p^2 \), is slightly more than the typical small effect according to Cohen (1988).

Table 4.3 Means and standard deviations for monolingual and bilingual children and effect sizes

<table>
<thead>
<tr>
<th>Measures</th>
<th>Monolingual (N=65)</th>
<th>Bilingual (N=70)</th>
<th>( \eta_p^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( M )</td>
<td>( SD )</td>
<td>( M )</td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>31.2</td>
<td>5.5</td>
<td>29.2</td>
</tr>
<tr>
<td>WAT</td>
<td>21.7</td>
<td>4.1</td>
<td>18.7</td>
</tr>
<tr>
<td>C-WAT Accuracy</td>
<td>47.8</td>
<td>3.5</td>
<td>46.8</td>
</tr>
<tr>
<td>C-WAT RT</td>
<td>2326.1</td>
<td>524.8</td>
<td>2395.5</td>
</tr>
<tr>
<td>Word Decoding</td>
<td>166.3</td>
<td>26.5</td>
<td>172.4</td>
</tr>
</tbody>
</table>

*Note: \( \eta_p^2 \) = partial eta squared; * significant at the 0.05 level.

4.3.3 Relationships with reading comprehension

Reading comprehension scores correlate with both measures of availability of semantic word knowledge (\( r = .56 \) and \( r = .25 \), respectively; see Table 4.4). The two measures of availability are correlated, as expected (\( r = .42 \)). The correlation between reading comprehension and availability as measured in the C-WAT is relatively weak due to the low reliability of the C-WAT accuracy measure (see Descriptives). Reading comprehension is inversely correlated to speed of access to semantic word
knowledge, showing that better reading comprehension is related to faster identification of semantically related targets. Word decoding shows no significant correlation with reading comprehension. Fast word decoding is part of accessibility as is indicated by the significant (negative) correlation.

Table 4.4 Intercorrelations for all measures

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reading Comprehension</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. WAT</td>
<td>.56**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. C-WAT Accuracy</td>
<td>.25**</td>
<td>.42**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. C-WAT RT</td>
<td>-.25**</td>
<td>-.16</td>
<td>.07</td>
<td></td>
</tr>
<tr>
<td>5. Word Decoding</td>
<td>.08</td>
<td>.04</td>
<td>.05</td>
<td>-.33**</td>
</tr>
</tbody>
</table>

Note: N=135. **. Correlations significant at the 0.01 level (2-tailed).

4.3.4 Contributions to reading comprehension

A regression analysis was performed to determine whether there is a (unique) contribution of accessibility of semantic word knowledge to reading comprehension over and above the roles of availability of that knowledge and word decoding. Measures of availability of semantic word knowledge and word decoding were entered first in the regression model; accessibility of semantic word knowledge was entered in a second step. In total, 31.8% of variance in reading comprehension was accounted for by word decoding and the two measures of availability of semantic word knowledge (WAT and C-WAT accuracy) most of which was attributable to the WAT. Adding speed of access (C-WAT RT) to the model explained an additional 2.5% resulting in a total of 34.3% of variance accounted for. This increase in variance due to speed of access was small but significant given that the variance accounted for by the three other tests was already taken into account (∆R²= .025;
Table 4.5 displays the contribution of aspects of word knowledge to the prediction of reading comprehension.

### Table 4.5. Variance accounted for by aspects of word knowledge in the prediction of reading comprehension

<table>
<thead>
<tr>
<th>Model</th>
<th>Predictors</th>
<th>Total $R^2$</th>
<th>$F$ Regression $\left(\text{df}_1/\text{df}_2\right)$</th>
<th>$R^2$ Change</th>
<th>F Change $\left(\text{df}_1/\text{df}_2\right)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C-WAT Accuracy, Word Decoding</td>
<td>.318</td>
<td>20.32* (3/131)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2</td>
<td>+ C-WAT RT</td>
<td>.343</td>
<td>16.96* (4/130)</td>
<td>.025</td>
<td>5.01* (1/130)</td>
</tr>
</tbody>
</table>

*Note. * $p < .05.$

In section 4.1.4 it was suggested that the reading comprehension delays of bilingual children could be related to differences in the availability of semantic word knowledge and possibly the accessibility of this knowledge. We explored this suggestion in a regression analysis of reading comprehension. In the first step availability (WAT) and speed of access (C-WAT RTs) were entered, in the second step Language background (monolingual versus bilingual) was entered. It turned out that Language background no longer showed an effect on reading comprehension ($\Delta R^2 = .000; F_{\text{change}}(1,131)<1; p = .83$). Another regression analysis showed that the effect of Language background also disappears when in the first step only availability is entered. When Language background is entered in addition to accessibility, it can still make a small, but non-significant, contribution to reading performance.

A path analysis conducted in Lisrel (Jöreskog & Sörbom, 1996) shows that the effect of language background indirectly affects reading comprehension, that is, via children’s semantic word knowledge. In this analysis, the two measures for the
available semantic knowledge are combined, accommodating the differences in reliability of the two measures. The model as depicted in Figure 4.2 fits the data fairly well ($\chi^2(9)=11.08$, p=.27, CFI=.98, RMSEA=.039). The effect of Language background runs via children’s semantic word knowledge. An additional direct effect of language background on reading comprehension is redundant and does not significantly improve the model ($\Delta\chi^2(1)=0.24$, p=.62).
Figure 4.2 Path model for the interrelationships of word decoding, language background, availability of semantic word knowledge, speed of access to semantic word knowledge, and reading comprehension.
Subgroups

To determine whether the contribution of availability and accessibility of semantic word knowledge to reading comprehension is the same for monolinguals and bilinguals, two additional regression analyses were carried out, each with in the final step a term for the interaction between language background and one of the predictor variables. In none of the analyses did the interaction term significantly explain any additional variance in reading comprehension.

A division in subgroups can also be made in terms of reading proficiency. Children scoring at or below the median on the reading comprehension test were categorized as less-proficient readers (M=25.8, N=70, 57% bilingual); children scoring above the median were considered proficient readers (M=34.9, N=65, 46% bilingual). A multi-group analysis shows that the path model (Figure 4.2) fits the data of both proficiency groups with equal regressions across groups ($\chi^2(29)=29.41$, $p=.44$, CFI=.99, RMSEA=.00). However, the model significantly improved when the explained variance in reading was allowed to differ between the groups ($\Delta\chi^2(1)=5.12$, $p=.02$). Comparing the standardized solutions of the two groups indicated that for the proficient comprehenders, measures of availability and accessibility explain 31% of variance (the standardized regression weights being .52 and -.20 respectively), whereas for the less-proficient comprehenders the measures of availability and accessibility explain only 20% of the individual differences in reading (the standardized regression weights being .41 and -.16 respectively).

The results can be summarised as follows. For measures of reading comprehension and availability of semantic word knowledge, monolingual children outperformed bilingual children. No statistically significant group differences were found for accessibility of semantic word knowledge. Bilingual children scored higher on the word decoding test, but this difference was not significant. Reading comprehension performance was correlated to both measures of availability of semantic word

---

9 The error terms were allowed to differ between the two groups, accommodating differences in reliability. One error-term estimation was slightly negative, and was subsequently fixed at zero.
knowledge; it was inversely correlated to accessibility of semantic word knowledge, showing that better reading comprehension is related to faster identification of semantically related targets. Word decoding also showed an inverse correlation with the accessibility measure: reading more words per minute was related to faster identification of semantically related targets, as could be expected.

A regression analysis showed that the availability of semantic word knowledge, word decoding and speed of access to semantic word knowledge together explain over 34% of individual differences in reading comprehension scores. Speed of access in terms of response times accounts for (a small amount of) variance in reading comprehension that cannot be accounted for by scores on word decoding and availability of semantic word knowledge. Differences in reading comprehension between monolingual and bilingual children ‘disappear’ when we take into account individual differences in availability of semantic word knowledge. Although the general pattern of relationships holds for both proficient and less-proficient readers, multi-group analyses show that the predictive power of the lexical-semantic measures is, relatively speaking, larger for the proficient than for the less-proficient comprehenders.

4.4 Discussion and conclusion

The present study set out to assess differences in availability and accessibility of semantic word knowledge of Dutch monolingual and bilingual minority children (research questions 3 and 4). It related those differences to differences in reading comprehension to investigate some of the mechanisms underlying reading comprehension differences (research question 6).

Research shows that superficially knowing frequent words is not enough to fully comprehend texts: it is also important to have semantic knowledge of those seemingly familiar words (Qian, 1999; Schoonen & Verhallen, 1998). In this study, we corroborated that availability of semantic knowledge is an important factor in reading proficiency. Moreover, the ability to access that semantic knowledge quickly as proposed by Nagy & Scott (2000) and Perfetti & Hart (2002) may have added value. The importance of availability and to a small extent accessibility of
semantic word knowledge for reading comprehension is evidenced by their correlation with and their distinguishable contribution to reading comprehension differences (research question 6). Both lexical-semantic knowledge (availability) and fluency (accessibility) seem to play a role in comprehending text: having semantic knowledge about words and being able to activate that knowledge quickly. As such, reading comprehension is supported by accessibility as a separate, measurable component. However, the importance of fast access of semantic word knowledge seems to be limited in the sample of children we studied.

The finding that bilingual children lag behind their monolingual peers in reading comprehension and in availability of semantic word knowledge (WAT scores) (research question 3) confirms earlier findings (August et al., 1999; Schoonen & Verhallen, 2008). The bilingual children have less semantic knowledge about words available than their monolingual peers. However, no significant differences between the two groups were found for accessibility of semantic word knowledge (research question 4). A path analysis showed that group differences in reading could be perceived as differences mediated by semantic knowledge differences. Controlling for semantic knowledge made the language background differences ‘disappear’. Evaluating a potential difference in predictive power of availability and accessibility in each group, we did not find significant interaction effects of language background and the semantic variables.

Furthermore, we compared path models for proficient and less proficient readers. This comparison demonstrated that, in absolute terms, regression weights for availability and accessibility of semantic knowledge were equal, but that due to a difference in the variance of reading comprehension, the standardized regression weights were different for the two reader groups. For the proficient comprehenders, these parameters were slightly higher. The less proficient group was somewhat more heterogeneous in reading proficiency. This larger variance in reading might be due to factors other than the ones we have measured, such as comprehension monitoring or inferencing skills.

Although there were no statistical differences for word decoding, the bilingual children were found to outperform monolingual children slightly
For monosyllabic words bilingual children were significantly faster than their monolingual peers, but not for multisyllabic, orthographically more complex words. This is in line with previous research reporting comparable or faster decoding by bilingual minority children. Droop and Verhoeven (2003) tested children across grade three and four: bilingual minority low-SES children decoded at the same level as Dutch low-SES children, for multisyllabic words. For monosyllabic (CVCC) words, the bilingual minority and the monolingual Dutch children obtained comparable scores in grade three, but the bilingual minority children obtained higher scores in grade four. For monosyllabic (CVC) words bilingual minority children obtained higher scores in both grades (2003: 86). An explanation for this finding may be that bilingual children read the words more superficially and hence faster. Such differences warrant our inclusion of word decoding as a control variable. However, the predictive power of word decoding for reading success at a later age (Mancilla-Martinez & Lesaux, 2010; Proctor et al., 2005) does not seem to pay off for the bilingual children, since their decoding fluency does not result in high levels of reading proficiency. As Nakamoto et al., (2008) showed, it is the interaction between decoding and ‘oral language’ (including vocabulary knowledge) that is beneficial to reading. Good word decoding skills might be a prerequisite to be able to benefit from vocabulary knowledge while reading.

The distinguishable relations that availability and accessibility of semantic word knowledge have with reading comprehension strengthens the importance of depth as a theoretical construct in its own right, next to a dimension such as breadth (Schoonen & Verhallen, 1998; Qian, 1999, 2002; Ouellette, 2006). In addition to the unique contribution of availability of semantic word knowledge, the present study indicates the unique contribution of accessibility of semantic word knowledge. Altogether this seems to add up to what Daller, Milton and Treffers-Daller (2007) call the ‘lexical space’, a three dimensional concept of vocabulary size (breadth), depth of word knowledge (availability) and fluency (accessibility). Distinguishing the three dimensions is sometimes difficult, possibly because the three dimensions probably grow at a similar rate due to language exposure.
Additional data from longitudinal research and training studies are necessary to explore how these dimensions relate to each other and might be affected by language input and eventually contribute to children’s reading proficiency. One example is training studies focusing on quality of word knowledge (e.g., Carlo et al., 2004) and on access to word relations the relevance of which was already suggested by Beck and colleagues (2002). Furthermore, future studies testing other modalities could provide insights. For example, Betjemann and Keenan (2008) tested children with reading problems using an auditory measure of semantic word knowledge. Their results suggest that semantic deficits are not confined to reading but are a more general comprehension problem. Finally, psycholinguistic studies using subconscious paradigms such as (masked) semantic priming are needed to firmly establish the role of accessibility of semantic word knowledge in reading comprehension.

In conclusion, our data show that semantic word knowledge and to a small extent fast access to it underlie differences in reading comprehension. This supports the pivotal role of semantics between word identification and comprehension mentioned in section 4.1.3. Easy access to semantic word knowledge seems to allow that knowledge to be used efficiently in higher order reading processes such as monitoring, integrating and inferencing information. However, more research is needed to determine to what extent, and how, the relations between semantic word knowledge, fast access to semantic word knowledge and reading comprehension can be interpreted in terms of truly causal mechanisms.
Chapter 5

Processing semantic relations: word knowledge and reading comprehension in monolingual and bilingual children

The findings of Chapter 4 identified speed of access as a factor in children’s reading performance. In this chapter, we will again address availability of semantic word knowledge (research question 3) and speed of access (research question 4). Speed of access will be investigated in terms of lexical decision speed and semantic classification speed. In addition, we will go one step further and investigate the activation of abstract, semantic word knowledge as reflected by semantic priming (research question 5). Finally, the contributions of semantic word knowledge, speed and priming to reading comprehension will be analysed (research question 6). The chapter starts with a discussion of relevant previous research (section 5.1), followed by a discussion of the research methodology used (section 5.2). The results are presented in section 5.3 and discussed in section 5.4.

5.1 Background
Although there is evidence for a close link between children’s reading comprehension and their semantic word knowledge (Nation & Snowling, 2004; Ouellette, 2006; Proctor, Uccelli, Dalton, & Snow, 2009), the mechanisms underlying this relationship are less obvious. How do we process or compute the

---

10 This chapter is an adapted version of Cremer, M. & Schoonen, R. Processing semantic relations: word knowledge and reading comprehension in monolingual and bilingual children. Manuscript submitted.
meaning of the words we read? Subtle differences in people’s understanding of a word’s meaning, i.e., their underlying semantic representations, which may not be immediately obvious ‘on the surface’, can cause differences in task performance.

In a definition task, Verhallen and Schoonen (1993) found significant differences between children in their knowledge of the meaning of seemingly familiar words. Children’s knowledge of word meaning has been shown to be related to their reading comprehension: poor comprehenders were found to perform poorer on tasks of expressive vocabulary - define words and provide multiple meanings - than normal readers matched for decoding ability (Nation & Snowling, 1998a). Consistent with this, in an association task, success at identifying abstract meaning aspects of stimulus words was shown to be a unique contributor in the prediction of reading comprehension levels beyond the contribution of vocabulary size (Qian, 1999; Schoonen & Verhallen, 1998). Similarly, children’s knowledge of definitions and synonyms was found to predict their reading comprehension (Ouellette, 2006). Not only conscious or declarative knowledge of word meaning, but also fast access to word meaning has been suggested to play a role in reading comprehension (Perfetti, 2007). The study in Chapter 4 showed that the speed with which children (consciously) identify abstract meaning aspects of words makes a contribution to reading comprehension beyond word decoding and semantic word knowledge. Differences in accessing semantically related words may point to different underlying semantic networks and relations between words. Yet, tasks requiring conscious processing do not allow us to make claims about the status of children’s underlying semantic representations and they do not tell us whether words automatically co-activate one another, nor whether this is beneficial to reading performance. Is knowledge of and access to abstract meaning aspects important for reading because of the automatic activation of those meaning aspects? In that case automatic activation may be one of the explanatory mediating factors between reading comprehension and semantic word knowledge.

In this study we investigate the nature of children’s word knowledge in relation to their reading comprehension using a paradigm that assesses the activation of semantic relations between words. Semantic priming effects can reveal to what
extent individual differences in the unconscious activation of semantic word knowledge (and in the quality of semantic representations) are predictive of differences in reading comprehension.

5.1.1 Semantic representations and processing

Children are thought to have functional knowledge about words from an early age (Mandler, 1994), while abstract semantic knowledge develops later (Petrey, 1977). As explained in section 2.1.2, word knowledge starts out idiosyncratic and bound by the specific contexts in which words are encountered. Nelson (2007) hypothesizes that children abstract category information from this initial, situational knowledge. Data from word association tasks show that young children give contextually related responses rather than abstract semantic responses, which are given more commonly after age nine (Petrey, 1977). Kindergarteners’ contextual responses to *dark* (*sleep, bed*) have been replaced by the semantic response *light* by third grade. In a cued recall task, children have been found to be better at recalling pairs of words related through function (*chair–living room; airplane–sky*) than pairs of words belonging to the same category (*chair–bed; airplane–train*) (Blewitt & Toppino, 1991), but the developmental aspect is less clear in this study.

The shift from functional to more semantic knowledge of words is a shift in abstraction. A seemingly related development is children’s gradual learning of abstract categories. In a clothing categorisation task, Jerger and Damian (2005) investigated adults’ and 4-14-year-olds’ recognition of pictures that were more or less related to the category. Whereas adults could classify typical (*pants*) and atypical category objects (*glove*) equally well, children were more accurate with typical category objects. Moreover, children were significantly less accurate than adults in classifying out-of-category related items (*necklace*). The data also showed more age-related improvement in accuracy for atypical objects than for typical objects and for related out-of-category objects than for unrelated out-of-category objects (*soup*). This was taken as an indication of children’s semantic category fine-tuning.

Weak semantic processing skills in children have been linked to reading
comprehension problems (Nation & Snowling, 1998a, 1998b, 1999, 2004; Ricketts, Nation, & Bishop, 2007). Nation and Snowling (1998a) compared 10-11-year-old children with specific reading comprehension difficulties to children without reading problems on similarity judgment, verbal fluency and speeded word reading. Poor comprehenders were weaker than skilled comprehenders on a synonym judgment task (Do boat and ship mean the same thing?) but not on a rhyme judgment task (Do rose and nose rhyme?); this held in particular for low-imageability words. When asked to generate as many examples of category members (animals, modes of transport, jobs) in 60 seconds, poor comprehenders produced significantly fewer words than the control group, while the two groups did not differ for rhyme fluency (generate as many rhymes to three spoken words in 60 seconds). Moreover, poor comprehenders were slower than controls at reading words that are typically read with support from semantics (irregular words and low-frequency words), even though the two groups were matched for decoding ability (as assessed by non-word reading). This shows that poor comprehenders also have word recognition weaknesses relative to normal readers (for similar results see Ricketts, Nation and Bishop (2007)). In general, tasks requiring semantic processing take longer than mere lexical tasks (cf. Bueno & Frenck-Mestre, 2008). The finding that poor comprehenders have specific difficulty making semantic judgments, generating a semantic set and reading aloud irregular and low-frequent words points to a deficit in semantic but not phonological processing.

Poor comprehenders have also been shown to be less sensitive to (category) abstract semantic relations between words than normal readers. Nation and Snowling (1999) used an on-line measure of the effects of semantic similarity on lexical decision performance. As explained in section 2.3, skilled adults are faster at deciding whether a target item (e.g. nurse) is a word or a non-word if they have previously encountered a semantically related prime (e.g. doctor) relative to an unrelated prime (Meyer & Schvaneveldt, 1971). Semantic priming effects result from unconscious processing and are taken to reflect underlying semantic representations. In their study, Nation and Snowling found that both good and poor comprehenders (mean age 10;7) showed priming for function-related words (e.g.
broom–floor; shampoo–hair), but in the case of category coordinates (e.g. cat–dog; aeroplane–train) poor comprehenders only showed priming if the category pairs were also strongly associated (i.e. commonly co-occurred in language use). In the study by Weekes and colleagues (2008) children studied spoken words that were semantically related (e.g., bed, rest, and awake) or phonologically related (e.g., pole, bowl, and hole). Children were then tested with free recall and a recognition test that contained non-studied critical words (e.g., sleep and roll). The results showed that poor comprehenders were poorer at recalling and recognising the studied spoken words in the semantic condition but not in the phonological condition. The authors take this to show a reduced tendency in poor comprehenders to infer themes from studied words in the semantic task. They conclude that poor comprehenders are less sensitive to abstract semantic associations between words because of reduced gist memory, which suggests that poor comprehenders are less skilled at retrieving patterns and relations in meaning across events.

Betjemann and Keenan (2008) investigated priming in children with reading disability (poor oral vocabulary and decoding) and compared their performance to age-matched controls (mean age 11;5). In contrast to the control group, the children with reading disability showed no significant priming effects for semantic pairs (ship – boat), and showed smaller priming effects for phonological/graphemic (goat – boat) and combined pairs (float – boat), both in visual and auditory lexical decision tasks. These findings suggest semantic priming deficits for the poor readers and they suggest that these deficits are not restricted to the visual modality. The finding that the poor readers also showed less priming than reading-age matched controls suggests that their semantic processing deficits are not due to a lower reading level but are a more fundamental semantic weakness. This semantic weakness may contribute to both the poor readers’ word reading and comprehension problems.

The exact role of abstract semantic word knowledge in reading comprehension remains unclear. Is it enough to understand the more abstract, general aspects of a word’s meaning or is it the automatic activation of this knowledge that contributes to reading comprehension? The lexical quality
hypothesis claims that variation in the quality of word representations has consequences for reading skill (Perfetti, 2007; Perfetti & Hart, 2001, 2002). Low-quality representations lead to specific word-related problems in comprehension. High lexical quality includes well-specified and partly redundant representations of form (orthography and phonology) and flexible representations of meaning, allowing for rapid and reliable meaning retrieval (Perfetti, 2007). Perfetti and Hart state that high-quality lexical representations not only speed up processing but are also “responsible for automaticity (or at least efficiency) of word identification”, which allows processing resources to be devoted to higher level comprehension (2001: 76). Note that word identification here includes semantic identification. The supportive role of semantic word knowledge may thus lie in its automatic activation. With less efficient meaning activation, poor comprehenders may fail to quickly detect relations between words and hence they may retrieve a more superficial, less useful gist.

The study in Chapter 4 shows that readers differ in the accessibility of their semantic word knowledge as measured in a conscious, timed semantic choice task and that this has its own effect on reading comprehension. It is less sure whether differences in semantic decisions reflect underlying differences in the quality of semantic word representations. Is it because activation spreads automatically to semantically related words that reading comprehension is supported? To investigate whether differences between children in reading comprehension can be accounted for by differences in underlying semantic representations and their interconnectedness, we adopted a semantic priming paradigm. For this we used a semantic classification task and a lexical decision task.

5.1.2 Monolingual and bilingual children

Persistent differences between monolingual and bilingual children in both word knowledge (August, Carlo, Lively, Lippman, McLaughlin & Snow, 1999; Scheele, Leseman, & Mayo, 2010) and reading comprehension (August, Carlo, Dressler, & Snow, 2005; Nakamoto, Lindsey, & Manis, 2007; Netten, Droop, & Verhoeven, 2011; Proctor et al., 2005) have been found. It is still an open question to what
extent these comprehension problems are due to failing lexical-semantic processing. Qian (1999, 2002) has empirically shown how important depth of vocabulary knowledge (in terms of knowledge of synonymy, polysemy and collocations) is for reading comprehension in young adult second language learners. Depth of vocabulary knowledge made a unique contribution to the prediction of reading comprehension levels, in addition to the prediction afforded by vocabulary size scores. Semantic word knowledge was also found to be a predictor of reading comprehension in the study with 9-to-11 year old children by Schoonen and Verhallen (1998). In contrast to our word association study in Chapter 3, Schoonen and Verhallen compared Dutch monolingual and bilingual minority children on a receptive task. They used the Word Association Test (Schoonen and Verhallen, 2008). Whereas bilingual children would associate the Dutch equivalent of banana with nice, monolingual children would more frequently connect banana to the abstract associate fruit. Semantic word knowledge scores contributed unique variance to reading comprehension, beyond the variance attributed to differences in vocabulary size (cf. Read, 1993, 2000). Chapter 4 shows that differences in reading proficiency between monolingual and bilingual children ‘disappear’ when differences in lexical-semantic knowledge are taken into account. In work comparing Spanish-English bilinguals with their English monolingual counterparts, Proctor and colleagues (2009) found that semantic word knowledge (‘semantic depth’) among bilinguals and monolinguals was predictive of English reading comprehension, after controlling for oral English language proficiency. While a dichotomous language-status variable was not predictive of reading comprehension when oral language was included in the final model, children with average and above-average oral language skills were more likely to benefit from increased semantic awareness, and bilingual participants were underrepresented in those categories.

Bilingual minority children are a special group of bilinguals. They differ in a number of respects from so-called balanced bilinguals who learn their languages from birth or soon afterwards (Meisel, 2007). Bilingual minority children often do not learn Dutch, their second language until they enter school at the age of four and
they are less exposed to Dutch input. Generally, their first language has low socioeconomic status in the Netherlands. The reported language delays of these bilingual children are considerable and research shows that bilingual minority children do not easily catch up with their monolingual age mates (Farnia & Geva, 2011).

The study presented in this chapter evaluated differences between Dutch monolingual and bilingual minority children in reading comprehension, semantic word knowledge (research question 3), lexical decision speed and semantic classification speed (research question 4) and semantic priming (research question 5) and the study assessed the respective roles of these variables in reading comprehension (research question 6). On the basis of the literature, we expect differences between monolingual and bilingual minority children in semantic word knowledge and reading comprehension. Second, we expect group differences in processing factors such as lexical decision speed, semantic classification speed, and priming effects for semantically related words, because of a possible link between semantic word knowledge, reading and semantic processing. Third, we expect priming effects to discriminate between good and poor comprehenders. The research by Nation and Snowling (1999) suggests that such processing differences do exist. Finally, we expect individual priming scores to be a factor in the prediction of reading comprehension. Larkin and colleagues (1996) found that individual priming scores for sixth-graders were a factor in the prediction of reading comprehension.

5.2 Method
To investigate differences between learners in reading comprehension and in semantic processing, and the relationship between these constructs, several tasks were administered to monolingual and bilingual minority children. For reading comprehension, semantic word knowledge and word decoding, (standardised) paper-and-pencil tests were used; for recognition speed and semantic priming a lexical decision task and a semantic classification task were designed measuring reaction times. A word decoding test was included as a control variable because word-decoding fluency affects speed of processing in lexical decision and semantic
classification tasks. Children’ language background was assessed through an interview and a questionnaire.

5.2.1 Participants

Data were collected from a total of 169 children. None of the children had taken part in the two preceding studies. Of the original sample, 39 children were excluded due to dyslexia (N=17), or because they had just arrived in the Netherlands (N=3), followed an adapted program (N=6), were bilingual in European languages (N=5), or because there was no complete test data for them (N=8). This resulted in a final sample of 130 children (N=50 girls; N=80 boys) whose mean age upon measurement was 11;3 years old (SD 6 months). Children were tested at the end of grade 5 (grade 7 in the Dutch system) at one of six regular primary schools in different towns and cities in the western part of the Netherlands. Schools were selected that did not work with the reading comprehension test used in this study or with any specific semantic word knowledge training programme. Teachers were asked to screen out all children who had any known behavioural, emotional, or learning difficulties. All children spoke Dutch fluently. Of the children, 83 spoke only Dutch at home (monolingual children) and 47 spoke no Dutch or some Dutch at home (bilingual minority children). Among bilingual minority children a variety of non-European languages was spoken at home as a first language, the most common being Turkish and Moroccan (Arabic and Berber). Within the two language groups, the proportion of girls and boys was comparable to that in the participant group as a whole. Schools varied in socioeconomic status and in language background of children. Some schools had children from a predominantly monolingual background; others had more multilingual populations. Seven children were born outside the Netherlands but had started Kindergarten in the Netherlands. For each student permission to take part was obtained through the school.
5.2.2 Materials

Reading comprehension

Reading comprehension was measured, as in the study in Chapter 4, with the standardised Dutch reading test for children in the final years of primary education (grades 4, 5 and 6), ‘Begrijpend Lezen 678’ (Aarnoutse & Kapinga, 2006). The test has been normed nationally on a sample of 42 schools. Reliability is reported as satisfactory ($\alpha = .83$). Scores on the test correlate strongly with other standardised tests of reading comprehension. The test comprises an answer sheet and a booklet containing reading passages on different topics, ranging in length from 122 words to 288 words. Each passage is accompanied by six or seven questions: three or four multiple-choice questions and two to four true/false statements. In this study, we decided to use a slightly shorter version of the test\(^{11}\); we excluded two passages that required additional instruction. This resulted in a total of five reading passages and 32 questions. Three questions are word-level questions; the other questions are above word level, including literal and inferential questions and overview questions of a larger scope. In this sample, the distribution of reading scores was slightly skewed, -.545 (se .212), with normal kurtosis, -.195 (se .422). Internal consistency reliability was satisfactory (Cronbach’s $\alpha = .77$).

Semantic word knowledge

As in Chapter 4, the multiple-choice Word Association Task (WAT, Schoonen & Verhallen, 2008) was administered as a test of semantic word knowledge. The task is based on Read’s (1993) word associate format and has been developed for children from 9 to 12 years old. Each of the 30 items in this written task consists of a stimulus word (e.g., vegetable) surrounded by six words. Three of the surrounding words are targets and are semantically related to the stimulus word (e.g., plant, lettuce, food); three are distracters and are only indirectly related or unrelated to the stimulus word (e.g., plate, warm, strong). Children are asked to identify the three

---

\(^{11}\) The original reading comprehension test used in the study in Chapter 4 and the slightly shorter version used here were highly correlated ($r = .95, p = .000$). This correlation is based on an analysis of the data from Chapter 4.
semantically related words by drawing lines from the stimulus word to the three words that really belong to it. The distinction between targets and distracters is a gradual one: children have to compare semantically related pairs to indirectly related pairs (vegetable – plant vs. vegetable – strong), reject the indirect or context-specific relations and recognise which words are related in meaning. A response preference for words related in meaning is seen as an indication of well-developed semantic word knowledge and of the prominence of abstract, semantic (rather than context-specific) connections between words in the child’s lexicon. The WAT was group administered in class. As part of the instruction two examples were worked out and explained. To prevent cheating two versions of the WAT were created (with a different ordering of the items) and randomly assigned to the children. For the task, the same scoring method was used as in Chapter 4. WAT scores in this sample were slightly skewed: -.822 (se .212); kurtosis did not deviate significantly from normality: .484 (se .422). Internal consistency reliability was .82.

Word decoding
To control for differences between children in word decoding skill the ‘Drie Minuten Test’ was administered, as in Chapter 4. It has a good reliability and validity (Verhoeven, 1992). As in Chapter 4, children were asked to read two word lists out loud. The lists consist of regular, non-related words increasing in decoding difficulty. The first list consists of 150 monosyllabic words; the second list consists of 120 multisyllabic words: 60 two-, 30 three-, and 30 four-syllable words. Children were given a break between the two lists. For each list the score is the number of words read correctly in one minute. In this study, the scores for the two word decoding lists are strongly correlated (Pearson’s $r = .822$) and show a high split-half reliability of $r_{1.2} = .90$; therefore, the two list scores are combined (averaged) to derive a single score for word decoding which is used in further analyses. Mean word decoding scores are normally distributed: skewness is -.156 (se .212) and kurtosis -.054 (se .420).
Lexical decision

To measure word recognition speed and semantic priming, a visual lexical decision task was designed. Children were instructed to decide as fast as possible for each stimulus whether it was an existing word or not. Accuracy as well as response times were recorded. A basic assumption of the lexical decision task is that a correct response to a target word requires access to a corresponding mental representation of that word. Although it has been doubted whether semantic processing is required for making lexical decisions, research has shown that the lexical decision task is indeed sensitive to semantic information (Vigliocco, Vinson, Lewis, & Garrett, 2004). To ensure that all words were attended to, we used single-word presentation whereby participants respond to primes as well as targets.

The critical stimuli consisted of 24 semantically related, non-associated prime-target pairs (see Appendix B and C for the Dutch stimuli and an English translation). Ten were coordinately related (e.g., nose – ear) and fourteen were subordinately related (e.g., knife – cutlery). For the construction of test pairs, the word association norms collected in the study reported on in Chapter 3 were used. The 24 prime words were stimulus words from the word association study; the 24 targets had never occurred as response. Thus, association strength for prime-target pairs was close to zero. To ensure children’ familiarity with the stimuli, all stimuli used were high-frequent nouns checked in the child school-language corpus (Schrooten & Vermeer, 1994). Word frequency and word length in terms of numbers of letters were balanced across conditions and versions (see Table 5.1). Care was taken to avoid pairings in which prime and target were orthographically or phonologically related.

---

12. The corpus provides cumulative word frequency for every two grades. As our participants were tested at the end of grade 5 (age 11), cumulative frequency for grade 5/6 was used. This strongly correlated with cumulative frequency until grade 5.
Table 5.1 Frequency (fre) and length (lett) of stimuli by condition and version in the lexical decision task

<table>
<thead>
<tr>
<th></th>
<th>PRIMED</th>
<th></th>
<th>CONTROL</th>
<th></th>
<th>Non-words</th>
<th>Fillers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>fre</td>
<td>lett</td>
<td>fre</td>
<td>lett</td>
<td>fre</td>
<td>lett</td>
</tr>
<tr>
<td>Primes (12)</td>
<td>310</td>
<td>4.5</td>
<td>266</td>
<td>5.3</td>
<td>300</td>
<td>5.5</td>
</tr>
<tr>
<td>Targets (12)</td>
<td>266</td>
<td>5.3</td>
<td>310</td>
<td>4.5</td>
<td>156</td>
<td>6.6</td>
</tr>
<tr>
<td>Primes (12)</td>
<td>300</td>
<td>5.5</td>
<td>156</td>
<td>6.6</td>
<td>310</td>
<td>4.5</td>
</tr>
<tr>
<td>Targets (12)</td>
<td>156</td>
<td>6.6</td>
<td>266</td>
<td>5.3</td>
<td>300</td>
<td>5.5</td>
</tr>
<tr>
<td>version 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>version 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.9</td>
<td>124</td>
</tr>
</tbody>
</table>

Note. In version 1, related targets are more frequent than unrelated targets, in version 2 vice versa.

Twenty-four unrelated control pairs were produced from the related pairs by randomly reordering the primes. Two versions of the task were created, each containing 12 related pairs and 12 unrelated pairs, such that each target word appeared once in each version: in one version in primed condition, i.e. with its semantically related prime (e.g., *nose* – *ear*), and in the other version in control condition, i.e. with an unrelated prime (e.g., *park* – *ear*). The 12 target words that appear in primed condition in version 1 and in control condition in version 2 will be referred to as targets A; the other 12 targets will be referred to as targets B. An overview of the arrangement of the stimuli into versions can be found in Appendix B.

A set of 51 orthographically legal non-words (e.g., *bref*) was created and added to each version. Non-words were created in *WordGen* (Duyck, Desmet, Verbeke, & Brysbaert, 2004), which uses the Celex lemma database. Bigram frequency boundaries and number of orthographic neighbours were constrained such that all non-words were ‘word-like’ in Dutch. None of the non-words were existing...
words in Turkish or Moroccan, the home languages of the majority of the bilingual children.

In addition, 27 filler words were added to each version. Fillers were matched for letter length and as closely as possible for frequency to the target words. This made a total of 126 items in each version (12 test pairs + 12 control pairs + 51 non-words + 27 fillers); 40.5% of the items were non-words; the proportion of semantically related trials was 9.5%. Response speed to fillers was used as a measure of general lexical decision speed. Figure 5.1 shows the experimental set up.

For both versions three random orders were created to prevent order effects. The six versions of the task were checked and corrected by the experimenter to ensure that test pairs were flanked by fillers, that no unintended relations occurred between stimuli and that consecutive filler words broke up the pattern of test words appearing consecutively. Each version started with six filler items. Children were randomly assigned a version of the task.

**Figure 5.1** Experimental set up: lexical decision task. The interstimulus interval (ISI) runs from one stimulus to the next, including a 500 ms fixation cross.
The task was run using the computer programme *E-prime* v2.0 (Schneider, Eschman, & Zuccolotto, 2002). Two identical Toshiba Satellite A110 laptop computers were used. Stimulus display was synchronised to the screen refresh rate (17 ms). All stimuli were presented in the centre of the screen in bold, black Courier New 24-point font against a white background. A lexical decision response was made to every stimulus by pressing either a ‘yes’ or ‘no’ button (the Alt keys, marked with a green or red sticker). Children used their dominant hand for a positive response. Each trial began with the presentation of a fixation cross for 500 ms, followed by a stimulus in lowercase letters presented until a response was detected or until 4000 ms had passed since stimulus onset (cf. Martens & de Jong, 2006). Following a response, a 2000 ms blank screen interval preceded the next trial. This interval was adapted from an initial 1000 ms, which, in a pilot with 11-year-old children, was found to be too short. The resulting interstimulus interval (ISI) was 2500 ms. Response times were measured from stimulus onset until a key was pressed. Children were required to keep their index fingers above the red and green button throughout the task. A practice set of nine stimuli (five words and four non-words) preceded the main experiment to familiarise participants with the task. For these practice trials on-screen feedback was given on accuracy and response times. During the experiment no feedback was given. The accuracy data from the lexical decision task were slightly skewed and peaked (*words* skewness -1.861, se .212; kurtosis 3.808, se .422; *non-words* skewness -1.161, se .212; kurtosis 1.602, se .422). The RT data for words were slightly skewed (.517, se .212), with normal kurtosis (-.254, se .422). RT data for non-words were normally distributed (skewness .392, se .212; kurtosis -.728, se .422). There was no speed-accuracy tradeoff for lexical decision (*r* = .001, *p* = .990).

**Semantic classification**

To measure semantic priming, a visual semantic classification task was designed, in addition to the lexical decision task. Semantic classification requires the use of semantic information and has been assumed to be more sensitive to semantic processing than lexical decision. The task used was animal classification. Children
were instructed to decide as quickly as possible for each stimulus whether it was an animal name or not. Again, a basic assumption of the task is that a correct response to a target word requires access to a corresponding mental representation of that word. To ensure that all words were attended to, we used single word presentation whereby primes as well as targets were responded to.

The critical stimuli consisted of 20 semantically related, non-associated prime-target pairs (see Appendix D and E for the Dutch stimuli and an English translation). Half were animal names (exemplars); half were object concepts (non-exemplars). For both the animal and the non-animal related pairs, primes and targets were related in that they shared their superordinate category (e.g., duck – goose; taxi - bus). Priming could occur for animal pairs as well as for non-animal pairs. Non-animal related pairs were included to prevent children from using their detection of a relationship between prime and target as a cue to respond ‘yes’. For the construction of test pairs, the word association norms collected in the study reported on in Chapter 3 were used. The 20 prime words were stimulus words from the word association study; the 20 targets had never occurred as response. Thus, association strength for prime-target pairs was close to zero. To ensure children’s familiarity with the stimuli, all stimuli used were high-frequent nouns checked in the child school-language corpus (Schrooten & Vermeer, 1994). Word frequency and word length, in terms of numbers of letters, were balanced across conditions and versions (see Table 5.2). Care was taken to avoid pairings in which prime and target were orthographically or phonologically related.
Table 5.2 Frequency (fre) and length (lett) of stimuli by condition and version in the semantic classification task

<table>
<thead>
<tr>
<th></th>
<th>PRIMED</th>
<th>CONTROL</th>
<th>Fillers (66)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(10)</td>
<td>(10)</td>
<td>(10)</td>
</tr>
<tr>
<td>fre</td>
<td>lett</td>
<td>fre</td>
<td>lett</td>
</tr>
<tr>
<td>version 1</td>
<td>382</td>
<td>4.8</td>
<td>171</td>
</tr>
<tr>
<td>version 2</td>
<td>161</td>
<td>5.0</td>
<td>85</td>
</tr>
</tbody>
</table>

*Note.* In version 1, related targets are more frequent than unrelated targets, in version 2 vice versa.

Twenty unrelated control pairs were produced from the related pairs by randomly reordering the primes. Two versions of the task were created, each containing 10 related pairs and 10 unrelated pairs, such that each target word appeared once in each version: in one version in primed condition, i.e. with its semantically related prime (e.g., *lion - tiger*) and in the other version in control condition, i.e. with an unrelated prime (e.g., *brush - tiger*). The 10 target words that appear in primed condition in version 1 and in control condition in version 2 will be referred to as targets A; the other 10 targets will be referred to as targets B. An overview of the arrangement of the stimuli into versions can be found in Appendix D.

Sixty-six filler words were added to each version to keep the proportion of related trials low. Fillers were matched for length (letters and syllables) and frequency to the target words. Half of the fillers were animal names. This resulted in a total of (10 related pairs + 10 unrelated pairs + 66 fillers=) 106 words in each version; 50% of the words were animals and the proportion of semantically related trials was 9.4%. Response speed to exemplar fillers was used as a measure of general semantic classification speed. Figure 5.2 shows the experimental set up.
For both versions three random orders were created to prevent order effects. The six versions of the task were checked and corrected by the experimenter to ensure that test pairs were flanked by fillers, that no unintended relations occurred between stimuli and that consecutive filler words broke up the pattern of test words appearing consecutively. Each version started with six filler items. Children took version 1 or 2 of the task depending on which version they had been assigned for the lexical decision task.

Figure 5.2 Experimental set up: semantic classification task. The interstimulus interval (ISI) runs from one stimulus to the next, including a 500 ms fixation cross.

The general task procedure was the same as for the lexical decision task. Children responded to every stimulus by pressing the ‘yes’ or ‘no’ button (marked with a green or red sticker) depending on whether the stimulus was an animal name or not. A practice set of eight stimuli (four animal and four non-animal names) preceded the main experiment to familiarise children with the task. The semantic classification accuracy data were somewhat skewed and peaked (skewness -1.357, se .212; kurtosis 1.899, se .422). The RT data were normally distributed (exemplars
skewness .387, se .212; kurtosis -.476, se .422; non-exemplars skewness .419, se .212; kurtosis -.366, se .422). As for lexical decision, there was no speed-accuracy tradeoff for semantic classification ($r = -.086, p = .332$).

**Language background**

To determine the language background of each student, one of two trained test assistants or the principal investigator conducted an individual interview with each student and filled out a questionnaire on the student’s language use at home and at school (outside the classroom). Teachers were also interviewed about the language background of their children. In the individual interviews children were asked, for example, which language is spoken to parents, siblings, other family members and friends. Through the questionnaire background information was also obtained about such things as country and date of birth and length of residence in the Netherlands (for those born elsewhere).

**5.2.3 Procedure**

Children were tested in their schools, on a regular school day. Two tasks were administered in class; three were completed individually. Testing in class started in the morning and lasted approximately 45 minutes. Individual testing took place in a quiet room and lasted approximately 30 minutes. Testing was done by one of two trained test assistants or the principal investigator. Administration followed a strict protocol. Tasks were administered to all children in the same order. Testing started in class with the WAT. When all children had finished, the written reading comprehension task was administered. Children who had finished were asked to leave the classroom with one of the test assistants to take the three individual tasks.

Individual sessions started with the lexical decision task, followed by the word decoding task, which lasted about three minutes and which was followed by the semantic classification task. Children were randomly assigned a version of the lexical decision and the semantic classification task. At the end, the participant was interviewed and the language questionnaire was filled out.
5.2.4 Data handling and analyses

For the reading comprehension task and the WAT, outliers were defined as scores that were three standard deviation units from the mean. This led to the removal of one student for both of these tasks. There were no outlying scores for word decoding.

The lexical decision and the semantic classification data sets were examined for outliers and missing data (i.e., no response detected). Accuracy scores and recognition speed were based on filler trials, those being the same for all children. Accuracy was calculated as the percentage of correct trials of a student’s total number of valid trials (i.e. response detected). The treatment of response time (RT) data was as follows. Any trial on which a participant made an error (incorrect trial), invalid trials (i.e. no response detected), and outliers were set to missing and replaced by imputed (estimated) values (see below). Outliers were defined as data points beyond three standard deviation units from the general mean \(^{13}\) and responses faster than 250 ms (see Betjemann & Keenan, 2008).

For the lexical decision task, incorrect trials \((1061 \ (335 \ w + 726 \ nw) = 6.2\%)\), invalid trials \((42 \ (10 \ w + 32 \ nw) = 0.2\%)\) and outliers \((358 \ (213+145) = 2.1\%)\) amounted to 8.6\% missing RT data. No participants or items were removed due to missing data. The RT data from the semantic classification task show a similar picture: incorrect trials \((885 = 6.4\%)\), invalid trials \((31 = 0.2\%)\) and outliers \((248 = 1.8\%)\) amounted to 8.4\% missing RT data. One participant had more than 30\% missing data (31\%) and was removed \((N=130)\); four stimuli (fillers) yielded more than 25\% missing data and were removed \((\text{number of fillers}=62)\). Missing observations were multiply imputed (5 times) in SPSS using constraints and were replaced by the mean of the imputed values.

Psychometric properties of the measures were established for the entire sample \((N=130)\). We established priming effects using analysis of variance procedures. Individual priming scores were calculated after equating RTs to control

\(^{13}\) We used the mean response time of all RT observations, see McDonough & Trofimovich (2009). For the semantic classification task this led to an upper cut off value of 2016 ms. For the lexical decision task, we did the same but for words and non-words separately (cut offs of 1845 and 2748 ms respectively).
for differences between target word sets. Differences between monolingual and bilingual minority children were established for all measures and correlations among the variables were computed. Hierarchical regression analyses were conducted to explore the role of recognition speed and semantic priming in explaining reading comprehension over and above the contribution made by word decoding and semantic word knowledge. First, we tested a baseline model including decoding and semantic word knowledge. Then, we tested the effect of recognition speed and priming. Finally, we tested for interactions between the variables in the final model. In a path analysis we further explored the relations between reading and the constituent variables, and investigated whether these relationships hold across different subgroups (monolingual vs. bilingual; high vs. low proficient readers).

5.3 Results

5.3.1 Descriptives

Mean scores and standard deviations for the measures for the total sample (N=130) are shown in Table 5.3. Accuracy scores for lexical decision and semantic classification are reported on in the text below. For lexical decision, response times to non-words and for semantic classification, response times to non-exemplars are also reported on in the text. All measures are reliable (between .77 and .91) except for the accuracy measure in the lexical decision and in the semantic classification task (.51 and .55). Not surprisingly, accuracy scores for those tasks are at ceiling, which shows that children could easily do the tasks. For lexical decision, mean accuracy scores (and standard deviations) were 96.4% (4.9) for words and 90.0% (8.3) for non-words. For semantic classification, mean accuracy was 94.8% (4.1). Mean RTs for lexical decisions to words were 828 ms (134) and for non-words 1127 ms (234); mean RTs for semantic decisions to exemplars were 862 ms (150) and for non-exemplars 938 ms (161). This shows that on average, ‘no’ responses took longer than ‘yes’ responses. Table 5.3 shows that semantic decisions took longer than lexical decisions, as expected. Scores for non-words will not be used in further analyses. Priming scores are described in detail in the following section.
Table 5.3 Descriptive statistics and effects of language background on performance, by measure

<table>
<thead>
<tr>
<th>Measures</th>
<th>Total (N=130)</th>
<th>Monolingual (N=83)</th>
<th>Bilingual (N=47)</th>
<th>η² p*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>23.3</td>
<td>4.6</td>
<td>24.9</td>
<td>3.8</td>
</tr>
<tr>
<td>WAT</td>
<td>20.7</td>
<td>4.9</td>
<td>22.3</td>
<td>3.7</td>
</tr>
<tr>
<td>Word Decoding</td>
<td>86.7</td>
<td>15.2</td>
<td>86.2</td>
<td>14.7</td>
</tr>
<tr>
<td>Lexical decision speed</td>
<td>828</td>
<td>134</td>
<td>805</td>
<td>114</td>
</tr>
<tr>
<td>Semantic classification speed</td>
<td>862</td>
<td>150</td>
<td>825</td>
<td>130</td>
</tr>
<tr>
<td>PrimingLDT</td>
<td>39.8</td>
<td>88.2</td>
<td>38.7</td>
<td>80.6</td>
</tr>
<tr>
<td>PrimingsCT</td>
<td>70.7</td>
<td>142.2</td>
<td>59.7</td>
<td>136.6</td>
</tr>
</tbody>
</table>

Note: * significant at the 0.05 level; LDT = lexical decision task; SCT = semantic classification task

5.3.2 Semantic priming

Semantic priming was established by comparing response times in primed (related) and control (unrelated) condition. We will refer to participants with version 1 as group 1 and to participants with version 2 as group 2. Groups 1 and 2 are the same participants across the two tasks. First, we calculated priming within targets A and within targets B (across groups), to see whether the primed vs. control condition had worked. For this, we checked whether group 1 and 2 were equally fast, on the basis of their response times to the filler items. Using this baseline speed as a covariate, priming effects for targets A and for targets B were calculated. Once priming effects were established, we calculated priming for individual participants, across targets A and B. For this, differences between targets A and B were taken into account. ANOVAs were conducted for all group comparisons.
**Lexical decision task**

Baseline speed, in terms of responses to the filler trials (which had been the same for all children), differed slightly but not significantly between groups: group 2 took 18 ms longer to respond to filler trials than group 1 (see Table 5.4), $F(1, 129) = .625$, $p = .431$, effect size partial eta squared ($\eta^2_{p}$) is .005. With baseline speed as a covariate, priming for targets A and for targets B can be calculated with more precision.

Including baseline speed as a covariate, for both target sets there was a priming effect with corrected RT values to words in primed condition being faster than to words in control condition: facilitation was 25 ms for targets A, which approached significance, $F(1,127) = 3.828$, $p = .053$, $\eta^2_{p} = .029$, and 37 ms for targets B, which was significant, $F(1,127) = 8.740$, $p = .004$, $\eta^2_{p} = .064$. These are reassuring priming scores, as compared with Nation and Snowling (1999: B8) who report 34 ms for normal readers. Betjemann & Keenan (2008) report 23 ms for poor readers (‘RD’ reading disability) and 37 ms for normal readers.
Table 5.4 Lexical decision (LDT) and semantic classification (SCT): original and equated RT means and standard deviations to targets A and B

<table>
<thead>
<tr>
<th></th>
<th>Primed condition</th>
<th>Control condition</th>
<th>Fillers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Primed condition</td>
<td>Control condition</td>
<td>Fillers</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>coor</td>
<td>coor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sub</td>
<td>sub</td>
<td></td>
</tr>
<tr>
<td><strong>LDT</strong></td>
<td><strong>Group 1</strong></td>
<td><strong>Group 2</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>original</td>
<td>original</td>
<td></td>
</tr>
<tr>
<td></td>
<td>equated</td>
<td>equated</td>
<td></td>
</tr>
<tr>
<td></td>
<td>787 (143)</td>
<td>785 (135)</td>
<td>819</td>
</tr>
<tr>
<td></td>
<td>793 (137)</td>
<td>854 (152)</td>
<td>(115)</td>
</tr>
<tr>
<td></td>
<td>733 (139)</td>
<td>785 (135)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>834 (142)</td>
<td>854 (152)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>743 (160)</td>
<td>839 (165)</td>
<td>837</td>
</tr>
<tr>
<td></td>
<td>850 (169)</td>
<td>825 (136)</td>
<td>(151)</td>
</tr>
<tr>
<td></td>
<td>743 (160)</td>
<td>804 (170)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>850 (169)</td>
<td>876 (184)</td>
<td></td>
</tr>
<tr>
<td><strong>SCT</strong></td>
<td><strong>Group 1</strong></td>
<td><strong>Group 2</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>original</td>
<td>original</td>
<td></td>
</tr>
<tr>
<td></td>
<td>equated</td>
<td>equated</td>
<td></td>
</tr>
<tr>
<td></td>
<td>795 (173)</td>
<td>861 (175)</td>
<td>884</td>
</tr>
<tr>
<td></td>
<td>902 (175)</td>
<td>966 (183)</td>
<td>(142)</td>
</tr>
<tr>
<td></td>
<td>790 (169)</td>
<td>861 (175)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>963 (210)</td>
<td>966 (183)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>835 (189)</td>
<td>845 (182)</td>
<td>922</td>
</tr>
<tr>
<td></td>
<td>997 (218)</td>
<td>973 (210)</td>
<td>(160)</td>
</tr>
<tr>
<td></td>
<td>835 (189)</td>
<td>905 (193)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>997 (218)</td>
<td>996 (191)</td>
<td></td>
</tr>
</tbody>
</table>

* coor = coordinate; sub = subordinate, ex. = exemplar; non = non-exemplar

To compute individual priming scores, across targets A and B, we must ensure that target set A and B are equally difficult. This was checked in the control condition. Again, baseline speed was taken as a covariate. In the control condition, RT values to targets A and B differed slightly, but not significantly. On average, response times to targets A were 10 ms faster than to targets B, $F(1,127) = .650, p = .422, \eta^2_p = .005$, which was not significant.
To make response times to targets A and B comparable, response times to targets A and B were equated. For this, a standard regression method (internal-anchor design) was used with fillers as anchor items (Engelen & Eggen, 1993). In brief, RT scores for targets A were re-scaled to those for targets B. This was done separately for coordinate and subordinate targets and for primed and control condition. The standard formula used for this uses RTs to fillers as well as the correlation between the mean RT to fillers and the mean RT for targets A or B. Table 5.4 shows the original and equated RT values.

After equating response times, individual priming scores were calculated for coordinate and subordinate targets: response times to targets in primed condition were compared to response times to targets in control condition. There is a significant priming effect for coordinate and for subordinate pairs. Children were 57 ms faster responding to a target word preceded by a coordinately related word (paired t-test, \( t \) (129) = 5.097, \( p < .001 \)) and 23 ms faster to a target word preceded by a subordinately related word (paired t-test, \( t \) (129) = 2.053, \( p = .042 \)), in comparison to control condition. A priming variable was created for coordinate and for subordinate priming by subtracting response times to targets in control and related condition. Since response times to coordinate and subordinate targets in control condition and in primed condition are correlated (\( r = .69, p = .000 \) and \( r = .67, p = .000 \)), a general priming variable for lexical decision is created reflecting the average priming effect. This general priming variable is reported in Table 5.3 and is used in further analyses.

**Semantic classification task**

Baseline speed, in terms of responses to filler trials (which had been the same for all children), differed slightly but not significantly between groups. As in the lexical decision task, in the semantic classification task group 2 showed longer response times than group 1. Group 2 took 38 ms longer to respond to fillers than group 1 (see Table 5.4), \( F(1, 129) = 2.033, p = .156, \eta^2_p = .016 \). Baseline speed was taken as a covariate to increase precision in determining priming for targets A and for targets B.
Correcting for baseline speed, we see a priming effect with RTs to words in primed condition being faster than to words in control condition: facilitation was 23 ms for targets A, which was not significant, $F(1,127) = 2.659, p = .105, \eta^2_p = .021$, and 37 ms for targets B, which was significant, $F(1,127) = 6.904, p = .010, \eta^2_p = .052$.

To compute individual priming scores, across targets A and B, we must ensure that target set A and B are equally difficult. This was checked in the control condition. Again, baseline speed was taken as a covariate. In the control condition, RT values to targets A were 43 ms faster than to targets B, $F(1,127) = 9.485, p = .003, \eta^2_p = .069$, so targets A may have been easier.

Response times to targets A and B were made comparable by equating response times to targets A and B. This was done separately for exemplars and non-exemplars and for primed and control condition. For this, a standard regression method (internal-anchor design) was used with fillers being anchor items (Engelen & Eggen, 1993). Table 5.4 shows the original and equated RT values.

After equating response times, individual priming scores for exemplar and non-exemplar targets were calculated: response times to targets in primed condition were compared to response times to targets in control condition. Children were 71 ms faster responding to an exemplar target preceded by a related prime (paired t-test, $t(129) = 5.672, p < .001$) than to an exemplar target preceded by an unrelated prime. There was no priming effect, on average, in the non-exemplar condition (paired t-test, $t(129) = .079, p = .937$). A priming variable was created for exemplar and non-exemplar priming by subtracting response times to targets in control and related condition. Since there was priming only for exemplars, the exemplar priming variable for semantic classification is reported in Table 5.3 and is used in further analyses.

5.3.3 Monolingual and bilingual children

Regarding the comparison between monolingual and bilingual minority children Table 5.3 shows means, standard deviations and effect sizes for all relevant measures. Analyses of variance revealed significant differences for measures of reading comprehension, semantic word knowledge, and lexical decision and
semantic classification speed. Monolingual children performed significantly better at both the reading comprehension task, $F(1,128) = 35.575$, $p < .001$, $\eta^2_p = .217$, and the semantic word knowledge task (WAT), $F(1,128) = 30.175$, $p < .001$, $\eta^2_p = .191$. For word decoding, scores on the combined measure showed that the bilingual minority children performed slightly better, but this difference was not statistically significant, $F(1,128) = .233$, $p = .630$, $\eta^2_p = .002$. Regarding lexical decision and semantic classification speed, monolingual learners were significantly faster than bilingual minority children at making semantic classifications, $F(1,128) = 15.450$, $p < .001$, $\eta^2_p = .108$, and lexical decisions, $F(1,128) = 6.836$, $p = .010$, $\eta^2_p = .051$. The groups showed comparable priming effects. When comparing groups in an ANCOVA with decoding as a covariate, group differences hardly change: speed differences become slightly more pronounced and priming differences remain the same.

### 5.3.4 Relationships with reading comprehension

Reading comprehension scores correlated strongest with semantic word knowledge ($r = .62$, see Table 5.5) and with lexical and semantic recognition speed ($r = -.36$ and $r = -.47$, respectively). The correlation with word decoding was weak ($r = .20$), as is appropriate for fluent readers. Lexical decision speed and semantic classification speed were also correlated to semantic word knowledge ($r = -.37$ and $r = -.48$, respectively) and to word decoding ($r = -.49$ and $r = -.39$). There was no significant correlation between reading comprehension and semantic priming. Moreover, the correlation is negative, whereas a positive correlation between reading comprehension and semantic priming would be expected. To explore thresholds or non-linear relations, separate analyses were computed for children with higher and lower priming scores or for children with higher and lower reading comprehension scores. None of these revealed significant correlations between reading comprehension and priming.
Table 5.5 Intercorrelations for all measures (N=130)

<table>
<thead>
<tr>
<th>Measures</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Reading Comprehension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Word Decoding                           .20*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 WAT                                    .62*</td>
<td>.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Lexical decision speed                 -.36*</td>
<td>-.49*</td>
<td>-.37*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Semantic classification speed          -.47*</td>
<td>-.39*</td>
<td>-.48*</td>
<td>.74*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Priming&lt;sub&gt;LDT&lt;/sub&gt;                  -.14</td>
<td>.00</td>
<td>.09</td>
<td>.05</td>
<td>.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Priming&lt;sub&gt;SCT&lt;/sub&gt;                  -.06</td>
<td>-.02</td>
<td>.13</td>
<td>.06</td>
<td>-.01</td>
<td>-.14</td>
<td></td>
</tr>
</tbody>
</table>

Note: * correlation significant at the 0.05 level; LDT = lexical decision task; SCT = semantic classification task

5.3.5 Contributions to reading comprehension

**Lexical decision task**

Linear regression analyses were conducted to determine the contribution of speed and priming scores to reading comprehension. In a baseline model we included decoding skill and semantic word knowledge (WAT); together these variables accounted for 40% of variance in reading comprehension scores; both variables contributed significantly. Adding speed or priming as measured in the lexical decision task could not improve the baseline model significantly (see Table 5.6a). Using separate priming variables for coordinate and subordinate priming made no difference. Inclusion of the interaction term between semantic word knowledge and lexical decision speed or priming did not account for additional variance in reading comprehension scores either, showing that the effect of speed or priming was not mediated by children’ level of semantic word knowledge.
Table 5.6a  Additional variance in reading comprehension accounted for by lexical decision speed and priming

<table>
<thead>
<tr>
<th>Model</th>
<th>Predictors</th>
<th>Total R²</th>
<th>R² Change</th>
<th>F Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Decoding, WAT</td>
<td>.404</td>
<td>--</td>
<td>43.093*</td>
</tr>
<tr>
<td>2</td>
<td>+ Lexical decision speed</td>
<td>.412</td>
<td>.008</td>
<td>1.684</td>
</tr>
<tr>
<td>3</td>
<td>+ Priming_{LDT}</td>
<td>.419</td>
<td>.007</td>
<td>1.579</td>
</tr>
</tbody>
</table>

*Note: * significant at the 0.05 level.

**Semantic classification task**

Semantic classification speed improved the baseline model significantly, explaining an additional 2.5% of the variance in reading scores \((p = .020)\) (see Table 5.6b). Semantic priming as measured in the semantic classification task could not improve the model significantly. Inclusion of the interaction term between semantic word knowledge and semantic classification speed or priming did not account for additional variance, showing that the effect of speed or priming was not mediated by children’s level of semantic word knowledge.
Table 5.6b Additional variance in reading comprehension accounted for by semantic classification speed and priming

<table>
<thead>
<tr>
<th>Model</th>
<th>Predictors</th>
<th>Total $R^2$</th>
<th>$R^2$ Change</th>
<th>F Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Decoding, WAT</td>
<td>.404</td>
<td>--</td>
<td>43.093*</td>
</tr>
<tr>
<td>2</td>
<td>+ Semantic classification speed</td>
<td>.430</td>
<td>.025</td>
<td>5.594*</td>
</tr>
<tr>
<td>3</td>
<td>+ Priming$_{SCT}$</td>
<td>.447</td>
<td>.017</td>
<td>3.841</td>
</tr>
</tbody>
</table>

Note: * significant at the 0.05 level

Language background

To examine to what extent differences between monolingual and bilingual minority children are accounted for by differences in speed and priming scores, we conducted additional regression analyses of reading comprehension. When entered into the regression equation after decoding, semantic word knowledge, semantic classification speed and priming, the effect of language background (i.e., monolingual or bilingual) was still significant (see Table 5.7) but considerably reduced. On its own, language background accounted for 21.7% of explained variance ($\Delta R^2 = .217; F_{change}(1,128)=35.58; p = .000$).
Table 5.7 Additional variance in reading comprehension accounted for by language background

<table>
<thead>
<tr>
<th>Model</th>
<th>Predictors</th>
<th>Total $R^2$</th>
<th>$R^2$ Change</th>
<th>$F$ Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Decoding, WAT, Semantic classification speed, Priming$_{SCT}$</td>
<td>.447</td>
<td>.447</td>
<td>25.220*</td>
</tr>
<tr>
<td>2</td>
<td>+ Language background</td>
<td>.479</td>
<td>.033</td>
<td>7.787*</td>
</tr>
</tbody>
</table>

*Note: * significant at the 0.05 level

A regression analysis with interaction terms was carried out to determine whether the contribution of speed and priming to reading comprehension is different for monolingual and bilingual minority children (i.e., whether speed and priming interact with language background). Inclusion of the interaction term between language background and speed or priming, either with the lexical decision or semantic classification variables, in the final step of the model did not account for additional variance, showing that the effect of speed or priming is not significantly different for the language groups.

We also examined whether there is an interaction effect of language background and semantic word knowledge, i.e., whether the role played by semantic word knowledge in reading comprehension is different for monolingual and bilingual minority children. To the baseline model with decoding and semantic word knowledge we first added language background. Adding to that the interaction term between language background and semantic word knowledge did not account for additional variance, showing that semantic word knowledge makes a comparable contribution to reading scores for the monolingual and bilingual minority children.

To further explore the relationships between reading and the lexical-semantic variables we conducted path analyses in Lisrel (Jöreskog & Sörbom, 1996). We did this separately for the lexical decision variables (speed and priming)
and the semantic classification variables (speed and priming). The relationships are modeled as is depicted in Figure 5.3, with minor differences between the two sets of variables. The basic model consists of three predictors of reading comprehension: semantic word knowledge, speed and priming, and we assume that decoding affects speed, and that language background affects semantic word knowledge. It turned out that language background in both analyses had a separate, direct effect on reading comprehension, and also on speed. So, contrary to our study in Chapter 4, the language background effect is not fully mediated by semantic word knowledge.

In the analysis with the lexical decision variables, speed and priming do not make a significant contribution to reading comprehension, and the corresponding regressions (arrows) could be dropped. The model fits the data fairly well ($\chi^2(9)=12.75$, $p=.17$; $CFI=.98$; $RMSEA=.053$). Semantic word knowledge and language background are the sole predictors of reading comprehension and can explain 44% of the variance.

In the analysis with the semantic classification variables, the same models were fitted and it turned out that only priming was not significant; semantic word knowledge and speed (and language background) had to be retained. Again the model fits fairly well ($\chi^2(7)=10.08$, $p=.18$; $CFI=.98$; $RMSEA=.059$). The three predictor variables (semantic word knowledge, speed and language background) can explain 46% of the variance in reading comprehension, of which 2% is uniquely attributable to speed.
Figure 5.3 The path model for the interrelationships of word decoding, language background, semantic word knowledge, speed, priming and reading comprehension. The model shows standardized path coefficients for the variables from the lexical decision task/semantic classification task (LDT/SCT), respectively. Signs are as expected (minus for Speed, and Language Background, which was coded 0=monolingual, 1=bilingual). Priming has no significant effect on Reading, and Speed only has a significant effect in terms of the SCT speed variable. In the LDT analyses, WAT and Speed had a significant residual correlation; in the SCT analyses WAT and RT, and WAT and Priming. Residual correlations are not shown in the figure.
Semantic word knowledge is the main predictor of reading comprehension. Language background predicts semantic word knowledge, and so indirectly affects reading comprehension. In addition, there is a direct link from language background to reading comprehension. In both analyses, there is a (residual) correlation between semantic word knowledge and speed. In the analysis with the semantic classification variables, there also is a significant (residual) correlation between semantic word knowledge and priming.

To investigate whether for different subgroups different path models should be postulated we conducted multigroup analyses, that is, the model is fitted to subgroup data simultaneously and it is tested whether parameters are the same across groups. We tested the model for proficient (higher than the median of 24) (N=59) and less proficient readers (median and lower) (N=71). It turned out that the regressions in the path model are invariant across groups. A multivariate test for differences between models with variant versus invariant regressions showed no significant difference (lexical decision: $\chi^2(7)=12.63$, $p=.08$; semantic classification: $\chi^2(7)=9.27$, $p=.23$). The two groups showed some difference in variance in reading comprehension and semantic word knowledge, but were highly comparable otherwise.

Similar results were found in a comparison between monolingual and bilingual minority children (N = 83 and 47 respectively). The general model now obviously no longer includes language background as a predictor. A multivariate test for differences between models with variant versus invariant regressions showed no significant difference (lexical decision: $\chi^2(4)=3.80$, $p=.43$; semantic classification: $\chi^2(4)=.25$, $p=.99$). Again, there were some differences in the variances between the groups (especially for the lexical decision variables), but the general model holds across the two groups with no additional predictive value (beyond semantic word knowledge) for speed or priming from the lexical decision data and with additional predictive value beyond semantic word knowledge for speed from the semantic classification data, as was already indicated in the whole-group analysis.
5.4 Discussion and conclusion

This study examined the contribution of semantic processing factors in the prediction of reading comprehension scores among 11-year-old monolingual and bilingual minority children. Consistent with the findings in Chapter 4 and with previous research comparing literacy outcomes among monolingual and bilingual minority children, the present study shows a significant advantage for monolingual children for reading comprehension and semantic word knowledge (research question 3), whereas the groups had comparable decoding skills. Regarding processing factors, monolingual children were significantly faster at making lexical decisions and semantic classifications (research question 4). The semantic priming effect that was found in the two tasks was comparable for monolingual and bilingual children (research question 5). Priming effects were not significantly different for reader groups either. Whereas children’ reading comprehension was correlated to their semantic word knowledge and to both lexical decision speed and semantic classification speed, their reading comprehension was only weakly related to decoding. Both speed measures were correlated to decoding and to semantic word knowledge. Surprisingly, reading comprehension and semantic priming were not significantly correlated.

Regarding the prediction of reading comprehension (research question 6), not all expectations were borne out by the data. As in Chapter 4, semantic word knowledge was a significant predictor of reading comprehension. In line with Chapter 4, in which categorization speed contributed to reading comprehension, is the finding that semantic classification speed contributed significant variance to reading comprehension. At the same time, lexical decision speed and semantic priming did not contribute significantly to reading comprehension. In this study, semantic classification speed was a more robust factor in reading performance than priming. Language background still contributed significantly to reading comprehension after the other variables had been taken into account. Path models in LISREL showed that language background has its own direct effect on reading comprehension. This shows that, contrary to the results in Chapter 4, the language background effect is not fully mediated by semantic word knowledge. The absence
of significant interaction effects with language background shows that the contribution of speed or priming and of semantic word knowledge to reading comprehension was comparable for monolingual and bilingual children. This was corroborated by a subsequent multigroup analysis. In line with these findings is an outcome reported for 13-14-year-olds by Van Gelderen, Schoonen, De Glopper, Hulstijn, Snellings, Simis and Stevenson (2003) who found differences in Dutch first language and second language reading comprehension and in English second language and third language reading comprehension but comparable regression models for the language groups. They found no differences between language groups in the patterns of regression weights on the constituent skills linguistic knowledge, speed of processing and metacognition.

We found priming effects for semantically-related, non-associated words in both the semantic classification and the lexical decision task, but we found no overall group difference in priming. Thus, we cannot provide evidence for a difference in underlying semantic representations between the monolingual and bilingual minority children in that respect. We did find a contribution of semantic classification speed beyond semantic word knowledge in explaining reading comprehension scores. This shows that such a processing variable plays a measurable role in reading comprehension. Other studies did find group differences in semantic priming for differently skilled reader groups (Betjemann & Keenan, 2008; Nation & Snowling, 1999). Important here is the fact that studies that investigate the relation between priming and reading comprehension mostly compare strongly differing reader groups, i.e., they compare children with reading comprehension disability to controls (Betjemann & Keenan, 2008). The children in our sample had differing comprehension levels but none of our children had comprehension disorders and all were more or less fluent readers.

Alternatively, our priming tasks may not have been sensitive enough to pick up differences in semantic activation between groups. As a result of piloting the design, we chose to use an interstimulus interval of 2500 ms, as children reported feeling rushed with the shorter interval used in the piloting of the tasks. The interval used here is longer than in some other studies. Nation and Snowling (1999) used a
500 ms interval in a continuous, auditory lexical decision task with children (mean age 10;7). It is possible that, with a shorter interval between prime and target, differences between monolingual and bilingual minority children, if present, could be found. Also, the time frame used in our measurement may capture other activation processes than the ones elicited by Nation and Snowling (1999). In addition, the number of critical priming trials used from the semantic classification task was rather small (five exemplar trials, in the absence of non-exemplar priming). This reduces the reliability of the measure. These factors make it hard to be conclusive about differences in semantic representations for our learners other than that we did not observe a significant group difference. Although the priming effects we found for the monolingual and bilingual minority children were comparable, the bilingual children were slower at making general lexical decisions and semantic classifications than the monolingual children. This points to slower word identification in the face of comparable semantic priming, which may suggest that our priming measurement reflects only one aspect of the “[h]igh-quality [lexical] representations” emphasized by Perfetti and Hart as being required for efficient word identification (2001: 76).

Finally, the priming effect observed is also dependent upon the stimuli used. The children in our sample showed priming for the semantically related, non-associated words. The words we used were high frequent words in a primary-school corpus. It is possible that the words were too simple in the sense that all children had comparable representations for the words. Priming words from a lower frequency range may possibly have elicited more (relevant) individual differences in semantic activation. Considering the three issues discussed, it is difficult, on the basis of our results, to make strong claims about the extent to which the poorer reading comprehension performance of the bilingual children is related to weaker semantic representations. The effect of speed of access to semantic information, however, seems robust enough.

We found small differences between types of priming. It is unclear why we did not find non-exemplar priming. Some studies have reported non-exemplar priming in semantic categorisation tasks for broad categories (animals) but not for
narrow categories (planets, months of the year) (Quinn & Kinoshita, 2008). Furthermore, we found less priming for subordinately related words (bread – food, M=25 ms) than for coordinately related words (nose – ear, M=60 ms) in the lexical decision task. Most studies have found semantic priming effects for words from the same semantic category, i.e. coordinates. Evidence from hierarchically related words is sparse (Hantsch, Jescheniak, & Schriefers, 2005). Hantsch and colleagues showed that lexical competition among semantically related words is not restricted to representations stemming from the same level of abstraction. They show that when an object is named at the basic level (e.g., fish), semantically related words from the subordinate level (e.g., carp) become activated and compete for selection with the target, and vice versa. To the best of our knowledge, there have been no studies that compared same category relations (category coordinates) to hierarchical relations (subordinate relations).

Response speed as measured in the semantic classification task contributed additional variance to reading comprehension beyond decoding and semantic word knowledge. The fact that response speed as measured in the lexical decision task did not contribute to reading performance points to a stronger semantic component in the semantic classification task, and to the importance of fast semantic access for reading comprehension. Both speed measures may well contain a semantic as well as a speed component: they are correlated to both semantic word knowledge and decoding. And while the monolingual and bilingual minority children did not differ in word decoding, they did differ for the two speed measures. At the same time, the absence of an interaction effect with semantic word knowledge shows that semantic classification speed does not depend on children’ level of semantic word knowledge.

To what extent may we expect differences in reading comprehension between monolingual and bilingual children who were nearly all born in the Netherlands and who have all completed at least five years of Dutch primary school? Our correlation and regression results corroborate that semantic word knowledge is an important factor in explaining reading comprehension differences for these fifth-grade children, while word decoding is only of minor influence at this stage. The absence of significant word decoding differences for the groups shows
that reading comprehension differences between these monolingual and bilingual minority children are not due to differences in decoding. This is consistent with research showing that word decoding and reading comprehension become less strongly correlated as children progress through primary school (Curtis, 1980; Sticht & James, 1984; Verhoeven, 1990). The current study also shows that there are still differences between the language groups in reading comprehension after differences in semantic word knowledge and recognition speed have been taken into account. In the study reported in Chapter 4 aspects of semantic word knowledge ‘explained away’ differences between monolingual and bilingual minority children in reading performance. In the current study, there is still room for other factors such as world knowledge, intelligence, or inferencing skill. Netten and colleagues (2011) found differences between four-to-sixth-grade monolingual and bilingual minority children in non-verbal reasoning ability and home language resources, both of which were identified as factors contributing to reading comprehension development.

The main contribution of this study is also its main challenge. Estimating semantic priming effects at the individual level with developing readers - part of whom were second language learners - is relatively new. At the same time, testing semantic priming in a school setting is not as neat a measurement as is obtained in a lab setting. Most psycholinguistic studies into priming use average group scores as a robust way of calculating priming effects. A positive exception to this is work by Larkin and colleagues (1996) who did relate individual priming scores to children’s reading comprehension. Individual priming scores bring along more noise and are less robust than group scores. Whereas we observed a significant priming effect in both tasks, individual scores revealed only small priming differences, and subsequently, small effects on reading comprehension. Studies elaborating this perspective and testing other modalities are needed. Moreover, to be able to draw conclusions about the causality of the interaction between processing speed or priming and reading performance, training studies are required that train children’ word recognition and their ability to understand how words are related (cf. Beck, McCaslin, & McKeown, 1980; Beck & McKeown, 1983; Fukkink, Hulstijn & Simis, 2005; McKeown, Beck, Omanson, & Perfetti, 1983).
In conclusion, the study in this chapter shows that semantic classification speed is a factor in the prediction of reading comprehension scores, in addition to the contribution made by semantic word knowledge. This points to semantic processing as an explanatory intermediating factor between reading comprehension and semantic word knowledge. At the same time, further lab and field studies are necessary to clarify the role of semantic activation in online reading comprehension.
Chapter 6

Semantic word knowledge in reading comprehension: conclusions

6.1 Introduction
The empirical studies in this thesis attempted to enhance the understanding of differences in word knowledge and reading between children and of the processes underlying these differences as well as to provide an insight into the relation between individual differences in word knowledge and reading comprehension. To this end, a series of three studies was conducted in which measures of free word association, reading comprehension, word decoding, semantic word knowledge, semantic categorisation, lexical decision speed, semantic classification speed and semantic priming were administered to several samples of Dutch monolingual and bilingual minority children. For many bilingual minority children in the Netherlands, a language other than Dutch is spoken at home and it is at school that they are first submersed in a primarily Dutch language environment. The specific research interest for this group springs from the substantial and persistent disadvantages that are found for these children with regard to their Dutch language and literacy development in comparison with their native Dutch peers (Cito, 2007), disadvantages that offset the cognitive, cultural and social advantages of growing up as bilinguals. Central to the reported disadvantages of these children are their vocabulary and reading skills. In this final chapter, the main findings of the three studies are summarised and their theoretical and educational implications are
discussed. In the last section, methodological considerations and recommendations for future studies will be provided.

6.2 Summary of main findings
6.2.1 Word associations as a measure of semantic knowledge

Chapter 3 reported on a free word association study that was aimed at eliciting word associations for the target population to be used in the studies in the succeeding chapters. In addition to that, the study investigated qualitative differences between monolingual and bilingual minority children and adults in their word knowledge as reflected by their word association responses to high frequent stimulus words. Findings from previous word association studies provide less clear outcomes than such a straightforward task may lead one to expect. Some studies show more heterogeneous responses for bilingual than for monolingual participants; other studies find phonologically-based responses for young participants, and predominantly semantic responses for native speakers; still other studies fail to obtain clear response preferences at all. Our data came from 422 children (age 8-13) and 54 adults (age 17-59). We found only small differences between language and age groups with respects to the degree of dispersion of their word associations across response categories (research question 1). We found similar patterns of responses for all participants: most association responses of monolingual and bilingual minority children and adults were meaning-related (research question 2). Loglinear model fitting showed that differences related to age were more pronounced than differences related to language background (monolingual vs. bilingual). We found an age effect in the sense that adults gave more indirect meaning-related responses than children. This difference is partly due to the relatively many ‘other’ responses of the (bilingual minority) children, which suggests less familiarity with stimuli and which implies a prominent role for exposure. Within the 17 response subcategories we distinguished, we found some response differences. Within the Direct Meaning-related category, the adults tended to respond more with subordinates and antonyms, while the children preferred partonyms, functionally related words and superordinates. Within the Indirect
Meaning-related category, children mentioned context-dependent characteristics more often than adults. The few Form-based associations were mostly given by the children; the few ‘Other’ associations were mostly given by the bilingual minority children.

Interestingly, the word association data showed no line of development from form-based associations to more abstract, meaning-based responses. Because the adults did not mainly give associations of the most abstract, conceptually meaningful (taxonomic) sort, we cannot consider form-based associations, indirect meaning-related associations and direct meaning-related associations as stages of word knowledge development. This lack of a clear link between response type and proficiency level renders the free word association task not suitable as a proficiency test: respondents’ free associations may reflect exposure and recency rather than semantic capacity. More controlled (e.g., multiple-choice) formats may prove more valuable for testing purposes.

6.2.2 Semantic word knowledge: availability and accessibility

Chapter 4 investigated whether Dutch monolingual and bilingual minority children differ in the availability of their semantic word knowledge and in the speed with which they access lexical and semantic information about (individual) words (research questions 3 and 4). The data showed the importance of availability and to a lesser extent accessibility of semantic word knowledge for reading comprehension through their correlation with and their distinguishable contribution to reading comprehension differences. A regression analysis indicated that both semantic word knowledge and speed of access made unique contributions to reading scores (research question 6). While semantic word knowledge and word decoding (as a control variable) together explained over 31% of individual differences in reading comprehension scores, speed of access accounted for a small but significant amount of extra variance in reading comprehension that could not be accounted for by scores on decoding and semantic word knowledge, bringing the total amount of explained variance to 34%. This shows that both semantic word knowledge (availability) and the ability to access that semantic knowledge quickly (accessibility) play a role in
comprehending text, as proposed by Meara (1996), Nagy & Scott (2000) and Perfetti & Hart (2002). Thus, accessibility supports reading comprehension as a separate, measurable component. At the same time, the role of speed of access to semantic word knowledge for reading comprehension seems to be modest in our sample of children.

In the introduction it was proposed that the reading comprehension delays of bilingual minority children might be related to differences in the availability of semantic word knowledge as well as the accessibility of that knowledge. We found that bilingual minority children lagged behind their monolingual peers in reading comprehension and in availability of semantic word knowledge (WAT scores), which is consistent with earlier findings (August et al., 1999; Schoonen & Verhallen, 1998). The bilingual minority children had less semantic knowledge about words available than their monolingual peers. However, bilingual minority children were only slightly and not significantly slower than monolingual children on speed of access. Importantly, a path analysis showed that differences between monolingual and bilingual minority children in reading comprehension could be perceived as differences mediated by differences in semantic word knowledge. A regression analysis showed that, when controlling for semantic knowledge and, to a lesser extent, speed of access, language background effects on reading comprehension ‘disappear’. Controlling only for speed of access reduces the effect of language background to non-significant. The contribution of availability and accessibility was comparable for the two language groups: we found no significant interaction effects of language background and availability or accessibility, suggesting that the relation between availability, accessibility and reading comprehension can be modelled in the same way for monolingual and bilingual minority children.

When splitting our sample on the basis of reading proficiency, path models showed that the general pattern of relationships holds for both proficient and less-proficient readers. Yet the predictive power of availability and accessibility of semantic word knowledge is, relatively speaking, larger for proficient than for less-proficient comprehenders. A comparison of path models for both reader groups
demonstrated that, in absolute terms, regression weights for availability and accessibility of semantic knowledge were equal, but that, due to a difference in the variance of reading comprehension, the standardized regression weights were different for the two reader groups. In the less proficient group, there was more unexplained variance in reading proficiency than in the proficient group, which may have to do with reading skills other than the ones we have measured, such as inferencing or comprehension monitoring.

In contrast to knowledge of word meaning, word decoding is not generally a problem area for bilingual minority children (cf. Droop & Verhoeven, 2003). Our study corroborated this: we found no statistical differences for word decoding; the bilingual minority children even decoded slightly faster than the monolingual children ($\eta_p^2=.014$), a difference that was only significant for orthographically simple, monosyllabic words. This may suggest that bilingual children read those words more superficially and hence faster. Although fluent word decoding may be a requirement for benefiting from word knowledge (cf. Nakamoto, Lindsey, & Manis, 2008), this small decoding advantage of the bilingual minority children does not make up for their weaker lexical-semantic skills.

### 6.2.3 Semantic representations and activation

Most of the evidence that shows that individual differences in children’s semantic word knowledge are related to differences in reading comprehension (Qian, 1999; Ricketts, Nation, & Bishop, 2007; Schoonen & Verhallen, 1998) comes from offline tasks that require conscious processing. Such findings leave open the possibility that poor comprehenders’ lower performance is simply due to difficulties organising definitions for words or working out the correct answer from multiple options, rather than a reflection of limited on-line use. Thus, on the basis of such offline tasks we cannot draw conclusions about children’s automatic semantic processing. Chapter 5 focused on individual differences in speed of access and semantic priming (research questions 4 and 5). The study in Chapter 5 examined the nature of monolingual and bilingual minority children’ word knowledge using a paradigm that provides an online measure of the effects of semantic similarity on word recognition.
performance: semantic priming. Semantic priming allows for an assessment of the extent to which readers show a normal pattern of priming for different types of semantic relation. Normal priming patterns would suggest that underlying semantic representations function adequately, despite poorer performance on offline tasks such as word definitions.

In addition to lower scores for the bilingual children for reading comprehension and semantic word knowledge - consistent with the data from Chapter 4 - we found that bilingual children were slower both at making lexical decisions and semantic classifications. Semantic classification speed explained additional variance in reading comprehension beyond decoding and semantic word knowledge, but lexical decision speed did not. This confirms the semantic component in the semantic classification task. Contrary to lexical decision, semantic classification contains both a lexical access component and a semantic processing component. The significant contribution of semantic classification speed to reading indicates the importance of fast semantic access for reading comprehension. For reading comprehension, the contributions of word decoding, semantic word knowledge and semantic classification speed together explain 43% of variance. Addition of the interaction term with semantic word knowledge shows that there is no differential effect: the contribution of semantic classification speed is comparable for children with different levels of semantic word knowledge. The variance left in reading scores is attributable to factors other than the ones measured.

The speed difference observed between the monolingual and bilingual children is interesting when compared to the study reported in Chapter 4, where no significant differences between language groups were found for the speed measure. In that study, monolingual and bilingual children were equally accurate and fast at choosing between the two options in the semantic choice task. There are important differences between the semantic choice task in Chapter 4 and the lexical decision and semantic classification task in Chapter 5. The semantic choice task in Chapter 4 involves lexical access - of three words: one stimulus and two targets - as well as a conscious weighing of the two options in relation to the stimulus. As such, it provided a less pure measurement of speed of access. The differences in lexical
access as found in Chapter 5 may simply not have been picked up in the semantic choice task: the task time measured may have consisted mostly of the conscious weighing and decision process. While the monolingual and bilingual minority children appear equally good at choosing between two options, bilingual minority performance drops significantly in a semantic word knowledge task with six options (WAT).

For monolingual as well as bilingual minority children, we found a significant priming effect for semantically-related, non-associated words in our lexical decision task and in our semantic classification task; there was no significant difference in the magnitude of the priming effect for the language groups. Therefore, we cannot provide evidence for a difference in semantic activation between the monolingual and bilingual minority children in that respect. Although we found no differences for language groups, classification speed made a significant contribution to explaining variance in reading comprehension beyond the contributions of decoding and semantic word knowledge. This confirms the findings in Chapter 4. Generally, studies that did find group differences in semantic priming (Betjemann & Keenan, 2008; Nation & Snowling, 1999) compared strongly differing reader groups, such as children with reading comprehension disability versus controls (Betjemann & Keenan, 2008). The children in our sample were all more or less fluent readers, not suffering from specific comprehension disorders. It is possible that for reader groups differing more strongly, we may have found a group difference in priming scores. At the same time, the absence in our study of a link between priming and reading scores, in the face of a study that did find this relationship (e.g., Larkin, Woltz, Reynolds & Clark, 1996), is remarkable. Although it is possible that semantic activation works well for all our children and that lexical-semantic differences concern mainly semantic word knowledge and (conscious) speed of access (speed of lexical decision and semantic classification), it is plausible that there are priming differences that we did not pick up.

We found significant facilitation for lexical decisions to coordinately related targets (nose – ear) and to subordinately related targets (cow – cattle, bread – food) as opposed to unrelated targets. In the semantic classification task we also
found significant facilitation for semantic classifications to coordinately related targets. Facilitation was less for subordinately related words (M=23 ms) than for coordinately related words (M=57 ms) in the lexical decision task. In the semantic classification task, average facilitation for coordinately related words was 71 ms. It is unclear why coordinate relations primed better than subordinate relations. Most studies report semantic priming effects for words drawn from the same semantic category, i.e. coordinates. Fewer studies have investigated hierarchically related words. Hantsch, Jescheniak and Schriefers (2005) found that when an object is named at the basic level (e.g., fish) semantically related words from the subordinate level (e.g., carp) become activated and compete for selection with the target, as well as the other way around. As far as we know, no studies have compared the magnitude of facilitation for same category relations (coordinates) and hierarchical relations (subordinates). Coordinates are not per se closer to each other semantically than subordinates (cf. banana – pear, banana – fruit). Semantic relatedness or meaning overlap is different for each individual word pair.

We found no priming in the semantic classification task for non-exemplar related pairs (‘no’ decisions) (taxi – bus). This may be because only the animal category is activated and decisions are made in relation to this category. The categorical relation of a non-exemplar word to the previous word may be irrelevant, even if it is activated.\(^{14}\) However, in a recent semantic classification study Quinn and Kinoshita (2008) found facilitated ‘no’ decisions for adults when the task instruction focused on animals, which the researchers considered a broad category. They did not find facilitated ‘no’ decisions for what they termed narrow categories (Planets, Months, Relatives). In their discussion, Quinn and Kinoshita suggest that categorisation decisions are made by monitoring the semantic features activated following lexical access, and that broad and narrow categories differ in the semantic features that are selected to be monitored. Our semantic classification task required ‘yes’ decisions to words that were animal names. Priming could have occurred between word pairs that were not animals. It is unclear why for our ‘broad’ category non-exemplar priming was absent.

\(^{14}\) This was suggested by an anonymous reviewer.
We found that reading comprehension differences between the children could be attributed to word decoding only to a small extent. The absence of significant word decoding differences for the groups shows that reading differences between these fifth-grade monolingual and bilingual minority children were not due to differences in decoding. Both Chapters 4 and 5 showed that the contribution of semantic word knowledge to reading comprehension was comparable for monolingual and bilingual minority children, which contrasts with a study by Droop and Verhoeven (2003) in which vocabulary contributed more to reading for Dutch bilingual minority children than for monolingual children. Unlike the data in Chapter 4, the data in Chapter 5 showed that there were still differences between the monolingual and bilingual minority children in reading comprehension after taking into account differences in decoding, semantic word knowledge, semantic categorization speed and priming. The possibility that semantic priming differences underlie the differences in reading comprehension between these monolingual and bilingual children was not borne out by our data. In Chapter 4, word knowledge components did ‘explain away’ differences between monolingual and bilingual minority children in reading performance. A comparison of mean scores and standard deviations for reading comprehension, semantic word knowledge and decoding in Chapters 4 and 5 indicates that the student samples are comparable. However, in Chapter 5, there is still room for factors other than lexical-semantic skills. These may include children’s world knowledge, intelligence or inferencing skill. Netten and colleagues (2011) found differences between four-to-sixth-grade monolingual and bilingual minority children in non-verbal reasoning ability and home language resources. Both of these contributed to reading comprehension development.

In conclusion, we tested the effects of different components of word knowledge and processing in the prediction of reading comprehension differences. Our data show that semantic word knowledge and fast access to semantic word knowledge contribute to differences between children in reading comprehension. These findings indicate the importance of semantics for both word identification and comprehension.
6.3 Theoretical and educational implications

Our results show that semantic word knowledge and fast access to semantic depth have distinguishable relations with reading comprehension. This strengthens the role of depth as a theoretical construct in its own right, next to vocabulary breadth (Schoonen & Verhallen, 1998; Qian, 1999, 2002; Ouellette, 2006). Furthermore, this adds a third dimension to vocabulary, namely the accessibility of semantic word knowledge. The data from Chapters 4 and 5 show that having fast access to semantic word knowledge contributes to reading proficiency. This supports the multidimensional model by Daller, Milton and Treffers-Daller (2007), discussed in Chapter 2.

The three dimensions are not always easy to distinguish, as correlational studies show (Tannenbaum, Torgesen, & Wagner, 2006). It is likely that they grow at a similar rate because of language exposure. Despite high correlations between breadth, depth and fluency, theoretically these dimensions may develop separately. This may be the result of external factors such as a person’s background or education. An uneven vocabulary profile may emerge when someone is taught large quantities of words without considerable meaning elaboration or embedding into existing semantic networks, as was often the case in early educational systems, resulting in a large but superficial vocabulary. Then, the size dimension may well exceed that of depth and fluency. Alternatively, one can imagine someone with a limited vocabulary, perhaps from one specific domain such as animals or the outdoors, with a deep understanding and flexible use of those words they know. This might apply for example, to someone who works with animals or someone who knows a lot about the outdoors. In sum, it is important to be aware of the fact that the three dimensions do not necessarily develop synchronously in all learners.

Having multiple components of word knowledge implies that these can be trained separately and separately affect reading proficiency. Education can choose to target a particular dimension of word knowledge. Enhancing children’s understanding of word meaning and how words are connected requires more attention for rich and varied input at both preschools and primary schools. Valuable class time for word learning might focus on how words relate: how are words
different, how can they be substituted by synonyms or specified through subordinate terms or defining features. In this way, words are not presented in isolation but become embedded in a semantic network (Haastrup & Henriksen, 2000; Read, 2004). Methods using word webs and semantic as well as thematic groupings may be useful here. In this way breadth and depth of word knowledge feed into one another. An example of this is the programme ‘Met Woorden in de Weer’ by Van den Nulft and Verhallen (2001), which focuses on depth of word knowledge and in particular on developing children’s knowledge of how words are embedded in semantic networks. Instruction is aimed at a deeper level of processing. Importantly, talking about and practising with words in different contexts gives teachers a better insight into the lexical-semantic development of children (Droop & Verhoeven, 2003: 101) and allows for delays in word knowledge to surface that would normally be seen. In the English as a foreign language (EFL) context, vocabulary development is also dependent on rich lexical input and interaction (e.g. Ellis & Heimbach, 1997). Gatbonton and Segalowitz (2005) propose a communicative language teaching methodology designed to enhance students’ speaking fluency and emphasize that this teaching method be evaluated in systematic testing. Fluency training in education focuses on productive skills mainly. This raises the question whether fluency training in receptive skills is less beneficial or less measurable. This question deserves serious research attention. Finally, with the widespread presence of computers in schools, the use of online speed measures in (diagnostic) school testing is something that schools may well incorporate into their assessment practices.

This thesis showed that components of lexical-semantic knowledge and processing can explain variance in reading comprehension. Other studies have also shown the contribution of sub skills to higher-order language performance. A Dutch EFL study found that knowledge tests and productive speed tests explained 80% of the variance in writing scores for 13-year-old students (Schoonen, Van Gelderen, De Glopper, Hulstijn, Simis, Snellings & Stevenson, 2003). Knowledge tests and receptive speed tests explained 83% of the variance in EFL reading for the same population (Van Gelderen, Schoonen, De Glopper, Hulstijn, Simis, Snellings &
Stevenson, 2004). Speed of lexical access was a weaker predictor of bilingual minority reading performance than measures of grammatical and vocabulary knowledge and metacognition.

This leaves open the question of causality: is training component skills such as depth or speed beneficial to higher order language performance? Intervention studies have shown that depth and fluency can be trained separately although not all studies have found transfer to higher-order skills. McKeown, Beck, Omanson and Perfetti (1983) investigated the effects of long-term vocabulary instruction on reading comprehension: fourth graders were taught 104 words over a five-month period. The researchers found substantial advantages for the trained group as opposed to the control group on accuracy of word knowledge, speed of lexical access, and comprehension of stories containing taught words. Carlo and colleagues (2004) showed the feasibility of improving reading comprehension scores through systematic bilingual minority vocabulary instruction among fourth- through fifth-grade Spanish-speaking English language children (ELLs). Over the course of a two-year study, those ELLs who received English vocabulary instruction focusing on depth of vocabulary knowledge and word comprehension strategy use performed as well as or better than an English-only control group in areas of word knowledge and reading comprehension. In a study with 13-14-year-old Dutch students with an intermediate level of L2 English, Fukkink, Hulstijn and Simis (2005) investigated the benefits of improving speed of lexical access in L2 English for reading comprehension. The results did show faster lexical access but no transfer of acceleration to reading speed or to higher-order text comprehension. As an explanation, Fukkink and colleagues suggest that “the role of fluent lexical access in L2 reading is too small to expect significant changes in higher-order reading comprehension in view of the complex nature of the reading process” (p. 71). A study with 14-15-year-old Dutch students by Snellings, Van Gelderen and De Glopper (2004) focused on the effects of trained lexical retrieval on EFL writing and shows a similar picture. Snellings and colleagues reported an improvement in lexical retrieval but no transfer to global text quality in short text writing. The training programmes were all relatively short (4 weeks). More time may be needed for
lexical retrieval to transfer to writing. At the same time, trained words were used more in students’ writing, which might be taken as a first step of transfer processes.

6.4 Methodological considerations and suggestions for future research

From the studies in this thesis a number of methodological observations can be made. When investigating semantic word knowledge through spontaneous word associations there is the risk that association responses reflect recent exposure rather than semantic knowledge. The role of instruction is important here. An instruction as open as ‘give the first word that comes to mind’ for a given stimulus word may well elicit words that happen to be still activated in the lexicon at that particular moment rather than words that qualitatively bear the strongest link to the stimulus word. Perhaps other instructions (e.g., ‘Give a (near) synonym for this word, as soon as possible’), making the free association task less ‘free’ may elicit responses that bear a semantic link to the stimulus word. For this reason, responses to free word association tasks cannot be measured along a certain ruler of increasing quality or semantic-relatedness. Hence, free word association tasks are not ideally suited for (language) proficiency testing. To assess differences in semantic word knowledge, more controlled procedures appear more useful, for example assessing whether test takers recognize related words that are presented to them.

Our data showed no differences in semantic priming between monolingual and bilingual minority children or between proficient and less proficient comprehenders. We used an interstimulus interval of 2500 ms, on the basis of a pilot. This interval is somewhat longer than the interval used in other studies (cf. Nation & Snowling, 1999). It would be interesting to administer our tasks with a shorter interval between stimuli. It is possible that with such a stricter measurement window, priming differences between (individual) children may be found. There were differences in priming between our monolingual and bilingual groups, but the standard deviations were rather large (see Table 5.3). Also, our window of measurement may have picked up other activation processes than those measured by Nation and Snowling (1999), who did find differences between readers in semantic
priming. Therefore, we cannot be conclusive about this aspect of individual differences in semantic representations for our children. Alternatively, it may be that our tasks have tapped into processes of co-activation between semantically related words but that reading comprehension differences in this population are due to processes other than lexical-semantic activation.

In addition to the duration of stimulus presentation, the modality in which words are presented is important. Cross-modal priming studies show semantic effects across the auditory and visual modality (Balota, Watson, Duchek, & Ferraro, 1999). Future studies with monolingual and bilingual minority children could extend the current findings by investigating other modalities. For example, Betjemann and Keenan (2008) tested children with reading problems using an auditory measure of semantic word knowledge. Their results suggest that semantic deficits are not only related to the visual mode but constitute a more general semantic problem.

Response behavior in any task is also dependent upon the stimuli used. The stimuli we used were high frequent words from a primary-school corpus. The children in our sample showed priming for semantically related, non-associated words. It is possible that the words were too simple in the sense that the children all had comparable representations or showed the same level of interconnectedness for the words. Priming words from a lower frequency range may possibly have elicited more differences between children in semantic priming as well as speed. In terms of the lexical space metaphor (see Figure 2.1), we may draw different figures for high and low frequent words: for high frequent words the three arrows may be equally long, whereas for low frequent words the depth and fluency arrows may not be fully developed yet since low frequent words take longer to become deeply known and fluently used. Alternatively, for low frequent words the lexical space may look different for different individuals.

Another relevant question is to what extent differences may be expected within our student population. All of these children grew up in the Netherlands, they have become alphabetized in the same way, their literacy development has had more or less the same written input, although the input at home will have been different. The results in this thesis replicate earlier findings that Dutch bilingual minority
children’ reading comprehension is not limited by poor decoding skills. Instead, the acquisition of word meaning and the speed of access to meaning are predictive of comprehension. An important difference between these monolingual and bilingual minority children is the amount of oral language input they receive. Language exposure is important not only purely linguistically but also in the process of abstraction that is part of cognitive development (Nelson, 2007). There are strong indications that reduced Dutch linguistic input is a major factor in explaining differences between monolingual and bilingual minority children in lexical-semantic skills (Messer, 2010).

A question that also deserves mention here concerns the extent to which lexical-semantic weaknesses of Dutch bilingual minority children are a linguistic problem or a matter of cognitive development. To what degree does knowledge of word meaning overlap with what psychologists call verbal intelligence? According to Sternberg, “vocabulary is the best single indicator of one’s overall intelligence” (1987: 89). However, most theories of intelligence comprise many different types of mental abilities, including analytical, problem-solving skills as well as linguistic skills such as verbal comprehension and reasoning (Gardner, 1983; Thurstone, 1938). As such, the verbal subtests of Wechsler’s intelligence scale for children (Wechsler, 1974) include components of world knowledge as well as vocabulary knowledge (In what ways are a lion and a tiger alike? What does ‘flexible’ mean?). It seems undeniable that the measures of word knowledge used in this thesis, which test the recognition of abstract relations between words, tap into some aspects of verbal intelligence and that performance on the WAT and the C-WAT in Chapter 4 is influenced by children’ ability to analyze abstract information. Both tasks do target what may be called an academic skill (Schleppegrell, 2001). This seems less likely for word recognition speed. To target the delays of children with poor semantic word knowledge and reading skill, it matters little whether we speak of lexical or conceptual delays. Understanding of concepts is expressed in relations between words. Language can facilitate children’s understanding of conceptual relations. Whether semantic word knowledge is part of intelligence or not, it is
important to discern which measurable (and trainable) aspects of (linguistic) comprehension feed into children’s reading performance.

Finally, much still needs to be learned about the causal mechanisms that explain the relation between word knowledge and reading comprehension. This thesis shows that speed of access is a small part of that explanation. As mentioned in section 6.2.3, up till now we do not know exactly which sub processes, under which conditions, contribute to higher-order performance. This is even more true for bilingual children, for whom many additional processes are taking place. Data from longitudinal research and training studies are needed to explore how word knowledge and the psycholinguistic processes involved relate to each other, are affected by language input, and eventually contribute to children’s reading performance.
Bibliography


Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19-35). Hillsdale, NJ: Erlbaum.


## Appendix A

Stimulus words used in the free word association test (Chapter 3)

<table>
<thead>
<tr>
<th>Stimuli list (I)</th>
<th>Stimuli list (II)</th>
<th>Stimuli list (II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>optreden</td>
<td>eend</td>
<td>angst</td>
</tr>
<tr>
<td>pen</td>
<td>muur</td>
<td>regen</td>
</tr>
<tr>
<td>computer</td>
<td>bijl</td>
<td>hout</td>
</tr>
<tr>
<td>kever</td>
<td>stelen</td>
<td>afbeelding</td>
</tr>
<tr>
<td>fiets</td>
<td>heelal</td>
<td>eenvoudig</td>
</tr>
<tr>
<td>bloem</td>
<td>schor</td>
<td>brood</td>
</tr>
<tr>
<td>oorbel</td>
<td>zeldzaam</td>
<td>sterk</td>
</tr>
<tr>
<td>wortel</td>
<td>slager</td>
<td>loon</td>
</tr>
<tr>
<td>mus</td>
<td>afval</td>
<td>brommer</td>
</tr>
<tr>
<td>metro</td>
<td>wapen</td>
<td>vis</td>
</tr>
<tr>
<td>kledingstuk</td>
<td>opleiding</td>
<td>geld</td>
</tr>
<tr>
<td>monteur</td>
<td>prijs</td>
<td>sport</td>
</tr>
<tr>
<td>lijm</td>
<td>doos</td>
<td>roest</td>
</tr>
<tr>
<td>agenda</td>
<td>voet</td>
<td>sigaar</td>
</tr>
<tr>
<td>helikopter</td>
<td>tube</td>
<td>zeilboot</td>
</tr>
<tr>
<td>neus</td>
<td>zebra</td>
<td>konijn</td>
</tr>
<tr>
<td>appel</td>
<td>geweldig</td>
<td>dalf</td>
</tr>
<tr>
<td>woestijn</td>
<td>oven</td>
<td>metselaar</td>
</tr>
<tr>
<td>kwast</td>
<td>dal</td>
<td>bedoefd</td>
</tr>
<tr>
<td>nauwelijks</td>
<td>metselaar</td>
<td>schaar</td>
</tr>
<tr>
<td>nep</td>
<td>dal</td>
<td>zwaan</td>
</tr>
<tr>
<td>koelkast</td>
<td>vulkaan</td>
<td>portier</td>
</tr>
<tr>
<td>geheugen</td>
<td>portier</td>
<td>skelet</td>
</tr>
<tr>
<td>oogst</td>
<td>orkest</td>
<td>klok</td>
</tr>
<tr>
<td>rolschaats</td>
<td>snoep</td>
<td>handdoek</td>
</tr>
<tr>
<td>jaargetijd</td>
<td>snoep</td>
<td>vork</td>
</tr>
<tr>
<td>gereedschap</td>
<td>snoep</td>
<td>klokk</td>
</tr>
<tr>
<td>banaan</td>
<td>snoep</td>
<td>supermarkt</td>
</tr>
<tr>
<td>maand</td>
<td>snoep</td>
<td>oceaan</td>
</tr>
<tr>
<td>kleur</td>
<td>snoep</td>
<td>regenboog</td>
</tr>
<tr>
<td></td>
<td></td>
<td>taxi</td>
</tr>
<tr>
<td></td>
<td></td>
<td>zolder</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gebit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>water</td>
</tr>
<tr>
<td></td>
<td></td>
<td>koe</td>
</tr>
<tr>
<td></td>
<td></td>
<td>potlood</td>
</tr>
</tbody>
</table>
Appendix B. Overview of stimuli used in the lexical decision task (Chapter 5). For an English translation, see Appendix C.

**VERSION 1**

<table>
<thead>
<tr>
<th>Primed</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>prime</td>
<td>target</td>
</tr>
<tr>
<td>A</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>neus</th>
<th>oor</th>
</tr>
</thead>
<tbody>
<tr>
<td>park</td>
<td>tuin</td>
</tr>
<tr>
<td>muur</td>
<td>wand</td>
</tr>
<tr>
<td>fiets</td>
<td>brommer</td>
</tr>
<tr>
<td>sigaar</td>
<td>peuk</td>
</tr>
<tr>
<td>brood</td>
<td><strong>voedsel</strong></td>
</tr>
<tr>
<td>boom</td>
<td><strong>plant</strong></td>
</tr>
<tr>
<td>hout</td>
<td><strong>materiaal</strong></td>
</tr>
<tr>
<td>mes</td>
<td><strong>bestek</strong></td>
</tr>
<tr>
<td>regen</td>
<td><strong>neerslag</strong></td>
</tr>
<tr>
<td>vulkaan</td>
<td><strong>berg</strong></td>
</tr>
<tr>
<td>egel</td>
<td><strong>dier</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>koelkast</th>
<th>vriezer</th>
</tr>
</thead>
<tbody>
<tr>
<td>hand</td>
<td>been</td>
</tr>
<tr>
<td>metselaar</td>
<td>schilder</td>
</tr>
<tr>
<td>schaats</td>
<td>slee</td>
</tr>
<tr>
<td>auto</td>
<td>vliegtug</td>
</tr>
<tr>
<td>bijl</td>
<td><strong>gereedschap</strong></td>
</tr>
<tr>
<td>oven</td>
<td><strong>apparaat</strong></td>
</tr>
<tr>
<td>koe</td>
<td><strong>vee</strong></td>
</tr>
<tr>
<td>zwaan</td>
<td><strong>vogel</strong></td>
</tr>
<tr>
<td>metro</td>
<td><strong>vervoer</strong></td>
</tr>
<tr>
<td>schoen</td>
<td><strong>kleding</strong></td>
</tr>
<tr>
<td>monteur</td>
<td><strong>beroep</strong></td>
</tr>
</tbody>
</table>

**VERSION 2**

<table>
<thead>
<tr>
<th>Primed</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>prime</td>
<td>target</td>
</tr>
<tr>
<td>B</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>koelkast</th>
<th>vriezer</th>
</tr>
</thead>
<tbody>
<tr>
<td>hand</td>
<td>been</td>
</tr>
<tr>
<td>metselaar</td>
<td>schilder</td>
</tr>
<tr>
<td>schaats</td>
<td>slee</td>
</tr>
<tr>
<td>auto</td>
<td>vliegtug</td>
</tr>
<tr>
<td>bijl</td>
<td><strong>gereedschap</strong></td>
</tr>
<tr>
<td>oven</td>
<td><strong>apparaat</strong></td>
</tr>
<tr>
<td>koe</td>
<td><strong>vee</strong></td>
</tr>
<tr>
<td>zwaan</td>
<td><strong>vogel</strong></td>
</tr>
<tr>
<td>metro</td>
<td><strong>vervoer</strong></td>
</tr>
<tr>
<td>schoen</td>
<td><strong>kleding</strong></td>
</tr>
<tr>
<td>monteur</td>
<td><strong>beroep</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
</tr>
</tbody>
</table>
Non-words (51)
rekald, kal, huuf, soulooi, koppe, ougel, preelik, arf, onaak, mamits, pruit, bruf, stuup, nuveler, rouler, nefersal, walerozer, saar, paluuk, gekee, colus, doisterijn, ekeest, schef, lopei, schowaal, bostaf, wazin, madiek, samelier, rexen, boens pakella, madioor, stofaar, zosje, ulijk, skelt, leeler, teuzer, werreg, honantie, korer, onterat, wepig, wrantaar, kron, spoen, plaak, broom, cangalier

Fillers (27)

muziek, snoep, familie, maaltijd, mens, vis, vloeistof, instrument, sieraad, groente, bloem, voorwerp, cultuur, vrucht, echo, nacht, rozijn, spreuk, kok, streep, tarwe, wonder, tunnel, feest, kasteel, volgorde, rotzooi

English translation of fillers (27)

music, candy, family, meal, person, fish, liquid, instrument, jewelry, vegetable, flower, object, culture, fruit, echo, night, raisin, spell, cook, line, wheat, wonder, tunnel, feast, castle, order, mess

Note: Version 1 and 2 contain the same 48 critical words and the same non-words and fillers. Both versions contain 48+51+27=126 stimuli. Words in bold print are category names.
Appendix C. Translation of stimuli used in the lexical decision task (Chapter 5). For the non-words and fillers and the Dutch stimuli, see Appendix B.

**VERSION 1**

<table>
<thead>
<tr>
<th>Primed</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>prime</strong></td>
<td><strong>target</strong></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>nose</td>
<td>ear</td>
</tr>
<tr>
<td>park</td>
<td>garden</td>
</tr>
<tr>
<td>wall</td>
<td>[thin wall]</td>
</tr>
<tr>
<td>bike</td>
<td>moped</td>
</tr>
<tr>
<td>cigar</td>
<td>fag</td>
</tr>
<tr>
<td>bread</td>
<td>food</td>
</tr>
<tr>
<td>tree</td>
<td>plant</td>
</tr>
<tr>
<td>wood</td>
<td>material</td>
</tr>
<tr>
<td>knife</td>
<td>cutlery</td>
</tr>
<tr>
<td>rain</td>
<td>[rainfall]</td>
</tr>
<tr>
<td>volcano</td>
<td>mountain</td>
</tr>
<tr>
<td>hedgehog</td>
<td>animal</td>
</tr>
</tbody>
</table>

**VERSION 2**

<table>
<thead>
<tr>
<th>Primed</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>prime</strong></td>
<td><strong>target</strong></td>
</tr>
<tr>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>fridge</td>
<td>freezer</td>
</tr>
<tr>
<td>hand</td>
<td>leg</td>
</tr>
<tr>
<td>bricklayer</td>
<td>painter</td>
</tr>
<tr>
<td>skate</td>
<td>sledge</td>
</tr>
<tr>
<td>car</td>
<td>airplane</td>
</tr>
<tr>
<td>axe</td>
<td>tool</td>
</tr>
<tr>
<td>oven</td>
<td>[machine]</td>
</tr>
<tr>
<td>cow</td>
<td>cattle</td>
</tr>
<tr>
<td>swan</td>
<td>bird</td>
</tr>
<tr>
<td>metro</td>
<td>transport</td>
</tr>
<tr>
<td>shoe</td>
<td>clothing</td>
</tr>
<tr>
<td>mechanic</td>
<td>profession</td>
</tr>
</tbody>
</table>

*Note: Version 1 and 2 contain the same 48 critical words and the same non-words and fillers. Both versions contain 48+51+27=126 stimuli. Words in bold print are category names. [ ] = approximate translation*
Appendix D. Overview of stimuli used in the semantic classification task (Chapter 5). For an English translation, see Appendix E.

**VERSION 1**

<table>
<thead>
<tr>
<th>Prime</th>
<th>Target</th>
<th>Prime</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>eend</td>
<td>gans</td>
<td>oceaan</td>
<td>giraf</td>
</tr>
<tr>
<td>leeuw</td>
<td>tijger</td>
<td>schip</td>
<td>rups</td>
</tr>
<tr>
<td>kat</td>
<td>panter</td>
<td>kaars</td>
<td>merel</td>
</tr>
<tr>
<td>vlinder</td>
<td>mug</td>
<td>banaan</td>
<td>stier</td>
</tr>
<tr>
<td>konijn</td>
<td>haas</td>
<td>potlood</td>
<td>vos</td>
</tr>
<tr>
<td>boek</td>
<td>krant</td>
<td>hert</td>
<td>kano</td>
</tr>
<tr>
<td>kwast</td>
<td>penseel</td>
<td>slak</td>
<td>peer</td>
</tr>
<tr>
<td>trein</td>
<td>boot</td>
<td>olifant</td>
<td>vijver</td>
</tr>
<tr>
<td>appel</td>
<td>aardbei</td>
<td>mus</td>
<td>stift</td>
</tr>
<tr>
<td>taxi</td>
<td>bus</td>
<td>koe</td>
<td>lamp</td>
</tr>
</tbody>
</table>

**VERSION 2**

<table>
<thead>
<tr>
<th>Prime</th>
<th>Target</th>
<th>Prime</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td></td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>olifant</td>
<td>giraf</td>
<td>trein</td>
<td>gans</td>
</tr>
<tr>
<td>slak</td>
<td>rups</td>
<td>kwast</td>
<td>tijger</td>
</tr>
<tr>
<td>mus</td>
<td>merel</td>
<td>boek</td>
<td>panter</td>
</tr>
<tr>
<td>koe</td>
<td>stier</td>
<td>taxi</td>
<td>mug</td>
</tr>
<tr>
<td>hert</td>
<td>vos</td>
<td>appel</td>
<td>haas</td>
</tr>
<tr>
<td>schip</td>
<td>kano</td>
<td>eend</td>
<td>krant</td>
</tr>
<tr>
<td>banaan</td>
<td>peer</td>
<td>kat</td>
<td>penseel</td>
</tr>
<tr>
<td>oceaan</td>
<td>vijver</td>
<td>leeuw</td>
<td>boot</td>
</tr>
<tr>
<td>potlood</td>
<td>stift</td>
<td>vlinder</td>
<td>aardbei</td>
</tr>
<tr>
<td>kaars</td>
<td>lamp</td>
<td>konijn</td>
<td>bus</td>
</tr>
</tbody>
</table>
Fillers (66)

bed, snor, suiker, mand, raadsel, oever, kaas, heelal, toneel, zwerver, ballon, luik, karton, nar, liefde, volk, boer, natuur, tent, riem, strijder, indiaan, oogst, veld, thee, oma, huwelijk, bliksem, koning, droom, bal, herfst, sjaal, kreeft, haan, mol, hagedis, eland, haai, pinguin, kraai, zwijn, muis, kip, tekker, paard, parkiet, tor, meeuw, varken, pauw, krokodil, duif, dolfijn, kameel, specht, wolf, wasp, krekel, aal, schaap, vlo, poedel, baviaan, snoek, kikker

English translation of fillers (66)

bed, moustache, sugar, basket, riddle, bank, cheese, universe, stage, tramp, balloon, hatch, cardboard, jester, love, people, farmer, nature, tent, belt, warrior, indian, harvest, field, tea, grandma, marriage, lightning, king, dream, ball, autumn, scarf, crab, rooster, mole, lizard, elk, shark, penguin, crow, boar, mouse, chicken, dachshund, horse, parakeet, beetle, seagull, pig, peacock, crocodile, pidgeon, dolfijn, camel, woodpecker, wolf, wasp, cricket, eel, sheep, flee, poodle, baboon, pine, frog

Note: Version 1 and 2 contain the same 40 critical words. Both versions contain 40+66 = 106 stimuli. Words in bold print are exemplar names.
Appendix E. Translation of stimuli used in the semantic classification task (Chapter 5). For the fillers and the Dutch stimuli, see Appendix D.

### VERSION 1

<table>
<thead>
<tr>
<th>Primed</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>prime</strong></td>
<td><strong>target</strong></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>duck</td>
<td>goose</td>
</tr>
<tr>
<td>lion</td>
<td>tiger</td>
</tr>
<tr>
<td>cat</td>
<td>panther</td>
</tr>
<tr>
<td>butterfly</td>
<td>mosquito</td>
</tr>
<tr>
<td>rabbit</td>
<td>hare</td>
</tr>
<tr>
<td>book</td>
<td>newspaper</td>
</tr>
<tr>
<td>[brush]</td>
<td>paint brush</td>
</tr>
<tr>
<td>train</td>
<td>boat</td>
</tr>
<tr>
<td>apple</td>
<td>strawberry</td>
</tr>
<tr>
<td>taxi</td>
<td>bus</td>
</tr>
</tbody>
</table>

### VERSION 2

<table>
<thead>
<tr>
<th>Primed</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>prime</strong></td>
<td><strong>target</strong></td>
</tr>
<tr>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>elephant</td>
<td>giraffe</td>
</tr>
<tr>
<td>snail</td>
<td>caterpillar</td>
</tr>
<tr>
<td>sparrow</td>
<td>blackbird</td>
</tr>
<tr>
<td>cow</td>
<td>bull</td>
</tr>
<tr>
<td>deer</td>
<td>fox</td>
</tr>
<tr>
<td>ship</td>
<td>canoe</td>
</tr>
<tr>
<td>banana</td>
<td>pear</td>
</tr>
<tr>
<td>ocean</td>
<td>pond</td>
</tr>
<tr>
<td>pencil</td>
<td>marker</td>
</tr>
<tr>
<td>candle</td>
<td>lamp</td>
</tr>
</tbody>
</table>

*Note: Version 1 and 2 contain the same 40 critical words. Both versions contain 40+66 = 106 stimuli. Words in bold print are exemplar names. [ ] = approximate translation*
Summary in English

Word knowledge is one of the factors that determine children’s reading proficiency and by extension their school success. In order to understand text, children not only need a sufficiently large vocabulary, they also need a good understanding of the meaning of the words they read. Word meaning can be viewed as a collection of meaning aspects, connected in a mental network. Of those meaning aspects there are semantic aspects that belong to the word’s core meaning and there are more context-specific aspects that only go with the word in a certain context. For example, for the meaning of the word teeth concepts such as jaw and mouth are important. Candy and toothbrush can also be connected with teeth but are not so much semantically but rather contextually related to it. Although the difference between semantic and contextual word knowledge is a gradient one there are indications that an understanding of the semantic meaning aspects of words may be particularly relevant for reading comprehension. This may be because semantic aspects (mouth) are relevant in any (con)text as opposed to context-specific aspects (candy). In this thesis it was investigated whether and to what extent semantic knowledge of concrete words and fast and automatic access to that knowledge contributes to children’s reading comprehension.

Considerable differences have been found with respect to children’s knowledge of word meaning. Children’s word knowledge is investigated with instruments such as definition tasks or with written tasks in which children have to connect words to other words. Differences are found in particular between monolingual Dutch children and children from immigrant families who speak another language at home than the language used at school. Data from the Dutch annual nation-wide educational assessments by Cito show that the Dutch language proficiency of for example Dutch pupils who speak Arabic or Turkish at home is below the national average at the end of primary school. Linguistic research has shown that bilingual minority children mention fewer meaning aspects per word in definition tasks and that the meaning aspects they mention are more contextual than semantic (Verhallen & Schoonen, 1993). For example, to the question, ‘What is a
nose?’ a typical answer from the bilingual minority group was, ‘You can smell with a nose’, whereas a typical answer from the monolingual group was, ‘A nose is a body part with which you can smell and breathe’. The use of semantic meaning aspects such as body part would be indicative of a more developed semantic word knowledge and a certain kind of network of meaning relations. A standardised word association task (WAT, Schoonen & Verhallen, 2008) showed similar differences. In this task, for a given stimulus word children were asked to choose from six words the three words that were most strongly related in meaning to the stimulus. This means they had to compare the meanings of the six words to that of the stimulus word. Differences between children in recognizing the related words in this task turned out to correlate with differences between the children in reading comprehension (Schoonen & Verhallen, 1998).

Although the importance of word knowledge for reading comprehension seems obvious, the exact relationship between the two constructs is not very clear. Is it sufficient for children to know the meaning of words or is it also important for them to understand relations between words? Does the speed with which children access their word knowledge have any added value? There are only a few studies available that relate individual differences in the speed with which meaning is activated to reading performance. Moreover, it is not known to what extent differences in the processing of word knowledge play a role in reading delays at school. The research reported in this thesis focuses on the question to what extent differences between children in semantic word knowledge and in the speed with which they use this knowledge can explain differences in reading comprehension. Differences in semantic word knowledge and in the accessibility of this knowledge have been investigated in this thesis for Dutch monolingual and Dutch bilingual minority children. An answer to this question contributes to our understanding of the exact relation between word knowledge and reading comprehension. Beside this theoretical relevance, understanding the relation between word knowledge and reading comprehension can contribute to the design of programs targeting reading delays in primary education.
To answer the question outlined above, in this thesis the word knowledge of monolingual and bilingual minority children was investigated in three studies. The following questions guided this investigation:

I. Are there structural differences in the word associations of monolingual and bilingual minority children and adults?

II. Does fast access to semantic word knowledge play a role in children’s reading comprehension?

III. Does automatic activation of the so-called semantic network contribute to children’s reading comprehension?

Part I: Are there structural differences in the word associations of monolingual and bilingual minority children and adults?

Differences in word knowledge have to do with the plurality of meanings and meaning aspects of a word that are known to somebody (for example dog, Snuggles, animal, pet, barking, doggie bag, etc.). When acquiring their mother tongue children’s word knowledge starts out context-specific and personal (dog – Snuggles). As children encounter words more often and in more contexts they will start seeing the connections between words and abstract the more general meaning aspects, which are expressed in connection with other words (dog – animal). At school, children’s word knowledge is deepened through the addition of new meanings, concepts and relations. Delays in word knowledge are found for bilingual minority children who speak another language than Dutch at home. They arrive at school with a Dutch vocabulary that is too limited to sufficiently deepen their word knowledge. If the semantic word knowledge of bilingual minority children is indeed less well developed, this may show in the kinds of words they associate with each other. As a preparation for the experiments in Chapters 4 and 5 of this thesis we collected word associations of children. These word associations allow us to investigate differences between children in semantic and contextual word knowledge. A word association task was administered to 422 monolingual Dutch and bilingual Turkish-Dutch children in grades 4-6 of several primary schools (age 8-13). Participants wrote down the first word that came to mind for 59 simple Dutch
words such as *cat*, *rainbow* and *taxi*. We expected that monolingual Dutch children would give more abstract, semantic meaning aspects than bilingual minority children. For comparison the word association task was also administered to 54 monolingual Dutch and bilingual Turkish-Dutch adults (age 17-59). The association responses of the children and adults were classified into three main association categories that were subdivided into 17 subcategories. The main categories were: semantic (directly related in meaning), contextual (indirectly related in meaning) and form based. To investigate the homogeneity of responses and possible response patterns and preferences, the responses were analysed using a concentration index and log linear models.

The results of this first study, reported in Chapter 3, show that there is a little more variation in the associations of the children than in those of the adults, which is mainly attributable to the children’s more frequent use of form-based associations (*olifant* – *olie* [elephant – oil]). The monolingual children and adults gave slightly more responses that were semantically related than the bilinguals did. The adults as a group gave more context-specific associations than semantic associations whereas the children gave more or less comparable proportions of both types. This is surprising and seems to indicate that semantic associations are not per se the optimum in terms of word knowledge development. The different response categories - semantic, contextual, form based - are not necessarily indicative of a scale of development. At the same time, the absence of semantic associations (in the adult group) need perhaps not surprise us since other factors than conceptual development (for example recent word use) may influence the meaning aspects that are activated in the mental lexicon and that are triggered in a spontaneous association task. This renders the free association task less suitable as a test of word knowledge development and emphasizes the need for more structured tasks that specifically target (the recognition of) semantic meaning aspects of words.
Part II: Does fast access to semantic word knowledge play a role in children’s reading comprehension?

To be able to use word knowledge efficiently it is important to have fluent access to it. When recognizing a word during reading we activate this word in our mental lexical-semantic network. The better words are integrated in the network (i.e., the better connected to other words), the easier and faster we have access to the knowledge that is attached to it. Studies into semantic fluency that ask children to name as many words of a certain category - animals, modes of transport - as fast as possible show that differences in fluency can be related to reading comprehension scores. This points to fast access to semantic word knowledge as a supportive factor in reading comprehension. However, most research into word knowledge, especially studies in an educational context, uses written tasks – for example definition tasks - that do not include a time measurement. Research that does not include a time dimension says little about how fast or easily children process or use word knowledge. Because of this it is unclear whether differences between children on such tasks have to do with differences in fast access to word knowledge – possibly due to differences in underlying semantic networks. Are children who score well on word knowledge tasks that do not include a time dimension simply better at consciously figuring out word tasks and formulating definitions? Or do these children score better due to faster access to their word knowledge, possibly because their semantic networks are different causing related concepts to be co-activated more easily?

To investigate the role of semantic word knowledge and speed of access in reading comprehension, in the second study, described in Chapter 4, we administered several word knowledge and reading tasks to 135 Dutch monolingual and bilingual minority children in grade 5 (age 10-11) of several Dutch primary schools. To measure speed of access to semantic word knowledge, a semantic choice task was designed and administered individually with laptops. In this task, for each stimulus word (e.g., mus [sparrow]), children had to decide as fast as possible which of two presented words (boom - vleugel [tree – wing]) goes best with the stimulus. Both options could be associated with the stimulus word but only one was
semantically related. In the task we measured whether and how fast children chose the semantic meaning aspect (vleugel [wing]). Beside this computer task, a written semantic word knowledge task (the WAT) and a reading comprehension task were administered. As a control measure for word decoding a well-known word decoding task (the Drie-Minuten-Toets) was administered to each child individually. The contributions of word decoding, semantic word knowledge and speed of access to explaining variance in reading comprehension scores were investigated with regression analyses and path models that reflect the relationships between all constructs.

The results, reported in Chapter 4, show that the bilingual children indeed scored lower than the monolingual children on reading comprehension, although their decoding scores were slightly higher. Their scores on the written semantic word knowledge task were also lower than for the monolingual children. In the computer task the two groups of children were equally accurate and fast in identifying the semantically related word. Differences between the children in word decoding, semantic word knowledge and speed together explained 34% of the variance in reading scores. Inclusion of the interaction term between language background and semantic word knowledge or speed did not explain extra variance in reading comprehension, which indicates that the contributions of semantic word knowledge and speed were comparable for the monolingual and bilingual children. The factor language background could not explain additional variance in reading after semantic word knowledge had been taken into account. This means that differences in reading comprehension scores between the monolingual and the bilingual minority children can be accounted for by semantic word knowledge. A path model shows that the effect of language background indeed runs via semantic word knowledge. These results show that knowledge of semantic meaning aspects is relevant and that fast access to it also contributes, albeit to a small extent, to reading comprehension. A division of the children in proficient and less proficient comprehenders demonstrates that semantic word knowledge and accessibility explain a little more variance in reading for the proficient comprehenders.
Part III: Does automatic activation of the so-called semantic network contribute to children’s reading comprehension?

After establishing the importance of semantic word knowledge for reading comprehension and the small added value of speed, we further investigated the role of accessibility in the third study. This study is described in Chapter 5. We wanted to know whether good comprehenders not only have fast access to semantic word knowledge but also automatically activate it based on word relatedness. The notion of what is called ‘semantic priming’ is relevant here. Priming studies measure the extent to which one word as it were co-activates another word. A faster reaction to a word (oor [ear]) preceded by a related word (neus [nose]) than to that word preceded by an unrelated word (fiets [bike]) is referred to as a priming effect. Psycholinguistic research with children shows that children who score low on reading comprehension on average demonstrate less co-activation of semantically related words in priming tasks. This is taken as an indication that these semantic relations are less developed in the children’s semantic networks. However, there is little research available that has measured to what extent individual differences in semantic priming can account for individual differences in reading comprehension. As such, it is unclear to what extent the reading delays of bilingual minority children are related to differences in the use of semantic word knowledge, such as the automatic activation of that semantic knowledge. Moreover, it is not known whether the relation between semantic word knowledge, automatic activation and reading comprehension is the same for monolingual children and bilingual minority children.

In the third study, we investigated the activation of semantic word knowledge with two priming tasks: a lexical decision task and a semantic classification task. These were administered on laptops so that response speed could be measured and semantic priming could be calculated. In the lexical decision task children had to indicate for each stimulus that appeared on the screen whether it was an existing Dutch word or not. Hence, this task also included non-words. In the semantic classification task children had to indicate for each presented word whether it was an animal name or not. Both tasks were designed not so much to determine
whether the children made the right choice or not, but rather to establish the extent to which they showed priming for words that appeared consecutively and that were semantically related (e.g., nose – ear or knife – cutlery). Speed in these two tasks was also included as a variable. In this third study five tasks were administered to a new group of 130 monolingual and bilingual minority children in grade 5 of several Dutch primary schools: two priming tasks, the Drie-Minuten-Toets for word decoding, the WAT for semantic word knowledge, and the reading comprehension task that was also used in the study in Chapter 4. The contributions of decoding, semantic word knowledge, lexical decision speed, semantic classification speed and semantic priming to explaining variance in reading comprehension were investigated using regression analyses and path models.

The results show a semantic priming effect for both the monolingual and bilingual children: both groups responded faster to words preceded by semantically related words than to words preceded by unrelated words. As in the second study, in this study the bilingual children scored lower on reading comprehension and semantic word knowledge than the monolingual children, but the priming scores of the two groups were comparable. Again, differences in reading were not attributable to word decoding: the two groups had a comparable decoding performance. As in the previous experiment, semantic word knowledge explained a considerable amount of variance in reading comprehension and speed could add some unique extra variance. This held for both the monolingual and bilingual children. However, only semantic classification speed, not lexical decision speed, explained variance in reading scores. Semantic priming did not explain variance in reading. In this third study, language background could still explain some variance in reading comprehension, after semantic word knowledge had been taken into account. A path model confirms this: language background has an effect on semantic word knowledge and speed as well as a direct effect on reading comprehension. Differences in reading comprehension between the two language groups can thus not be completely explained by semantic word knowledge. It appears that in this sample there were still other differences between the language groups that played a role. The results of this third study confirm the importance of semantic word
knowledge for reading comprehension and the small contribution of speed of access. A contribution of priming effects could not be indicated.

Conclusions and implications
The data in this thesis show that semantic word knowledge is important for the reading comprehension of both monolingual and bilingual minority children and that the speed with which meaning is recognised plays a role in this. Where bilingual children lag behind in knowledge of word meaning and in speed, they also lag behind in reading comprehension. Children show comparable activation of semantically related words. In this study semantic priming did not explain variance in reading comprehension. Future research with a more controlled design, also including other modalities (for example auditory priming), may say more about the role of automatic activation in explaining differences in reading comprehension in this population. It is important for both monolingual and bilingual children that they know not only the meaning of a word but also the meaning relations that this word has with other words and that they, when encountering words, quickly see which words are related.
Woordkennis is een van de factoren die de leesvaardigheid en daarmee het schoolsucces van kinderen bepalen. Voor een goed leesbegrip moeten kinderen niet alleen voldoende woorden kennen, ze moeten de betekenis ook goed kennen. Woordbetekenis kun je zien als een verzameling betekenisaspecten, in de vorm van een netwerk en opgeslagen in ons geheugen. Binnen die verzameling zijn er semantische betekenisaspecten die tot de kernbetekenis van het woord horen en meer context-specifieke aspecten die alleen in een bepaalde context bij het woord horen. Zo zijn voor de betekenis van het woord *gebit* begrippen als tand, kaak, en mond belangrijk. Snoepen en tandenborstel hangen ook samen met *gebit* maar zijn er niet (zozeer) semantisch maar meer contextueel mee verbonden. Hoewel het verschil tussen semantische en contextuele woordkennis gradueel is, zijn er aanwijzingen dat een goed begrip van juist de semantische betekenisaspecten van woorden relevant is voor leesbegrip. Dit kan zijn omdat semantische betekenisaspecten (tand) in elke (con)text relevant zijn, in tegenstelling tot context-specifieke informatie (snoepen). In dit proefschrift is onderzocht of en in welke mate semantische kennis van concrete woorden en het snel en automatisch paraat hebben daarvan bijdraagt aan het leesbegrip van kinderen.

Er blijken behoorlijke verschillen te zijn tussen kinderen in wat ze van de betekenis van woorden weten. De woordkennis van kinderen wordt bijvoorbeeld onderzocht met definitietaken of schriftelijke taken waarin kinderen woorden met elkaar moeten verbinden. Verschillen worden met name gevonden tussen eentalige Nederlandse kinderen en kinderen uit migrantengezinnen die thuis een andere taal spreken dan de taal die op school wordt gebruikt. De gegevens uit de periodieke onderwijspeilingen van het Cito laten bijvoorbeeld zien dat de Nederlandse taalvaardigheid van leerlingen die thuis Arabisch of Turks spreken aan het eind van de basisschool lager ligt dan het landelijk gemiddelde. Taalkundig onderzoek wijst uit dat tweetalige kinderen uit migrantengezinnen in definitietaken minder betekeniselementen per woord noemen en dat de betekeniselementen die ze noemen meer context-specifiek dan semantisch zijn (Verhallen & Schoonen, 1993).
Bijvoorbeeld bij de vraag, ‘Wat is een neus?’ was een typisch antwoord uit deze groep tweetaligen ‘Met een neus kan je ruiken’, terwijl een typisch antwoord van een eentalig kind was ‘Een neus is een lichaamsdeel waarmee je kunt ruiken en ademen’. Het gebruik van abstracte betekeniselementen zoals *lichaamsdeel* zou wijzen op een beter ontwikkelde semantische woordkennis en een bepaald soort netwerk van betekenisrelaties. Met een gestandaardiseerde woordassociatietaak (WAT, Schoonen & Verhallen, 2008) worden vergelijkbare verschillen gevonden. Kinderen moesten in de woordassociatietaak voor een bepaald stimuluswoord uit zes omringende woorden de drie sterkst gerelateerde woorden kiezen en dus bewust betekenissen afwegen. De verschillen in deze taak in het herkennen van de woorden die in betekenis overeenkwamen met het stimuluswoord bleken samen te hangen met verschillen in leesbegrip (Schoonen & Verhallen, 1998).

Hoewel het belang van woordkennis voor begrijpend lezen logisch lijkt, is de exacte relatie niet geheel duidelijk. Is het voldoende dat kinderen de betekenis van woorden kennen of moeten ze ook verbanden tussen woorden zien? Heeft de snelheid waarmee kinderen woordbetekenis paraat hebben nog een meerwaarde? Er zijn slechts enkele studies beschikbaar die individuele verschillen in de snelheid waarmee woordbetekenis geactiveerd wordt relateren aan lezen. Bovendien is niet bekend in hoeverre het verwerken en gebruiken van woordkennis een rol speelt bij leesachterstanden in het onderwijs. Het onderzoek in dit proefschrift richt zich dan ook op de vraag *in hoeverre verschillen tussen kinderen in semantische woordkennis en in de snelheid waarmee ze deze kennis paraat hebben een verklarende factor zijn voor verschillen in begrijpend lezen*. Verschillen in semantische woordkennis en in de snelheid of automatische toegang tot deze kennis zijn in dit proefschrift onderzocht voor eentalige Nederlandse kinderen en tweetalige kinderen uit migrantengezinnen. Een antwoord op deze vraag draagt bij aan ons begrip van de exacte relatie tussen woordkennis en leesbegrip. Naast dit theoretisch belang kan een beter begrip van de relatie tussen woordkennis en leesbegrip ook bijdragen aan het remediëren van leesachterstanden in het basisonderwijs.
Om bovenstaande vraag te kunnen beantwoorden is in dit proefschrift de woordkennis van eentalige kinderen en tweetalige kinderen uit migrantengezinnen onderzocht door middel van een driedelig onderzoek. De volgende vragen lagen aan dit onderzoek ten grondslag:

I. Zijn er structurele verschillen in de woordassociaties van een- en tweetalige kinderen en volwassenen?

II. Speelt snelle toegang tot semantische woordkennis een rol in het leesbegrip van kinderen?

III. Draagt automatische activatie van het zogenaamde semantisch netwerk bij aan het leesbegrip van kinderen?

Deel I: Zijn er structurele verschillen in de woordassociaties van een- en tweetalige kinderen en volwassenen?

Verschillen in woordkennis hebben te maken met de veelheid en verscheidenheid aan betekenissen of betekenisaspecten van een woord die iemand kent (bijvoorbeeld hond, fikkie, dier, huisdier, blaffen, viervoeter, hondenweer, etc.). Bij het leren van hun moedertaal verwerven kinderen in eerste instantie context-specifieke woordkennis en is hun woordkennis nog erg persoonlijk (hond - fikkie). Naarmate zij woorden vaker en in meer contexten tegenkomen leggen ze steeds meer verbanden tussen woorden en destilleren ze als het ware de algemenere betekenisaspecten, die in de samenhang met andere woorden tot uitdrukking komen (hond - dier). Op school wordt de aanwezige woordkennis steeds verder uitgediept door de toevoeging van nieuwe betekenissen, concepten en verbanden. Achterstanden in woordkennis worden gesignaleerd bij tweetalige kinderen met een andere thuistaal dan het Nederlands. Zij komen op school met een te kleine woordenschat om hun woordkennis voldoende uit te diepen. Als semantische woordkennis minder ontwikkeld is bij tweetalige kinderen zou dit kunnen blijken uit de woorden die zij met elkaar associëren. Ter voorbereiding op de experimenten in Hoofdstuk 4 en 5 van dit proefschrift zijn woordassociaties van kinderen verzameld. Deze woordassociaties stellen ons in de gelegenheid om de verschillen te onderzoeken tussen kinderen wat betreft semantische en contextuele woordkennis.
Een woordassociatietaak werd voorgelegd aan 422 eentalige Nederlandse en tweetalige Turks-Nederlandse kinderen in groep 6, 7 en 8 van verschillende basisscholen (leeftijd 8-13 jaar). Zij schreven het eerste woord op dat bij hen opkwam voor 59 stimuluswoorden, zoals *kat, regenboog* en *taxi*. De verwachting was dat eentalige kinderen meer abstracte, semantische betekenisaspecten zouden noemen dan tweetalige kinderen. Ter vergelijking hebben we de taak ook aan 54 Nederlandse en Turks-Nederlandse volwassenen voorgelegd (leeftijd 17-59 jaar). De associatieresponsen van de kinderen en volwassenen werden geclassificeerd in drie hoofdtypen associaties met daarbinnen 17 subcategoriën. De hoofdtypen waren: semantisch (direct betekenisgerelateerd), contextueel (indirect betekenisgerelateerd) en vormgerelateerd. Om zicht te krijgen op de spreiding van de responsen en op patronen en voorkeuren werden de associatieresponsen geanalyseerd met behulp van een concentratie-index en loglineaire modellen.

De resultaten van dit eerste onderzoek staan beschreven in Hoofdstuk 3. Ze laten zien dat er iets meer variatie zit in de associaties van de kinderen dan in die van de volwassenen, met name omdat de kinderen vaker vormgerelateerde associaties gaven (*olifant – olie*). De eentalige kinderen en volwassenen gaven iets meer semantisch gerelateerde associaties dan de tweetaligen. De volwassenen als groep gaven meer context-specifieke associaties dan semantische associaties terwijl de kinderen ongeveer evenveel van beide gaven. Dit is verrassend en lijkt aan te geven dat semantische associaties niet per se het optimum zijn in termen van woordkennisontwikkeling. We kunnen de verschillende typen associaties (semantisch, contextueel en vormgerelateerd) niet zonder meer als indicatie van een schaal van ontwikkeling zien. Tegelijkertijd is het niet noemen (bij de volwassenen) van semantische associaties misschien niet verrassend aangezien andere factoren dan conceptuele ontwikkeling (bijvoorbeeld recent gebruik van woorden) van invloed kunnen zijn op welke betekenisaspecten geactiveerd zijn in het lexicon en naar boven komen in een spontane associatietas. Dit maakt de vrije woordassociatietas niet zonder meer geschikt als toets van woordkennisontwikkeling en het pleit voor meer gestructureerde taken die specifiek gericht zijn op (het herkennen van) semantische betekenisaspecten.
Deel II: Speelt snelle toegang tot semantische woordkennis een rol in leesbegrip?

Voor een optimaal gebruik van woordkennis is snelle toegang ertoe belangrijk. Bij het herkennen van een woord tijdens lezen activeren we dit woord in ons mentale lexicaal-semantische netwerk. Hoe beter woorden ingebed zijn in het netwerk (hoe beter verbonden met andere woorden), hoe makkelijker en sneller we de er mee samenhangende kennis paraat hebben. Studies naar semantische vloeiendheid, waarbij kinderen zo snel mogelijk zoveel mogelijk woorden van een bepaalde categorie moeten noemen, zoals dieren of vervoersmiddelen, hebben verschillen tussen kinderen in vloeiendheid gerelateerd aan lees(begrips)scores. Dit wijst op een rol van snelle toegang tot semantische woordkennis voor lezen. Het meeste onderzoek naar woordkennis echter, met name onderzoek in een onderwijscontext, gebruikt schriftelijke taken zonder tijds meting, zoals definitietaken. Dergelijk onderzoek waarbij tijd niet wordt meegenomen in de meting zegt weinig over hoe snel of gemakkelijk kinderen woordkennis verwerken of gebruiken. Het is dus onduidelijk of verschillen tussen kinderen op zulke taken te maken hebben met verschillen in snelle toegang tot woordkennis - wellicht ten gevolge van verschillen in hun onderliggende semantische netwerken. Zijn de kinderen die goed scoren op semantische woordkennistaken zonder tijds meting gewoon beter in het bewust oplossen van woordtoren en het formuleren van definities? Of scoren deze kinderen beter door snellere toegang tot hun woordkennis, mogelijk doordat hun semantische netwerken anders zijn waardoor verwante concepten elkaar makkelijker activeren?

Om de rol van semantische woordkennis en snelheid van toegang tot die kennis te onderzoeken in relatie tot leesbegrip hebben we in het tweede onderzoek, dat beschreven staat in Hoofdstuk 4, een aantal lees- en woordkennistaken afgenomen bij 135 Nederlandse eentalige kinderen en tweetalige kinderen uit migrantengezinnen in groep 7 (leeftijd 10-11 jaar) van verschillende basisscholen. Om snelheid van toegang tot semantische woordkennis te meten werd een semantische keuzetaak ontworpen en individueel via laptops afgenomen. In deze taak moesten kinderen voor elk gegeven stimuluswoord (mus) zo snel mogelijk
kiezen welk van twee aangeboden woorden (boom – vleugel) er het best bij past. Beide opties konden worden gerelateerd aan het stimuluswoord maar slechts een ervan was semantisch gerelateerd. In de taak werd gemeten of kinderen de juiste betekenisaspecten koz en hoe snel ze dat deden. Verder werden een schriftelijke semantische woordkennistaak (de WAT) en een begrijpend leestach afgelopen en werd elk kind individueel getest met de Drie-Minuten-Toets als controlemaat voor technisch lezen. De bijdragen van technisch lezen, semantische woordkennis en snelheid aan het verklaren van variantie in begrijpend lezen zijn onderzocht in regressiesanalyse en een padmodel dat de onderlinge relaties tussen de constructen weergeeft.

De resultaten staan beschreven in Hoofdstuk 4. Ze laten zien dat de tweetalige kinderen inderdaad lager scoorden dan de eentalige kinderen op leesbegrip, hoewel ze het iets beter deden op technisch lezen. Hun scores op de semantische woordkennistaak waren ook lager dan die van de eentalige kinderen. Wat betreft de computertask waren de een- en tweetalige kinderen even goed en even snel in het identificeren van het semantisch gerelateerde woord. Verschillen tussen de kinderen in technisch lezen, semantische woordkennis en snelheid op de computertask konden samen 34% van de variantie in leescores verklaren. Er was geen extra effect van de interactie van taalachtergrond en semantische woordkennis of snelheid, hetgeen er op wijst dat de bijdragen aan lezen van woordkennis en snelheid vergelijkbaar waren voor de een- en tweetaligen. De factor taalachtergrond kan geen variantie in lezen meer verklaren nadat we rekening hebben gehouden met semantische woordkennis. Dit betekent dat de verschillen in leescores tussen de een- en tweetalige kinderen verklaard kunnen worden door semantische woordkennis. Een padmodel laat zien dat het effect van taalachtergrond inderdaad via semantische woordkennis verloopt. Deze resultaten laten zien dat kennis van semantische betekenisaspecten relevant is en dat snelle toegang tot die woordkennis, in mindere mate, ook bijdraagt aan leesbegrip. Een deling van de kinderen in goede en minder goede lezers gaf aan dat semantische woordkennis en toegankelijkheid iets meer variëerde in leesbegrip konden verklaren bij de goede lezers.
Deel III: Draagt automatische activatie van het semantisch netwerk bij aan leesbegrip?

Na het vaststellen van het belang van semantische woordkennis voor leesbegrip en van de kleine toegevoegde waarde van snelheid, is de rol van toegankelijkheid van semantische woordkennis verder onderzocht in het derde onderzoek. Dit onderzoek staat in Hoofdstuk 5 beschreven. Centraal stond de vraag of goede lezers kennis van woordbetekenis niet alleen snel beschikbaar hebben maar ook automatisch activeren op basis van woordverwantschap. Onderzoek naar wat ‘semantic priming’ genoemd wordt kan hier iets over zeggen. In priming onderzoek wordt gemeten in welke mate het ene woord het andere als het ware mede-activert. Er is sprake van priming bij een snellere reactie op een woord (oor) wanneer dit is voorafgegaan door een gerelateerd woord (neus) dan wanneer dit is voorafgegaan door een ongerelateerd woord (fiets). Taalpsychologisch onderzoek bij kinderen laat zien dat kinderen die laag scoren op leesbegrip gemiddeld ook minder co-activatie van semantisch gerelateerde woorden laten zien zoals gemeten met semantic priming. Dit is een indicatie dat deze semantische relaties minder ontwikkeld zijn in hun semantische netwerk. Er is echter weinig onderzoek beschikbaar dat meet in hoeverre individuele verschillen in semantic priming individuele verschillen in leesbegrip kunnen verklaren. Zo is het vooralsnog niet duidelijk in hoeverre de lesachterstanden van twee talige kinderen of migrantengezinnen te maken hebben met verschillen in het gebruik van semantische woordkennis, zoals het automatisch activeren van die semantische kennis. Bovendien is niet bekend of de relatie tussen semantische woordkennis, automatische activatie en leesbegrip hetzelfde is voor eentalige kinderen en twee talige kinderen uit migrantengezinnen.

In het derde onderzoek is activatie van semantische woordkennis onderzocht aan de hand van twee priming taken: een lexicale decisietaak en een semantische classificatietaak. Deze werden afgenomen op de computer zodat responsiesnelheid gemeten werd en daarmee priming berekend kon worden. In de lexicale decisietaak moesten kinderen voor elke stimulus die op het scherm verscheen aangeven of het een bestaand woord was of niet. Er zaten dus ook niet-
bestaande woorden in deze taak. In de semantische classificatietaak moesten kinderen voor elk stimuluswoord aangeven of het een dier was of niet. Bij beide taken ging het er niet zozeer om of ze de juiste keuze maakten maar om de mate waarin ze priming vertoonden voor opeenvolgende woorden die semantisch gerelateerd waren (bijvoorbeeld neus - oor of mes - bestek). Snelheid op deze taken werd ook meegenomen als variabele. In totaal werden voor dit derde onderzoek vijf toetsen afgenomen bij een nieuwe groep van 130 een- en tweetalige kinderen in groep 7 van verschillende basisscholen: de twee priming taken op de computer, de Drie-Minuten-Toets, de WAT voor semantische woordkennis en de begrijpend leestoets die ook gebruikt is in Hoofdstuk 4. De bijdragen aan verklaarde variantie in begrijpend lezen van technisch lezen, semantische woordkennis, lexicale decisie snelheid, semantische classificatiesnelheid en semantic priming zijn onderzocht in regressieanalyses en een padmodel.

De resultaten laten een semantic priming effect zien voor zowel de eentalige als de tweetalige kinderen: beide groepen reageerden sneller op een woord dat voorafgegaan was door een semantisch gerelateerd woord dan wanneer dit niet het geval was. Ook in dit onderzoek scoorden de tweetalige kinderen lager op leesbegrip en semantische woordkennis dan de eentalige kinderen, maar de priming scores van de eentalige en de tweetalige kinderen waren vergelijkbaar. De leesverschillen tussen de twee taalgroepen waren ook dit keer niet te wijten aan technisch lezen; de beide groepen scoorden hier vergelijkbaar op. Wederom verklaarde semantische woordkennis een significant deel van de variantie in begrijpend leesscores en net als in het onderzoek in Hoofdstuk 4 kon snelheid nog een kleine hoeveelheid variantie extra verklaren na de bijdrage van semantische woordkennis. Dit gold voor zowel een- als tweetalige kinderen. Echter, alleen snelheid gemeten in de semantische classificatietaak en niet in de lexicale decisietaak droeg bij aan leesbegrip. Semantic priming droeg niet bij aan de verklaarde variantie in leesbegrip. In dit derde onderzoek kan taalachtergrond nog wat variantie in leesscores verklaren nadat we rekening hebben gehouden met semantische woordkennis. Een padmodel bevestigt dit: taalachtergrond heeft naast een effect op semantische woordkennis en snelheid een rechtstreeks effect op
begrijpend lezen. Verschillen in begrijpend lezen tussen de beide taalgroepen in dit onderzoek kunnen dus niet volledig verklaard worden door de woordkennisvariabelen. Blijkbaar speelden er in deze onderzoeksgroep nog andere verschillen tussen de kinderen een rol. De uitkomsten van dit derde onderzoek bevestigen het belang van semantische woordkennis in begrijpend lezen en de kleine toegevoegde waarde van toegankelijkheid (snelheid) hiervan. Een bijdrage van priming effecten kon niet aangetoond worden.

Conclusies en implicaties
Dit proefschrift laat zien dat semantische kennis van woordbetekenis belangrijk is voor het leesbegrip van zowel eentalige kinderen als tweetalige kinderen uit migrantengezinnen en dat de snelheid waarmee betekenis herkend wordt hier een rol in speelt. Waar tweetalige kinderen achterblijven in kennis van woordbetekenis en in snelheid, blijven ze ook achter in leesbegrip. Kinderen vertonen evenveel activatie van semantisch gerelateerde woorden. De mate van activatie droeg in dit onderzoek niet bij aan het verklaren van variantie in leesbegrip. Toekomstig onderzoek dat gebruik maakt van een meer gecontroleerde opzet en dat ook andere modaliteiten betrekt (bijvoorbeeld auditieve priming) kan mogelijk meer uitwijzen over de rol van automatische activatie in het verklaren van verschillen in leesbegrip bij deze populatie. Het is voor zowel eentalige als tweetalige kinderen belangrijk dat ze niet alleen de betekenis van een woord kennen maar ook de betekenisrelaties die het woord heeft met andere woorden en dat ze, wanneer ze woorden tegenkomen, snel begrijpen welke woorden met elkaar samenhangen.
Curriculum Vitae

Marjolein Cremer was born on 10 July 1983 in Zuidlaren (The Netherlands). In 2001, she enrolled in the degree programme English Language and Culture at the Rijksuniversiteit Groningen, chose a specialisation in linguistics and graduated (cum laude) in 2006. Her fascination for language learning was expressed in her BA dissertation, which she wrote on bilingual children’s understanding of verbs. She obtained a Harting scholarship and spent a year in Wales (United Kingdom) studying linguistics and teaching Dutch at Bangor University. She wrote her MA dissertation on English accent varieties and linguistic attitudes. In September 2006, she started her PhD research at the Amsterdam Center for Language and Communication, which resulted in the present work. She currently works as a teacher trainer at the Centrum voor Nascholing and teaches in the English department of the University of Amsterdam.