Developing semantic networks

Individual differences in Dutch monolingual and bilingual children's semantic knowledge and reading comprehension

Tessa Marie Spätgens

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111

e-mail: lot@uu.nl
http://www.lotschool.nl

Cover illustration © 2016: 'Energie' by Irene Laumans

ISBN: 978-94-6093-290-8
NUR 616

Developing semantic networks
Individual differences in Dutch monolingual and bilingual children's semantic knowledge and reading comprehension

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. ir. K.I.J. Maex

ten overstaan van een door het College voor Promoties ingestelde

commissie, in het openbaar te verdedigen in de Agnietenkapel

op woensdag 20 juni 2018, te 12.00 uur

door Tessa Marie Spätgens

geboren te Roermond

Promotiecommissie:

Promotor:	prof. dr. J.J.M. Schoonen	Radboud Universiteit Nijmegen
Copromotor:	prof. dr. J.H. Hulstijn	Universiteit van Amsterdam
Overige leden:	prof. dr. A.F.J. Dijkstra	Radboud Universiteit Nijmegen
	prof. dr. T. Fitzpatrick	University of Swansea
	prof. dr. P.F. de Jong	Universiteit van Amsterdam
	prof. dr. F. Kuiken	Universiteit van Amsterdam
	prof. dr. J.E. Rispens	Universiteit van Amsterdam
	prof. dr. J.C. Schaeffer	Universiteit van Amsterdam
Faculteit:	Faculteit der Geesteswetenschappen	

Table of contents

Author contributions

Chapters 1 and 5

Tessa Spätgens wrote these chapters and made revisions based on feedback from Rob Schoonen and Jan Hulstijn.

Chapter 2

A slightly modified version has been submitted as:
Spätgens, T. & Schoonen, R. (submitted). The structure of developing semantic networks: evidence from single and multiple word associations in young monolingual and bilingual readers.

Chapter 3

A slightly modified version has been previously published as:
Spätgens, T. & Schoonen, R. (2018). The semantic network, lexical access and reading comprehension in monolingual and bilingual children: An individual differences study. *Applied Psycholinguistics*, *39*(1), 225-256.

Chapter 4

A slightly modified version has been submitted as:
Spätgens, T. & Schoonen, R. (submitted). Individual differences in reading comprehension in monolingual and bilingual children: the influence of semantic priming during sentence reading.

For chapters 2, 3, and 4, Tessa Spätgens was the principal investigator, responsible for task design, data collection with help of student-assistants, analyses, writing and revision. Rob Schoonen wrote the original project proposal that served as the basis for the research questions which the chapters aimed to answer, and acted as supervisor, providing feedback on methodology, analyses and text. Jan Hulstijn, also in the role of supervisor, provided feedback on methodology and text.

Chapter 1
Introduction

Acquiring sufficient vocabulary knowledge is undoubtedly one of the main challenges school-age children face: it is a prerequisite for early literacy and school success, yet the individual variability in vocabulary knowledge is considerable (Biemiller, 2006). In addition, vocabulary knowledge is not a one-dimensional construct involving only the sheer number of words to be acquired, but also includes other aspects: an individual's knowledge of words needs to be high in quality and easily accessible, so that it can be used flexibly and efficiently during language processing. A lexical representation that is high in quality is multi-faceted, involving both form knowledge and extended semantic content that covers the full range of meaning dimensions associated with it. Effortless access to these extensive representations helps to ensure fluent processing during complex linguistic tasks such as reading comprehension (Perfetti & Hart, 2002; Perfetti, 2007). Therefore, in addition to vocabulary size, other aspects of lexical knowledge such as depth (cf. Anderson & Freebody, 1981; Henriksen, 1999), including semantic network organization (cf. Meara & Wolter, 2004), and fluency (cf. Segalowitz, 2010) need to be developed by language learners.

Apart from the generally observed variability in groups with homogeneous language backgrounds, bilingual minority children speaking another language at home in addition to the majority language, are especially at risk of disadvantages in vocabulary acquisition. Due to the importance of frequency of use for the acquisition of vocabulary, and the relatively reduced exposure experienced by bilinguals in each of their two languages, vocabulary knowledge may be especially susceptible to differences between monolinguals and bilinguals (Bialystok, 2009; Michael & Gollan, 2005). Indeed, research shows that bilinguals typically command smaller vocabularies, and various aspects of vocabulary depth are less developed in this group (Bialystok, 2009; Cremer, 2013; Melby-Lervåg & Lervåg, 2014; Michael & Gollan, 2005; Raudszus, Segers, & Verhoeven, 2018; Schoonen & Verhallen, 2008; Schwartz & Katzir, 2012).

Given the central role that vocabulary plays in reading comprehension processes, the variability in individuals' vocabulary knowledge may also exert a strong influence on individual differences in reading comprehension (Perfetti & Stafura, 2014). The increasing prominence of written texts as sources of content knowledge throughout the school years, means that a detailed understanding of the subskills that contribute to reading comprehension for both monolingual and bilingual children is not only of theoretical, but also educational significance. The role of vocabulary size is by now well-established (e.g. Alderson, 2005; Grabe, 2009; Stæhr, 2008; Stanovich, 2000), but the additional contributions of various aspects of vocabulary depth and fluency[1] are still under study (e.g. Cain & Oakhill, 2014; Cremer, 2013; Richter, Isberner, Naumann, & Neeb, 2013; Swart, Muijselaar, Steenbeek-Planting, Droop, De Jong, & Verhoeven, 2017a). Measures of connection strength in the semantic network structure of the mental lexicon and fluency of access to semantic information may contribute to reading comprehension skill, but evidence for the role of these vocabulary aspects for individual differences in reading comprehension is not yet conclusive (Betjemann & Keenan, 2008; Cremer, 2013; Nation & Snowling, 1999; Richter et al., 2013; Swart et al., 2017a).

The research presented in this dissertation attempts to further our understanding of the role of semantic network structure and semantic activation for individual differences in monolingual and bilingual minority children's reading comprehension. Three empirical studies sought to map the semantic network and its activation in different ways: in Chapter 2, organization in the semantic network is studied using word associations. Two online measures of semantic knowledge are used in Chapters 3 and 4: single-word auditory semantic priming and semantic priming during reading. Hereby, access to semantic information and priming of semantic relations in the network can be studied. The contribution of these aspects of vocabulary knowledge to reading comprehension are assessed in each study, and differences between monolinguals and bilingual minority

[1] The term fluency is not used here to refer to technical reading skill, i.e. decoding speed, but to the fluent access to lexical representations. More details on the definitions used in this dissertation are provided in section 1.1.2.

children are examined. This introductory chapter will first discuss these three focal aspects of vocabulary knowledge in more detail, before turning to the relevant literature and theoretical background pertaining to the participating language groups and reading comprehension. Finally, an outline of the dissertation will be provided.

1.1. Definitions of aspects of depth and fluency: the semantic network, semantic relations and semantic access

The aspects of vocabulary in focus in this dissertation, semantic relations in the semantic network and access to semantic knowledge, are all subject to considerable variation in definitions and operationalizations in the literature. In this section, the relevant background and definitions used in this dissertation are discussed.

1.1.1. Organization in the semantic network and types of semantic relations

The distinction between vocabulary breadth, i.e. size, and depth (Anderson & Freebody, 1981) is pervasive in the literature researching vocabulary knowledge itself (both Read, 2004, and Schmitt, 2014, have provided extensive overviews) and also in studies targeting the role vocabulary plays for other higher-order skills such as reading comprehension (e.g. Cain & Oakhill, 2014; Ouellette, 2006; Qian, 1999; 2002; Swart et al., 2017a). The construct of depth has been defined and operationalized in many different ways however, one of which is an approach emphasizing the organization within the network of semantic relations between words (Meara & Wolter, 2004). Language users have declarative knowledge of these relations, which can be measured using tasks that ask respondents to identify related words or word pairs that share specific relations (Qian, 1999; Schoonen & Verhallen, 2008). Additionally however, the mental lexicon is organized as a network in which words are connected through semantic relations[2] of various types (Aitchison, 2012). The relative importance of different types of semantic relations in an individuals' mental lexicon can be measured

[2] Note that connections in the network can also be of a phonological or orthographical nature, however, semantic relations are in focus here.

using the free word association task, in which participants are given single word stimuli and asked to provide the first word or first few words that come to mind in response to these stimuli (e.g. Aitchison, 2012; De Deyne & Storms, 2008a; Fitzpatrick, 2013). Furthermore, priming studies have shown that relations in the network are activated during language use due to spreading activation (Bock & Levelt, 1994; Collins & Loftus, 1975). In this dissertation, both the organization in the participants' semantic networks, i.e. the various types of semantic relations they encompass, and automatic activation of these relations, semantic priming, are examined.

The terminology used to describe the different types of semantic relations is varied. In this dissertation, a continuum from context-dependent to context-independent knowledge will be employed, based on previous dichotomies stemming from language acquisition research. Knowledge of semantic relations and their place in the network structure need to be developed during language acquisition, and there is evidence for a shift in importance of different types of semantic relations during this development. Nelson (1982; 1985; 1991; 2007) has published extensively on pre-linguistic conceptual development in children, and the continuation of this development into the early stages of first language acquisition. Young children initially conceptualize the world in terms of events: in categorization tasks, four-year-olds are more likely to categorize words according to common co-occurrence in scripts or events (i.e. *breakfast – milk*) than according to taxonomic relations such as subordination or coordination (i.e. *breakfast – dinner*), contrary to adults. Furthermore, in word association tasks, children provide more answers that reflect contextual, functional and perceptual relations, and actions, while adults supposedly primarily respond with taxonomically related words. Although some contrasting evidence has also been reported, the majority of studies do suggest that as children grow older, taxonomic relations, which are driven by shared intrinsic features, become more important (Mirman, Landrigan, & Britt, 2017). Over the years, this shift has received different names and varying definitions: from syntagmatic to paradigmatic knowledge (K. Nelson, 1977; Zareva, 2007), from episodic to categorical knowledge (K. Nelson, 2007), or from thematic to taxonomic knowledge (e.g. Lin & Murphy, 2001; Mirman et al., 2017). Importantly however, it is clear that initially, children focus on relations driven by common co-occurrence, and then abstract from these episodic experiences to more generalized representations of concepts.

In the dichotomies discussed above, co-occurrence in context is juxtaposed to shared intrinsic features, and especially co-occurrence subsumes a variety of relations that have different characteristics (Estes, Golonka, & Jones, 2011). In this dissertation, the full range of semantic relations is conceptualized as occupying spaces on a continuum from more context-dependent to more context-independent. The notion of semantic knowledge becoming more decontextualized or context-independent throughout development has been used previously by Verhallen and Schoonen (1993; 1998), Schoonen and Verhallen (2008) and Cremer (2013) and is in many ways analogous to the distinctions discussed above. However, by using the concept of a continuum, additional subdivisions of the range of semantic relations can be made.

On the most context-independent extreme of this scale are taxonomically related words such as *squirrel – animal* or *squirrel – horse*: these concepts are related by virtue of their shared intrinsic features, whether they regularly co-occur or not. Next, in a word pair such as *squirrel – fur*, one word describes an intrinsic feature of the other, which also holds true independent of context, although now only one shared characteristic is evoked. In a pair such as *squirrel – forest*, very few if any intrinsic features are shared, but these concepts do typically co-occur. Finally, a word pair such as *squirrel – cute* represents a subjective relation which is specific to an individual's personal context specifically. The examples are provided in Table 1.1 below.

More elaborate definitions of these subcategories will be discussed in the context of the word association data presented in Chapter 2. However, the general notions of context-dependent and context-independent knowledge will be important throughout this dissertation. As will become clear in the sections on bilingualism and reading comprehension below, especially context-independent knowledge may be most relevant for differences between monolinguals and bilinguals, and between more and less advanced comprehenders.

Table 1.1

Examples of subcategories on the continuum of context-dependent to context-independent relations

Context-dependent	←	→	**Context-independent**
Subjective	Situational	Feature	Taxonomic
squirrel – cute	*squirrel – forest*	*squirrel – fur*	*squirrel – animal*

A final remark needs to be made here regarding a specific type of relation identified in studies on the semantic network, the *associative* relation. Typically, associative relations are empirically determined by gathering word association data from groups of participants: resulting stimulus-response pairs have a certain association strength, depending on the number of participants that provided that particular response for that particular stimulus. However, as pointed out by McRae, Khalkhali and Hare (2011), many researchers contrast associative relations directly with 'semantic' relations, mostly defined as category relations. This dichotomy is problematic, firstly because the term 'semantic' relations encompasses many more relations than category relations alone, as discussed above, and is therefore ambiguous. Secondly, associative relations as determined by empirical evidence from word association tasks may include any imaginable type of semantic relation, including category relations (De Deyne & Storms, 2008a). Finally, the reverse also holds: word pairs may be related despite not being associated, such as for example *dog – wolf* and *desert – snake*.[3] The definitions in this dichotomy are thus confounded: associative relations can be of many semantic types, and semantic relations may or may not lead to an associative relation. Indeed, there is abundant evidence indicating that association has its own separate effect on semantic priming, in addition to semantic relatedness of various types (Lucas, 2000). In this dissertation, the term *associated* is therefore explicitly used to refer to

[3] Word pairs checked against the University of South Florida free word association norms (D. L. Nelson, McEvoy, & Schreiber, 2004) and the preliminary English language data from the Small World of Words project, which is the source of the Dutch norms by De Deyne and Storms (De Deyne & Storms, 2008b): preliminary data retrieved from https://smallworldofwords.org/nl/project/explore on 10 January 2018.

word pairs that are associated according to empirical evidence, that is, found to be related in free word association tasks.

1.1.2. Semantic access

In addition to the size and depth aspects of vocabulary knowledge, researchers have pointed out that *fluency*, the automaticity with which lexical information can be accessed, is another important dimension of vocabulary knowledge (e.g. Beck, Perfetti, & McKeown, 1982; Perfetti & Hart, 2002; Perfetti, 2007; Segalowitz, 2010). This term is not used here to refer to decoding speed, but to the fluency of access to lexical representations, and specifically, the semantic information in these representations. Lexical representations are multifaceted, including phonological and orthographic form, and semantic content. Accessing lexical information can imply these different aspects: for example, Richter and colleagues (2013) designed separate tasks testing access to phonological, orthographical and semantic representations. Again focusing on semantic knowledge, this dissertation will consider the role of access to semantic information specifically and hence use the term *semantic access*.

1.2. Differences between monolingual and bilingual minority children

In the Netherlands, about 17% of the children in sixth grade (i.e., age 11-12) currently have a non-Western immigrant background and within this group, about 90% have at least one parent who was born in the Netherlands (CBS, 2016). A majority of children raised with another language speaks Dutch at home in addition to their L1 (Berkel, Schoot, Engelen, & Maris, 2002; Heesters, Berkel, Schoot, & Hemker, 2007; Sijtstra, Schoot, & Hemker, 2002). Consistent with evidence from international studies (e.g. Hoff et al., 2012; Mancilla-Martinez & Vagh, 2013; Melby-Lervåg & Lervåg, 2014), these bilingual minority children are often found to show disadvantages compared to monolinguals on both vocabulary and reading comprehension (e.g. Cremer, 2013; Heesters et al., 2007; Raudszus et al., 2018). Thus, despite the fact that these children are exposed to Dutch from an early age, their proficiency in Dutch appears to be lagging behind.

According to the weaker links hypothesis (Gollan & Silverberg, 2001; Gollan, Montoya, Cera, & Sandoval, 2008; Michael & Gollan, 2005), delays in bilinguals are likely caused by differences in frequency of exposure: because bilinguals' language input is divided across two languages, the exposure to each language is reduced relative to monolinguals' input. Especially vocabulary, in part because it is already subject to considerable individual variability, is hypothesized to be affected by this difference in language experience (Michael & Gollan, 2005).

Indeed, evidence for differences between monolinguals' and bilinguals' vocabulary beyond size abounds. Studies assessing declarative knowledge of context-independent relations consistently find disadvantages for bilinguals: in definition tasks, bilinguals are found to provide less context-independent information such as category names and intrinsic features (Raudszus et al., 2018; Schwartz & Katzir, 2012; Verhallen & Schoonen, 1993). In forced-choice tasks requiring participants to select context-independent relations over unrelated and context-dependently related distractors, monolinguals also outperform bilinguals (Cremer, 2013; Schoonen & Verhallen, 2008). Knowledge of context-independent relations may thus be reduced in bilinguals, perhaps an indication that they are lagging behind their monolingual peers on the developmental path towards a more context-independently organized lexicon.

However, studies targeting spontaneous measures of semantic organization do not necessarily appear to corroborate this idea. Cremer, Dingshoff, De Beer and Schoonen (2011) analyzed Dutch monolingual and bilingual children's word associations and only found a higher number of null responses in the bilingual group, reflecting the stimulus was unknown, but no clear differences in specific semantic response categories between the two groups. Fitzpatrick (2006; 2007; 2009) studied L2 speakers from various language backgrounds and monolingual L1 speakers of English, and concluded that word association patterns in both language groups are highly personal, and similar between the two languages of the L2 speakers.

Similarly, the findings on declarative context-independent knowledge in bilinguals do not directly translate to semantic priming. A recent study involving bilingual minority children found no priming in the L2 in a lexical decision task for various semantic relations, but did obtain facilitation in a looking-while-listening paradigm, recording gaze time to pictures following auditorily presented primes (Goodrich & Lonigan, 2018). The authors conclude that evidence for semantic priming was limited in this group, and suggest that this may be due to relatively low proficiency levels. In contrast, Cremer (2013) specifically compared semantic priming of context-independent relations in monolingual and bilingual children, both category coordinates and subordinate – superordinate pairs, and found no significant differences between the language groups. Both showed similar priming effects, despite delays for the bilinguals on other language proficiency tasks testing reading comprehension and declarative context-independent knowledge. Despite clear disadvantages in declarative knowledge of semantic relations, evidence for the amount of automatic activation of these relations displayed by bilingual children compared to monolingual children is still mixed.

Finally, many studies show that lexical access in bilinguals is slower than in monolinguals (Bialystok, 2009; Michael & Gollan, 2005), but subtle differences appear depending on the specific tasks that are used, especially when semantic access is required. For example, Cremer (2013) found bilingual children were slower at making both lexical and semantic decisions on single words, suggesting their access to both word form and semantic content was at a disadvantage compared to their monolingual peers. However, in a speeded forced-choice task requiring children to identify context-independently related words, monolinguals and bilinguals were equally fast. Similarly, Gollan, Montoya, Fennema-Notestine and Morris (2005) demonstrated that picture naming speed in bilingual adults was slower compared to monolinguals, while semantic classification speed of pictures was similar in both language groups. The first task requires retrieval of the language-specific word-form in addition to the semantic representation, while for the second task only access to the semantic representation is required, similar to the semantic decision task to words in Cremer’s work. Although subtle differences in task characteristics thus appear to cause mixed findings, access to semantic representations based on word form is likely to be slower in bilinguals.

To summarize, evidence concerning semantic network organization, semantic priming and semantic access in bilingual minority children compared to monolinguals is somewhat mixed, even though we would expect disadvantages for bilinguals based on evidence for delays in vocabulary size and declarative knowledge of semantic relations. Given the importance of vocabulary knowledge for comprehension, if delays do exist, these may be a source of delays in reading comprehension. Each of the three aspects of vocabulary knowledge will be assessed in relation to language group and reading comprehension in this dissertation.

1.3. Connections to reading comprehension

Various frameworks of reading comprehension propose that word knowledge is a primary component of this complex higher-order skill. For example, in the Reading Systems Framework (RSF, Perfetti & Stafura, 2014) some of the most important subprocesses of comprehension are considered to be fluent access to a word's semantic content and the retrieval of related semantic information. The Lexical Quality Hypothesis or LQH (Perfetti & Hart, 2002; Perfetti, 2007) posits that word knowledge needs to be high in quality for reading comprehension processes to be able to run smoothly. In the LQH, the semantic aspects of word representations are considered to be high in quality if they are rich in semantic information, including the various meanings a word may have depending on different contexts, and information involving related words. Additionally, this information needs to be easily accessible, which frees up cognitive resources that can in turn be devoted to higher-order comprehension processes.

Both of these propositions mean that vocabulary knowledge is not only required to be extensive in terms of size, but also in the amount of semantic information that is associated with individual lexical entries and the automaticity with which this information can be accessed. Many studies addressing the influence of depth and fluency of vocabulary knowledge on reading comprehension confirm these propositions: even though size, depth, and fluency are strongly correlated and are likely measuring at least partially overlapping parts of the same construct (Swart et al., 2017a; Vermeer, 2001), various measures of depth and fluency are found to predict additional variance over vocabulary size in many studies.

Starting with studies assessing the influence of fluent access to semantic information, a range of operationalizations of fluency are in use. Many studies employ tasks requiring participants to produce as many category members as possible within a certain time limit, typically 60 seconds. This measure, often combined with other measures of semantic knowledge into one composite score, is found to contribute to reading comprehension in children (Nation & Snowling, 2004; Tannenbaum, Torgesen, & Wagner, 2006) but it does not purely measure fluent retrieval of the semantic representation of words: what is measured also contains a component of categorical knowledge. However, there are also examples of studies finding a relation between comprehension and access to semantic information of lexical entries specifically. Cremer (2013) found a small contribution to reading comprehension of semantic categorization speed after knowledge of context-independent relations was already controlled for, but no measures of vocabulary size were included. Richter and colleagues (2013) measured both accuracy and speed of retrieval of semantic representations in a categorization task, and showed that speed contributed to reading comprehension in addition to accuracy itself. Access to semantic representations thus contributes to comprehension, but studies combining semantic access measures with vocabulary size measures appear to be limited in the literature at present.

Among the various operationalizations of depth, many focus on context-independent knowledge. Numerous studies use definition tasks, which are typically scored on context-independent knowledge aspects such as superordinates and intrinsic features. Significant contributions to individual differences in children's reading comprehension are reported for these tasks (Cain & Oakhill, 2014; Ouellette, 2006; Swart et al., 2017a; Tannenbaum et al., 2006; Tilstra, McMaster, Van den Broek, Kendeou, & Rapp, 2009). Forced-choice tasks aimed at the identification of related words, either all context-independent such as category relations and intrinsic features (Cremer, 2013) or a mixture of synonymy, polysemy and collocational connections (Qian, 1999; 2002) show similar results.

A more specific effect of semantic network organization and activation can also be hypothesized, for example based on theories of reading comprehension such as the Reading Systems Framework (Perfetti &

Stafura, 2014) and the Landscape Model (Van den Broek, Young, Tzeng, & Linderholm, 1999; Van den Broek, Rapp, & Kendeou, 2005). Both of these frameworks suggest that related semantic representations that are automatically activated during reading, can be more easily integrated into the mental model of a text. Through this swift linkage of related concepts without effortful processing, coherence can be established more easily, improving comprehension. Studies specifically targeting the semantic network as an instantiation of depth, indeed show interesting connections to reading comprehension, both in offline and online tasks. For example, Swart and colleagues (2017a) administered a word association task to monolingual children and scored their responses on similarity to adult word association data. They found that these scores were significant predictors of reading comprehension scores: the more similar a child's associations were to adults' associations, the higher their reading comprehension scores. This contribution was significant in addition to measures of vocabulary size and a definition task. These results are promising, but since no classification of the semantic types of associations was performed, they cannot inform us on the role of different types of semantic preferences for reading comprehension.

Regarding online measures of activation in the semantic network, a number of group studies shows that differences in reading comprehension may be associated with differences in semantic priming: poor readers may show reduced sensitivity to semantic priming, and this may be dependent on the type of semantic relation. Betjemann and Keenan (2008) compared young poor readers to age-matched controls and found that poor readers showed reduced priming of associative pairs, which included a range of semantic relations from context-dependent to context-independent. In an ERP study targeting priming of associated and non-associated categorically related word pairs, adult poor comprehenders were found to show smaller differences between related and unrelated word pairs in the P200 and N400 components, compared to skilled readers (Landi & Perfetti, 2007). Because these ERP components are sensitive to semantic processing, the authors conclude that the findings are an indication of less proficient semantic processing in poor comprehenders. Finally, both Nation and Snowling (1999) and Bonnotte and Casalis (2010) examined semantic priming of category coordinates and functional relations which were both associated and non-associated, and found interesting differences between the four semantic types in their child participants. In both studies, poor comprehenders showed reduced priming or no priming on the non-associated category pairs. Additionally, Bonnotte and Casalis observed associated functional priming in their poor readers only, but no functional priming in the normal readers, while both poor readers and normal readers showed associated and non-associated functional priming in the study by Nation and Snowling. It appears then, that activation in the semantic network as measured by semantic priming may contribute to reading comprehension as well, and again, context-independent relations may play a special role: both of the above studies found the same relative lack of categorical priming in poor readers.

Because these studies all compare groups, they cannot inform us on contributions of priming to individual differences in comprehension, although they do lead us to expect that such a connection exists. In a study that used a synonym judgement task and synonym word pairs, participants showing more facilitation were found to show higher scores on reading comprehension as well (Larkin, Woltz, Reynolds, & Clark, 1996). However, in a design more close to Nation and Snowling's (1999), Cremer (2013) studied non-associated context-independent priming using both category coordinates and subordinate – superordinate pairs and found no significant contribution to reading comprehension on an individual level. Given the differences that were observed in group studies and the theoretical role of activation of related words in comprehension, replication and extension of these findings is necessary to determine if and how semantic priming contributes to individual differences in comprehension.

1.4. Goals and outline

A number of open questions for research arise in light of the above discussion of theoretical and empirical findings regarding semantic knowledge in monolinguals and bilinguals and the relation between semantic knowledge and reading comprehension. Firstly, despite delays in reading comprehension, vocabulary size and declarative knowledge of context-independent relations that are consistently found in bilinguals, evidence for differences in semantic organization, semantic priming and semantic access compared to monolinguals is more mixed. This leads us to the first research question: *How do monolingual and bilingual minority children's semantic network organization, semantic access, and sensitivity to semantic priming differ?* The present dissertation will attempt to map and compare the semantic networks and activation within these networks of monolingual and bilingual minority children to study each of these aspects.

Secondly, there is ample evidence that vocabulary size and declarative knowledge of context-independent relations are positively related to reading comprehension, but for the other aspects of vocabulary discussed above, there are still open questions. Semantic network organization as measured by word associations may also be related to reading comprehension, but results from a semantically oriented analysis of response behavior have not yet been related to reading comprehension. Group studies of semantic priming in poor comprehenders compared to normal comprehenders are promising, but individual differences appear to be less clearly related to reading comprehension. Finally, the number of studies targeting semantic access specifically is currently still limited. The relation between each of these aspects of word knowledge will be examined in this dissertation, in order to answer the second research question: *How do individual differences in semantic network organization, semantic access, and semantic priming relate to differences in reading comprehension?*

To answer these questions, semantic organization as measured by word associations will be targeted first in Chapter 2. Based on previous findings in adults, a multiple response approach is used (cf. De Deyne & Storms, 2008a; 2008b). By requiring three responses rather than a single response, more detailed data can be gathered, and in adults semantic types have been shown to be distributed differently across first, second and third responses. Additionally, a detailed classification system is used to be able to analyze the semantic characteristics of the associative responses. The chapter takes on a methodological focus, assessing the advantages of the multiple association task compared to single associations, but simultaneously considers differences between monolinguals and bilinguals, and the relation of association preferences to reading comprehension.

Chapter 3 considers both semantic access and activation in a single-word semantic priming experiment involving non-associated context-dependently related and context-independently related word pairs. By using an auditory semantic categorization task, the influence of semantic access and priming on reading comprehension can be assessed independent of modality: any effect must then be located at the semantic level specifically, and cannot be related to advantages in decoding speed. Individual priming scores are calculated to assess whether these can predict scores on a reading comprehension task. Monolingual and bilingual minority children's sensitivity to the two semantic relations is compared.

Finally, Chapter 4 takes a step further and examines both associated and non-associated word pairs, context-dependently and context-independently related, in a self-paced reading paradigm. Semantic priming of these relations during the process of reading can be approximated more closely in this task than in the single-word priming study in Chapter 3. Reading times on targets and words in the post-target region following related and unrelated prime words are compared, so that the degree of facilitation due to the four semantic relations can be examined. Individual facilitation scores resulting from the self-paced reading task are related to scores on a reading comprehension task, to determine the influence of semantic activation of related words during reading. In this chapter as well, monolingual and bilingual children's facilitation of semantic relations during reading is compared. Finally, in the general discussion in Chapter 5, the findings from the three studies are brought together to answer the overarching research questions.

Chapter 2

The structure of developing semantic networks: evidence from single and multiple word associations in young monolingual and bilingual readers[4]

Abstract

The present study focuses on the effect of an important methodological choice in word association studies: the elicitation of single versus multiple responses, which has been shown to affect the numbers and types of associations adults produce (e.g. De Deyne & Storms, 2008a; 2008b). A total of 11,725 associations from 207 monolingual and bilingual minority children were classified according to a detailed coding system, and differences between the word-level and speaker-level semantic networks resulting from first, second and third responses were examined. We show that in children as well, the multiple association task elicits more and qualitatively different responses, resulting in more dense and diversified semantic networks on the word level. On the speaker level, reading comprehension scores were related differently to initial and later responses, suggesting that a more complex measure of semantic knowledge is obtained from the multiple association task. No differences were found between monolinguals' and bilinguals' associative preferences. We argue the multiple association task produces richer and more detailed data on language user's semantic networks, and suggest a number of ways in which this task could be used to enhance our knowledge of the development of the semantic network throughout childhood and its relation to other language skills.

[4] A slightly modified version of this chapter has been submitted as:
Spätgens, T. & Schoonen, R. (submitted). The structure of developing semantic networks: evidence from single and multiple word associations in young monolingual and bilingual readers.

2.1. Introduction

The network metaphor for the storage of semantic knowledge in the mental lexicon is evoked in many studies on lexical knowledge, and the studies using word associations to probe the exact nature of this network are numerous (Fitzpatrick, 2013). Assumed to reflect the strongest connections in the mental lexicon's network structure (e.g. Schmitt, 1998), word associations provide us with valuable information on both the networks specific words reside in, and the organization of individual language users' semantic knowledge. In other words, word associations give us insight in semantic networks on two levels: on the word level and on the level of the mental lexicon of the individual language user. Word association data on the word level are widely used in psycholinguistic studies on lexical knowledge and retrieval (most notably semantic priming studies, cf. McNamara, 2005), while word association data on the level of the language user have been employed with varying success as a tool to measure vocabulary knowledge (e.g. Cremer et al., 2011; Fitzpatrick, Playfoot, Wray, & Wright, 2013) and distinguish between different age groups (Borghi & Caramelli, 2003; Fitzpatrick et al., 2013), and different types of language learners such as L1 and L2 learners or bilinguals (e.g. Cremer et al., 2011; Fitzpatrick, 2006; 2007; 2009; Fitzpatrick & Izura, 2011; Söderman, 1993; Wolter, 2001; Zareva, 2007; Zareva & Wolter, 2012).

Most word association studies require their participants to produce a single association to each stimulus word that is presented. It has been argued that only these initial associations are unbiased, while later associations suffer from chaining and retrieval inhibition caused by the first response (McEvoy & Nelson, 1982). However, eliciting single responses creates associative networks that only represent very strong associates, while evidence for weaker links is unreliable or even missing altogether (D. L. Nelson, McEvoy, & Dennis, 2000). This means that although single associations seem to be the default option in many studies and are used in the larger association databases such as the Edinburgh Associative Thesaurus (Kiss, Armstrong, Milroy, & Piper, 1973) and the University of South Florida association norms (D. L. Nelson et al., 2004), they may not provide sufficiently extensive information about a word's associative connections.

The qualitative and quantitative consequences of the use of single versus multiple responses has been investigated in great detail in a series of recent studies by De Deyne and colleagues (De Deyne & Storms, 2008a; 2008b; De Deyne, Navarro, & Storms, 2013; De Deyne & Verheyen, 2015, among others). Using their word association data base, which is the largest collected to date (De Deyne et al., 2013), they analyzed associative networks on the word level. Comparing these networks including only first responses versus first, second and third responses, they discovered that the average set size of non-unique responses, i.e. responses that were provided by more than one person, increased when multiple associations were included (De Deyne & Storms, 2008b). As a result, more semantic links can be uncovered, building more extensive semantic networks around stimulus words and thereby obtaining more accurate association norms. Furthermore, by categorizing their data using an extensive semantic coding scheme, the researchers showed that these additional associations are also qualitatively different from first associations (De Deyne & Storms, 2008a). Most notably, taxonomic associations such as category subordinates, superordinates and coordinates show a clear decline across responses, while words associated through context become more prominent across responses. The authors argue that this shows that the multiple association task provides more accurate association norms, since more semantic links are uncovered (De Deyne & Storms, 2008b) and that the resulting associative networks indicate that semantic organization is much more contextually[5] than categorically oriented than previously thought (De Deyne & Verheyen, 2015).

2.1.1. The single versus multiple association task in children

The present study aims to replicate and expand upon these findings in a number of ways. Firstly, knowing that age plays an important role in associative behavior (Borghi & Caramelli, 2003; Cremer et al., 2011; Fitzpatrick et al., 2013), we will investigate the effect of eliciting single

[5] The authors use the term *thematic*, which relates to the *thematic-taxonomic* distinction (e.g. Jones & Golonka, 2012; Markman & Hutchinson, 1984; Mirman & Graziano, 2011).

versus multiple word associations from children. It has been suggested that as children's semantic knowledge develops, they increasingly abstract from concrete experience to more context-independent semantic knowledge (Elbers, Van Loon-Vervoorn, & Van Helden-Lankhaar, 1993; K. Nelson, 1977; 1982; 1985; 1991), even though both types of semantic knowledge are present from an early age (Blewitt & Toppino, 1991) and remain important in the adult mental lexicon (De Deyne & Verheyen, 2015). Typically, taxonomic semantic links are considered to be the most abstract, context-independent knowledge (Cremer, 2013; Verhallen & Schoonen, 1993) and since De Deyne and Storms (2008a) demonstrated that especially these semantic links showed different behavior when comparing initial and later associative responses, it will be interesting to see which patterns across responses turn up in children's word associations. It is conceivable that there is a similar but overall reduced pattern, with taxonomic responses being most prominent as first responses, but less prominent overall than in adults. Alternatively, taxonomic links may be less prominent in children than in adults but already developing, and showing up more as second and third associations instead of as first associations. Additionally, the behavior of other types of semantic links such as associations pertaining to features of the stimulus and associations that are related through context may also show different patterns in children than in adults.

Furthermore, it is highly likely that as in adults, the number of associative links that can be found using a multiple response procedure will be higher than with a single response procedure. Both these qualitative and quantitative differences are consequential for the use of association data on the word level to control for other linguistic measures. We expect that as in adults, more and different types of semantic links will be found in children in a multiple word association task compared to a single association task.

2.1.2. Differences between monolingual and bilingual children

Secondly, we want to look more closely at the effects of eliciting multiple responses on the speaker level, both regarding the possibility of distinguishing learner groups using this type of data and regarding its use as a vocabulary measurement tool that may be predictive of other language skills. To start with the learner groups, a large number of studies has attempted to compare L1 and L2 speakers' and monolingual and bilingual speakers' word associations, to see whether their associative behaviors are different (e.g. Cremer et al., 2011; Fitzpatrick, 2006; 2007; 2009; Fitzpatrick & Izura, 2011; Söderman, 1993; Wolter, 2001; Zareva, 2007; Zareva & Wolter, 2012). The findings in these studies have been mixed, which is likely due to the variation in types of L2 or bilingual speakers involved but also the classification methodologies used.

For example, initial L1/L2 association studies such as Söderman (1993), examined differences between L1 and L2 learners using the traditional syntagmatic/paradigmatic classification scheme. In this system, associations are mainly classified according to word class. Syntagmatic associations belong to a different word class than their stimulus word, while paradigmatic associations belong to the same word class. Söderman (1993) found that both syntagmatic and paradigmatic responses were frequent in L1 and advanced L2 speakers and that as L2 proficiency increased, learners produced more paradigmatic responses and fewer phonologically related and other associations.

However, other researchers have argued that this distinction is not sufficiently detailed, since using word class exclusively ignores the range of semantic relations that may exist between concepts. More fine-grained categorization systems have been proposed and resulted in different findings. For example, Fitzpatrick (2006; 2007; 2009) devised a classification system with a semantic category, a position-based category and a form-based category to compare L1 and L2 speakers. She concluded that there may be no such thing as an L2 or even an L1 word association profile. Instead, she found that speakers had personal association profiles, which reflected similar behavior in both the L1 and the L2. Based on Fitzpatrick's work, Cremer et al. (2011) also developed their own classification system, which divides the semantic category in a direct meaning-related and indirect meaning-related category based on the distinction between context-dependent and context-independent semantic knowledge (Schoonen & Verhallen, 2008; Verhallen & Schoonen, 1993). They found that both bilingual children and adult L2 speakers produced more 'other' and form-based associations than their monolingual peers, but provided roughly equal proportions of answers in the semantically oriented categories in comparison with the respective monolingual groups.

In short, while some researchers found differences between monolingual and bilingual or L2 speakers, which may or may not disappear as L2 proficiency improves, others did not. The studies that employed more fine-grained semantic categories did not find clear differences between monolingual and bilingual or L2 speakers in those categories. This is contrary to what we might expect based on other studies that have found differences in knowledge of various types of semantic relations, such as for example Verhallen and Schoonen (1993), who found that L1 children produced more context-independent relations than L2 children in a definition task. If some studies find differences between monolingual and bilingual associative behavior and others do not, it may be the case that these differences are simply very subtle. Potentially, the multiple association approach could be a useful tool to detect such subtle differences, which could for example turn up only in later associations rather than early ones. Furthermore, a comprehensive classification system with categories that are motivated from a semantic and acquisitional perspective may allow for a more detailed comparison between the two groups.

Using the multiple association approach and an extensive coding scheme based on De Deyne and Storms (2008a), and Cremer et al. (2011), we will try to shed more light on these issues. We have developed a classification system largely based on De Deyne and Storm's (2008a) coding scheme with slight adjustments that are related to the distinction between context-dependent and context-independent semantic knowledge (cf. Cremer et al., 2011; Cremer, 2013; Schoonen & Verhallen, 2008; Verhallen & Schoonen, 1993). The classification system and the rationale behind it will be discussed in more detail in the Method section.

2.1.3. Word associations and other linguistic skills: reading comprehension

Finally, word association tasks may be useful as a measure of vocabulary knowledge that may be predictive of other linguistic skills, and initial and later responses may fare differently in this regard as well. Cremer et al.

(2011) argue that word associations are not restricted enough to be used as an assessment tool for vocabulary knowledge. They maintain that if some type of association is considered to be the most important or most advanced, for example superordinates, then participants should be asked to produce superordinates in order to test whether they have this semantic knowledge. Nevertheless, word associations do provide insight into the types of semantic links that are primary in a language user's mental lexicon. Indeed, Fitzpatrick et al. (2013) demonstrated that when tested on separate occasions, individuals produce similar types of associations over time, which means that the prominence of those semantic links is stable. As such, word associations can be a useful tool to investigate the relation between semantic knowledge or preferences and other language skills.

In the present study, we will investigate how individual learners' word associations relate to their reading proficiency, and whether there is a difference in the predictive potential of initial and later responses. Since it has been suggested that knowledge and activation of semantic relations and perhaps especially context-independent relations contributes to reading comprehension (Bonnotte & Casalis, 2010; Cremer & Schoonen, 2013; Nation & Snowling, 1999; Ouellette, 2006), we expect that learners who produce more context-independent associations such as taxonomic associations, are also better comprehenders. Depending on the associative patterns that children will show, for example if they, like adults, produce more taxonomic associations in the first instance than in later responses, the response position of these context-independent associations may be of particular importance.

In summary, the goals of the present study are to examine in what ways the semantic networks arising from word association tasks are affected by the elicitation of single versus multiple associations. On the word level, we expect to find more and more semantically diverse links in a multiple association task compared to a single association task, which is of relevance for the measurement of association norms. On the level of the language user, we anticipate that multiple associations may be more sensitive to differences between monolingual and bilingual minority children, and may be more accurate as a predictive tool for other measures of language proficiency, in this case reading comprehension.

2.2. Method

2.2.1. Participants

Participants were recruited through schools in five different cities in the Netherlands. Only schools in neighborhoods with a mixed population of monolingual and bilingual speakers of Dutch were approached. Eight schools gave their permission to administer the test in one or two fifth-grade groups, i.e. children aged 10-11. A passive informed consent procedure was applied, and two children did not participate in the tasks because their parents objected. In total, 232 children participated in the study. Seven children with disorders such as dyslexia or autism were identified by the teachers and excluded from the analyses. The data from sixteen children were removed because their data were incomplete as they could not be present for all tasks, and one child who misunderstood the instructions was also excluded. Finally, one child was excluded because he had not gone to school in the Netherlands from grade 1 onwards.

The final data set included the results from 207 children. The mean age of the children was 11;2 (*SD* = 0;6) and ages ranged from 9;10 to 12;10. Based on a language questionnaire, the participants' language status was determined. The majority of the children in the bilingual group spoke both Dutch and another language at home (133 children), and a few children used another language exclusively at home (10 children). A wide range of languages was reported, with a majority of the children speaking Moroccan Arabic, Berber or Turkish. The descriptives of the participants are provided in Table 2.1.

Table 2.1
Age and gender by language group

Language group	*N*	Gender F/M	Age (*SD*)
Monolingual	64	33/31	10;11 (0;6)
Bilingual	143	69/74	11;3 (0;6)
Total	207	102/105	11;2 (0;6)

2.2.2. Materials

2.2.2.1. Word association task

In this task, children were required to produce three associations to twenty words. A total of 80 nouns was investigated, spread over four lists of 20 words each. Each list was randomized into three orders, to avoid order effects and cheating. All words were selected from the 5000 most frequent words in Schrooten and Vermeer's corpus of school language, which includes words used in Dutch elementary education, based on sources such as school books and classroom interaction recordings (Schrooten & Vermeer, 1994). Frequent words were used to ensure all participants were familiar with the stimuli. The four lists are provided in Appendix A. To avoid chaining effects as much as possible, a format based on that used by De Deyne and Storms (2008a; 2008b) was designed. Following De Deyne and Storms, the three associations were filled out from top to bottom, rather than from left to right. In addition, all words were printed three times in a small table, with an open cell under each word for the children's answers. This was done to remind the children of the stimulus word before they made each association.

2.2.2.2. Reading test

The standardized reading test 'Begrijpend Lezen 678' by Aarnoutse and Kapinga (2006) was used to measure reading comprehension skill. It includes a total of 44 multiple choice questions and has been normed in a sample of 42 schools across the Netherlands. It is designed to measure reading comprehension in grades 4, 5 and 6. None of the schools had used this reading test before with the children who participated. The scores on the reading test were somewhat skewed but otherwise normally distributed with skewness -0.720 (SE = 0.169) and kurtosis -0.511 (SE = 0.337).

2.2.2.3. Language questionnaire

In this short questionnaire, children were asked which language they had acquired first and which languages they spoke at home, and whether they had gone to school in the Netherlands from grade 1 onwards.

2.2.3. Procedure

The word association task was administered first. In total, 12 versions of the task were devised (4 lists x 3 versions), and these were handed out randomly. In each group, each version was given to at least one child. A short instruction and an example were printed on the booklets and read out by the experimenter. It was stressed that there were no right or wrong answers, and that the test would not be checked on spelling mistakes. The participants were urged to write down the first words that came to mind in response to the stimulus words, and to use single words as much as possible. They were also asked to put down a cross for missing associations if only one or two associations came to mind easily, rather than thinking it through too much. Any unfamiliar words could also be crossed out. A maximum of 25 minutes was necessary for all children in a group to complete the task.

Subsequently, the children were administered the reading test. The use of the answer sheets was explained and an example text with questions was read out and discussed by the experimenter. All children finished the test within the set time limit of 45 minutes. Finally, the children were given the language questionnaire, which took about five minutes to complete.

2.2.4. Analysis

2.2.4.1. Classification of word association data

Based on findings on lexical organization in acquisition and on the relation between different types of semantic knowledge and reading comprehension, it was decided to use a classification system that incorporates the distinction between context-independent to context-dependent semantic relations (Cremer, 2013; Schoonen & Verhallen, 2008; Verhallen & Schoonen, 1993), but as a continuum. On this scale, the most context-independent semantic relations exist between concepts that are related because they share many intrinsic features, and are thus inherently related independent of context. These relations are included in the taxonomic category. A step further down are feature relations, where one concept expresses a necessary or prototypical feature of the other. This is again independent of context, but contrary to taxonomically related items, means only one characteristic is shared. Situational relations exist between concepts that co-occur in given contexts, but are not necessarily or prototypically related. Note that these include features that are non-prototypical. Finally, the most context-dependent category subsumes subjective relations, which are thus tied to an individual's personal context, including for example subjective evaluations of concepts. Many studies on first language acquisition have shown that children start out mainly focusing on context-dependent semantic relations between concepts that are bound through context, and abstracting from this situation-specific knowledge to more decontextualized knowledge later on (e.g. K. Nelson, 1977; 1982; 1985; 1991; 2007; Petrey, 1977; Elbers et al., 1993; Lin & Murphy, 2001). Therefore, this continuum is a useful way of studying knowledge of semantic relations.

Indeed, although the ordering discussed above is not used, to a large extent these categories are present in the extensive coding scheme from De Deyne and Storms (2008a)[6], which in turn is an extension of classifications of semantic categories by McRae and Cree (2002) and Wu and Barsalou (2009). The classification scheme used by De Deyne and Storms was somewhat simplified for our purposes by combining small subcategories into larger ones, and we incorporated the explicit ordering from context-independent to context-dependent relations. The only change we made relating to the content, is that we switched the functions subcategory from the situational category to the feature category, with an emphasis on those functions being necessary or prototypical. This was done because the function of an object is inherent to its meaning in the same way an animate beings' behavior is. As is customary in word association studies, a lexical category was included to cover purely form-related associations, which do not fall into the range of semantically oriented associations. Finally, the 'other' category includes indirect and unclear links. The full classification system with examples is provided in Table 2.2.

[6] There are a few terminological differences: we prefer the term *features* for their *entity* category, and *subjective* for their *introspective* category.

Table 2.2
Classification system of word associations

Semantic (ranging from context-independent to context-dependent)	
Taxonomic	Coordinate: *crocodile - alligator, king - queen* Superordinate: *dog - animal, trumpet - instrument* Subordinate: *dog - chihuahua, king - Willem-Alexander* Synonym: *castle - fort, cook - chef*
Feature	Context-independent characteristics (necessary or prototypical: normally, X is Y): a. Physical (colour, shape, taste, texture): *banana - yellow, ocean - deep* b. Internal (ability, traits, etc.): *snail - slow, cucumber - edible* Behavior (typical behavior or action of a stimulus): *dog - bark, thief - steal* Function (typical function of the stimulus): *cot - sleep, trumpet - music* Partonym (whole-part and part-whole): *coat - sleeve, apple - peel*
Situational	Co-occurrence (stimulus and response occur together in given possible contexts): *princess - hat, butterfly - sun* Context (response is an event, location, time which forms a context for stimulus or vice versa): *ocean - rocks, cook - restaurant* Action (action performed by or with stimulus in a situation or script): *bakery - choose, sea - swim* Context-dependent attributes (potential, X can be Y, but need not be): a. Physical: *car - blue, cook - fat* b. Internal: *king - calm, island - deserted*
Subjective	Emotion/evaluation: *cheese - eew, bride - beautiful* Personal: *hair dresser - my mother, dog - Layla*
Non-semantic	
Lexical	Compound (forward and backward): *crocodile - leather* Orthographic/phonological similarity (similar form, other meaning): *house - hose* Morphemic change: *cat - catty, dog - dogs*
Other	Mediation (linked via intermediate concept named earlier): *teacher - easy* (via *lesson*), *cucumber - fruit* (via *vegetable*) Non-classifiable (including responses to stimuli that have been misinterpreted as different words): *skirt - loud* (interpretation of *rok* 'skirt' as *rock*) Repetition No response

In some cases, multiple classifications could apply, for example when stimuli and responses formed a compound. In such cases, a code from the semantic categories was applied if it was most probable, similar to the methodology applied in Cremer et al. (2011).

The classification was done by the first author and another coder. To make sure there was sufficient agreement, the associations to two sets of words were classified by both coders. A first set of sixteen words coded in two rounds of eight words each was used as an initial training set and in order to identify potential problematic categories. In this set, 82.6% of associations were classified into the same main categories by both coders. Disagreement was resolved through discussion and the category definitions were adjusted where necessary. Another set of sixteen words, again in two rounds of eight words each was used to check whether there was sufficient agreement between both coders to code the rest of the data independently. In this second set, 94.4% were classified in the same main categories by both coders. Disagreement was resolved through discussion. Each coder then classified the associations to half of the remaining 48 words, which were randomly assigned.

2.2.4.2. Data cleaning and statistical analysis

Before the statistical analyses, all lexical associations (68 associations, 0.5% of the data) and all 'other' associations (5.0% of the data) were removed. In this way, only semantic associations remain in the final data set, consisting of 11,725 associations. For the reading task, the raw scores were used.

In order to determine the effect of response position on each of the semantic association types, each association was coded one or zero for each main association category and coded 1, 2 or 3 for response position. Separate binomial mixed effects analyses were performed for each of the four semantic association types using the lme4 package for multilevel and mixed effects analyses (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2016). It was chosen to do separate binomial analyses, so that each association type could be inspected in a mixed effects model that takes into account the multilevel and crossed nature of the data.

For the second part, individual scores for each semantic type were created by extracting the random intercepts for each participant from mixed effects models of each association category, to be able to study the relation of the association types with reading comprehension scores. These random intercepts are controlled for item and school variation and represent the individual's preference for a certain semantic type compared to the group as a whole. This resulted in four association scores for each child, one for each semantic type. Linear mixed effects models were estimated to determine the effect of each association type on the reading scores, controlled for variation between schools. The specifications of each model are discussed in more detail in the Results section.

2.3. Results

2.3.1. Quantitative differences between initial and later responses

To establish whether a larger number of different associations is obtained in the multiple association task, the average set size per stimulus for each response position and the total set was calculated. Only non-unique responses were counted, that is, responses provided by at least two participants. Note that if a particular response occurs in each response position for a particular stimulus, it will be included in each set size. Table 2.3 shows the set sizes of non-unique associations, indicating that on average, the full response set for a particular stimulus contains about six additional non-unique responses compared to the first response alone. More associative links can thus be identified using the multiple association task.

Table 2.3
Average set size of non-unique responses per stimulus

	Min	Max	*M* (*SD*)
Response 1	4	19	11.6 (3.2)
Response 2	9	21	14.1 (2.7)
Response 3	8	20	14.2 (2.6)
Total	11	28	18.1 (3.5)

2.3.2. Qualitative differences between initial and later responses

A first impression of the qualitative differences between associations across responses is provided in Table 2.4 below. Overall, the number of semantically related associations decreases since the number of ‘other’ associations, mainly null responses, increased across response position. The number of taxonomic and feature associations appears to drop, while the number of situational and subjective responses seems to increase, mostly between the first and second response, and to a lesser extent between the second and third response.

Table 2.4
Counts and percentages of association types by response position

	Counts				
	Taxonomic	Features	Situational	Subjective	Total
Response 1	586	1613	1558	252	4009
Response 2	361	1224	2045	312	3942
Response 3	289	1072	2096	317	3774
Total	1236	3909	5699	881	11725
	As percentage of response position				
	Taxonomic	Features	Situational	Subjective	Total
Response 1	14.6	40.2	38.9	6.3	100.0
Response 2	9.2	31.1	51.9	7.9	100.0
Response 3	7.7	28.4	55.5	8.4	100.0
Total	10.5	33.3	48.6	7.5	100.0

To determine the effect of response position statistically, four separate binomial mixed effects models were made with each binary coded semantic association type as dependent variable. The estimated main effect of response position in these models is therefore the increase or decrease depending on response position of the probability, in log odds, that a response is of this semantic type. Random intercepts for item, participant and school were included to take into account the multilevel structure of the data, where participants are nested in schools and crossed with items. The random effects structure was kept maximal with regards to the grouping variables of interest, items and participants (c.f. Barr, Levy, Scheepers, & Tily, 2013; Linck & Cunnings, 2015). Random slopes for response position for schools were not estimated, since we do not have enough schools to reliably determine these variances. We did estimate the more restricted models excluding random slopes for participants and items as well, and verified that for each category, the maximal model indeed was the best fit to the data. Response position was included as a continuous variable.[7] The estimates for the maximal models for each association type are provided in Table 2.5.

The main effects of response position confirm the patterns observed in Table 2.4. The probability of a response being taxonomic or a feature association decreases across response position (β = -0.206 (0.095), z = -2.18, p = 0.030, meaning a 19% decrease between each response position and β = -0.268 (0.070), z = -3.82, p < 0.001, a 24% decrease, respectively), while the probability of situational and subjective associations increases (β = 0.433 (0.060), z = 7.24, p < 0.001, a 54% increase, and β = 0.421 (0.081), z = 5.17, p < 0.001, a 52% increase, respectively). Thus, even taking into account variation on the item and participant levels, the effect of response position is significant for all four semantic categories.

An interesting exploratory observation is that for each category, the random intercept and slope variation between items is larger than between participants, with random intercept variance for items ranging from 2.459 to 3.888, compared to the random intercept variance for participants which ranges from 0.201 to 1.629. This suggests that some items are more conducive to associations of a particular type than others, and while participants also vary in how likely they are to provide certain types of associations, this variation is comparatively smaller. This pattern holds for each semantic type.

[7] We also considered inclusion of response position as a categorical variable, but this lead to convergence issues in the models which include random slopes for response position. We compared the models without random slopes with the categorical and continuous response position variables, and these showed the same patterns across response position for taxonomic, feature and situational associations. Only for subjective associations, there was no increase between response 2 and response 3 when response position was included as a categorical variable.

Table 2.5

Mixed effects model estimates for the maximal models of each associative category (random slopes by response position)

	Taxonomic			Features			Situational			Subjective		
Random effects												
Items												
Intercept	3.888			3.372			3.479			2.459		
Slope	0.382			0.297			0.206			0.126		
Participants												
Intercept	1.629			0.201			0.396			1.400		
Slope	0.186			0.052			0.053			0.039		
School												
Intercept	0.004			0.017			0.000			0.050		
Fixed effects	Estimate (*SE*)	*z*	*p*	Estimate (*SE*)	*z*	*p*	Estimate (*SE*)	*z*	*p*	Estimate (*SE*)	*z*	*p*
Intercept	-2.341 (0.267)	-8.76	0.000	-0.362 (0.223)	-1.63	0.010	-0.954 (0.222)	-4.30	0.000	-4.099 (0.269)	-15.26	0.000
Response position	-0.206 (0.095)	-2.18	0.030	-0.268 (0.070)	-3.82	0.000	0.433 (0.060)	7.24	0.000	0.421 (0.081)	5.17	0.000

Table 2.6

Descriptives for reading comprehension scores and raw number of 'other' answers by language groups

	Total (*N* = 207)		Monolingual (*N* = 64)		Bilingual (*N* = 143)				
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*(*df*)	*p*	*d*
Reading comprehension	29.99	6.08	32.06	6.43	29.06	5.69	3.210 (109)	0.002	0.49
Number of 'other' responses	3.03	4.25	1.98	2.99	3.50	4.64	-2.808 (179)	0.006	0.39

In the same vein, comparing the intercept variances between models, i.e. the variances of the different association types, we see that the by-participant random intercept variance estimates for the taxonomic (1.629) and subjective items (1.400) are noticeably larger than for the feature (0.201) and situational associations (0.396). Children thus differ more in how likely they are to provide taxonomic and subjective associations, than with respect to feature and situational associations. To a certain extent, this may well be linked to the fact that these categories are simply smaller, meaning that variation in log odds is likely to be larger. For example, producing three more associations in a category in which most children produce only a small number, makes a larger difference than producing three more associations in a category in which most children produce many associations. However, given that this principle should also apply to the variation between items, which appear more similar by comparison in terms of intercept variance (i.e. 3.888 for taxonomic associations, 3.372 for features, 3.479 for situational associations, and 2.459 for subjective associations), this cannot be the only explanation for this pattern.

2.3.3. Differences between language groups

To determine the effect of language status, we added a main effect of language group by means of a binary factor [0 monolingual, 1 bilingual] and the interaction with response position to the mixed effects models testing the effect of response position, but neither was significant for any of the four association types (extended models not shown). In our sample, the monolingual and bilingual children thus do not differ significantly in terms of preferences for any of the four association types, independent of response position. We checked whether the monolingual and bilingual children in our sample did differ in terms of language skill as far as possible by comparing their scores on the other two measures of language knowledge we had at our disposal, namely reading comprehension and the number of 'other' responses. Table 2.6 shows that there were small but significant differences for both measures: bilingual children produced more 'other' responses (Cohen's d = 0.39, a small effect), suggesting that their vocabularies were smaller, and bilingual children performed worse on the reading comprehension scores (Cohen's d = 0.49, close to a moderate effect size). In other words, although the bilingual children showed worse performance in reading comprehension and a very rough measure of vocabulary size, they did not show differences in their associative preferences.

2.3.4. Associative behavior and reading comprehension

To study the effect of individual children's inclination to produce certain associative types on their reading comprehension scores, we extracted the random intercepts for each participant from a model including only the intercept in the fixed part of the model, and random intercepts for item, participant and school. The participant intercepts serve as a score for the tendency of the individual children to produce an association of the semantic type in question, compared to the group as a whole. In short, the random intercepts for participants are the individual deviations in log odds from the overall mean log odds (the fixed intercept) of producing a certain association type, controlled for item and school variance (the random intercepts for item and school).

We extracted these scores from the full dataset, i.e. responses 1, 2 and 3, and from two reduced data sets including only response 1 on one hand, and responses 2 and 3 on the other hand. These were entered into a series of mixed effects models as continuous predictors for the reading comprehension scores, with school as a random intercept. In order to compare the predictive value of the later responses relative to the initial responses, we estimated models including the joint score for responses 1, 2 and 3, two models including the two separate scores for response 1 and responses 2 and 3, and a model including both separate scores together.

The situational and subjective scores did not have any effect on the reading comprehension scores, regardless of whether the full set or either of the subsets were considered, and therefore these will not be discussed any further here. However, the taxonomic and feature associations did show an effect in each or some of the analyses, which we will discuss below. The models for the taxonomic association scores are presented in Table 2.7.

Comparing the four models, an interesting effect of the different response positions becomes apparent. The combined taxonomic score including all responses in model 1 has a significant positive effect on the reading comprehension scores: β = 2.854 (0.892), $t = 3.20$, $p = 0.001$. When children's log odds of associating taxonomically compared to the whole group increased by 1, their odds of producing taxonomic associations increased by a factor 2.718), their reading score increased by 2.854. In other words, having a log odds score of 0.35, that is, producing about 42% more taxonomic associations than the average child, results in a 1 point increase on the reading score, which is about 1/6 of the standard deviation for the reading scores.

When considered separately in models 2 and 3, the scores on the first response and the second and third responses show a significant positive effect: β = 1.460 (0.627), $t = 2.33$, $p = 0.020$, and β = 3.929 (1.239), $t = 3.17$, $p = 0.002$, respectively. However, when we enter both separate scores in the same analysis, model 4, the effect of the first response is no longer significant: β = 0.800 (0.675), $t = 1.19$, $p = 0.236$. In this case, only the second and third responses are predictive of the reading scores: β = 3.290 (1.343), $t = 2.45$, $p = 0.014$, apparently including variance explained by the first response in model 2. This corresponds to a log odds score of 0.30, or producing about 36% more second and third taxonomic responses compared to the group as a whole for a 1 point increase in reading scores. We checked whether there was a significant interaction between the scores from response 1 and responses 2 and 3, which was not the case. This suggests that in principle, the children who provide more taxonomic links overall are better comprehenders. However, especially the children who manage to produce taxonomic associations as second and third responses, and thus presumably know more taxonomic links for individual words, are better comprehenders. This illustrates an interesting difference between initial and later associations.

Table 2.7

Mixed effects model estimates for the effects of taxonomic association scores on reading comprehension (random intercepts for school)

	1			2			3			4		
Random effects												
School	2.10			2.21			1.46			1.72		
Residual	33.72			34.47			33.98			33.81		
Fixed effects	Estimate (*SE*)	*t*	*p*	Estimate (*SE*)	*t*	*p*	Estimate (*SE*)	*t*	*p*	Estimate (*SE*)	*t*	*p*
Intercept	29.869 (0.661)	45.19	0.000	29.860 (0.674)	44.29	0.000	29.871 (0.597)	50.03	0.000	29.853 (0.624)	47.85	0.000
Taxonomic score responses 1, 2, 3	2.854 (0.892)	3.20	0.001									
Taxonomic score response 1				1.460 (0.627)	2.33	0.020				0.800 (0.675)	1.19	0.236
Taxonomic score responses 2, 3							3.929 (1.239)	3.17	0.002	3.290 (1.343)	2.45	0.014

Table 2.8

Mixed effects model estimates for the effects of feature association scores on reading comprehension (random intercepts for school)

	1			2			3			4		
Random effects												
School	1.73			1.71			1.80			1.71		
Residual	35.32			34.83			35.53			34.99		
Fixed effects	Estimate (*SE*)	*t*	*p*	Estimate (*SE*)	*t*	*p*	Estimate (*SE*)	*t*	*p*	Estimate (*SE*)	*t*	*p*
Intercept	29.911 (0.630)	47.45	0.000	29.914 (0.626)	47.76	0.000	29.905 (0.639)	46.81	0.000	29.914 (0.627)	47.69	0.000
Feature score responses 1, 2, 3	-2.800 (2.356)	-1.19	0.235									
Feature score response 1				-3.852 (1.854)	-2.08	0.038				-3.871 (1.872)	-2.07	0.039
Feature score responses 2, 3							-0.305 (1.824)	-0.17	0.867	0.148 (1.824)	0.08	0.936

For the feature associations, the same models were estimated, which are presented in Table 2.8. Here, we see that the combined score in model 1 does not have a significant effect (β = -2.800 (2.356), *t* = -1.19, *p* = 0.235), nor does the score for the second and third responses in model 3 (β = -0.305 (1.824), *t* = -0.17, *p* = 0.867). Only the separate score for the first response contributes significantly to the reading score, whether entered separately in model 2 (β = -3.852 (1.854), *t* = -2.08, *p* = 0.038), or together with the second and third responses in model 4 (β = -3.871 (1.872), *t* = -2.07, *p* = 0.039). The significant effect is negative, indicating that children who produced more feature associations as first responses, perform worse on the reading comprehension test. In model 2 for example, having a log odds score of 0.26, that is, producing about 30% more feature associations than the average child, was associated with a one point decrease in the reading scores. Again, including an interaction between the scores for the initial and later responses did not improve the model.

A possible explanation is that the feature associations are in a complementary relation with the taxonomic associations. In this case, the children who do not produce many taxonomic associations may produce feature associations instead, the next category on the context-independent/context-dependent continuum. It seems that this points to a less well-developed semantic network, which negatively affects reading comprehension. This explanation is supported by the fact that taxonomic associations are especially prominent in the first response position, and it is the feature associations in this position that are negatively correlated with reading comprehension.[8]

[8] Note that it is not possible to confirm or reject this hypothesis by calculating the correlation between the taxonomic and feature scores. Since we created the associative variables by recoding each association into four binary types, they are by definition negatively correlated.

2.4. Discussion

2.4.1. Quantitative and qualitative differences between words' semantic networks resulting from initial and later associations

In this study, we set out to examine how our impression of associative networks on the word level and the individual mental lexicon level are affected by the use of a single versus multiple word association task, in monolingual and bilingual minority children. On the word level, we expected to find more and more semantically diverse associations, based on the findings by De Deyne and Storms (2008a; 2008b), which was indeed borne out by the data. Regarding the number of associations, when including second and third associations, over 1.5 times more non-unique associations, i.e., associations provided by at least two participants, were found per stimulus on average. This means that second and third associations do not represent exclusively idiosyncratic semantic links, which are specific to the individuals producing them. On the contrary, the later associations add useful information about the stimulus word's semantic network. Such information is especially relevant when word association data are used as a control measure for other tasks, such as priming experiments. When association strength between word pairs needs to be zero or as close to zero as possible, using single word association norms means that word pairs may be included which are actually associated, even if it is in the second or third instance. Such additional links are especially relevant when a stimulus word has a very strong first associate, which may drown out other associations, for example in the case of *blood* - *red* (De Deyne & Storms, 2008b). In and of itself, this finding is a strong argument for using multiple instead of single word association data, even if that is currently not common practice.

As for the types of semantic links that are found in initial versus later associations, interesting differences were also observed. Recall that for adults, De Deyne and Storms (2008a) observed a clear decline in taxonomic responses in the second and third position, while feature, situational and subjective associations became more prominent. In our data, these patterns were mimicked by all categories except the feature category, which actually decreased across responses, similar to the taxonomic category. On the word level, this suggests that not only do second and third associations add more semantic links, these links are also qualitatively different. Again, this is highly relevant when word association norms are used as control measures. For example, when situational semantic links are studied and need to be controlled for association strength, using norms derived from single associations may miss a substantial amount of associative links that would be found in a multiple association task (cf. Spätgens & Schoonen, 2018).

2.4.2. Single versus multiple associations in individual language users' mental lexicons

2.4.2.1. Differences between children and adults

Our findings also have implications for our understanding of language users' semantic networks in general, and more specifically children's mental lexicons. Firstly, as was already found for adults by De Deyne and Verheyen (2015), semantic relations in the mental lexicon are much more context-dependently oriented than previously thought, and this is also the case for the children in our study. Only a minority of associations is taxonomic (10.5% across responses in our study), while most associations are situational in nature (48.9%). As Borghi and Caramelli (2003) have pointed out, it makes sense that the number of context-dependent associations would exceed taxonomic associations, given that they include a far larger set of possible ties, including locations, actions, and a variety of types of entities that may co-occur together in a host of different contexts.

Secondly, we see some interesting similarities and differences between the behavior of our 10-11-year-old participants and the adult data presented by De Deyne and Storms (2008a). Starting with the taxonomic associations, it is interesting to see that children exhibit the same decreasing pattern across responses, showing that as in adults, if a taxonomic association is produced, it is most likely produced as a first response. In terms of absolute numbers however, we do see that overall, the children produce fewer taxonomic associations: 10.5% in our study, compared to 19.6%[9] for the adults in De Deyne and Storms' work. We can conclude that although the children in this age group possess fewer taxonomic links, the available taxonomic knowledge is already organized similarly to that of adults, since it is more prone to being produced early as well.

Another interesting difference shows up in the feature category, the only category where the association pattern across responses is actually different from that observed in adults. Mirroring the taxonomic associations, children produce more feature associations in the first instance, and across the second and third responses, this number drops. We would like to tentatively suggest that this also relates to the development from context-dependent to context-independent knowledge: features may be the 'next best thing' on the way to a more context-independently organized semantic network. When context-independent knowledge is not yet fully fledged, feature relations may be more available to adhere to the apparent preference for producing context-independent knowledge earlier. In terms of situational and subjective associations, the children in our study show similar patterns to the adults in De Deyne and Storms (2008a).

On the individual mental lexicon level, the multiple word association data thus provide us with additional information to compare associative behavior in different age groups and examine the process of children's mental lexicons developing towards the adult patterns. In terms of the development from a main focus on context-dependent knowledge toward the extension to context-independent knowledge (c.f. Elbers et al., 1993; K. Nelson, 1977; 1982; 1985; 1991), the children in our age group seem to be fairly close to the adult behavior, but are still expanding their context-independent knowledge and preferences in various ways. An interesting path for future research would be to compare more age groups in the same way and see how not only the proportions of answers in different categories develop during childhood, but also across response positions.

[9] We recalculated the percentages provided by De Deyne and Storms (2008a) to reflect the percentage of taxonomic associations among the semantic categories, i.e. by leaving the lexical associations out in the same way we did in our study. Including lexical associations, they arrived at 18.9% taxonomic associations.

2.4.2.2. Differences between monolinguals and bilinguals

We also set out to test whether the multiple association task would be more suited than the single task to study differences in monolingual and bilingual associative behavior. Since previous studies have not always found clear patterns (e.g. Cremer et al., 2011; Fitzpatrick, 2006; 2007; 2009; Söderman, 1993), we hypothesized that potentially, there are very subtle differences between monolingual and bilingual speakers' semantic networks, which may turn up in secondary and tertiary associations if not in initial associations. However, language status was not predictive of any of the four types of semantic associations and did not interact with response position, suggesting that language status did not play a role in our sample's associative behavior.

Of course, there is no such thing as *the* bilingual or *the* L2 learner, and factors such as age of acquisition and amount of exposure will play an important role in any measure of bilingual language skills. It is therefore not surprising that studies with different populations find different results. Indeed, the bilingual children in this study were almost all born in the Netherlands, all started Dutch elementary school at age four, and the majority spoke Dutch at home in addition to their L1. As such, these bilingual learners have had extensive Dutch language exposure and may therefore have similar Dutch language knowledge compared to their monolingual counterparts. However, looking at the two other measures of language skill that we had at our disposal, the reading comprehension scores and the number of 'other' i.e. null or non-classifiable responses, we did find small but significant differences between the two groups, with monolingual children outperforming bilingual children. This suggests that there are at least subtle differences in language skill. However, these differences did not extend to the association task, meaning that for those words that the bilingual children do know, their associative behavior is similar to that of monolingual children.

Our findings mirror those obtained in the single word association task reported by Cremer et al. (2011), who did not find differences between their monolingual and bilingual learners in terms of production of associations in the various semantic categories. The subtle differences in preferences for different semantic categories we expected to turn up in our multiple task were not borne out by the data. A potential explanation may be that the L2 semantic network structure is more a property of the L2 learner's existing knowledge, than of the individual words being learned. As was proposed by Wolter (2006), L2 learners may use the semantic network from their L1 to structure their L2 knowledge. Perhaps especially in the case of bilinguals and advanced L2 learners, new words may be integrated into this structure fairly quickly, resulting in the apparent incongruity of the number of words known being lower in bilinguals while the association behavior is similar. Furthermore, the stimulus words were all concrete concepts and high frequency words, selected to be known to all children as much as possible. It may be the case that with more abstract and less frequent words, the monolingual and bilingual children would have shown different associative behavior, an effect that has also been observed in L1 adults who responded to frequent and infrequent words with different types of associations (W. S. Stolz & Tiffany, 1972).

2.4.2.3. Single versus multiple associations and reading comprehension

Finally, to assess how the denser semantic networks arising from multiple association tasks may relate to other language skills, we looked into their effect on reading comprehension scores. We found that situational and subjective associations were not related to reading comprehension, independent of response position. This was expected based on previous studies that have shown that context-independent knowledge and especially taxonomic knowledge may be predictive of reading comprehension skill (Bonnotte & Casalis, 2010; Cremer & Schoonen, 2013; Nation & Snowling, 1999; Ouellette, 2006). Taxonomic and feature associations were related to the reading comprehension scores, and each showed interesting differences between initial and later responses.

The taxonomic associations' relation to reading comprehension scores showed an interesting difference between earlier and later responses. Examining the full set, the first response set and the second and third response set all resulted in positive effects on the reading comprehension scores, suggesting that overall, knowledge of taxonomic relations is associated with better reading comprehension. However, when entered together, the effect of the first response set scores disappeared, and only the taxonomic scores of the second and third responses had a significant positive effect on the reading scores. Regarding the relation between reading comprehension and knowledge of semantic relations, this suggests that especially children who have a stronger preference for taxonomic links and can produce not only one, but also more taxonomic relations, are better at reading comprehension. It is interesting to see that even this spontaneous and unconstrained vocabulary task can be related to reading comprehension skill. However, the effects appear to be fairly modest, with children having to produce about 42% more primary, secondary and tertiary taxonomic responses compared to the group as a whole to achieve a 1 point or about 1/6 standard deviation increase in their reading score.

The initial but not the secondary and tertiary feature associations were also related to the reading scores, but negatively so. This seems to go against findings from other studies in which defining characteristics were included among other semantic aspects in measures of vocabulary depth, which were positively related to reading comprehension (Cremer & Schoonen, 2013; Tannenbaum et al., 2006). As we proposed in the Results section, the presence of this negative relation despite the fact that feature associations are fairly context-independent in nature, may be the result of a complementary relation between taxonomic and feature associations. This hypothesis is strengthened by the observation that the feature associations showed the same decline across responses as the taxonomic associations, a characteristic of context-independent knowledge which in adults is reserved for the taxonomic responses only. In the development towards a more context-independently oriented network structure, feature associations may form a stepping stone. Our data suggest that children who are still prone to produce these feature associations early, i.e. whose semantic network is not as far developed yet, are also those who perform worse in terms of reading comprehension. This reinforces the idea that especially the most context-independent knowledge is associated with good performance in reading comprehension. The difference between our findings and those obtained by Cremer and Schoonen and Tannenbaum et al. is probably also related to the fact that in our study, the participants were not explicitly asked to produce or select defining characteristics. The preferred routes in the children's semantic networks that are spontaneously accessed in the association task appear to bear a different relation to reading comprehension.

In these analyses of reading comprehension, using only single responses would have provided us with seemingly similar outcomes, namely a positive effect of taxonomic associations and a negative effect of feature associations. However, the fact that we gathered multiple associations allowed us to discover a different and more complex relation between these two types of semantic knowledge and reading comprehension. The shape of the effects provide an interesting glance at the relation between the development of the semantic network structure and reading comprehension.

2.4.3. Limitations

There are a few limitations to the present study that should be taken into account in the interpretation of the results and for future studies. Firstly, the number of words and participants included was relatively small, compared to the enormous association data bases that exist. This means that for example the numbers of additional associations that we found in the second and third answer sets, should be interpreted with some caution. However, it is not the case that because of this, some words only appeared in the second and third instance which would have turned up as first associations in a larger sample: De Deyne and Storms (2008b) found similar set sizes for the first, second and third responses, but a much larger set size for the collapsed set, namely 31 non-unique responses compared to 18 in our sample. If anything, our small sample might therefore lead us to underestimate the additional variability that can be detected using a multiple task.

The number of items each child responded to may also affect the validity of the association scores for the different semantic types, since more accurate associative profiles could be gained from larger samples of items. We controlled for this as much as possible by including random intercepts for items in our analyses, meaning that if certain items were more conducive to certain types of associations, these differences were taken into account in the calculation of the personal association scores.

As was mentioned earlier, the stimuli were selected to be highly frequent, concrete concepts, which may be a reason why we did not find any differences between the monolingual and bilingual children. It is also likely that with more difficult stimuli, a different effect of associative behavior on the reading task might have emerged. This would be an interesting question for future research.

Finally, the inclusion of a standardized vocabulary size measure to control for vocabulary size effects[10] on the reading comprehension scores would have strengthened our conclusions regarding the relation between associative behavior and reading comprehension. This would be a meaningful additional measure for studies exploring the relation between associative behavior and other linguistic skills in more detail.

2.5. Conclusions

Despite the limitations discussed above, our study has shown that the use of the multiple association task leads to interesting insights on the structure of the mental lexicon. Similarly to the larger adult studies by De Deyne and Storms (2008a; 2008b), considering multiple responses resulted in larger and qualitatively different semantic networks compared to single responses, which is highly relevant for our interpretation and use of association norms. The differences and similarities between adults' and children's association patterns across responses provide valuable information on the development of the semantic network structure. An interesting topic for future research would be to examine these patterns in more different age groups and with a larger variety of stimulus words such as less frequent or more abstract words, to map the development of semantic links alongside vocabulary growth.

Expanding to other learner groups and linguistic skills, we found that young monolinguals and bilinguals show similar associative behavior, while differences in reading comprehension skill do appear to be affected by associative preferences in various ways. Again, the initial and later

---

[10] We considered using the 'other' associations as a rough measure of vocabulary size, but these scores are not reliable since over a third of the children did not produce any 'other' associations.

associations appear to foreground different types of knowledge, relating to reading comprehension skill in different ways. These results open up a fresh perspective for research on the use of associations as measures of vocabulary knowledge. As Fitzpatrick (2013) has noted, the large number of studies devoted to this issue remain inconclusive, and our results suggest the multiple association task could be a useful approach. The fact that the results from this unconstrained task are related to such a complex skill as reading comprehension, are also encouraging for the more detailed study of the role of knowledge of semantic relations in reading comprehension. Examining this relation further using more focused tasks such as priming experiments (e.g. Cremer, 2013; Nation & Snowling, 1999) or self-paced reading tasks involving different types of semantic relations, is an interesting avenue for future research which is explored in Chapters 3 and 4.

Chapter 3

The semantic network, lexical access and reading comprehension in monolingual and bilingual children: An individual differences study[11]

Abstract

Using a semantic priming experiment, the influence of lexical access and knowledge of semantic relations on reading comprehension was studied in Dutch monolingual and bilingual minority children. Both context-independent semantic relations in the form of category coordinates, and context-dependent semantic relations involving concepts that co-occur in certain contexts were tested in an auditory animacy decision task, along with lexical access. Reading comprehension and the control variables vocabulary size, decoding skill and cognitive processing speed were tested by means of standardized tasks. Mixed effects modelling was used to obtain individual priming scores and to study the effect of individual differences in the various predictor variables on the reading scores. Semantic priming was observed for the coordinate pairs but not the context-dependently related pairs, and neither context-independent priming nor lexical access predicted reading comprehension. Only vocabulary size significantly contributed to the reading scores, emphasizing the importance of the number of words known for reading comprehension. Finally, the results indicate that the monolingual and bilingual children perform similarly on all measures, suggesting that language status was not highly predictive of vocabulary knowledge and reading comprehension skill.

---

[11] A slightly modified version of this chapter has been previously published as: Spätgens, T. & Schoonen, R. (2018). The semantic network, lexical access and reading comprehension in monolingual and bilingual children: An individual differences study. *Applied Psycholinguistics*, *39*(1), 225-256.

3.1. Introduction

As children's school careers progress, more and more emphasis is placed on the acquisition of knowledge from written texts, making reading comprehension a fundamental skill for school success for both monolingual and bilingual minority children. It is therefore of paramount importance to understand how reading comprehension functions, and to tease apart the components of language competence that feed into it. One component that has received considerable attention is vocabulary size, and its importance for reading comprehension is already well-established (e.g. Alderson, 2005; Grabe, 2009; Stæhr, 2008; Stanovich, 2000). However, apart from its size, the quality of word knowledge may well be highly relevant as well, as posited in Perfetti's Lexical Quality Hypothesis (Perfetti & Hart, 2002; Perfetti, 2007) and confirmed by a number of empirical studies investigating lexical fluency and the semantic network (Cremer, 2013; Qian, 1999; Tannenbaum et al., 2006). This could also be a source of differences in reading comprehension between monolingual and bilingual children, since young bilinguals have been found to lag behind their monolingual peers both in terms of knowledge of semantic relations (e.g. Cremer et al., 2011; Verhallen & Schoonen, 1993) and reading comprehension (e.g. CBS, 2014; Droop & Verhoeven, 2003; Smits & Aarnoutse, 1997). For these reasons, this study focuses on the influence of the quality of the semantic network and lexical access on reading comprehension in Dutch monolingual and bilingual children.

In the mental lexicon, lexical items are organized in a semantic network structure (Aitchison, 2012) in which they are linked through various types of semantic relations. These semantic connections are part of our word knowledge and develop over time. Following Verhallen and Schoonen (1993), Schoonen and Verhallen (2008), and Cremer (2013), we will focus on the difference between context-dependent and context-independent semantic relations. Whereas the former hold between words or concepts which occur together in context, the latter are more intrinsically motivated, existing between words that are related independent of context, and which often share inherent qualities (Cremer, 2013; Schoonen & Verhallen, 2008; Verhallen & Schoonen, 1993). The two terms represent the extremes of a continuum, on which we can place different types of relations. Example word pairs are *squirrel - cute*, a subjective and therefore highly context-dependent relation; *squirrel - forest*, which are related through frequent co-occurrence and therefore also context-dependent, but the relation is more semantically oriented; and *squirrel - animal*, which share many intrinsic qualities and are related independent of context.

The relevance of this distinction can be observed in both monolingual and bilingual language acquisition. In monolingual vocabulary acquisition, context-dependent knowledge precedes context-independent knowledge, since children need to abstract from direct experience to more generalized, decontextualized knowledge (K. Nelson, 1977; 1982; 1985; 1991; 2007; Petrey, 1977; Elbers et al., 1993; Lin & Murphy, 2001). Bilinguals have been found to have generally less extensive semantic knowledge in their L2 compared to monolinguals in their L1, for example providing fewer semantically oriented word associations (Cremer et al., 2011), which implies a smaller vocabulary but also more limited semantic connections to familiar words. A definition task and structured interview by Verhallen and Schoonen (1993) also showed that bilinguals were especially behind in terms of their context-independent knowledge. For instance, in defining common Dutch words, the bilinguals produced fewer words that bore a context-independent relation to the target items.

Various studies have already established that knowledge of semantic relations contributes to reading comprehension. For example, Tannenbaum, Torgesen and Wagner (2006) found that in monolinguals aged 9-10, the ability to provide synonyms and multiple attributes such as category, function and location for nouns, i.e. both context-independent and more context-dependent semantic knowledge, was associated with higher reading scores. Combined with data from a sentence production task and a category generation task, where subordinates were produced in response to category labels, these measures were able to account for unique variance in the children's reading comprehension scores, over and above vocabulary size. Similarly, Ouellette (2006) found that the ability to produce synonyms, unique semantic features and category superordinates contributed to reading comprehension in monolingual children, even more so than vocabulary size. The relevance of the contrast between context-independent and context-dependent semantic knowledge for reading comprehension has also been targeted specifically, by Cremer and Schoonen (2013). They used the Word Associates Test (Schoonen & Verhallen, 2008), which required their 10-11-year-old participants to distinguish subordinates, superordinates, synonyms, meronyms and defining characteristics from contextually related distractor items, such as *banana – slip*. The children who were better at selecting the context-independently related items also obtained higher reading scores, suggesting that these items may be particularly important for reading comprehension.

It is important to note that in all these reading comprehension studies, the vocabulary and reading tasks were unrelated, that is, the words used in the vocabulary tasks were not selected from the texts in the reading tasks. This means that generally more extensive semantic knowledge contributes to reading comprehension, and we would like to argue that this may be due to the working and structure of the semantic network. The various tests used to assess word knowledge involve different types of semantic relations, which are represented in the semantic network structure and activate each other through spreading activation (cf. Bock & Levelt, 1994; Collins & Loftus, 1975). It could be exactly this spreading activation in a well-developed semantic network that helps reading comprehension, for example by allowing the reader to connect related concepts within the text more quickly and easily, thus helping interpretation of the text by establishing coherence (cf. Van den Broek et al., 1999; Van den Broek et al., 2005).

This explanation is supported by a few studies that have used online tasks to test spreading activation for various semantic relations and have found that there is a connection with reading comprehension skill. For instance, Nation and Snowling (1999) compared groups of poor and proficient monolingual comprehenders aged 10-11, using an auditory semantic priming experiment which involved categorically and functionally related word pairs, i.e. context-independent and slightly more context-dependent relations. They found that in the absence of associative relations between words, poor comprehenders showed no priming for the categorically related word pairs, while the groups were comparable for the functionally related items. These results suggest a special role for knowledge of category relations compared to functional relations, and thus for context-independent compared to more context-dependent knowledge. Bonnotte and Casalis (2010) performed a similar study with a visual instead of an auditory task, and found similar results for categorical priming, but a different pattern for the functional items. Skilled readers did not exhibit functional priming, and poor readers only showed functional priming for pairs that were also associatively related. The authors argue that the longer SOA, 800ms in their study, compared to an ISI of 500ms in Nation and Snowling's study, might be responsible for the different results. An additional difference is that Bonnotte and Casalis used paired presentation, while Nation and Snowling used single presentation, where participants responded to all items. What both studies show, however, is that differences in sensitivity to priming of various types of semantic relations may be associated with differences in reading skill.

An interesting question that remains is whether this relation between reading comprehension and online measures of the interconnectedness of the semantic network can also be found on the individual level, since this could have important implications for vocabulary instruction as a means of improving reading comprehension skill. The studies by Nation and Snowling (1999) and Bonnotte and Casalis (2010) have compared groups of poor readers and skilled readers, who were selected to be quite far apart in terms of reading competence. When we look at average readers, can the strength of individuals' semantic networks predict their reading comprehension? Indeed, Larkin, Woltz, Reynolds and Clark (1996) used a semantic priming experiment involving a synonym judgement task, where the semantic relation between primes and targets was also always synonymy, i.e. a context-independent semantic relation. Words were presented in pairs such as *big – huge*, which would be a prime for the pair *large – giant* at a lag of zero to two intervening items. The priming scores were positively associated with reading comprehension in sixth-graders, even explaining 26% of the variance in the reading comprehension scores. Conversely, using a semantic classification task, Cremer (2013) investigated individual differences in categorical, i.e. also context-independent, priming, and reading comprehension and found no relation, even though her stimuli, like Nation and Snowling's (1999), were category coordinates.[12] Therefore, while the study by Larkin and colleagues suggests that there is a connection between semantic priming and reading comprehension on the individual level, the findings by Cremer suggest that differences between average readers may be too small to detect such a relation. The different semantic relations that were used, synonymy versus category membership, may have caused the different findings.

In this study, our main aim is to partially replicate and extend the studies by Cremer (2013) and Nation and Snowling (1999) to further examine the connection between reading comprehension and context-dependent and context-independent priming on the individual level. Based on Nation and Snowling's findings, we predict that higher reading comprehension scores will be associated with higher context-independent priming scores, reflecting the advantage for children with more developed semantic networks in reading comprehension.

A second aim of this study is to look at a third dimension of vocabulary knowledge alongside size and network structure, namely fluency of retrieval of semantic knowledge (Beck et al., 1982). We will use the term semantic access, or access for short, since fluency has been used to refer to the automaticity of a variety of subprocesses in reading, such as word attack, word identification, and comprehension (Wolf, Miller, & Donnelly, 2000), but also generating category members and producing meaningful sentences involving target words (Tannenbaum et al., 2006). The Lexical Quality Hypothesis (Perfetti & Hart, 2002; Perfetti, 2007) posits that reading comprehension depends on the quality of word representations, where representations that are high in quality can be accessed effortlessly, which leaves more processing capacity available to be devoted to higher-order comprehension processes. Indeed, Cremer (2013) found that semantic access as measured by response times in a semantic classification task

[12] Nobre & Salles (2016) also looked at the relation between individual differences in semantic priming and reading comprehension. However, the semantic relations included were not specified and the set-up of the experiment was conducive to strategic processing, making it difficult to draw conclusions on the effect of subconscious processing of different types of semantic relations on reading comprehension.

could explain a small amount of variance in the reading comprehension scores of monolingual and bilingual readers, namely 2.5%, in addition to the variance already explained by knowledge of context-independent semantic relations and decoding. We therefore predict that children who can access their semantic knowledge faster, will also show better reading comprehension.

The current study's third aim is to compare Dutch monolingual and bilingual minority children in terms of knowledge of semantic relations, semantic access and reading comprehension. A number of studies in the Netherlands have found that bilinguals lag behind their monolingual peers in terms of reading comprehension and various types of vocabulary measures (Berkel et al., 2002; Cremer, 2013; Heesters et al., 2007; Sijtstra et al., 2002; Smits & Aarnoutse, 1997). These consistent delays are found despite the fact that most elementary school children from a minority background in the Dutch context are second or third generation immigrants (CBS, 2016) and mostly speak Dutch at home in addition to their L1 (Berkel et al., 2002; Heesters et al., 2007; Sijtstra et al., 2002).

The weaker links hypothesis (Gollan & Silverberg, 2001; Gollan et al., 2005; Gollan et al., 2008; Michael & Gollan, 2005) provides an explanation for these perhaps counterintuitive findings. According to the hypothesis, bilinguals are at a disadvantage due to reduced exposure and use of each of their languages, compared to monolinguals who receive all exposure in a single language. This has been found to negatively affect productive vocabulary in bilinguals compared to monolinguals (Gollan et al., 2005). Since the availability and automaticity of semantic connections in the mental lexicon can also only develop through experience with these semantic connections, reduced exposure is likely to affect the semantic network of bilinguals as well. In addition, since the school environment is where a large amount of decontextualized semantic knowledge is transmitted, the children with less well-developed knowledge of the language of instruction are likely additionally negatively affected in the development of context-independent semantic knowledge.

Indeed, as was discussed earlier, there is evidence from previous studies that bilinguals have more limited knowledge of semantic relations in their L2, especially context-independent knowledge (cf. Berkel et al., 2002;

Verhallen & Schoonen, 1993) and may access semantic information more slowly (Cremer, 2013).[13] These findings, combined with the fact that especially context-independent knowledge may be particularly important for reading comprehension (cf. Nation & Snowling, 1999), lead us to expect that these lower-order vocabulary knowledge components may be a source of the often lower reading comprehension scores also found in bilinguals in the Dutch context. This hypothesis was also put forward and confirmed by Cremer and Schoonen (2013), who found that differences in reading comprehension between monolingual and bilingual children were mediated by offline knowledge of semantic relations. However, Cremer (2013) did not find a contribution of online knowledge of semantic relations for either monolinguals or bilinguals, but did find that differences in semantic access were partially responsible for differences in monolingual and bilingual reading scores. In this study, we intend to partially replicate and extend these findings by comparing the effect of both context-dependent and context-independent knowledge and semantic access on reading comprehension in Dutch monolingual and bilingual minority children.

3.1.1. Description of current research

To test the predictions put forward in the previous section, we designed a semantic priming experiment involving both context-dependent and context-independent word pairs, which is an extension of the visual semantic classification task used by Cremer (2013) and is similar to the auditory lexical decision task used by Nation and Snowling (1999). Monolingual and bilingual minority children aged 10-11 took part in the experiment, a standardized reading comprehension task, and various control tasks for vocabulary size, word decoding and cognitive processing speed. The priming experiment and its stimuli were designed to maximize context-independent and context-dependent semantic processing, as

[13] Gollan et al. (2005) found similar semantic classification times for pictures in monolingual and bilingual adults in their dominant language, suggesting that semantic access to concepts based on words or pictures may carry different bilingualism effects. For our purposes however, access to semantic knowledge triggered by lexical items is the most important, since this is the same type of semantic access required during reading.

opposed to orthographic, strategic, or associative processing. We will shortly discuss the most important design choices below.

To make sure participants were required to access the semantics of the stimuli, we used a semantic classification task, namely animacy decision, in which children were required to decide for each word whether it represented an animate or inanimate concept. This is thus in opposition to a lexical decision task, which can be performed by simply retrieving the word form without accessing word meaning (McNamara, 2005). Also, this allows for the use of response times to filler items as a measure of access to semantic knowledge. Furthermore, the stimuli were presented aurally to be able to make a stronger claim that any effect of the priming scores on reading comprehension is at the semantic level and not, for example, at the orthographic level. Finally, we used continuous presentation, i.e. participants responded to all items one by one. This minimizes strategic processing, since participants are not made aware that stimuli are paired, as is the case with a paired presentation style (McNamara, 2005).

As for the selection of the stimuli, the context-independent pairs were category coordinates, which is similar to both Cremer's (2013) and Nation and Snowling's (1999) test items. However, the context-dependent pairs were designed to be located slightly more towards the context-dependent end of the continuum than the functional pairs used by Nation and Snowling. This allowed us to make a sharper contrast between the two types of semantic relations, since an object's function can be quite integral to its conceptualization. The pairs are *location – person or animal often found at this location* and *person – object or location that is often linked to this person*. These pairs were inspired by studies on thematic priming such as Hare et al. (2009) and are related through frequent co-occurrence in the same context. Note that some of the pairs Nation and Snowling deemed functional have the same format, but we avoided a functional connection between our pairs. Although subjective relations would be even more context-dependent, they are also too individual to be tested reliably across participants. All pairs were strictly controlled for association strength, so that the relation was only semantic and not associative. More details on the selection of the pairs and examples are provided in the Method section.

3.2. Method

3.2.1. Participants

All participants were recruited through their schools. The participating schools were all located in mixed neighborhoods with both residents with uniformly Dutch language backgrounds, and speakers of other mother tongues. SES in these neighborhoods was average to low (SCP, 2015). Parents were informed through a passive informed consent procedure, and all agreed to their child's participation.

A total of 151 children participated in the study. Teachers were asked to indicate whether children had serious oral language impairments or other disabilities such as dyslexia or ADHD. One child had been diagnosed with both ADHD and dyslexia, a further twelve children had been diagnosed with dyslexia. The data from these children were removed. No other cases were reported. A further nine children were not able to participate in all tasks or had missing data on some of the tasks. Finally, seven participants with extreme scores on the animacy decision task were removed. More details on the outlier criteria are discussed in the Data handling section.

This leaves a final sample of 122 children, 64 girls and 58 boys. 36 children spoke only Dutch at home, and 86 used other languages at home. Of this bilingual group, 82 children indicated they spoke Dutch at home in addition to their L1. Mean age was 11;3 (*SD* = 0;6), ranging from 10;4 to 12;6. Table 3.1 shows the age and gender distributions across the monolingual and bilingual groups.

Table 3.1
Age and gender by language group

	N	Girls	Boys	Age (*SD*)
Monolingual	36	19	17	11;1 (0;6)
Bilingual	86	45	41	11;3 (0;6)
Total	122	64	58	11;3 (0;6)

3.2.2. Materials

The participants completed six tasks in total. The main tasks were a standardized reading comprehension task and the priming experiment using a semantic decision task. Two tasks were included to control for abilities that may mediate the hypothesized effect of knowledge of semantic relations and fluency on reading comprehension, namely receptive vocabulary size and cognitive processing speed. In addition, a word decoding task was administered to control for a potential influence of decoding skill on the reading comprehension scores. Finally, in a short language interview, the children were asked about which languages they speak, and with whom, in order to establish language status and language dominance.

3.2.2.1. Reading comprehension task

To test reading comprehension skill, a shortened version of the standardized test '*Begrijpend Lezen 678*' [Reading Comprehension grades 456] by Aarnoutse and Kapinga (2006) was used, which was the same as used by Cremer (2013, chapter 5) in her priming study. Time constraints necessitated the shortening, as the test battery as a whole was quite extensive. The final test consisted of 32 questions on five short texts, testing both superficial and in-depth comprehension. None of the participating schools had administered this test to the children before.

3.2.2.2. Priming experiment

As was discussed earlier, an auditory semantic decision task was used to measure activation of context-dependent and context-independent semantic relations. 40 prime-target pairs were made for the experiment, 20 for each semantic relation. The word pairs can be found in Appendix B.

Stimuli. For the context-independent pairs, coordinates were used. Out of the various types of context-independent meaning relations, such as sub- and superordinates and synonyms, coordinates were found to be most suitable for the selection of a sufficiently large number of items. In addition, the items are on the same level in the semantic hierarchy, making the semantic decision to both items more similar compared to sub- and superordinate pairs such as *dog – animal*. Both animate and inanimate coordinates were used, again to be able to include more items. The animate items were all animal pairs, and the inanimate items were object pairs.

Context-dependent semantic relations have been investigated far less than context-independent meaning relations, and are generally less clearly defined. Because the difference between context-independency and context-dependency is gradient, we focused on relations that were as context-dependent as possible, avoiding functional and definitional pairs. To again have both animate and inanimate targets, two formats were used for the context-dependent condition. The first was *location – animal or person often found at this location*. Examples include *forest – squirrel* and *train station – conductor*. The second format was *person – object or location linked to this person*. Possible pairs include *teacher – classroom* and *thief – purse*. Note that, in contrast to the context-independent pairs, the primes and targets are always dissimilar in terms of animacy in these context-dependent subsets. Each subset contained 10 pairs, which amounts to 20 pairs per semantic relation.

Table 3.2
Stimulus pairings per condition

Condition	Context-independent		Context-dependent	
Prime - target type	Animate coordinates	Inanimate coordinates	Location - animate	Person - inanimate
Semantic decision	*creature - creature*	*thing - thing*	*thing - creature*	*creature - thing*
Related	duck - goose	**guitar** - piano	*forest* - squirrel	thief - purse
Control	thief - goose	*forest* - piano	**guitar** - squirrel	duck - purse

Two fully counterbalanced versions of the experiment were made, in which one half of the targets appeared in the related condition, and the other half in the unrelated condition and vice versa. Each participant thus encountered each target once. Unrelated control pairs were formed by re-pairing primes and targets across the two semantic relations, such that the animacy pattern remained the same. Thus, the primes preceding a given target in the related or unrelated condition were always either both animate or both inanimate. In this way, there can be no confound because of an answer 'switch' between target and prime, which is not present in the control pair or vice versa. The design is shown in Table 3.2, with primes marked typographically to clarify the re-pairing to form unrelated control pairs.

To control for association strength, data from a previous study were used, in which multiple word associations were gathered from 207 children from the same target population (Spätgens & Schoonen, submitted). 80 stimulus words were divided into four 20-word lists, and each child provided up to three associations for each word, resulting in association data from at least 50 children per item. Since adults show different word association patterns than children, it is important to use children's norms to control for the present experiment. Furthermore, using multiple association data allowed us to control for associations that are maybe not as immediate as those resulting from a single association task, but still prevalent.

To form the prime-target pairs for the present experiment, the stimulus words were used as primes. The related targets never occurred as first associations in the dataset, and some targets occurred as second or third associations at most once, indicating that they were only weak, idiosyncratic associations.

Relatedness of all prime-target pairs was checked by means of a questionnaire among 33 adult native speakers of Dutch. They were asked to rate all prime-target pairs and an equal number of unrelated distractor pairs on a five-point Likert scale ranging from 'no or almost no relation' to 'strong relation'. The pairs included in the experiment had an average relatedness score of 4.11 and average relatedness of the four subsets ranged from 3.93 to 4.31. There were no phonological similarities between primes and targets in either the related and unrelated conditions and none of the critical pairs formed compounds.

Care was taken to make sure all sets were as similar as possible in terms of frequency and duration in milliseconds. For frequency, the word list based on reading materials for elementary schools by Schrooten and Vermeer (1994) was used. It was not possible to match individual primes and targets, however, we made sure pair relatedness strength, average frequency and duration for both primes and targets did not differ between the halved subsets which are compared in the related and unrelated conditions. Mean pair relatedness and mean frequencies and durations of primes and targets by subsets and halved subsets can be found in Appendix C.

In addition to these prime-target pairs, 120 fillers were included. Since the stimuli were presented as single items to minimize strategic processing (McNamara & Altarriba, 1988), this puts the relatedness proportion at 10%. Half were animate and half inanimate, and they were similar to the critical stimuli in frequency and length. Combined, the fillers and critical stimuli included a roughly equal number of animal, person, object and location items. The experiment was preceded by an additional 12 practice items, again including even numbers of animals, persons, objects and locations. In total, participants thus responded to 212 items.

All stimuli were recorded by a female native speaker of Dutch with a neutral accent.

Presentation. For each of the two versions of the experiment, three pseudorandomized lists were compiled, to minimize a potential influence of order effects. Participants were randomly assigned one of the six lists. Care was taken to avoid unintended semantic or phonological relations between consecutive items, and animacy was varied such that between one and five consecutive items were of the same animacy type. Each critical pair was flanked by one to four filler items. The lists were divided in three parts, to allow for two short breaks during the experiment. A pilot test with 16 children in the same age group had shown that performance in terms of speed and accuracy improved with a second break. The first eight items at the start of the experiment and after each break were fillers, to allow participants to get used to the task each time before critical items came up. Within each part, the number of animate and inanimate items, divided across the four word types (animals, persons, objects, locations), was roughly equal. Finally, the inter stimulus interval or ISI was 1000ms. After a response, there was a blank screen for 500ms, followed by a screen with a fixation point (+) for 500ms, and then the screen went blank again and at the same time, the auditory stimulus was played.

The experiment was run using E-Prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002) on two identical laptops. Participants were required to indicate whether each item was animate or inanimate by means of the Alt keys. These were marked with stickers with small symbols to help participants remember which was which: a heart for the animate items, and a building block for the inanimate items. Participants used their dominant hand for the 'animate' answer. Both accuracy and response time (RT) were recorded. Response times were measured from the onset of the stimulus, as some words may be recognized before they have been pronounced fully. No correction for word duration was applied since duration was carefully controlled across the stimulus sets.

3.2.2.3. Vocabulary size

For receptive vocabulary size, the Cito *Leeswoordenschat* [Reading vocabulary] test by Verhoeven and Vermeer (1995) was used. This standardized task consists of 32 multiple choice items, requiring children to select the correct meaning for words presented in neutral sentences.

3.2.2.4. Word decoding

Word decoding skill was measured using the *Drie Minuten Toets* [Three Minutes Test] (Verhoeven, 1992) which is widely used in the Dutch school system and was therefore familiar to all participants. The test consists of three word lists including words of increasing length, of which only the two most difficult lists were used. Participants are required to read aloud as many words as possible within one minute, while making as few mistakes as possible. The resulting score is the number of words read, minus the number of errors made.

The two word lists correlated strongly ($r = 0.821$, $p < 0.001$) and were therefore combined into one measure by averaging the scores for each child.

3.2.2.5. Cognitive processing speed

Cognitive processing speed was measured using the Rapid Automatized Naming (RAN, Denckla & Rudel, 1974) and Rapid Alternating Stimulus (RAS, Wolf, 1986) tests. In these tests, participants are required to name a series of 50 items from a card as quickly as possible, while the time needed to complete the task is recorded. RAN tests consist of one type of character, and in this study, the letters edition was used. RAS tests include a mix of multiple types of stimuli and in this study, the letters, numbers and colors edition was used.

As a score, the time (in seconds) needed to name all 50 items is used - the number of errors made is thus not incorporated. The test developers consider over five errors or self-corrections to be 'excessive' (e.g. over 10%, Wolf & Denckla, 2005), and a potential reason for re-testing at a later time, which was not possible in this study. However, since only very few children produced just over five errors and self-corrections combined (three did so for the RAN test, one for RAS, all varying between six and eight errors and self-corrections combined), no corrective measure was taken.

3.2.2.6. Language interview

To establish language dominance, a short questionnaire on linguistic background was administered to each participant. The children were asked whether they were born in the Netherlands, from what age onwards they had gone to school in the Netherlands, which languages they spoke at home, how often and with whom they used these languages, and finally which language they used most.

3.2.3. Procedure

All tests were administered by the first author or one of two trained test assistants, according to a set protocol. The reading and vocabulary tests were administered in class, while all other tasks were done individually in a quiet room in school. Per group, testing lasted one or two school days, depending on group size. In the morning on the first day, testing began with the reading task, which lasted about 35 minutes including instruction, followed by the vocabulary task which took roughly 25 minutes including instruction. The reading comprehension task started with an example text with four questions. These were discussed by the experimenters with the class to familiarize the participants with the answer sheet and the different types of questions (multiple choice with four options and true/false statements). Similar to the reading task, the vocabulary test was preceded by two example questions which were discussed with the group. During both tasks, the experimenters were available for practical questions, but no information relating to the content of the tasks was provided.

For the individual tasks, the participants joined one of the experimenters in a quiet room. The same order of tasks was maintained for each child: first the semantic decision task, then word decoding, RAN, RAS and finally the language interview. In all, the individual sessions took around 25 minutes. Before starting the experiment, the participants received a verbal instruction which included a short discussion of the concept animacy and some examples. The importance of answering quickly and accurately was stressed. This was reinforced with a short written instruction. For the first twelve practice items, the children received feedback on the screen, which showed both whether they gave the right answer, and how fast they were in milliseconds. After the practice items, they could ask more questions if needed, and then the experiment began. During the experiment, no feedback was provided.

3.2.4. Data handling and analysis

RTs for inaccurate responses were set to missing (1360 responses, 5,6% of data). Then, the average RT for each child was calculated. RTs over 2.5 standard deviations from the mean (the individuals' means and standard deviations were used) were defined as outliers, and removed (614 responses, 0.03% of data). RTs under 350ms were removed so that only real responses and not accidental taps were recorded (15 responses, < 0.01% of data). 350ms instead of the commonly used 250ms (e.g. Betjemann & Keenan, 2008; Cremer, 2013) was used as the cut-off point because of the auditory and therefore linear nature of the stimuli. This means that we need to add at least some time onto this lower boundary, in which the participants have been exposed to some input. Since some words can be recognized even before they have been heard in their entirety, we chose to limit this extra time to 100ms.

Three children with accuracy scores under 85% and four children with mean RTs over 1700ms were identified as outliers not representative for the group as a whole and removed from the data set. Mixed effects analyses were performed to answer the various research questions. All analyses were done in R 3.1.3 (R Core Team, 2015), using the lme4 package for multilevel and mixed effects analyses (Bates et al., 2015).

3.3. Results

3.3.1. Descriptives

Skewness and kurtosis values for all main measures are reported below in Table 3.3. Overall, the measures are mostly normally distributed, except for the RAN scores which can be characterized as slightly skewed and peaked.

The internal consistency for the reading comprehension task in this sample was somewhat lower than in Cremer's study (2013) but not unsatisfactory: Cronbach's α = 0.634. Finally, the internal consistency of the vocabulary task was satisfactory, with Cronbach's α = 0.709.

Table 3.3
Skewness and kurtosis values for main measures

	Skewness (*SE*)	Kurtosis (*SE*)
Reading comprehension	-0.209 (0.219)	-0.423 (0.435)
Vocabulary size	-0.164 (0.219)	-0.193 (0.435)
Word decoding	0.060 (0.219)	-0.412 (0.435)
RAN (ms)	1.201 (0.219)	2.050 (0.435)
RAS (ms)	0.715 (0.219)	0.329 (0.435)
Access (mean RT to fillers)	0.347 (0.219)	-0.164 (0.435)

Table 3.4
Descriptives for task scores in the monolingual and bilingual groups

	Total (*N* = 122)		Monolingual (*N* = 36)		Bilingual (*N* = 86)		
	M	*SD*	*M*	*SD*	*M*	*SD*	*d*
Reading comprehension	22.47	3.89	23.06	4.37	22.22	3.67	0.21
Vocabulary size	16.86	4.60	17.64	4.74	16.53	4.53	0.24
Word decoding	91.34	13.37	91.19	13.66	91.40	13.32	0.02
RAN (ms)	22.05	3.65	21.96	3.14	22.10	3.86	0.04
RAS (ms)	30.24	5.63	29.74	5.79	30.44	5.59	0.12
Access (mean RT to fillers)	1149.70	160.67	1117.73	164.31	1163.09	158.17	0.28

Table 3.4 shows the descriptives of the scores on the various tasks for the monolingual and bilingual children, including effect sizes of the differences between the groups. Differences between the group means are in the expected directions for all measures: the monolingual children

perform slightly better on all tasks except decoding. For reading comprehension, vocabulary size and access, Cohen's *d* > 0.20, a small effect size (Cohen, 1969). However, none of these differences were found to be significant.

3.3.2. Overall semantic priming

To establish the effect of priming across the four sets of word pairs in the experiment, a mixed effects analysis was performed on the response times to the target items. Since the response times to the target items were skewed and peaked (skewness: 1.722 (*SE* = 0.037), kurtosis: 4.918 (*SE* = 0.074)), they were log transformed using the natural log (skewness: 0.592 (*SE* = 0.037), kurtosis: 0.885 (*SE* = 0.074)). In this data set, participants and items are crossed since all children responded to each word once, half in the related condition and half in the unrelated condition. Participants and items are nested under classes. For each of these levels, a random intercept was included to control for variation between classes, subjects and items.

The eight different types of targets are characterized by a 2x2x2 design: Relatedness [0 related, 1 unrelated] x Relationship Type [0 context-independent, 1 context-dependent] x Animacy [0 animate, 1 inanimate]. These three variables and their three-way and lower order interactions were entered as fixed effects. Access, the children's mean RTs to fillers, was entered as a covariate to control for the effect of differences in overall speed, since a slower participant may show a reduced priming effect and vice versa (e.g. Kliegl, Masson, & Richter, 2010). Table 3.5 shows the estimates from this model. As a rule of thumb, absolute t values over 2 are considered significant (Gelman & Hill, 2006).

As could be expected, access is a significant predictor of the response times to the individual targets. Since this variable was used as a control variable, it will not be discussed any further here. The results indicate that there are two positive main effects among the three dichotomous predictors. Firstly, there is a main effect of relatedness. This indicates that overall, response times to unrelated items were higher, and thus longer, than to related items. Hence, an overall priming effect seems to be present, but the shape of this effect will become clearer when looking at the interactions.

Table 3.5

Fixed and random effects estimates for the overall priming model (4415 items)

Random effects			
Variance between classes	0.00076		
Variance between subjects	0.00160		
Variance between items	0.00009		
Residual variance	0.05125		
Fixed effects	**Estimate (*SE*)**	***t***	
Intercept	6.958 (0.016)	424.4	*
Relatedness	0.059 (0.014)	4.3	*
Relationship Type	0.036 (0.022)	1.6	
Animacy	0.097 (0.023)	4.3	*
Access	0.812 (0.027)	30.5	*
Relatedness*Relationship Type	-0.058 (0.019)	-3.0	*
Relatedness*Animacy	0.001 (0.020)	0.0	
Relationship Type*Animacy	-0.034 (0.032)	-1.1	
Relatedness*Relationship Type*Animacy	0.007 (0.027)	0.2	

Secondly, there is a positive main effect of animacy. Here, the inanimate items yield a higher, and therefore longer RT. In other words, identifying inanimates took participants longer than identifying animates, despite the fact that the instruction for the experiment was focused on making the semantic decision for both types of items as similar as possible. If the semantic decision is based on a search through the sets of animate and inanimate items, searching the inanimate set may take longer since it is larger. Alternatively, the decision may be made based on activating or matching intrinsic features, which are likely more varied for inanimate compared to animate items, and may therefore also take longer. Alternatively, participants may have treated the semantic decision as a sort of yes/no task after all, asking themselves: ‘is it an animate being?’ rather than ‘is it animate or inanimate?’. In any case, we will see below that this main effect of animacy does not interact with the effect of relatedness, which means that it has not affected the priming scores.

The main effect of relationship type is not significant, indicating that the category to which items belonged did not matter for the response times. This suggests that, in accordance with the design of the experiment, children were not aware of the type of semantic relation that existed between primes and targets. Furthermore, the targets in the context-dependent and context-independent conditions were thus indeed very similar.

Of the four interactions that were tested, only the interaction between relatedness and relationship type was significant. Figure 3.1 shows that in fact, the overall priming effect is due to a large priming effect for the context-independent items, while the difference between unrelated and related items in the context-dependent condition is much smaller. Indeed, the parameter estimates also show that for the context items, the main effect of relatedness is essentially cancelled out: the overall effect is 0.059, and the interaction effect, for which context-dependent items are coded 1, is -0.058. Only the context-independent items thus elicited a priming effect.

The other two-way interactions between relatedness and animacy and relationship type and animacy are not significant, which again shows that the experiment worked as intended. Even though animacy did show a significant main effect, it is not the case that priming occurred more for either animate or inanimate items, or that animacy behaved differently in either of the semantic categories.

Finally, the three-way interaction is also not significant. This means that the various subcategories (animate and inanimate targets within the context-dependent and context-independent conditions) did not behave differently. Together with the interaction between relatedness and relationship type, this is evidence that the subcategories within the two semantic relations behaved similarly, and can thus be combined to establish context-dependent and context-independent priming effects. Therefore, the object coordinates and animal coordinates are taken together in the context-independent set, and the location-animate and person-inanimate items are combined to form the context-dependent set. Henceforth, we will use these two sets in separate analyses to study the two types of priming effects in more detail.

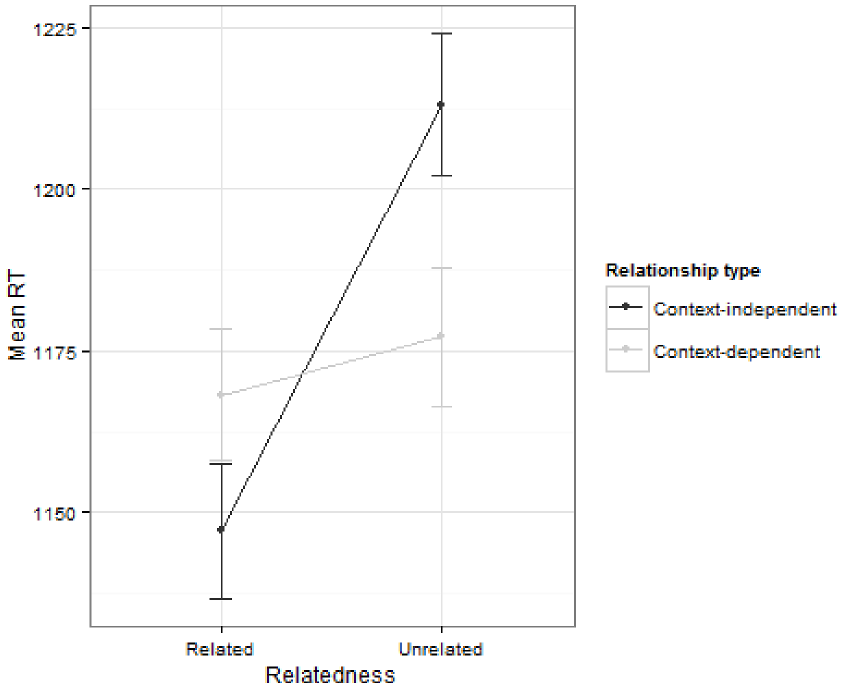

Figure 3.1. Mean response times by relatedness and relationship type (error bars represent standard error)

3.3.3. Semantic priming in monolingual and bilingual participants

Similar analyses were performed to see whether monolingual and bilingual children show different priming effects due to context-independently and context-dependently related primes. For this, the context-independent and the context-dependent items were analyzed separately.[14] Again, the log transformed response time was modeled with random intercepts for classes, subjects and items. As fixed effects, language group and relatedness and their interaction were included, and mean response time to fillers was entered as a covariate. Table 3.6 below shows the results for both models. As we can see, neither the context-independent items nor the context-dependent items show a significant interaction between

[14] We performed both these separate analyses and an analysis with the full target set and a three-way interaction between language group, relatedness and relationship type. Since the results were the same, we report the separate models for context-independent and context-dependent items for ease of interpretation.

relatedness and language group, meaning that the two language groups did not exhibit different priming effects, contrary to our expectations. For the context-dependent items, the main effect of relatedness again shows that there was no priming effect for the group as a whole, while the context-independent items did show an overall priming effect.

Table 3.6

Fixed and random effects estimates for the monolingual and bilingual priming models

	Context-independent priming (2136 responses)			Context-dependent priming (2279 responses)		
Random effects						
Variance between classes	0.00001			0.00037		
Variance between subjects	0.00043			0.00170		
Variance between items	0.00361			0.00322		
Residual variance	0.05160			0.05000		
Fixed effects	Estimate (*SE*)	*t*		Estimate (*SE*)	*t*	
Intercept	7.001 (0.019)	377.2	*	7.014 (0.020)	345.0	*
Language group	0.008 (0.016)	0.5		0.017 (0.017)	1.0	
Relatedness	0.085 (0.018)	4.7	*	0.000 (0.017)	0.0	
Access	0.834 (0.033)	25.2	*	0.787 (0.038)	20.6	*
Language group*Relatedness	-0.036 (0.022)	-1.7		0.006 (0.020)	0.3	

3.3.4. Calculating individual priming scores

In order to establish the individual priming scores for each of the two semantic relations, another mixed effects model was applied to both the context-independent and the context-dependent responses. The random structure was the same as in the overall priming model discussed above: with random intercepts for class, subject and item. In addition, a random slope for relatedness was included for the participants. In this way, individual priming scores can be established by extracting the estimates for the random slopes for each individual. These scores correspond to the difference between the individual's response times on the unrelated items, compared to the related items. Recall that the unrelated items were coded 1, so that a positive value for the individual slope means that there was a priming effect, since the participant exhibited longer and thus slower response times on the unrelated items. By estimating the priming scores in this way, rather than subtracting mean response times on the related items from mean response times on the unrelated items, differences between children, items and classes are taken into account. As such, more accurate individual priming scores can be obtained.

Even though there is no overall priming effect for the context-dependent items, we tried to estimate individual priming scores to capture the individual variation which may still be large enough to affect the reading scores. However, the model was not able to produce estimates for both the individual intercepts, i.e. the average response time on the related items, and the individual slopes, that is, how much the average response time to the unrelated items deviates from the average response time to the related items. This was evidenced by the fact that the model collapsed onto perfectly correlated random intercepts and slopes. Potentially, this is due to the fact that the context-dependent items did not show a consistent priming effect to begin with. This also means that the data from context-dependent items are not suitable for inclusion in the final step, and therefore they will not be discussed any further.

A summary of the estimates for the individual priming scores on the context-independent items are provided in Table 3.7 below. Note that the numbers are very small due to the log transformation of the response times. Table 3.7 also shows a summary of the individual scores when calculated by the same model but with untransformed response times, as an illustration of what the individual priming scores would be in that case. However, due to the skewness and peakedness of the response times, these numbers should be interpreted with caution.

Table 3.7

Summary of individual priming scores on the context-independent items, with and without transformation (N = 122)

	1st Quartile	*M* (*SD*)	3rd Quartile
Priming based on log transformed RTs	-0.00024	0.01359 (0.02596)	0.02773
Priming based on untransformed RTs	-1.61171	6.86388 (14.35792)	15.78807

3.3.5. Access, context-independent priming and reading scores

In the final step, the effects of the control tasks, context-independent priming and language group on the reading scores were determined by means of a series of mixed-effects models, shown in Table 3.8. For these analyses, the vocabulary scores, word decoding, RAN and RAS measures, and access were centered. In addition, word decoding and access had to be divided by 100 and 1000, respectively, to make sure the values of all variables were on comparable scales. Two children were removed from the dataset for this final step, because they turned out to be extreme bivariate outliers when it came to the relation between reading comprehension and context-independent priming, and strongly distorted the correlation between these measures. With these children in the data set, there was a negative correlation for the monolingual group, while excluding them meant the correlation became positive.[15] This brings the total number of children for these analyses down to 120, with 34 children in the monolingual group, and 86 in the bilingual group.[16]

[15] These two participants had respective ZxZy products of -5.580 and -5.584. The difference in the trend line with and without these children in the sample is illustrated in two scatterplots in Appendix D.

[16] We checked whether the absence of differences between the two groups was upheld after exclusion of these two participants, which was the case. No group differences were found in terms of reading score, control tasks or access, and the mixed-effects analyses performed for priming also yielded the same results, with no differences between the monolingual and bilingual children for either context-independent or context-dependent priming.

Table 3.8

Fixed and random effects estimates for the reading comprehension models (N = 120)

	1		2		3	
Random effects						
Variance between classes	0.5145		0.5649		0.4048	
Residual variance	10.0317		10.1424		10.4018	
Fixed effects	Estimate (*SE*)	*t*	Estimate (*SE*)	*t*	Estimate (*SE*)	*t*
Intercept	22.445 (0.389)	57.7 *	22.336 (0.429)	52.1 *	22.886 (0.823)	27.8 *
Vocabulary size	0.368 (0.068)	5.4 *	0.366 (0.070)	5.3 *	0.364 (0.073)	5.0 *
Word decoding	2.828 (2.766)	1.0	2.796 (2.783)	1.0	2.643 (2.824)	0.9
RAN	-0.138 (0.104)	-1.3	-0.136 (0.104)	-1.3	-0.13 (0.107)	-1.2
RAS	-0.023 (0.071)	-0.3	-0.024 (0.075)	-0.3	-0.041 (0.078)	-0.5
Access (mean RT to fillers)			0.599 (1.981)	0.3	-0.643 (3.634)	-0.2
Context-independent priming			8.122 (11.989)	0.7	-12.043 (25.884)	-0.5
Language group					-0.688 (0.882)	-0.9
Language group*Access					2.235 (4.309)	0.5
Language group*Context-independent priming					25.982 (29.538)	0.9
Deviance (-2*Log Likelihood)	616.55		616.06		614.69	
Difference			0.49 (*ns*)		1.86 (*ns*)	
Difference *df*			2		5	
Compared to model			1		1	

Random intercepts for class were included in each step to account for the hierarchical structure of the data. In the first step, Model 1, only the control tasks (vocabulary size, word decoding, processing speed) were entered as fixed effects. As we can see in Table 3.8, only the vocabulary size measure is associated with the reading scores in this case, with children who scored one point above the mean on vocabulary showing an increase of 0.368 in the reading comprehension scores. None of the other control tasks are significantly associated with the reading scores. In Model 2, we added access and context-independent priming, but neither has a significant main effect on the reading scores. Further exploration of the models (not shown here) revealed that also in the absence of the control variables, neither of the critical variables were significantly associated with the reading scores. Therefore, it is not the case that there is an effect of access or priming which is filtered out by the control tasks. Finally, to compare our results to Nation and Snowling's (1999), we divided the group in above average and below average readers, but found no contribution of context-independent priming to the reading scores in either group.

Even though the monolingual and bilingual children did not show differences in either priming, access or reading, the impact of priming and access on the reading scores may still differ between the two language groups. Therefore, language group and the interactions between language group and access and language group and context-independent priming were included in Model 3. Neither access nor context-independent priming show a significant interaction with language group, and thus neither group's reading scores benefited from higher access or priming scores. Both Model 2 and Model 3 failed to represent a significant reduction of the deviance score compared to Model 1, meaning that the best fit was achieved using only the control variables.

3.4. Discussion and conclusions

3.4.1. Context-independent and context-dependent semantic priming

The semantic priming experiment was designed to tap into both context-independent and context-dependent semantic connections in the participants' mental lexicons, and we hypothesized that as a group, children would show both types of semantic priming. However, across the whole group, context-independent priming was observed, but not context-dependent priming, even though both types were similar in terms of relatedness strength. Context-independent priming, especially using category coordinates, has been studied extensively (see Lucas, 2000, for an overview), and is known to occur with and without the presence of an additional associative relationship. It is therefore not surprising that as a group, the children showed non-associative context-independent semantic priming.

The fact that no overall context-dependent priming effect was found is likely to be due to the fact that we controlled very strictly for association strength to make sure that we were tapping into purely semantic connections. The word association data which were utilized (Spätgens & Schoonen, submitted) were gathered by means of a multiple association format, e.g. requiring three associations to each stimulus word, instead of the normally used single response format. In the present experiment, no targets were included which had occurred as first responses, nor any that occurred more than once as second or third responses to their primes. This is a more strict approach than has been employed by other studies, which typically use word association norms that consist of single responses only (for example Nation and Snowling (1999) but also Hare, Jones, Thomson, Kelly and McRae (2009) who tested very similar context-dependent pairs, including *location – person/animal*).

Indeed, in the word association data we used, context-independent associations were especially prominent as first responses, while context-dependent associations became more numerous in the second and third response sets. This spread of different types of semantic relations across response positions has also been observed by De Deyne and Storms (2008a). Controlling for the second and third responses has likely made a considerable difference in the types of pairs selected compared to other studies. It is thus likely that the absence of context-dependent priming in our study compared to other research is due to the more stringent word association criteria. This finding suggests that context-dependent semantic relations are mainly associative in nature, certainly compared to context-independent semantic relations, which is in tune with the fact that context-dependent relations are supported by the co-occurrence of concepts in experience.

A methodological point that may be of use for future studies on semantic priming is the fact that we were able to elicit priming for both animate and inanimate items in our animacy decision task. Cremer (2013) also used a semantic classification task in which participants were required to judge whether stimulus words referred to animals, and found that a priming effect only occurred for the animal items, i.e. the items to which the correct response was 'yes', while 'no' items did not elicit priming. By formulating the task in such a way that the answers are 'animate' and 'inanimate', we did find priming for both sets of stimuli, even though the inanimate items did yield a longer RT. Potentially, 'no' items are discarded quickly after initial superficial processing, and are therefore not processed in as much depth.

3.4.2. Differences between monolingual and bilingual children

Based on previous studies on the development of different types of semantic knowledge in monolingual and bilingual minority children, we hypothesized that the bilingual children would show less priming than monolingual children overall, and that they would especially show less context-independent priming. Regarding access, reading comprehension and the control tasks, we also expected lower scores for the bilinguals.

The differences between the groups were all in the expected direction, with monolingual children outperforming bilingual children on all measures except decoding. The effect sizes for reading comprehension, vocabulary size and access suggest that there are small differences between the groups on these measures. However, contrary to our expectations, neither the reading scores, nor the control tasks nor the priming and access measures showed significant differences between the two language groups. The small differences we found between the monolingual and bilingual groups are in line with the weaker links hypothesis (Gollan & Silverberg, 2001; Gollan et al., 2005; Gollan et al., 2008; Michael & Gollan, 2005), suggesting that the hypothesis does not only apply to vocabulary size (Gollan et al., 2005), but also knowledge of semantic relations. However, since the differences are not statistically significant, it is difficult to draw firm conclusions from them.

In a way, this is a positive finding since it suggests that contrary to previous findings, for example a recent national report on language ability at the end of elementary school (CBS, 2014), the bilingual children in this sample were not disadvantaged in the standardized measures of reading comprehension and vocabulary size. This may be due to the fact that all bilingual children in this study had gone to Dutch schools from age four onwards, and the vast majority was born in the Netherlands. In addition, the fact that all schools were in average to low SES neighborhoods may have played a role, meaning that in this specific population, bilingual children actually perform similarly to their monolingual peers.

Although Droop and Verhoeven (2003) found that low SES bilinguals showed worse performance on reading comprehension and vocabulary than low SES monolinguals, a recent Dutch national assessment report on children in grades 3 and 6 (Kuhlemeier et al., 2014) shows that when SES is taken into account, differences between monolinguals and bilinguals disappear, in line with our findings in grade 5. Furthermore, most of the bilinguals used Dutch in addition to their L1 at home, while only a small minority used the L1 exclusively at home. Indeed, large scale national studies examining reading comprehension and vocabulary of Dutch elementary school children have found that in grade 6, only bilingual children who do not use Dutch at home are lagging behind their monolingual peers in terms of reading comprehension when SES is controlled for (e.g. Heesters et al., 2007). This may explain the difference with the national report from CBS, which did not differentiate according to language use at home and did not control for SES.

Given this lack of significant differences in the standardized language measures, it is not highly surprising that the bilingual children performed similarly to the monolingual children on the priming tasks and the access measure. Apparently, in this sample, the bilingual children's Dutch competence is fairly close to that of the monolingual children, and their knowledge of semantic relations is no different. Our findings do not allow us to discern distinct bilingual patterns of context-independent and context-dependent priming, and also in terms of access to semantic knowledge, the bilingual children in this study perform similarly to their monolingual peers.

3.4.3. Reading comprehension and the influence of access and priming

The analyses of the reading comprehension scores were done in three steps, first examining the control variables, then adding access and context-independent priming, and finally examining the interaction between language group and access and language group and priming. In line with previous studies, there was a significant effect of vocabulary size on the reading comprehension scores in each of the three phases. Decoding did not have a significant influence on the reading scores, which is normal for both monolingual and bilingual children of this age in Dutch (Verhoeven & Van Leeuwe, 2008; Verhoeven & Van Leeuwe, 2012). The cognitive processing tasks (RAN and RAS) were mainly included since some studies have found that they affect reading comprehension in addition to word recognition, especially for children reading in their L2 (see for a large meta-analysis: Swanson, Trainin, Necoechea, & Hammill, 2003; and for L2 and bilingual readers: Erdos, Genesee, Savage, & Haigh, 2011; Olkkonen, 2013), and because they likely tap into overlapping abilities together with the access and priming measures. However, in this study, cognitive processing speed did not affect the reading comprehension scores on its own, which is in accordance with other studies that have shown that automatized naming mainly affects word recognition, but not reading comprehension (e.g. Di Filippo et al., 2005; Scarborough, 1998).

In the second step, adding access and context-independent priming did not improve the model for reading comprehension. Even when leaving out the control measures, access and context-independent priming could not contribute to the reading comprehension scores, meaning that it was not the case that the control variables filtered out some component of the variance which access or priming could have potentially explained. Finally, when we added interactions to examine possible differences in the contribution of access and priming for the two language groups, these could not explain any additional variance. Neither the monolingual nor the bilingual children showed an association between access and reading comprehension or priming and reading comprehension.

Given the similarity of our experiment to Nation and Snowling's (1999) and our additional focus on semantic processing, we would have expected a positive association between context-independent priming and reading comprehension, but even when looking at above average and below average comprehenders separately, we did not find such an effect. A potentially important difference is that in our experiment, an ISI of 1000 ms was used, after piloting showed that children in our target population experienced this as an already very fast pace for the task. With an ISI of 500 ms, Nation and Snowling's experiment may have been more sensitive to very early priming effects. Since our participants did not show context-dependent priming, we cannot compare our results to Bonnotte and Casalis (2010), who found a difference between poor and proficient readers in functional priming which might have also turned up in the context-dependent items used in the present study.

Our priming results do corroborate Cremer's findings (2013), who used an SOA of 2000 ms and included fewer critical word pairs, suggesting that also with our more strict experimental parameters, individual differences in context-independent priming do not contribute to reading comprehension. It has been demonstrated that semantic priming is inherently noisy, especially under circumstances where strategic processing of the stimuli is unlikely to occur (J. Stolz, Besner, & Carr, 2005; Yap, Hutchison, & Tan, 2016). Stolz et al. and Yap et al. argue that even though group-level semantic priming effects are very consistent, an individual's priming score may not reflect a stable characteristic of their semantic processing system. Both studies found individual priming scores to vary widely across test sessions and items, especially in experimental settings that encouraged automatic processing rather than strategic processing, which complicates relating individual priming scores to individual differences in other domains. Even though we used mixed effects modelling techniques to counter this issue by taking variation between items and participants into account when calculating priming scores (cf. Kliegl et al., 2010), our results suggest that indeed, priming scores reflecting automatic processing may be too noisy for use in individual differences studies. The fact that our data were collected in a field setting and not in a lab, may additionally contribute to this. However, this latter argument cannot be a full explanation, since Nation and Snowling's experiment (1999) was administered in the same way.

Contrary to Cremer (2013), who found that semantic access in a classification task could explain a small but significant amount of variance in reading scores, namely 2.5%, we did not find a significant association between access and reading comprehension. In many respects, the participants and experiment were similar in her and our study, and it may simply be the case that because this effect is so small, it is more likely that it is not always detected. However, a potentially important difference is the modality in which the stimuli were presented: visual in Cremer's, aural in the present study. Cremer showed that access as measured by lexical decision did not explain any variance in the reading scores, whereas access measured by semantic classification did, which suggests that there is certainly some semantic component involved in the relation between access and reading. However, given that the stimuli were presented visually, some degree of decoding speed may be incorporated in the access measure, which may be responsible for the explained variance in the reading scores. In our auditory task, this cannot be the case, which would suggest that semantic access per se may not contribute to reading comprehension.

Possibly, more sensitive measures need to be used in order to find the relation between the semantic network and reading comprehension on an individual level. We would like to suggest the use of online measures of reading such as self-paced reading or eye-tracking, in which use of semantic relations during reading could be tracked. Incorporating semantic relations that represent cohesive ties in texts and studying how these relations are handled during reading could provide us with more information on the use of different types of semantic knowledge in reading comprehension.

It is important to note that the research presented here is correlational in nature, combining separate measures of reading, vocabulary knowledge and access. This means that the causal direction of any relation between reading comprehension and the various predictor variables cannot be determined with certainty. Indeed, the relation may be bidirectional to some degree. However, as we have argued in the introduction, there are many reasons to believe that a well-developed semantic network and semantic access contribute to reading comprehension. Online measures such as self-paced reading or eye tracking could be a fruitful direction for future research in this respect as well.

Chapter 4
Individual differences in reading comprehension in monolingual and bilingual children: the influence of semantic priming during sentence reading[17]

Abstract

Group differences in semantic priming between young readers with different comprehension levels have been reported, with poor readers showing reduced or no context-independent semantic priming compared to normal readers. However, other studies have not been able to replicate these effects on an individual differences level, even though the spreading of semantic activation is hypothesized to play a role in the reading comprehension process. In the present study, we investigated whether priming during sentence reading, rather than single word priming, could be related to children's reading comprehension scores. A self-paced reading experiment involving both associated and non-associated, context-dependent and context-independent semantic relations was administered to 137 Dutch monolingual and bilingual minority children. Delayed facilitative priming effects were observed for non-associated context-dependently and context-independently related word pairs, but these were not linked to individual differences in reading comprehension. Monolinguals and bilinguals showed similar performance on almost all language measures, including semantic priming and reading comprehension.

[17] A slightly modified version of this chapter has been submitted as:
Spätgens, T. & Schoonen, R. (submitted). Individual differences in reading comprehension in monolingual and bilingual children: the influence of semantic priming during sentence reading.

4.1. Introduction

Reading comprehension is a central skill taught to school-age children, serving as a means of knowledge transfer that becomes increasingly important throughout their school careers and beyond. Nevertheless, the levels of reading comprehension that children attain vary, even setting aside problems related to specific reading difficulties and more general language impairments. Studying the component skills that support comprehension may help explain these individual differences and serve as a starting point to guide intervention efforts. Vocabulary knowledge is a central element of the comprehension process (cf. Perfetti & Stafura, 2014), which consists of many subcomponents. One of these vocabulary knowledge components is the semantic network, or the structure connecting words in the mental lexicon through various types of semantic relations (Aitchison, 2012). The present study aims to extend existing research on the connection between comprehension and automatic activation, i.e. semantic priming, in the semantic network, adding to our understanding of the influence of this vocabulary component on reading comprehension.

In previous research attempting to tease apart the subskills underlying reading comprehension, vocabulary size is identified as an undisputed key player (e.g. Alderson, 2005; Grabe, 2009; Stæhr, 2008; Stanovich, 2000). Additionally however, other components of vocabulary knowledge have a role to play, such as depth, fluency and knowledge of semantic relations (e.g. Cremer & Schoonen, 2013; Ouellette, 2006; Swart et al., 2017a; Tannenbaum et al., 2006), but also the automatic activation of semantic relations within the semantic network (Betjemann & Keenan, 2008; Bonnotte & Casalis, 2010; Nation & Snowling, 1999). In the studies by Bonnotte and Casalis and Nation and Snowling specifically, group differences in semantic priming were found between normal readers and poor readers with reading comprehension delays of at least one year. These studies reported an absence of priming or reduced priming for unassociated category coordinates such as *cat – tiger* in poor comprehenders, and in addition Bonnotte and Casalis found that poor readers showed priming for associated functional pairs such as *broom – floor*, whereas normal readers did not exhibit such facilitation. In contrast, recent research attempting to relate individual differences – rather than group differences – in semantic activation to reading comprehension has found different results. Both Cremer (2013) and Spätgens and Schoonen (2018) studied normal readers and found no advantages in reading comprehension for children showing more semantic priming of non-associated category relations. It is conceivable that because normal readers are a more homogenous group than those juxtaposed in Nation and Snowling (1999) and Bonnotte and Casalis (2010), individual differences in semantic activation facilitating comprehension may also be more subtle.

A potential way of finding these more subtle effects may be to study priming during reading, rather than single word priming. If advantages in semantic activation would be causally related to reading comprehension skill, a likely mechanism for this contribution would be through facilitation of the incorporation of these semantic links in the mental model of a text. Such a use of semantic priming during reading comprehension is posited in the memory-based processing phase in the Landscape model of reading (e.g. Van den Broek et al., 1999; Van den Broek et al., 2005). The activation of semantic links observed in single-word priming may however not have a close enough correspondence to activation *during* reading. Correlating comprehension performance to semantic activation during reading may therefore be an additional way to test for the presence of individual differences similar to the group-level effects that have been found in the aforementioned studies by Nation and Snowling (1999) and Bonnotte and Casalis (2010).

The first goal of the present study is therefore to study semantic activation during reading of four types of semantic relations close to those studied by both Nation and Snowling (1999) and Bonnotte and Casalis (2010): context-dependent and context-independent semantic relations, both associated and non-associated. The first contrast distinguishes word pairs related through co-occurrence in context, such as *forest – squirrel*, from word pairs related through shared inherent qualities, such as *cat – tiger* (Cremer, 2013; Spätgens & Schoonen, 2018; Verhallen & Schoonen, 1993). This distinction is related to the developmental path from episodic to categorical relations observed in vocabulary acquisition (K. Nelson, 2007). The second contrast, association strength, is defined here as the prevalence of a target as a response to its prime in group data from a free word association task. The presence of an associative relation, which in this definition can occur *in addition* to various types of semantic relations, is known to cause additional activation compared to non-associated pairs in priming studies, the so-called ‘associative boost’ (see Lucas, 2000, for a comprehensive overview). The distinction between associated and non-associated pairs was crucial in the group differences reported by Nation and Snowling (1999) and Bonnotte and Casalis (2010).

Based on Nation and Snowling (1999) and Bonnotte and Casalis (2010), it appears that especially activation of non-associated, context-independent pairs is predictive of reading comprehension. This notion of the importance of context-independent knowledge is corroborated by evidence from declarative tasks of semantic knowledge, which have shown that children with more context-independent knowledge (Ouellette, 2006) or a better ability to distinguish context-independent relations from context-dependent relations (Cremer & Schoonen, 2013) performed better at reading comprehension tasks. We therefore specifically expect to find a positive relation between the automatic activation of non-associated context-independent semantic knowledge and reading comprehension. In addition, by including associated and context-dependent pairs, we involve a similar contrast to that used by Nation and Snowling (1999) and Bonnotte and Casalis (2010), so that we may be able to identify other effects of semantic activation as well.

The second goal of the present study is to compare monolingual and bilingual minority children in these respects. In the Dutch context, about 17% of children in the final year of elementary school are currently from a non-Western immigrant background (CBS, 2016). This minority group is often associated with poorer performance on a range of language measures (cf. CBS, 2014). On the other hand, differences between monolinguals and bilingual minority children in the final years of elementary school (age 9-12) appear to be diminishing in large scale educational reports: it may be the case that only bilinguals who do not speak any Dutch at home are at a disadvantage (e.g., compared to bilinguals who use both Dutch and another language at home, cf. Heesters et al., 2007), or that socio-economic status (SES) is a mediating factor (cf. Kuhlemeier et al., 2014, who controlled for

SES differences and found no significant differences between monolinguals and bilinguals on various language measures). Recent studies targeting knowledge of semantic relations and reading comprehension have also found mixed results in the same population: Cremer (2013) demonstrated delays in reading comprehension in bilinguals in two separate studies, and found disadvantages in offline knowledge of context-independent relations for bilinguals, but not in semantic priming. Our previous research found no significant differences in terms of semantic activation or reading comprehension (Spätgens & Schoonen, 2018). In the present study, we will address the difference between these groups again to determine whether our previous findings are confirmed in another sample from the same population.

4.1.1. Description of the current research

In order to study semantic activation during reading, a single sentence self-paced reading experiment was designed. The sentences contained prime-target word pairs that bore an associated or non-associated, context-dependent or context-independent semantic relation. Facilitation on each of these four semantic types was measured by contrasting sentences including related pairs with the same sentences in which the targets were preceded by unrelated primes. Individual facilitation effects emerging in this task were then used to predict reading comprehension scores, while controlling for vocabulary size, declarative knowledge of context-independent semantic relations and decoding speed. To the best of our knowledge, this is the first study attempting to relate semantic priming during reading to individual differences in reading comprehension.

4.2. Method

4.2.1. Participants

Participants were recruited through their schools, which were selected based on location: all seven schools were in neighborhoods with an overall average to low SES (SCP, 2015) and mixed populations of monolingual residents and multilingual residents with a migration background. All parents were informed via the schools on the contents and purpose of the study through a passive informed consent procedure approved by the ethics committee of the Faculty of Humanities of the University of Amsterdam. Thirteen children from the recruited groups did not participate in the tasks because their parents objected.

In total, 160 children took part.[18] Of this group, the data from 23 children were removed for either developmental reasons such as being diagnosed with dyslexia (nine children), ADHD (one child), or general language delays (three children); due to incomplete data or equipment failure (six children); because of low accuracy on the self-paced reading task (57%, one child); or because the participants had started elementary school in the Netherlands after grade 1 (three children).

The final sample of 137 children had a mean age of 11;2 (*SD* = 0;5) and included 61 monolingual and 76 bilingual participants, 48.9% girls. The participant descriptives are provided in Table 4.1.

Table 4.1
Age and gender by language group

	Total	Girls	Boys	Age (*SD*)
Monolingual	61	28	33	11;1 (0;5)
Bilingual	76	39	37	11;2 (0;5)
Total	137	67	70	11;2 (0;5)

4.2.2. Materials

4.2.2.1. Self-paced reading experiment

A single sentence self-paced reading experiment with a non-cumulative moving window paradigm was designed. In this paradigm, an initially fully masked sentence is presented, after which the first button press reveals the first word. Following a second button press, the first word is re-masked and the second word is revealed, etc. An example sentence of three words would be presented in the following four stages, each presented on a new screen after a button press:

[18] In addition, two groups participated in two separate pilot sessions following the same passive informed consent procedure, testing earlier versions of the self-paced reading experiment (36 children in total).

```
--- ----- -----.

Mia ----- -----.

--- loves -----.

--- ----- words.
```

In this way, reading times for each word can be measured. To ensure participants read the sentences for comprehension, a plausibility judgement was required after each sentence.

Word pair construction. To test for semantic priming, 56 prime-target word pairs were developed to be included in the test sentences. Four types of semantic relations in a 2x2 contrast were used: prime-target word pairs in the sentences were either context-dependently or context-independently related, and either associated or non-associated. Example related word pairs of each type are provided in Table 4.2.

Table 4.2
Example related word pairs by semantic type

	Context-independent		Context-dependent	
Non-associated	cat	tiger	desert	snake
	trumpet	flute	captain	harbor
Associated	banana	fruit	hunter	rifle
	ocean	sea	store	money

As becomes clear from the examples, context-independent pairs are represented by category relations, either subordinate – superordinate pairs, or coordinate pairs. Context-dependent pairs were either *location – animal/person often found at this location*, or *person – location/object this person often occurs at/with*. The same types of relations were used in Spätgens and Schoonen (2018), although subordinate – superordinate pairs were added to be able to create enough associated, context-independently related pairs. Primes and targets were selected from the 5000 most frequent words in the Dutch school language corpus from Schrooten & Vermeer (1994).

All non-associated pairs used in this experiment are from a pool of potential stimuli developed for the semantic priming experiment in Spätgens & Schoonen (2018). These pairs are controlled for relatedness through a questionnaire in which items were rated by 33 adult speakers of Dutch, so that context-independent and context-dependent pairs were equally strongly related. Association strength for these items was determined based on a previous study in the same target population (Spätgens & Schoonen, submitted). For this data base, three association responses to the primes were elicited from at least 50 children to determine association strength with potential target words. Non-associated targets in our test items were never provided as first responses, and at most once as second or third responses. In contrast, the associatively related pairs were selected based on their regular occurrence as association responses in this data base. Across the two sets, the associated targets appeared 6.7 times as first responses, 3.8 times as second responses and 3.0 times as third responses, on average.

All translated word pairs are provided in Appendix E, and length and frequency information for primes and targets by semantic category is provided in Appendix F. Related and control primes were matched for frequency and length within each of the semantic types.

Sentence construction. The related prime-target pairs were embedded in sentences, each paired with another sentence which was identical except for the replacement of the related prime by the semantically unrelated prime. The construction of the test sentences was done according to a number of criteria. Firstly and most importantly, both the related and unrelated sentences were required to be equally acceptable.

Secondly, various restrictions were placed on the positioning of the prime-target pairs in the sentence. Prime words were always preceded by at least one word. Prime-target pairs were separated by at least one and at most three words, which did not steer the meaning of the sentence toward the target. Finally, targets were followed by at least five words: three as potential spillover positions, and two more to make sure any wrap-up effects at the end of the sentence would not occur in the spillover region.

Thirdly, both plausible and implausible sentences were designed to be plausible up until the fourth word after the target, so that slowed reading due to an implausible turn of events would not be captured in the spillover region. Overall, sentence length was kept similar across plausible and implausible sentences (M = 12.8 words, SD = 1.53 and M = 12.6 words, SD = 1.49, respectively) so that length could not serve as a clue for the plausibility judgement. A plausible and implausible translated example are provided below.

Plausible example (context-dependent, unassociated prime-target pair)

The old	**captain**	looks at the	**harbor**	*while the big*	boat sails away.
The old	**lady**	looks at the	**harbor**	*while the big*	boat sails away.
	prime		target	spillover region	

Implausible example (context-independent, unassociated prime-target pair)

Dina draws a	**cat**	next to the	**tiger**	*on the small*	soft chewing gum.
Dina draws a	**sun**	next to the	**tiger**	*on the small*	soft chewing gum.
	prime		target	spillover region	

In all, 112 test items were constructed, 56 related and 56 unrelated sentences. These were divided equally across the four semantic types (14 pairs i.e. 28 sentences for each type) and within semantic types, half was plausible and half was implausible. All test sentences (in Dutch) are provided in Appendix G.

List compilation. To compile the experiment lists, the full set of 112 test sentences was divided in two halves, each including seven related and seven unrelated items from each of the four semantic types, with equal numbers of plausible and implausible items. Each list thus contained only one sentence of each of the pairs, meaning that within children, the halved subsets within each semantic type were used as related and control items.[19] Additionally, seventeen similarly structured plausible and implausible dummy sentences were used as practice trials and start-up trials at the start of the experiment and after the half-way break. In total, each participant therefore read and responded to 73 items, of which 38 were plausible and 35 were implausible. The startup trials contained a few more plausible items to create an easier start to the experiment and after the break.

Of each of the two lists, three pseudo-randomized versions were created with different orderings. Within each of these six versions, related and unrelated sentences of each of the four semantic types were spread evenly across the list, occurring roughly equally often in earlier and later trials both before and after the break. No more than four consecutive items were of the same plausibility type.

Presentation. The experiment was run on two identical laptops using E-Prime software (Schneider et al., 2002). Sentences were presented in a monospaced font (Courier New), and sized to fit a single line on the test computers. Each trial began with a white screen for 500ms, followed by the first sentence screen with a fully masked sentence. Masking was done by replacing all letters and commas with hyphens (-), leaving punctuation at the end of the sentences unchanged.

Participants pressed the space bar to reveal each individual word in turn. After the last word in the sentence was revealed, a final space bar press led to a screen with a green V and a red X for the plausibility judgement to be made. The 'C' and 'M' keys were marked with a green and red sticker respectively, to serve as the plausible and implausible answer options. After the answer was provided, a white screen appeared again for 500ms, automatically followed by the next masked sentence. Reading times per word, i.e., between space bar presses, and the speed and accuracy of the plausibility judgement were recorded.

---

[19] In the first pilot group, a design with two blocks was tested in which each child was administered both versions, with the order counterbalanced across children. However, in this setup the experiment became too long for the children to remain fully concentrated, and in addition all children reported recognizing the sentences in the second block, making the measurement unreliable.

4.2.2.2. Reading comprehension

Reading comprehension was tested using the same shortened version of *Begrijpend Lezen 678* [Reading Comprehension grades 4 5 6] (Aarnoutse & Kapinga, 2006) used in Spätgens and Schoonen (2018) and in Cremer (2013, chapter 5). Both superficial and in-depth comprehension is tested in 32 questions on five short texts. The test is preceded by an example text with one multiple choice question with four answer options, and three true/false statements. The internal consistency of the task in this sample was satisfactory (Cronbach's α = 0.73).

4.2.2.3. Vocabulary size

A standardized vocabulary size test was used, the Cito *Leeswoordenschat* [Reading vocabulary] test by Verhoeven and Vermeer (1995). In this test, target words are provided in a neutral sentence and participants are required to select the correct definition from four answer options. There are two practice items and 32 test items. Internal consistency for the vocabulary task was somewhat lower than for reading comprehension and the context-independent semantic knowledge task, but not unsatisfactory: Cronbach's α = 0.62.

4.2.2.4. Context-independent semantic knowledge

Declarative context-independent semantic knowledge was assessed using the Word Associates Test (WAT, Schoonen & Verhallen, 2008). This task consists of 30 test items plus two practice items, each including one target word and six answer options. Participants are required to select three answer options that are 'always related' (i.e. context-independently related) to the target word. Distractor items include unrelated and context-dependently related words. Items were scored correct, i.e. awarded one point, if all three context-independently related words were selected. Any other answer pattern received zero points. Two differently ordered versions were randomly distributed in the class test setting. Internal consistency in this sample was satisfactory with Cronbach's α = 0.71.

4.2.2.5. Word decoding

Decoding speed was measured by using the two most difficult lists from the standardized *Drie Minuten Toets* [Three Minutes Test] (Verhoeven, 1992). For each list, participants were required to read as many words as possible out loud within one minute. This results in two scores per participant: the number of words read minus the number of errors made on each respective list. The scores on the two lists correlated strongly ($r = 0.91$, $p < 0.001$) and were therefore averaged to create one score per participant.

4.2.2.6. Language interview

The participants' language status was determined by means of a short language interview including questions on their home language(s) and the age at which they started school in the Netherlands.

4.2.3. Procedure

Participants performed the tasks in the school setting, which required one to two school days per class. All tasks were administered by the first author or one of two trained test assistants. Testing started in class with reading comprehension, vocabulary size and the WAT. Afterwards, the children performed the self-paced reading experiment individually in a quiet room, followed by the word decoding task and the language interview.

As was discussed in the Materials section, each of the classroom-administered tasks contained a few example questions, which were discussed according to the protocols prescribed by the respective instruction manuals. Classroom testing lasted 90 minutes in total.

The individual sessions started with the self-paced reading task. Participants were explained the self-paced reading procedure, and told to read as quickly and accurately as possible. The plausibility judgement was defined as a decision on whether sentences were 'logical' and 'made sense'. The experimenter provided a few example sentences similar to the experimental items for the child to judge verbally and discussed the answers. A short written reinforcement of the instructions was provided at the start of the experiment, followed by the practice items, for which feedback was provided on-screen. After the practice trials, the experimenter asked whether everything was clear and again emphasized the importance of reading as quickly and as accurately as possible. No further feedback was provided during the experiment.

Next, the two word lists for decoding were administered, again according to the protocols set by the task's instruction manual, followed by the language interview. Individual sessions lasted about 20-30 minutes.

4.2.4. Data handling and analysis

All analyses were performed in R version 3.4.1 (R Core Team, 2017). The reading time data and the reading comprehension scores specifically were modelled using the lme4 package for mixed effects models (Bates et al., 2015).

For the reading time analyses, all dummy items were removed before the analyses. The high mean accuracy on the plausibility judgement task (90% correct) indicated that participants performed the task by reading for comprehension. None of the participants in the final data set had an accuracy score lower than 70%. Because the sentences were designed to become implausible after the spillover region and accuracy was not predictive of reading times in any of the four target sets (all *p*'s > 0.05), both correctly and incorrectly answered items were included in the reading time analyses. Three sentences had comparatively low average accuracy, between 55% and 60% (all other items were > 70% accurate), but were not outliers in terms of reading times on the targets.[20]

For reasons of brevity, the set of targets itself will be referred to as T0, and the following words in the spillover region will be referred to as S1 (target + 1), S2 (target + 2) and S3 (target + 3). Minimal data cleaning was performed on the reading times on each of these target positions: reading times shorter than 100 ms were removed, and personal outlier scores were removed by trimming the data that were +/- 2.5 *participant SD* from the *participant* mean for each position.

A strict protocol was used to identify the effects of relatedness for each of the semantic types in the targets and spillover regions. In the first step,

[20] We also checked that for all reported models, running on a trimmed dataset involving only reading times from correctly answered items led to the same conclusions, which was the case.

reading times for the target and each spillover position (T0 to S3), transformed using the natural log, were modelled separately using mixed effects models, which included the following covariates as fixed effects:

- The frequency (log transformed using the natural log) and length (in letters) of the target words – i.e. T0, S1, S2 or S3.
- The position of the sentence in which the target occurred in the experiment, taking into account that reading times speeded up consistently throughout the experiment (cf. Hofmeister, 2011).
- The reading time of the prime (log transformed using the natural log), controlling for difficulties reading the prime that may affect reading times on the targets (cf. Baayen, Davidson, & Bates, 2008).
- The participant's decoding speed score, to control for individual differences in reading speed which may affect the shape of the priming effect (cf. Kliegl et al., 2010).

The rest of the fixed part of the model consisted of the four-way interaction and all lower interactions between the variables of interest Language Group, Condition, Context-Dependency and Association, which were all contrast coded and centered binary factors. The four-way interaction was thus: Language Group [-.5 monolingual; +.5 bilingual] x Condition [-.5 related; +.5 unrelated] x Context-Dependency [-.5 context-independent; +.5 context-dependent] x Association [-.5 non-associated; +.5 associated]. This interaction simultaneously takes into account the influence of Condition on the 2 x 2 semantic contrast and differences between the language groups. [21]

Finally, the random part of the models included random intercepts for the participant, target and school levels. Random slopes for the variables of interest and their interactions were included as much as possible, following Barr, Levy, Scheepers & Tily (2013). This meant that the three-way and lower interactions between Condition x Context-dependency x Association were included as random slopes for participants, and where possible, a

[21] Because the number of words between primes and targets varied across sentences, we checked whether model fit was improved by including another factor representing this difference in the interaction, which was not the case. For most models there was no significant difference in deviance, and in all cases, deviance for the simpler model excluding the effect of distance between prime and target was lower.

random slope for Language Group by targets was also estimated. However, to attain convergence the latter slope needed to be omitted for some models.

To determine effects of Condition within the two semantic contrasts, these main models were examined for significant interactions involving the Condition factor and the Context-Dependency and/or Association factors. If a significant interaction was found, it was examined by breaking up the target set first by Context-Dependency or Association, depending on the interaction that was found to be significant. A separate model was then fitted for each subset with the same characteristics as described above, except that now a three-way interaction was the largest interaction involved (i.e. Language Group x Condition x Association if the data set were split by Context-Dependency). If an interaction between Condition and, in this example, Association was found to be significant in this step, the relevant data set was again split up and analyzed in the same way.

If a significant main effect of Condition without interactions was observed at any of these steps, a semantic priming score was calculated for the relevant semantic set. First, the log transformed reading times were residualized using a model including all control variables and random intercepts for items, participants and schools. The priming scores were then calculated based on these residualized reading times for each participant by subtracting the individual mean for related items from the individual mean for unrelated items.[22]

4.3. Results

4.3.1. Descriptives

The scores for the two language groups on the standardized tasks measuring reading comprehension, vocabulary size, context-independent semantic knowledge and decoding speed are presented in Table 4.3. All differences are in the expected direction, with monolinguals showing

[22] The method used in the previous chapters, using participants' random slopes for Condition, unfortunately was not possible in the present dataset, due to perfect correlations in the random effects structure in the simplified models used to estimate the scores.

slightly higher scores. However, independent samples t-tests showed that there were no significant differences for reading comprehension, vocabulary size and decoding (all *p*s > 0.05), and the effect sizes for these tasks were negligible. Only for context-independent semantic knowledge a small effect (Cohen, 1969) appears: Cohen's $d = 0.36$. This difference was also statistically significant: $t(133.22) = 2.1$, $p = 0.038$.

Table 4.3

Descriptives for standardized tasks by language group

	Total (*N* = 137)		Monolingual (*N* = 61)		Bilingual (*N* = 76)		
	M	*SD*	*M*	*SD*	*M*	*SD*	*d*
Reading comprehension	23.18	4.39	23.44	4.38	22.96	4.41	0.11
Vocabulary size	16.12	4.14	16.33	4.01	15.95	4.27	0.09
Context-independent semantic knowledge	19.21	4.37	20.07	4.06	18.53	4.51	0.36
Word decoding	93.05	16.08	93.70	14.81	92.53	17.12	0.07

4.3.2. Self-paced reading task

As discussed in the Method section, mean accuracy on the plausibility judgement task was 90%, and both correct and incorrect items were kept in the dataset for the analyses. Mean reading times at the target and spillover positions are provided in Table 4.4, and a more detailed visualization of the reading times at each position by condition, semantic set and language group is presented in Figure 4.1.

Table 4.4

Mean raw reading times at the target and spillover positions

	M	*SD*
T0 (*N* = 7463)	454.5	171.4
S1 (*N* = 7459)	453.1	158.3
S2 (*N* = 7450)	443.0	151.7
S3 (*N* = 7444)	446.5	152.1

Following the protocol described in the Method section, reading times of words at all four target and spillover positions were examined systematically, starting with models including all four semantic sets and the full four-way interactions. The model estimates for the fixed effects of the variables of interest (Language Group, Condition, Context-Dependency and Association) are provided in Table 4.5 for each of the four full models.[23]

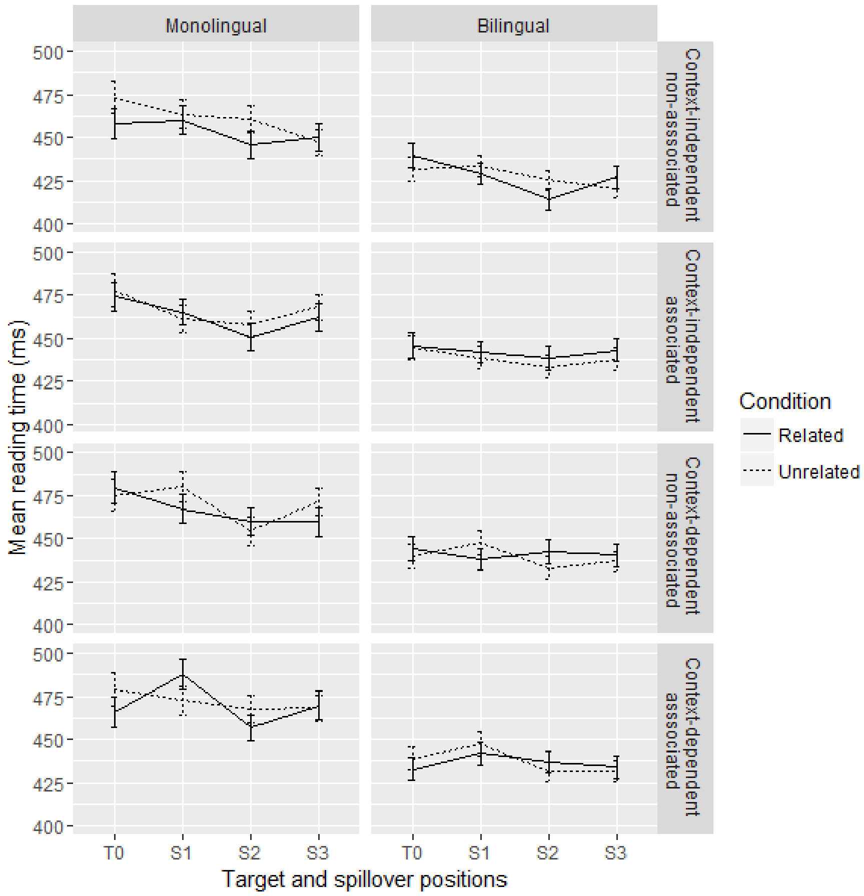

Figure 4.1. Raw reading times on targets and spillover words by semantic type, condition, and language group.

[23] Note that fixed effects for control variables and the random structure are omitted for reasons of brevity and clarity in all tables reporting estimates for the priming models in this chapter.

Table 4.5

Fixed effects of variables of interest for full models of log transformed reading times at the T0, S1 , S2 and S3 positions

	T0 (N = 7463)			S1 (N = 7459)		
	Estimate (SE)	t	p	Estimate (SE)	t	p
Language Group	-0.068 (0.035)	-2.0	0.050 *	-0.062 (0.031)	-2.0	0.046 *
Condition	0.001 (0.006)	0.2	0.812	0.003 (0.005)	0.5	0.613
Context-Dependency	-0.005 (0.006)	-0.7	0.466	0.012 (0.011)	1.1	0.254
Association	-0.010 (0.007)	-1.5	0.129	0.002 (0.010)	0.2	0.815
Language Group x Condition	-0.008 (0.011)	-0.7	0.482	0.014 (0.010)	1.4	0.172
Language Group x Context-Dependency	-0.008 (0.012)	-0.7	0.466	-0.014 (0.011)	-1.3	0.189
Condition x Context-Dependency	0.006 (0.011)	0.6	0.582	0.008 (0.010)	0.8	0.427
Language Group x Association	0.000 (0.012)	0.0	0.997	-0.001 (0.010)	-0.1	0.889
Condition x Association	0.005 (0.011)	0.5	0.601	-0.025 (0.010)	-2.4	0.016 *
Context-Dependency x Association	-0.013 (0.013)	-1.0	0.296	-0.007 (0.021)	-0.4	0.722
Language Group x Condition x Context-Dependency	0.017 (0.022)	0.8	0.428	0.024 (0.020)	1.2	0.231
Language Group x Condition x Association	0.031 (0.021)	1.5	0.141	0.026 (0.021)	1.3	0.209
Language Group x Context-Dependency x Association	0.001 (0.023)	0.1	0.956	-0.016 (0.020)	-0.8	0.428
Condition x Context-Dependency x Association	0.035 (0.022)	1.6	0.107	0.003 (0.020)	0.1	0.896
Language Group x Condition x Context-Dependency x Association	-0.061 (0.044)	-1.4	0.164	0.051 (0.040)	1.3	0.208

* $p < 0.05$.

Table 4.5 (continued)

Fixed effects of variables of interest for full models of log transformed reading times at the T0, S1 , S2 and S3 positions

	S2 (N = 7450)				S3 (N = 7444)			
	Estimate (SE)	t	p		Estimate (SE)	t	p	
Language Group	-0.059 (0.031)	-1.9	0.058		-0.069 (0.032)	-2.1	0.032	*
Condition	0.006 (0.005)	1.3	0.184		0.001 (0.005)	0.2	0.855	
Context-Dependency	0.007 (0.011)	0.7	0.485		0.016 (0.009)	1.8	0.066	
Association	0.003 (0.010)	0.3	0.779		0.006 (0.009)	0.7	0.518	
Language Group x Condition	-0.013 (0.010)	-1.4	0.172		-0.014 (0.010)	-1.4	0.167	
Language Group x Context-Dependency	0.008 (0.011)	0.7	0.466		-0.010 (0.010)	-1.0	0.344	
Condition x Context-Dependency	-0.023 (0.010)	-2.3	0.023	*	0.010 (0.010)	1.0	0.342	
Language Group x Association	0.005 (0.010)	0.5	0.625		-0.007 (0.010)	-0.7	0.472	
Condition x Association	-0.002 (0.010)	-0.2	0.865		-0.003 (0.010)	-0.3	0.784	
Context-Dependency x Association	-0.021 (0.020)	-1.0	0.301		-0.037 (0.018)	-2.1	0.034	*
Language Group x Condition x Context-Dependency	-0.016 (0.020)	-0.8	0.422		-0.002 (0.021)	-0.1	0.934	
Language Group x Condition x Association	-0.009 (0.019)	-0.5	0.642		0.006 (0.020)	0.3	0.773	
Language Group x Context-Dependency x Association	-0.044 (0.019)	-2.3	0.021	*	-0.020 (0.020)	-1.0	0.319	
Condition x Context-Dependency x Association	0.053 (0.020)	2.7	0.008	**	-0.012 (0.020)	-0.6	0.568	
Language Group x Condition x Context-Dependency x Association	-0.005 (0.040)	-0.1	0.907		0.050 (0.041)	1.2	0.223	

* $p < 0.05$; ** $p < 0.01$

Only reading times for the spillover positions at S1 and S2 were found to show significant interactions involving Condition, which will be discussed in more detail below. No significant interaction involving Condition was found at T0 and S3, suggesting that there were no significant differences between the related and unrelated condition at these positions. A significant interaction between Context-Dependency and Association does appear at S3, which suggests that the words at this spillover position were not comparable across the four semantic sets, independent of Condition. Apparently, there were inherent differences between the semantic sets in terms of the difficulty of reading these words, which persisted even after controlling for length and frequency. The significant main effects of language group at T0, S1 and S3 and to a lesser extent the marginal effect at S2 suggest that bilingual participants read the targets and spillover words faster. The significant interaction between Language Group, Context-Dependency and Association at S2 suggests that there is a difference in reading speed between the language groups, which depends on the semantic types. The exact shape of this interaction will however be dependent on the other significant interactions involving Context-Dependency and Association, which will be discussed further below.

4.3.2.1. Priming at S1

For reading times at the S1 position, the interaction between Condition and Association was significant in the full model. The S1 reading times were therefore split into two datasets, along the Association dimension. In the associated set, no significant effects of the variables of interest or their interactions was found, except for a significant main effect of Language Group ($\beta = -0.061$ (0.031), $t = 2.0$, $p = 0.046$), which is consistent with the significant main effect in the full S1 model. No significant effect of Condition was thus observed for the associated items. The model estimates for the non-associated set are presented in Table 4.6.

Table 4.6

Fixed effects of variables of interest for log-transformed reading times at S1 for the non-associated set (N = 3728)

	Estimate	(*SE*)	*t*	*p*	
Language Group	-0.069	(0.031)	-2.2	0.027	*
Condition	0.015	(0.007)	2.1	0.035	*
Context-Dependency	0.014	(0.014)	1.0	0.330	
Language Group x Condition	0.002	(0.014)	0.1	0.896	
Language Group x Context-Dependency	-0.005	(0.016)	-0.3	0.746	
Condition x Context-Dependency	0.007	(0.014)	0.5	0.638	
Language Group x Condition x Context-Dependency	-0.001	(0.029)	0.0	0.971	

** $p < 0.05$*

Again, a main effect of Language Group was found, with faster reading times for the bilingual group. In addition, a main effect of Condition with longer reading times for the unrelated items indicated a delayed facilitation effect of non-associatively related targets, whether context-independently or context-dependently related. The β estimate of 0.015 in log-transformed reading time translates to a facilitative priming effect of 1.5% of the mean reading times in this set (450.6), or about 6.8 milliseconds.

Table 4.7

Fixed effects of variables of interest for log-transformed reading times at S2 for the context-independent set (N = 3731)

	Estimate	(*SE*)	*t*	*p*	
Language Group	-0.062	(0.031)	-2.0	0.044	*
Condition	0.018	(0.007)	2.7	0.008	**
Association	0.013	(0.017)	0.7	0.462	
Language Group x Condition	-0.005	(0.014)	-0.4	0.729	
Language Group x Association	0.027	(0.014)	1.9	0.047	*
Condition x Association	-0.028	(0.013)	-2.1	0.033	*
Language Group x Condition x Association	-0.006	(0.027)	-0.2	0.813	

** $p < 0.05$; ** $p < 0.01$*

4.3.2.2. Priming at S2

The full model of the reading times at S2 showed significant interactions between Condition and Context-Dependency, and between Condition, Context-Dependency and Association and the S2 reading times were therefore firstly broken up along the Context-Dependency dimension. No significant effects of the variables of interest or their interactions were observed for the context-dependent items. The context-independent items however did show a significant interaction. The estimates for the priming model of the S2 context-independent items are provided in Table 4.7.

A main effect of Language Group was observed again, combined with a significant interaction between Language Group and Association. Furthermore, a significant main effect of Condition was found, but the significant interaction between Condition and Association suggests that this effect was dependent on the associative type. Both interactions were studied in more detail by splitting the data set one more time along the Association dimension. In the resulting associated, context-independent set no significant effects of the variables of interest were found. In contrast, in the non-associated, context-independent set, significant main effects did arise, which clarify the shape of the higher-order interactions in the previous model. Table 4.8 presents the estimates for this latter set. The significant main effects of both Language Group and Condition in absence of an interaction indicate that the bilinguals read words in this set faster than the monolinguals, and that there was a significant facilitation effect for words at S2 following non-associated, context-independently related targets. The β estimate of 0.033 for the log-transformed reading time translates to a facilitative priming effect of 3.3% of the mean reading times in this set (434.7), or about 14.3 milliseconds. This effect holds for both language groups, given the nonsignificant interaction.

Table 4.8
Fixed effects of variables of interest for log-transformed reading times at S2 for the non-associated, context-independent set (N = 1873)

	Estimate	(*SE*)	*t*	*p*	
Language Group	-0.069	(0.029)	-2.4	0.018	*
Condition	0.033	(0.009)	3.5	0.000	***
Language Group x Condition	0.000	(0.019)	0.0	0.999	

** $p < 0.05$; *** $p < 0.001$*

4.3.3. Semantic priming and reading

In the second step, semantic priming scores based on residualized reading times were calculated for the non-associated S1 targets and the non-associated, context-independent S2 targets. The descriptives by language group are provided in Table 4.9. As was also clear from the previous analyses, there are no significant differences between the language groups (both *p*'s > 0.05).

Table 4.9
Descriptives for priming scores based on residualized reading times by language group

	Total (*N* = 137)		Monolingual (*N* = 61)		Bilingual (*N* = 76)		
	M	*SD*	*M*	*SD*	*M*	*SD*	*d*
S1 non-associated priming	0.0148	0.0743	0.0136	0.0781	0.0158	0.0715	0.03
S2 non-associated, context-independent priming	0.0327	0.0998	0.0328	0.1066	0.0326	0.0946	0.00

The two individual scores were entered in separate mixed effects models as predictors of the reading comprehension scores, together with language group and their interaction, and the control variables vocabulary size, WAT and decoding. A random effect controlling for school variance was included. The fixed effects of these models are provided in Table 4.10 (S1 non-associated priming) and Table 4.11 (S2 non-associated context-independent priming). As expected, vocabulary size and context-independent semantic knowledge were significant predictors of the reading comprehension scores in both models, while Language Group did not make a significant contribution. Neither the S1 nor the S2 priming scores exert a significant influence on the comprehension scores or show a significant interaction with Language Group, but the main effect of S1 was marginal. However, inclusion of neither the S1 nor the S2 priming scores represented a significant improvement on a base model including only the control variables.

Table 4.10

Fixed effects estimates for reading comprehension scores, S1 non-associated priming model (N = 137)

	Estimate (*SE*)		*t*	*p*	
Intercept	11.22	(2.20)	5.1	0.000	***
Vocabulary size	0.33	(0.09)	3.7	0.000	***
Context-independent semantic knowledge	0.28	(0.09)	3.3	0.001	**
Decoding	0.01	(0.02)	0.5	0.625	
S1 non-associated priming	8.34	(4.43)	1.9	0.060	
Language Group	0.08	(0.68)	0.1	0.901	
S1 non-associated priming x Language Group	-0.68	(8.88)	-0.1	0.939	

*** p < 0.01; *** p < 0.001*

Table 4.11

Fixed effects estimates for reading comprehension scores, S2 non-associated, context-independent priming model (N = 137)

	Estimate (*SE*)		*t*	*p*	
Intercept	11.63	(2.18)	5.3	0.000	***
Vocabulary size	0.37	(0.09)	4.0	0.000	***
Context-independent semantic knowledge	0.26	(0.09)	3.0	0.003	**
Decoding	0.01	(0.02)	0.2	0.819	
S2 non-associated, context-independent priming	4.45	(3.32)	1.3	0.180	
Language Group	0.27	(0.70)	0.4	0.699	
S2 non-associated, context-independent priming x Language Group	-6.30	(6.57)	-1.0	0.338	

*** p < 0.01; *** p < 0.001*

4.4. Discussion and conclusions

4.4.1. Semantic priming during sentence reading

We observed two delayed significant effects of non-associated word pairs, but no facilitation due to associative relations. The absence of effects for associatively related word pairs appears counter-intuitive at first sight. Not only was the associative boost, i.e. additional facilitation for associated versus non-associated word pairs (Lucas, 2000), not borne out in our data, we did not find a significant effect for associated word pairs altogether. Other studies of semantic priming at the sentence level have typically used associated word pairs, with mixed results. While there are studies that report associative boost effects (e.g. Myers, Cook, Kambe, Mason, &

O'Brien, 2000) or facilitation of associated pairs but only a marginal effect for non-associated pairs (Carroll & Slowiaczek, 1986), others found no evidence for any associative facilitation (Morris, 1994; Traxler, Foss, Seely, Kaup, & Morris, 2000). Some researchers suggest that associative priming within sentences is highly susceptible to variation in the previous context (Camblin, Gordon, & Swaab, 2007; Hess, Foss, & Carroll, 1995) and may only be observed in early measures of reading, such as initial eye landing spots and first-pass reading times (Camblin et al., 2007; Lavigne, Vitu, & d'Ydewalle, 2000). These mixed findings suggest that associative priming potentially could have been elicited in our sample as well, given different stimuli or measures that are more sensitive to early processes.

Because semantic priming studies at the sentence level often equate association to semantic relation or compare associated to non-associated pairs, there is less evidence for pure priming of non-associated word pairs during sentence reading. Carroll and Slowiaczek (1986) found a marginal effect at their non-associatively related category coordinate targets, but did not examine the post-target regions. Our two effects of non-associated priming suggest that it may affect later measures especially, with a weaker general effect of both context-dependent and context-independent pairs at S1, and a stronger effect of context-independent pairs only at S2. Facilitation for context-dependent items was thus only observed in conjunction with the context-independent set. The effect of non-associative, context-independently related word pairs we found at S2 is in accordance with previous findings from single word priming studies in the same target population: Cremer (2013) and Spätgens and Schoonen (2018) found such facilitation in the written and auditory modality, respectively.

4.4.2. Semantic priming and reading comprehension

Following the findings by Nation and Snowling (1999), we hypothesized that especially non-associated, context-independent priming would be predictive of reading comprehension scores. However, although we found delayed facilitation of exactly this semantic type at S2, it was not a significant predictor of the reading comprehension scores. The marginal effect we did find, a positive relation between non-associated priming at S1 and reading comprehension, is tentative at best, and did not significantly improve the reading comprehension model. Clearly, the data do not warrant a strong interpretation of this effect.

It appears that although we expected to approximate the process of semantic activation during reading more closely by studying word pairs in a sentence context, this was not sufficient to identify a clear relation to reading comprehension on an individual differences level. This suggests that the group effects identified by Nation and Snowling (1999) and Bonnotte and Casalis (2010) are not borne out in individual differences between normal readers, as was also concluded in Cremer (2013) and Spätgens and Schoonen (2018).

If semantic activation of related words is used during reading comprehension, as is for example posited in the Landscape model (Van den Broek et al., 1999; Van den Broek et al., 2005), an absence of individual differences effects on comprehension is curious. However, priming effects are very subtle and subject to the influence of a wide range of experiment-internal and -external factors. Therefore, it may be the case that even if semantic activation contributes directly to comprehension, it is difficult to capture such a relation. More sensitive and early measures of reading processes such as those obtained in eye tracking may be a solution to this issue. However, Stolz, Besner and Carr (2005) and Yap, Hutchison and Tan (2016) showed that although group-level priming effects are generally robust, individuals do not show consistent facilitation across sessions. Translating individual differences from such a measure to individual differences in other skills may therefore be complicated altogether, especially when priming effects are obtained from a single session.

In sum, because no significant effects were found, we cannot draw definitive conclusions on the relation between semantic priming and reading comprehension from this data set. The results do confirm the principal role of vocabulary size and context-independent semantic knowledge for reading comprehension[24], and suggest that even if future studies are able to identify significant effects of semantic priming on

---

[24] It was also not the case that a potential effect of the priming measures was already captured by the vocabulary size influence, because bivariate correlations between the priming scores and reading comprehension were non-significant.

reading comprehension on an individual level, the influence of this factor will be much smaller than that of established vocabulary measures such as size and depth.

4.4.3. Differences between monolinguals and bilinguals

The only significant differences observed between monolinguals and bilinguals were lower scores on the WAT, i.e. declarative context-independent semantic knowledge, and faster reading times for bilinguals at most target and spillover positions in the self-paced reading task, independent of relatedness or semantic type. The lower performance on the WAT has been identified in multiple previous studies (Cremer, 2013; Schoonen & Verhallen, 2008), but may be all the more interesting in the present sample given that differences in vocabulary size and reading comprehension were not observed. Limitations in declarative context-independent knowledge may thus persist in bilinguals even when their vocabulary size is on par with that of their monolingual peers. Apparently then, building the number of vocabulary items does not necessarily coincide with building the quality or depth of knowledge on these items, which is in line with studies that have shown that vocabulary size is a measure separate from other dimensions of vocabulary knowledge such as depth (see Schmitt, 2014, for an overview). However, these limitations in declarative context-independent knowledge did not translate to differences in semantic priming between monolinguals and bilinguals. This would appear counter-intuitive at first sight, but it was also the case in the sample studied by Cremer (2013). Working with the concept of context-independency in a declarative task such as the WAT may be more difficult for bilinguals, even if automatic activation of this knowledge is already in place.

The faster reading times for bilinguals in the self-paced reading experiment may point to a subtle advantage in decoding speed, which would be in accordance with small advantages in decoding speed for that have been found in other studies (Melby-Lervåg & Lervåg, 2014; Raudszus et al., 2018). However, it may also be a performance difference specific to this task, since there were no group differences in the decoding task itself.

Most notable however, is the absence of significant group differences on most measures, including semantic priming, in line with our previous work (Spätgens & Schoonen, 2018). Our selection of schools in neighborhoods with low to average SES and the fact that the bilingual children mostly spoke Dutch at home in addition to another language, likely plays a role in these results. Recent national assessment reports on the Dutch school population (e.g. Kuhlemeier et al., 2014) also show that monolinguals and bilinguals in the current Dutch context show similar language performance when SES is controlled for.

4.4.4. Limitations and directions for future research

There are a few limitations to the self-paced reading experiment and the other tasks used in our study, which may be improved upon in future studies. Regarding the self-paced reading experiment, testing in a highly controlled lab session would have been preferred, because priming effects are concerned with such minute differences in behavioral measures. In our target population however, this was not a viable option, and testing in the school setting may have affected the potential of our experiment to capture these subtle effects, in addition to the general instability of individual priming effects discussed earlier.

Future studies may also be able to improve on some of the features that affected the strength of our experiment. Our attempt to study the facilitative effect of a 2 x 2 contrast of semantic relations did mean that the number of semantic pairs that satisfied all necessary criteria was somewhat limited within each set. With larger sets of related and unrelated pairs, additional priming effects or a clearer relation to reading comprehension might have been found.

Finally, the use of highly frequent words enabled us to disentangle priming effects from potential vocabulary size effects. However, because these words were so familiar, individual differences in activation between them may not have been sufficiently large in this age group to be predictive of reading comprehension differences at fifth-grade levels.

With regard to the other tasks, higher reliability scores would have been desirable. In addition, as Swart and her colleagues (2017a) have pointed out, reading comprehension is a complex process, which may best be assessed by means of multiple tasks to tap all its facets.

4.4.5. Concluding remarks

Despite these limitations, there is clear consistency between our findings and previous research, in which non-associated context-independent priming was also observed, but not found to be predictive of individual differences in reading comprehension. This suggests that either relating a relatively unstable measure such as priming to a complex skill such as reading comprehension is not feasible on an individual level, or that an individual's degree of semantic activation, even when captured during reading, simply has little correspondence to his or her reading comprehension performance.

Chapter 5
Discussion

The present thesis sought to examine the influence of individual differences in various measures of semantic knowledge and activation on reading comprehension in monolingual and bilingual minority children. The structure of the semantic network and the activation of knowledge within this network were studied in a number of ways, in an attempt to answer the following research questions:

1) How do monolingual and bilingual children's semantic network organization, semantic access, and sensitivity to semantic priming differ?
2) How do individual differences in semantic network organization, semantic access, and semantic priming relate to differences in reading comprehension?

In this chapter, the findings from the three empirical studies will be brought together to answer these main questions.

5.1. Semantic knowledge in monolinguals and bilinguals

The bilingual children participating in the studies in this dissertation mostly acquired Dutch in addition to a minority language at home. In contrast to the many studies both within and outside the Netherlands showing disadvantages for bilinguals on measures of vocabulary knowledge and reading comprehension (Bialystok, 2009; Cremer, 2013; Melby-Lervåg & Lervåg, 2014; Michael & Gollan, 2005; Raudszus et al., 2018; Schoonen & Verhallen, 2008; Schwartz & Katzir, 2012) only a few significant differences between the language groups were found in the three studies comprising this dissertation. Moreover, none of these significant differences concerned the semantic knowledge aspects in focus.

5.1.1. Semantic organization

Previous studies on word association in monolinguals and bilinguals, and L1 and L2 learners have reported mixed results. In a study involving participants from the same population as the children who contributed to the study in Chapter 2, Cremer and colleagues (2011) found differences between monolinguals and bilinguals in terms of the number of 'other' associations, reflecting that more stimulus words were unknown to the bilinguals. By extension, this resulted in fewer semantically related associations by the bilinguals. However, the language groups did not show clear preferences for semantic subtypes: within the semantic associations the participants did provide, the proportion of responses was similar across semantic subtypes between the language groups. An absence of differences in semantic preferences in the word association task appears curious when taking into account the differences that have been reported for monolinguals and bilinguals in vocabulary size, knowledge of semantic relations and other depth measures (Cremer, 2013; Hoff et al., 2012; Raudszus et al., 2018; Schoonen & Verhallen, 2008; Schwartz & Katzir, 2012; Verhallen & Schoonen, 1993). By using the multiple association task and a detailed acquisition-based classification system, an attempt was made to identify perhaps subtle differences between monolingual and bilingual children in terms of associative behavior.

Overall, the semantic organization of the monolingual and bilingual 10-11-year-olds in Chapter 2 showed many similarities to the data obtained for adults by De Deyne and Storms (2008a): similar to the adult study, taxonomic associations were most prominent as first responses, but became less frequent as second and third responses, while the reverse was true of situational and subjective associations. Only feature associations showed opposite patterns: decreasing across responses for the children in this study, while the adults in De Deyne and Storms (2008a) showed an increase of feature associations across responses. For the children in this sample, feature associations thus showed the pattern that in adults is reserved for the most context-independent semantic type, taxonomic relations. This effect was interpreted as a sign of the structure of the semantic network still being in development, and on the road to a system which involves context-independent relations in addition to context-dependent relations. When taxonomic associations are not yet as readily available, feature associations may serve as the next best thing.

The data did not show significant differences between the language groups in any of the four semantic categories. The bilinguals did produce more 'other' associations, similar to the findings by Cremer et al. (2011), and obtained lower scores on the reading comprehension task. Bilinguals' associative preferences were thus very similar to monolinguals', despite more limited language proficiency in terms of reading comprehension. Wolter (2006) suggested that the semantic structure of the L2 mental lexicon may well build on that established in the L1 and in the case of our bilingual language learners, it may even be the case that both structures are built simultaneously and are therefore very similar. Even if bilinguals' vocabulary size was comparatively smaller, the frequent words used in the word association study were likely familiar enough to have been firmly embedded in a structure that was similarly shaped to that of monolinguals. Indeed, as suggested by the weaker links hypothesis (Gollan & Silverberg, 2001; Gollan et al., 2008; Michael & Gollan, 2005) differences in vocabulary knowledge between monolinguals and bilinguals can be attributed to differences in exposure to a large extent, and for highly frequent items, such differences may thus be reduced.

5.1.2. Access to semantic information

Studies on access to semantic information have reported mixed results for monolinguals and bilinguals. Cremer (2013) demonstrated disadvantages for bilingual children in access as measured by lexical decision speed and semantic classification speed, but not for speed in a forced-choice task assessing knowledge of context-independent semantic relations. Similarly, Gollan, Montoya, Fennema-Notestine and Morris (2005) found that picture naming speed in bilingual adults was slower compared to monolinguals, while semantic classification speed of pictures was similar in both language groups.

In Chapter 3, access was operationalized as semantic categorization speed, similar to Cremer (2013). However, contrary to her data, the monolinguals and bilinguals did not show significant differences on this measure. The absence of significant group differences in terms of vocabulary size, decoding, rapid naming and reading comprehension in this sample suggests that the monolingual and bilingual groups were similar in terms of language proficiency. This may explain the discrepancy – the children in Cremer's study did show lower scores on a reading comprehension task and in knowledge of context-independent semantic relations.[25] Even though vocabulary size does not fully overlap with other aspects of vocabulary knowledge (such as depth and fluency, cf. Schmitt, 2014; Tannenbaum et al., 2006) and therefore will not necessarily coincide with a measure of semantic access, the consistent absence of significant group differences in all language measures in this sample is in line with the similar performance in terms of semantic access, and appears to point to comparable language proficiency in the two groups.

5.1.3. Single-word and sentence-level semantic priming

As previously discussed, evidence for vocabulary delays in bilinguals compared to monolinguals abounds. Additionally, bilinguals have been shown to be at a disadvantage in terms of declarative knowledge of especially context-independent semantic relations (Cremer, 2013; Raudszus et al., 2018; Schoonen & Verhallen, 2008; Schwartz & Katzir, 2012; Verhallen & Schoonen, 1993). Furthermore, semantic priming has been demonstrated to be dependent on proficiency level in L2 (Kotz & Elston-Güttler, 2004) and may not necessarily occur in bilingual minority children (Goodrich & Lonigan, 2018). Bilinguals were therefore expected to show reduced semantic priming, especially of context-independent semantic relations. However, Cremer (2013) did not find such a difference between her monolingual and bilingual participants, who did show differences in reading comprehension and declarative context-independent semantic knowledge, which indicates that the relation between semantic priming and vocabulary knowledge may not necessarily be straightforward.

The general patterns of activation in the two priming studies in this dissertation can be characterized as follows. The auditory single-word priming study reported on in Chapter 3 targeted both context-independent

[25] Gollan and colleagues (2005) did not report other language proficiency measures for their adult participants.

(*tiger* – *cat*) and context-dependent (*forest* – *squirrel*) semantic relations, which were not associated. Associative relations were defined in this dissertation as empirically determined relations based on group data from a word association task; non-associatively related word pairs involved targets that were at most idiosyncratic second or third responses to their primes in the word association task reported in Chapter 2. A group-level priming effect was obtained for the context-independent word pairs, but not for the context-dependently related word pairs. The latter, situationally related pairs, were defined somewhat differently from the functional pairs used in the studies by Bonnotte and Casalis and Nation and Snowling. However, priming for word pairs similar to those used in this dissertation has been demonstrated in adults (Hare et al., 2009). It was argued in Chapter 3 that the stringent control for association strength using the multiple association data obtained in Chapter 2 may have been responsible for the absence of facilitation of context-dependent pairs: situational semantic relations may be more dependent on association, and controlling for second and third association responses removed many possible pairs of this type. However, in the sentence reading study in Chapter 4, associated pairs were added for both semantic types, but failed to elicit facilitation. The presence or absence of context-dependent semantic priming is thus not unequivocally related to association strength. Furthermore, previous research shows that the consistent associative boost effects found in single-word priming (cf. Lucas, 2000) may not be as straightforward in sentence reading studies (e.g. Morris, 1994; Traxler et al., 2000), which may have affected the results for the associated pairs in the self-paced reading experiment. Consistent with Chapter 3 however, facilitation effects were found at the group level for non-associated context-independently related word pairs. These resulted in facilitation in the spillover region, namely at the second word after the target, and additionally, a general facilitation effect of non-associated pairs was observed at the first word after the target.

Interestingly, no significant differences in activation were found between the two language groups, for any of the semantic relations that were included in either the auditory single word semantic priming experiment or the self-paced reading experiment. Semantic priming effects were observed especially for non-associated, context-independent semantic pairs for both monolinguals and bilinguals, which is contrary to the findings by Goodrich and Lonigan (2018) but in line with Cremer (2013). Both of these studies also used category word pairs, but the participants in Cremer's study were from the same population as the participants in this study, which could explain the consistency with her findings.

5.1.4. Standardized measures

Notably, both language groups also performed similarly on most other language tasks that were included: decoding, vocabulary size and reading comprehension scores did not differ between the groups in both priming studies. The monolinguals only significantly outperformed the bilinguals on reading comprehension in Chapter 2 and the declarative context-independent knowledge task administered in Chapter 4. The effect was small but close to moderate for the reading comprehension task in Chapter 2 (d = 0.49) and small for the context-independent knowledge task in Chapter 4 (d = 0.36). Additionally, the bilinguals showed generally faster reading times in the self-paced reading experiment. With a few exceptions, language proficiency as measured in the samples participating in the priming studies was thus very similar for monolinguals and bilinguals.

The measure of declarative context-independent knowledge used, the WAT, has so far consistently shown significant differences between monolinguals and bilinguals (e.g. Cremer, 2013; Schoonen & Verhallen, 2008), and the findings from Chapter 4 suggest that this disadvantage is thus not necessarily dependent on a disadvantage in vocabulary size, but also not related to a difference in priming of similar semantic relations. Small advantages in decoding speed for bilinguals have been found by other studies as well (Melby-Lervåg & Lervåg, 2014; Raudszus et al., 2018), which may be the source of the faster self-paced reading times by bilinguals in Chapter 4.

5.1.5. Semantic knowledge in monolinguals and bilinguals: conclusions

Combined, the three studies showed few significant differences between monolingual and bilingual minority children, even though sample sizes were sufficiently large for medium effects to surface as significant results in the group comparisons. The only significant differences that were found concerned other measures than the focal semantic knowledge aspects: reading comprehension and the number of 'other' responses in the word association study, and declarative context-independent knowledge as measured by the WAT and reading times in the self-paced reading study. The comparable performance on reading comprehension and vocabulary size in the priming studies is perhaps most notable, given the many studies showing disadvantages for bilinguals on these measures, both across the world (e.g. Hoff et al., 2012; Mancilla-Martinez & Vagh, 2013; Melby-Lervåg & Lervåg, 2014; Stæhr, 2008) and in the Dutch context specifically (e.g. CBS, 2014; Heesters et al., 2007; Raudszus et al., 2018).

The difference between the priming studies and the above-mentioned research, and also between the priming studies and the word association study may be related to socio-economic status (SES) of the participants. A recent Dutch national assessment report (Kuhlemeier et al., 2014) showed that when SES is controlled for, differences between monolinguals and bilinguals disappear. In the two priming studies, the participating schools were selected based on neighborhood SES status more strictly than in the word association study: all participating schools were in average to low SES neighborhoods, while three out of eight schools in the word association study were in above-average SES neighborhoods. The numbers of monolingual and bilingual children in these schools were quite different: 42 monolinguals and 45 bilinguals in the higher SES schools, 22 monolinguals and 98 bilinguals in the lower SES schools. However, when the data from the association study are split by SES, although the differences between the language groups are no longer significant (likely due to the reduced group sizes), the effect sizes in especially the low SES schools remain similar to those observed in the full set (see Table 5.1). This means that SES differences cannot be a full explanation for the contrasting results.

Table 5.1

Differences between language groups in reading comprehension scores and raw number of 'other' responses split by SES in word association study

	Monolingual		Bilingual				
	M	*SD*	*M*	*SD*	*t*(*df*)	*p*	*d*
Reading comprehension							
All schools	32.06	6.43	29.06	5.69	3.210 (109.0)	0.002	0.49
Lower SES schools	31.14	5.48	28.58	5.42	1.980 (30.9)	0.057	0.47
Higher SES schools	32.55	6.89	30.11	6.18	1.732 (82.4)	0.087	0.37
'Other' responses							
All schools	1.98	2.99	3.50	4.64	-2.808 (179.4)	0.006	0.39
Lower SES schools	2.23	3.41	3.82	4.83	-1.816 (42.3)	0.076	0.43
Higher SES schools	1.86	2.78	2.80	4.16	-1.250 (77.2)	0.215	0.27

Nevertheless, the monolingual and bilingual children were mostly found to be close in terms of language proficiency. As Michael and Gollan (2005) rightly point out, the lack of significant differences could also be due to insufficiently sensitive tasks. Especially given that the majority of the participants grew up speaking both Dutch and another language at home, and that all participants received Dutch formal schooling and have become literate in Dutch, differences may be expected to be subtle at any rate, and may indeed be difficult to detect. The use of highly frequent stimuli in the tasks that were specifically designed for this dissertation could mean that they were less sensitive to such small differences, certainly when we take into account that bilinguals' disadvantages in vocabulary knowledge may be largely due to their comparatively reduced exposure in the two languages (Gollan et al., 2008; Michael & Gollan, 2005). More research with a more diverse range of stimuli is undoubtedly needed if we wish to map differences in semantic knowledge in similar participant groups in detail. However, the consistency between the three studies, including performance on standardized tasks, is still considerable: bilinguals' performance was mostly found to be on par with that of monolinguals.

5.2. Semantic knowledge and reading comprehension

Although studies targeting vocabulary size, declarative knowledge of context-independent relations and other aspects of vocabulary depth consistently find a positive relation with reading comprehension, evidence for contributions of the semantic knowledge aspects targeted in this dissertation was limited. No significant contributions of semantic access or semantic priming were found, but preferences for categorical and feature associations were found to be related to comprehension.

5.2.1. Semantic organization

Although it is not possible to measure proficiency directly based on a spontaneous and unconstrained task such as the free word association task (Cremer et al., 2011; Fitzpatrick, 2013), the task does provide valuable information on preferred or easily accessible routes in an individual's semantic network. Considering the previously reported relevance of declarative knowledge of semantic relations and activation of those relations (Betjemann & Keenan, 2008; Bonnotte & Casalis, 2010; Cremer & Schoonen, 2013; Nation & Snowling, 1999) and even specifically of the adult-likeness of associative behavior (Swart et al., 2017a), the semantic characteristics of an individual's word associations were expected to also be related to reading comprehension performance.

A modest positive effect of taxonomic association scores showed that this expectation was indeed borne out by the data. Children who provided more taxonomic associations, also obtained higher reading comprehension scores. Both children providing more taxonomic associations as first responses and as later responses scored higher on reading comprehension, but especially children providing taxonomic associations as second and third responses were more proficient readers. Conversely, children providing more feature associations as initial responses performed worse on the reading comprehension test, while later feature responses were not related to comprehension performance. Both taxonomic and feature relations are context-independent semantic relations, and tasks testing knowledge of both relations have been found to be positively related to reading comprehension (Cremer & Schoonen, 2013; Tannenbaum et al., 2006). The positive effect of taxonomic associations was therefore expected, but a negative effect of feature associations runs counter to previous results from declarative tasks.

As discussed above, the different association patterns obtained for features in the adult study by De Deyne and Storms (2008a) may be a sign that feature associations are more similar to taxonomic associations for the children in our study, while they behaved more like the situational and subjective associations in adults. Especially context-independent semantic relations are likely to still be in development in this age group, and for children for whom taxonomic associations are not yet as readily available, feature associations may take on a similar function. If this is indeed the case and especially children who are able to produce taxonomic associations are better readers, then children who show a preference for feature associations may perform worse at reading comprehension because their semantic network structure is at a less advanced stage in development. It could be argued that the results obtained by Swart and colleagues (2017a), who scored their child participants' associations based on adult norms and found that they contributed to reading comprehension, provide evidence to support this claim. Children who provided responses that were often provided by adults to the same stimuli performed better on comprehension than those that provided associations that were never or less frequently provided in adult norms. The children in Chapter 2 who performed similar to adults (providing more taxonomic associations), scored higher on reading comprehension, and those that did not (provided more features as initial responses) performed worse.

Nevertheless, the discrepancy between the negative relation of reading comprehension with feature associations and the positive relation found in studies targeting defining features by using definition tasks or selection tasks remains (e.g. Cremer & Schoonen, 2013; Tannenbaum et al., 2006). These differences may be due to the unconstrained nature of the free word association task. Association responses represent preferences for – or more easily accessible routes to – certain types of semantic relations, which do not necessarily correspond directly to knowledge that would be available when elicited directly, as was also pointed out by Cremer et al. (2011).

5.2.2. Access to semantic information

The Lexical Quality Hypothesis (Perfetti & Hart, 2002; Perfetti, 2007) predicts that lexical representations need to be high in quality in order for reading comprehension to be successful. The quality of lexical representations is not only concerned with their semantic, orthographic and phonological contents, but also with fast access to this information. Indeed, access to semantic information as measured by semantic categorization tasks specifically was found to be positively related to reading comprehension by Richter, Isberner, Naumann and Neeb (2013) and also Cremer (2013). Additionally, Cremer found a small contribution to reading comprehension of speed of access to context-independent semantic information in a forced-choice task, but no significant contribution of lexical decision speed. These findings suggest that access to semantic information contributes to reading comprehension, but this result was not replicated in the study described in Chapter 3.

The speed of semantic categorization in the auditory task in Chapter 3 did not bear a significant relation to individual differences in the reading comprehension scores. The effects in the studies by Cremer (2013) and Richter and colleagues (2013) were comparatively small: 2-2.5% additional variance could be explained by the semantic access measures after other vocabulary measures were taken into account. Given that these effects, although significant, are quite small, the likelihood of their presence remaining undetected in some studies is larger. Additionally however, both Richter and colleagues and Cremer required semantic categorization of written stimuli; Chapter 3 aimed to target a possible relation of semantic access to reading comprehension independent of the written modality, by using auditory stimuli. If only semantic access to visual stimuli is related to reading comprehension, this suggests that some type of decoding component may be involved in this process, which is important for its relation to reading comprehension. However, further studies investigating access to semantic information from auditory stimuli would be necessary to support this hypothesis.

5.2.3. Single-word and sentence-level semantic priming

Previous research identified group differences in semantic priming between poor and normal readers (Betjemann & Keenan, 2008; Bonnotte & Casalis, 2010; Nation & Snowling, 1999). In Nation and Snowling and Bonnotte and Casalis' work, normal readers showed more non-associated, context-independent priming than poor readers, and additionally Bonnotte and Casalis found reduced functional priming in normal readers compared to poor readers. The apparent advantage of well-developed context-independent knowledge for reading comprehension in this online task is in line with evidence from various offline tasks targeting this type of semantic knowledge in both group-level and individual differences studies (Cremer & Schoonen, 2013; Nation & Snowling, 1998; Nouwens, Groen, & Verhoeven, 2017; Ouellette, 2006; Swart et al., 2017a; Tannenbaum et al., 2006). A remaining question however, is whether these group differences in sensitivity to semantic priming also hold on an individual differences level within a group of normal readers: can an individual's degree of semantic priming predict his or her reading comprehension? Cremer (2013) found no such advantages in reading comprehension for children showing more context-independent semantic priming in an individual differences design with normal readers.

Firstly, the individual context-independent priming scores from the single-word priming study were not related to individual differences in reading comprehension in the normal readers that participated in the study in Chapter 3. This was thus in contrast with the expectations based on group differences between normal readers and poor readers obtained in previous studies (Betjemann & Keenan, 2008; Bonnotte & Casalis, 2010; Nation & Snowling, 1999), but in accordance with the individual differences data reported by Cremer (2013).

Based on these results, it could still be argued that it is necessary to examine priming during reading rather than in a single-word paradigm, because a likely route for semantic priming to exert an influence on comprehension is by helping to establish coherence during the reading comprehension process. Such a route is posited in the Landscape model of reading comprehension for example (Van den Broek et al., 1999; Van den Broek et al., 2005), which suggests that semantic relations are activated during reading and integrated into the representation of a text, which is built on linkage between words and concepts. Therefore, although single-word priming may not be sensitive enough to predict comprehension in normal readers, priming during sentence reading may be more informative since it can approximate this process more closely. However, this hypothesis was not borne out by the data from the self-paced reading study: although delayed facilitation effects were found at the group level for non-associated word pairs in general and for non-associated, context-independent pairs specifically, neither of these priming effects bore a significant relation to the reading comprehension scores on the individual level.

The non-associated, context-independent semantic priming observed at the group level is consistent between Chapters 3 and 4, with Cremer (2013) and with the normal reader groups from Nation and Snowling (1999) and Bonnotte and Casalis (2010): they also showed a group effect for these pairs, which is what set them apart from poor readers. The absence of a significant contribution to the reading scores, in accordance with Cremer's results, suggests that the group differences in sensitivity to non-associated context-independent priming simply do not hold in an individual differences context targeting normal readers. There is a counterexample to this: Larkin, Woltz, Reynolds, and Clark (1996) did manage to find a significant relation between individual priming scores of synonymy and reading comprehension. However, the results presented in this dissertation do represent a fairly close replication of the way the non-associated, context-independent pairs were tested in both of the group studies that served as examples (Bonnotte & Casalis, 2010; Nation & Snowling, 1999). Therefore, despite the evidence from those group studies, this type of priming appears to exert a limited contribution to comprehension on the individual level.

A few possible explanations for the incongruity between the results from group studies and the individual differences approach taken in this dissertation can be put forward. Firstly and most straightforwardly, even if group differences between normal and poor comprehenders exist, these may not translate to an effect in individual differences in normal readers. Differences between individuals in a more homogenous group of normal readers may be too subtle, for example because a certain baseline degree of priming necessary for comprehension at this age has already been reached.

Secondly, the gap between priming and comprehension as tested in this dissertation may be too large: small effects from a speeded measure targeting very subtle differences in processing are compared to comprehension scores derived from texts with different vocabulary and ample time to finish the task. In a group study, the small priming effects are pooled and may therefore show a connection to comprehension more easily, especially when normal comprehenders are juxtaposed to readers who have an explicit delay. On an individual differences level, the connection between priming and a comparatively untimed comprehension task may be a step too far, but if the priming tasks and the comprehension tasks were to pose more similar demands, results may be different.

Thirdly and finally, semantic priming effects are highly sensitive to a host of factors (e.g. McNamara, 2005), and there were subtle methodological differences between the priming studies in this dissertation and the studies by Nation and Snowling (1999), Bonnotte and Casalis (2010) and Cremer (2013). These differences have likely affected the resulting priming effects, which may in turn affect the connection to comprehension. And although we might expect a connection between priming and comprehension to turn up even under slightly different experimental circumstances if it were strong enough, there is evidence showing that priming effects on the individual level are inherently unstable. The generally robust group-level priming effects that are obtained in many studies, do not translate to an equally robust effect at the individual level: priming effects in individuals are inconsistent, and only group studies may be able to filter out the noise associated with this (J. Stolz et al., 2005; Yap et al., 2016). This could mean that if a relation between semantic priming and reading comprehension does exist at the individual level, classic semantic priming experiments may not allow us to reliably establish individual priming scores that could be used to study this connection.

5.2.4. Semantic knowledge and reading comprehension: conclusions

As was argued in Chapter 1, teasing apart the roles of vocabulary components for reading comprehension is an important research goal if we wish to support young readers in their acquisition of this indispensable skill. Many studies show that when it comes to predicting reading comprehension, there is more to vocabulary knowledge than size alone (e.g. Raudszus et al., 2018; Swart et al., 2017a; Tannenbaum et al., 2006; Verhoeven & Van Leeuwe, 2008), and researchers have also started to investigate how different vocabulary components affect different reading comprehension subprocesses (Cain & Oakhill, 2014). Understanding how individual differences in these vocabulary components are related to individual differences in reading comprehension, is an important starting point for potential intervention studies.

The joint discussion of the three studies appears to paint a bleak picture for the vocabulary components that were targeted in this dissertation: the initial hypotheses concerning the positive relation of semantic access and priming to reading comprehension had to be rejected. As pointed out by Richter and colleagues (2013) regarding the role of semantic access, even if a small contribution is present, this is likely to be far more limited than the contribution of knowledge of semantic representations themselves. This may be a reason why no effect of semantic access was found in Chapter 3.

The two semantic priming studies, in conjunction with the results obtained by Cremer (2013), suggest that the role of priming is also likely to be modest at best. The findings appear to indicate that, despite the initially promising results from group studies, automatic activation of semantic relations may not be the most fruitful vocabulary component for future individual differences research or intervention studies to target. Either an individual's sensitivity to semantic priming simply has little correspondence to reading comprehension performance, or relating this relatively unstable measure to a complex skill such as reading comprehension may not be feasible on an individual level.

However, the relevance of context-independent knowledge for reading comprehension was confirmed in this dissertation, by the contributions of both the spontaneous association preferences from Chapter 2 and the more formally measured scores from the WAT in Chapter 4. Due to the correlational nature of the studies, it is difficult to make hard claims on the directionality of these effects. They are likely to be at least bidirectional, as has been shown for vocabulary size measures (e.g. Swart, Muijselaar, Steenbeek-Planting, Droop, De Jong, & Verhoeven, 2017b; Verhoeven, Van Leeuwe, & Vermeer, 2011). For association behavior specifically, we know that association norms are related to collocational data from texts (e.g. Wettler, Rapp, & Sedlmeier, 2005), meaning that associative preferences are likely related to the amount of reading an individual engages in, which in turn of course affects reading comprehension performance. A meta-analysis on interventions targeting reading comprehension has concluded that improving word knowledge helps to improve comprehension (Edmonds et al., 2009), suggesting that word knowledge is indeed predictive of comprehension. However, studies targeting knowledge of semantic relations specifically did not find that gains in this type of semantic knowledge translated to gains in reading comprehension (Proctor et al., 2011; Verhallen, Schoonen, & Appel, 2001). Nevertheless, given the hypothesized importance of many aspects of vocabulary knowledge for comprehension (Perfetti & Hart, 2002; Perfetti, 2007; Perfetti & Stafura, 2014), it is likely that there is indeed a contribution to comprehension from the semantic components under study in this dissertation, but longitudinal studies or intervention research would be necessary to confirm this proposition.

5.3. Strengths, limitations, and suggestions for future research

The three studies presented in this dissertation all have their strengths and limitations affecting the possible reach of the conclusions, and most of these strengths and limitations are opposite sides of the same coin. The choice to test children in the school setting likely affected the measurement potential of especially the priming experiments: given the small scale, i.e. tens of milliseconds, and noisiness of priming effects, testing in a highly controlled lab setting is preferred. However, this choice also meant that larger participant groups could be recruited. Especially in a population that is typically difficult to reach, this approach was deemed necessary to obtain more robust results in the individual differences analyses.

Furthermore, in all three studies, a quite specific set of stimuli was used: all highly frequent, concrete nouns. By selecting frequent items, the potential for finding semantic priming was maximized and a confound with vocabulary size was avoided. As a result however, it may be the case that activation of these familiar words is already so well-entrenched that differences between quite proficient bilinguals and their monolingual peers are hard to find, and that it is no longer predictive of differences in reading comprehension at fifth-grade levels[26]. Finally, the numbers of critical stimuli in each of the conditions in the two priming experiments were relatively limited, because they were controlled quite stringently. Word association data were gathered from the target population, the relatedness of the non-associated word pairs was checked through an additional questionnaire and the sentences in the self-paced reading experiment were designed such that related and unrelated primes could be tested in the same context. These precautions strengthened the quality of the stimuli, but limited their number and make generalization to the rest of the lexicon difficult.

In all, the abovementioned limitations suggest an obvious first step for future research: replication and extension using more and more diverse stimuli, preferably with similarly large participant groups but in more controlled settings. Additionally, as suggested by Swart and colleagues (2017a), the use of multiple measures of reading comprehension might reveal more connections with semantic knowledge, especially if certain aspects of vocabulary knowledge are differently related to subprocesses of comprehension (Cain & Oakhill, 2014). Furthermore, other methodologies monitoring processing during reading could be employed, such as eye tracking, which allows for the monitoring of more aspects of processing than simple reading times per word. There is evidence to suggest that priming of associatively related words especially affects early measures such as skipping rates and first-pass reading (Camblin et al., 2007), which can be measured using this technique.

5.4. Concluding remarks

The overarching goals of this dissertation were to map the semantic networks of Dutch monolingual and bilingual minority children and to assess the relation of these networks to reading comprehension. Both language groups were found to be well on their way in the development of semantic relations as measured by associative behavior, and consistently showed priming for context-independent relations in both auditory single-

[26] Although frequency information was not reported explicitly in the group studies that served as examples, Nation and Snowling (1999) did use stimuli that were rated to be highly familiar to children in the target age. Frequency can therefore not be a full explanation for the discrepancy with their study.

word and sentence-reading semantic priming experiments. Although the absence of differences between the monolingual and bilingual children on the experimental tasks may also be related to the type of stimuli used, the consistency across the three studies, including on most of the standardized tasks, is noteworthy. The bilingual minority children, who predominantly spoke both Dutch and their L1 at home, were mostly able to keep up with their monolingual class mates. Especially in light of the often reported language delays in bilinguals, this is a positive nuance indicating that bilingual minority children are not unequivocally lagging behind in proficiency because of their language background.

Finally, contributions of the variables of interest to reading comprehension scores were limited. The connections that were found between associative preferences and comprehension, confirm that the organization of children's semantic networks is related to their comprehension skill, and provide interesting questions for future research. Conversely, the priming studies suggest that especially automatic activation of semantic information and semantic relations within the semantic network may not be clear contributors to reading comprehension, even though further research would be necessary to confirm the findings with more varied stimuli and different comprehension tasks. The results do support the overall importance of vocabulary size and knowledge of context-independent relations for reading comprehension.

References

Aarnoutse, C., & Kapinga, T. (2006). *Begrijpend lezen 678 [Reading comprehension grades 456]*. Ridderkerk: 678 Onderwijsadvisering.

Aitchison, J. (2012). *Words in the mind: An introduction to the mental lexicon* (4th ed.). Malden, MA: Wiley-Blackwell.

Alderson, J. C. (2005). *Assessing reading*. Cambridge: Cambridge University Press.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255-278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74*(4), 506-521.

Berkel, S. v., Schoot, F. v. d., Engelen, R., & Maris, G. (2002). *Balans van het taalonderwijs halverwege de basisschool 3 [Balance of language education halfway through elementary school 3]* (No. 20, PPON series). Arnhem: Cito.

Betjemann, R. S., & Keenan, J. M. (2008). Phonological and semantic priming in children with reading disability. *Child Development, 79*(4), 1086-1102.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3-11.

Biemiller, A. (2006). Vocabulary development and instruction: A prerequisite for school learning. In S. Neuman, & D. Dickinson (Eds.), *Handbook of early literacy research (Vol 2)* (pp. 41-51). New York: Guilford Press.

Blewitt, P., & Toppino, T. C. (1991). The development of taxonomic structure in lexical memory. *Journal of Experimental Child Psychology, 51*(2), 296-319.

Bock, K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984). New York: Academic Press.

Bonnotte, I., & Casalis, S. (2010). Semantic priming in French children with varying comprehension skills. *European Journal of Developmental Psychology, 7*(3), 309-328.

Borghi, A. M., & Caramelli, N. (2003). Situation bounded conceptual organization in children: From action to spatial relations. *Cognitive Development, 18*(1), 49-60.

Cain, K., & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'Année Psychologique, 114*(4), 647-662.

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory and Language, 56*(1), 103-128.

Carroll, P., & Slowiaczek, M. L. (1986). Constraints on semantic priming in reading: A fixation time analysis. *Memory & Cognition, 14*(6), 509-522.

CBS. (2014). *Jaarrapport integratie 2014 [Year report integration 2014]*. Den Haag/Heerlen: Centraal Bureau voor de Statistiek.

CBS. (2016). *Jaarrapport integratie 2016 [Year report integration 2016]*. Den Haag/Heerlen/Bonaire: Centraal Bureau voor de Statistiek.

Cohen, J. (1969). *Statistical power analysis for the behavioural sciences*. New York: Academic Press.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407-428.

Cremer, M. (2013). *Accessing word meaning: Semantic word knowledge and reading comprehension in Dutch monolingual and bilingual fifth-graders (Doctoral dissertation University of Amsterdam)*. Utrecht: LOT.

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.

Cremer, M., & Schoonen, R. (2013). The role of accessibility of semantic word knowledge in monolingual and bilingual fifth-grade reading. *Applied Psycholinguistics, 34*(6), 1195-1217.

De Deyne, S., Navarro, D. J., & Storms, G. (2013). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods, 45*, 480-498.

De Deyne, S., & Storms, G. (2008a). Word associations: Network and semantic properties. *Behavior Research Methods, 40*(1), 213-231.

De Deyne, S., & Storms, G. (2008b). Word associations: Norms for 1,424 dutch words in a continuous task. *Behavior Research Methods, 40*(1), 198-205.

De Deyne, S., & Verheyen, S. (2015). Using network clustering to uncover the taxonomic and thematic structure of the mental lexicon. Paper presented at the *Word Structure and Word Usage. Proceedings of the NetWordS Final Conference,* Pisa. 172-176. Retrieved from http://ceur-ws.org.

Denckla, M. B., & Rudel, R. (1974). Rapid “automatized” naming of pictured objects, colors, letters and numbers by normal children. *Cortex, 10*(2), 186-202.

Di Filippo, G., Brizzolara, D., Chilosi, A., De Luca, M., Judica, A., Pecini, C., . . . Zoccolotti, P. (2005). Rapid naming, not cancellation speed or articulation rate, predicts reading in an orthographically regular language (Italian). *Child Neuropsychology, 11*(4), 349-361.

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly, 38*(1), 78-103.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research, 79*(1), 262-300.

Elbers, L., Van Loon-Vervoorn, A., & Van Helden-Lankhaar, M. (1993). 'Contrastive usage' and the development of lexical organization and innovative labelling. In M. Verrips, & F. Wijnen (Eds.), *Amsterdam series in child language development 1: The acquisition of Dutch.* (pp. 5-19). Amsterdam: University of Amsterdam.

Erdos, C., Genesee, F., Savage, R., & Haigh, C. A. (2011). Individual differences in second language reading outcomes. *International Journal of Bilingualism, 15*(1), 3-25.

Estes, Z., Golonka, S., & Jones, L. L. (2011). Thematic thinking: The apprehension and consequences of thematic relations. *Psychology of Learning and Motivation, 54*, 249-294.

Fitzpatrick, T. (2006). Habits and rabbits: Word associations and the L2 lexicon. *EuroSLA Yearbook, 6*, 121-145.

Fitzpatrick, T. (2007). Word association patterns: Unpacking the assumptions. *International Journal of Applied Linguistics, 17*(3), 319-331.

Fitzpatrick, T. (2009). Word association profiles in a first and second language: Puzzles and problems. In T. Fitzpatrick, & A. Barfield (Eds.), *Lexical processing in second language learners* (pp. 38-52). Bristol, Buffalo, Toronto: Multilingual Matters.

Fitzpatrick, T. (2013). Word associations. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 6193-6199). Malden, MA: Wiley-Blackwell.

Fitzpatrick, T., & Izura, C. (2011). Word association in L1 and L2. *Studies in Second Language Acquisition, 33*(3), 373-398.

Fitzpatrick, T., Playfoot, D., Wray, A., & Wright, M. J. (2013). Establishing the reliability of word association data for investigating individual and group differences. *Applied Linguistics, 36*(1), 23-50.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58*(3), 787-814.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition, 33*(7), 1220-1234.

Gollan, T. H., & Silverberg, N. B. (2001). Tip-of-the-tongue states in Hebrew–English bilinguals. *Bilingualism: Language and Cognition, 4*(1), 63-83.

Goodrich, J. M., & Lonigan, C. J. (2018). Language-minority children's sensitivity to the semantic relations between words. *Journal of Experimental Child Psychology, 167*, 259-277.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.

Hare, M., Jones, M., Thomson, C., Kelly, S., & McRae, K. (2009). Activating event knowledge. *Cognition, 111*(2), 151-167.

Heesters, K., Berkel, S. v., Schoot, F. v. d., & Hemker, B. (2007). *Balans van het leesonderwijs aan het einde van de basisschool 4 [Balance of reading education at the end of elementary school 4]* (No. 33 PPON series). Arnhem: Cito.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition, 21*(2), 303-317.

Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General, 124*(1), 62-82.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language, 39*, 1-27.

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Processes, 26*(3), 376-405.

Jones, L. L., & Golonka, S. (2012). Different influences on lexical priming for integrative, thematic, and taxonomic relations. *Frontiers in Human Neuroscience, 6*, 1-17.

Kiss, G. R., Armstrong, C., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. In A. Aitken, R. Bailey & N. Hamilton-Smith (Eds.), *The computer and literacy studies* (pp. 153-165). Edinburgh: University Press.

Kliegl, R., Masson, M. E., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition, 18*(5), 655-681.

Kotz, S. A., & Elston-Güttler, K. (2004). The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related brain potentials. *Journal of Neurolinguistics*, 17(2-3), 215-235.

Kuhlemeier, H., Jolink, A., Krämer, I., Hemker, B., Jongen, I., Van Berkel, S., & Bechger, T. (2014). *Balans van de leesvaardigheid in het basis- en speciaal basisonderwijs 2 [Balance of reading skill in elementary and special education 2].* (No. 54, PPON series). Arnhem: Cito.

Landi, N., & Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language, 102*(1), 30-45.

Larkin, A. A., Woltz, D. J., Reynolds, R. E., & Clark, E. (1996). Conceptual priming differences and reading ability. *Contemporary Educational Psychology, 21*, 279-303.

Lavigne, F., Vitu, F., & d'Ydewalle, G. (2000). The influence of semantic context on initial eye landing sites in words. *Acta Psychologica, 104*(2), 191-214.

Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General, 130*(1), 3-28.

Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning, 65*(S1), 185-207.

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7*(4), 618-630.

Mancilla-Martinez, J., & Vagh, S. B. (2013). Growth in toddlers' spanish, english, and conceptual vocabulary knowledge. *Early Childhood Research Quarterly, 28*(3), 555-567.

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology, 16*(1), 1-27.

McEvoy, C. L., & Nelson, D. L. (1982). Category name and instance norms for 106 categories of various sizes. *American Journal of Psychology, 95*, 581-634.

McNamara, T. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York and Hove: Psychology Press.

McNamara, T., & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language, 27*, 545-559.

McRae, K., & Cree, G. S. (2002). Factors underlying category-specific semantic deficits. In E. M. E. Forde, & G. Humphreys (Eds.), *Category-specificity in mind and brain* (pp. 211-249). East Sussex, UK: Psychology Press.

McRae, K., Khalkhali, S., Hare, M., & McRae, K. (2011). Semantic and associative relations in adolescents and young adults: Examining a tenuous dichotomy. *The Adolescent Brain: Learning, Reasoning, and Decision Making, 18*, 39-66.

Meara, P., & Wolter, B. (2004). V_Links: Beyond vocabulary depth. *Angles on the English Speaking World, 4*, 85-96.

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological Bulletin, 140*(2), 409-433.

Michael, E. B., & Gollan, T. H. (2005). Being and becoming bilingual. In J. R. Kroll, & A. M. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389-407). New York: Oxford University Press.

Mirman, D., & Graziano, K. M. (2011). Individual differences in the strength of taxonomic versus thematic relations. *Journal of Experimental Psychology: General, 141*(4), 601-609.

Mirman, D., Landrigan, J., & Britt, A. E. (2017). Taxonomic and thematic semantic systems. *Psychological Bulletin, 143*(5), 499.

Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning Memory and Cognition, 20*(1), 92-103.

Myers, J. L., Cook, A. E., Kambe, G., Mason, R. A., & O'Brien, E. J. (2000). Semantic and episodic effects on bridging inferences. *Discourse Processes, 29*(3), 179-199.

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39*(1), 85-101.

Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition, 70*, B1-B13.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*(4), 342-356.

Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure? *Memory & Cognition, 28*(6), 887-899.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers, 36*(3), 402-407.

Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: A review of research and theory. *Psychological Bulletin, 84*(1), 93-116.

Nelson, K. (1982). The syntagmatics and paradigmatics of conceptual development. In S. Kuczaj (Ed.), *Language development volume 2: Language, thought and culture.* (pp. 335-364). Hillsdale, NJ: Erlbaum.

Nelson, K. (1985). *Making sense: The acquisition of shared meaning*. New York: Academic Press.

Nelson, K. (1991). Concepts of meaning in language development. In N. A. Krasnegor, D. M. Rumbaugh, R. L. Schiefelbusch & M. Studdert-Kennedy (Eds.), *Biological and behavioral determinants of language development* (pp. 89-115). Hillsdale, NJ: Erlbaum.

Nelson, K. (2007). *Young minds in social worlds: Experience, meaning, and memory*. Cambridge, MA/London, UK: Harvard University Press.

Nobre, A. P., & Salles, J. F. (2016). Lexical-semantic processing and reading: Relations between semantic priming, visual word recognition and reading comprehension. *Educational Psychology, 36*(4), 753-770.

Nouwens, S., Groen, M. A., & Verhoeven, L. (2017). How working memory relates to children's reading comprehension: The importance of domain-specificity in storage and processing. *Reading and Writing, 30*(1), 105-120.

Olkkonen, S. (2013). Speed in cognitive tasks as an indicator of second/foreign language reading and writing skills. *Eesti Rakenduslingvistika Ühingu Aastaraamat, 9*, 195-208.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554-566.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189-213). Amsterdam/Philadelphia: John Benjamins.

Perfetti, C. A., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*(1), 22-37.

Petrey, S. (1977). Word associations and the development of lexical memory. *Cognition, 5*, 57-71.

Proctor, C. P., Dalton, B., Uccelli, P., Biancarosa, G., Mo, E., Snow, C., & Neugebauer, S. (2011). Improving comprehension online: Effects of deep vocabulary instruction with bilingual and monolingual fifth graders. *Reading and Writing, 24*(5), 517-544.

Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review/La Revue Canadienne Des Langues Vivantes, 56*(2), 282-307.

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning, 52*(3), 513-536.

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

R Core Team. (2016). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

R Core Team. (2017). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Raudszus, H., Segers, E., & Verhoeven, L. (2018). Lexical quality and executive control predict children's first and second language reading comprehension. *Reading and Writing, 31*(2), 405-424.

Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined. In P. Bogaards, & B. Laufer (Eds.), *Vocabulary in a second language* (pp. 209-227). Amsterdam/ Philadelphia: John Benjamins.

Richter, T., Isberner, M., Naumann, J., & Neeb, Y. (2013). Lexical quality and reading comprehension in primary school children. *Scientific Studies of Reading, 17*(6), 415-434.

Scarborough, H. S. (1998). Predicting the future achievement of second graders with reading disabilities: Contributions of phonemic awareness, verbal memory, rapid naming, and IQ. *Annals of Dyslexia, 48*, 115-136.

Schmitt, N. (1998). Quantifying word association responses: What is native-like? *System, 26*(3), 389-401.

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning, 64*(4), 913-951.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Incorporated.

Schoonen, R., & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing, 25*(2), 211-236.

Schrooten, W., & Vermeer, A. (1994). *Woorden in het basisonderwijs. 15.000 woorden aangeboden aan leerlingen* [Words in elementary education. 15,000 words presented to pupils] Tilburg: Tilburg University Press.

Schwartz, M., & Katzir, T. (2012). Depth of lexical knowledge among bilingual children: The impact of schooling. *Reading and Writing, 25*(8), 1947-1971.

SCP. (2015). *SCP statusscores [SCP SES scores]*. Retrieved from http://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.

Sijtstra, J. M., Schoot, F. v. d., & Hemker, B. T. (2002). *Balans van het taalonderwijs aan het einde van de basisschool 3 [Balance of language education at the end of elementary school 3]* (No. 19, PPON series). Arnhem: Cito.

Smits, D., & Aarnoutse, C. (1997). Een longitudinaal onderzoek naar verschillen in taal- en leesprestaties van autochtone en allochtone kinderen. [A longitudinal study on the differences in language and learning achievement in autochtonous and minority children] *Nederlands Tijdschrift Voor Opvoeding, Vorming En Onderwijs, 13*(1), 33-52.

Söderman, T. (1993). Word associations of foreign language learners and native speakers – different response types and their relevance to lexical development. In B. Hammarberg (Ed.), *Problem, process, product in language learning: Papers from the Stockholm-Abo conference* (pp. 157-169). Stockholm: Stockholm University.

Spätgens, T., & Schoonen, R. (2018). The semantic network, lexical access and reading comprehension in monolingual and bilingual children: An individual differences study. *Applied Psycholinguistics, 39*(1), 225-256.

Spätgens, T., & Schoonen, R. (submitted). *The structure of developing semantic networks: Evidence from single and multiple word associations in monolingual and bilingual readers.* Unpublished manuscript.

Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal, 36*(2), 139-152.

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York, etc.: Guilford Press.

Stolz, J., Besner, D., & Carr, T. (2005). Implications of measures of reliability for theories of priming: Activity in semantic memory is inherently noisy and uncoordinated. *Visual Cognition, 12*(2), 284-336.

Stolz, W. S., & Tiffany, J. (1972). The production of “child-like” word associations by adults to unfamiliar adjectives. *Journal of Verbal Learning and Verbal Behavior, 11*(1), 38-46.

Swanson, H. L., Trainin, G., Necoechea, D. M., & Hammill, D. D. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlation evidence. *Review of Educational Research, 73*(4), 407-440.

Swart, N. M., Muijselaar, M. M., Steenbeek-Planting, E. G., Droop, M., De Jong, P. F., & Verhoeven, L. (2017a). Differential lexical predictors of reading comprehension in fourth graders. *Reading and Writing, 30*(3), 489-507.

Swart, N. M., Muijselaar, M. M., Steenbeek-Planting, E. G., Droop, M., De Jong, P. F., & Verhoeven, L. (2017b). Cognitive precursors of the developmental relation between lexical quality and reading comprehension in the intermediate elementary grades. *Learning and Individual Differences, 59*, 43-54.

Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading, 10*(4), 381-398.

Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading, 32*(4), 383-401.

Traxler, M. J., Foss, D. J., Seely, R. E., Kaup, B., & Morris, R. K. (2000). Priming in sentence processing: Intralexical spreading activation, schemas, and situation models. *Journal of Psycholinguistic Research, 29*(6), 581-595.

Van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes, 39*(2-3), 299-316.

Van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the online construction of a memory representation. In H. van Oostendorp, & S. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71-98). Mahwah, New Jersey: Lawrence Erlbaum.

Verhallen, M., & Schoonen, R. (1993). Lexical knowledge of monolingual and bilingual children. *Applied Linguistics, 14*(4), 344-363.

Verhallen, M., Schoonen, R., & Appel, R. (2001). Verdiepen van woordkennis: Een empirische studie naar de effecten van een trainingsprogramma [Deepening word knowledge: An empirical study into the effects of a training program]. *Pedagogische Studiën, 78*(4), 239-255.

Verhallen, M., & Schoonen, R. (1998). Lexical knowledge in L1 and L2 of third and fifth graders. *Applied Linguistics, 19*(4), 452-470.

Verhoeven, L. (1992). *Drie-minuten-toets: Handleiding.* [Three-Minute-Test]. Arnhem: Cito.

Verhoeven, L., Van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading, 15*(1), 8-25.

Verhoeven, L., & Vermeer, A. (1995). *Leerlingvolgsysteem 5678: Leeswoordenschat* [Monitoring and Evaluation System grades 3456: Reading Vocabulary]. Arnhem: Cito.

Verhoeven, L., & Van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology, 22*(3), 407-423.

Verhoeven, L., & Van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Reading and Writing, 25*(8), 1805-1818.

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics, 22*(2), 217-234.

Wettler, M., Rapp, R., & Sedlmeier, P. (2005). Free word associations correspond to contiguities between words in texts. *Journal of Quantitative Linguistics, 12*(2-3), 111-122.

Wolf, M., Miller, L., & Donnelly, K. (2000). Retrieval, automaticity, vocabulary elaboration, orthography (RAVE-O): A comprehensive, fluency-based reading intervention program. *Journal of Learning Disabilities, 33*(4), 375-386.

Wolf, M. (1986). Rapid alternating stimulus naming in the developmental dyslexias. *Brain and Language, 27*, 360-379.

Wolf, M., & Denckla, M. B. (2005). *The rapid automatized naming and rapid alternating stimulus tests: Examiner's manual*. Austin, Tex.: Pro-ed.

Wolter, B. (2001). Comparing the L1 and L2 mental lexicon. *Studies in Second Language Acquisition, 23*(1), 41-69.

Wolter, B. (2006). Lexical network structures and L2 vocabulary acquisition: The role of L1 lexical/conceptual knowledge. *Applied Linguistics, 27*(4), 741-747.

Wu, L., & Barsalou, L. W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica, 132*(2), 173-189.

Yap, M. J., Hutchison, K., & Tan, L. C. (2016). Individual differences in semantic priming performance: Insights from the semantic priming project. In M. N. Jones (Ed.), *Big data in cognitive science* (pp. 203-226). New York: Psychology Press.

Zareva, A. (2007). Structure of the second language mental lexicon: How does it compare to native speakers' lexical organization? *Second Language Research, 23*(2), 123-153.

Zareva, A., & Wolter, B. (2012). The ‘promise’ of three methods of word association analysis to L2 lexical research. *Second Language Research, 28*(1), 41-67.

Appendices

Appendix A
Word association stimuli (Chapter 2)

Appendix B
Critical stimuli in the semantic priming experiment (Chapter 3)

Appendix C
Stimulus characteristics by subsets and halved subsets in the semantic priming experiment (Chapter 3)

Appendix D
Scatter plots with trend line including and excluding bivariate outliers (Chapter 3)

Appendix E
Related word pairs in the self-paced reading experiment (Chapter 4)

Appendix F
Length and frequency information for primes and targets in the self-paced reading experiment (Chapter 4)

Appendix G
Test sentences in the self-paced reading experiment (Chapter 4)

Appendix A: Word association stimuli (Chapter 2)

Table A.1

Stimulus words and translation equivalents word association task

List 1		List 2		List 3		List 4	
bruid	*bride*	uil	*owl*	karper	*carp*	hamer	*hammer*
hert	*deer*	dief	*thief*	kaas	*cheese*	markt	*market*
duin	*dune*	woestijn	*desert*	pop	*doll*	olifant	*elephant*
rok	*skirt*	sloot	*ditch*	vlinder	*butterfly*	camping	*camp site*
krokodil	*crocodile*	eiland	*island*	wieg	*cot*	trui	*sweater*
kapper	*hair dresser*	potlood	*pencil*	visser	*fisherman*	schilder	*painter*
kasteel	*castle*	appel	*apple*	jas	*coat*	taxi	*taxi*
eend	*duck*	jager	*hunter*	keuken	*kitchen*	kat	*cat*
winkel	*shop*	ridder	*knight*	baby	*baby*	kok	*cook*
stoel	*chair*	bakkerij	*bakery*	regen	*rain*	zee	*sea*
gitaar	*guitar*	trompet	*trumpet*	hond	*dog*	gang	*hallway*
hamster	*hamster*	konijn	*rabbit*	piloot	*pilot*	zebra	*zebra*
juf	*teacher*	chocola	*chocolate*	toerist	*tourist*	jam	*jam*
komkommer	*cucumber*	koe	*cow*	tuin	*garden*	station	*station*
oceaan	*ocean*	bos	*forest*	banaan	*banana*	slak	*snail*
kast	*cupboard*	kapitein	*captain*	restaurant	*restaurant*	boer	*farmer*
cola	*coke*	heks	*witch*	stal	*stable*	prinses	*princess*
dokter	*doctor*	fiets	*bike*	kanarie	*canary*	auto	*car*
chauffeur	*driver*	krekel	*cricket*	koning	*king*	egel	*hedgehog*
leeuw	*lion*	zwaan	*swan*	mus	*sparrow*	indiaan	*indian*

Appendix B: Critical stimuli in the semantic priming experiment (Chapter 3)

Table B.1

Critical stimuli in the semantic priming experiment

Context-independent pairs							
Animal - animal				Object - object			
kat	tijger	*cat*	*tiger*	chocola	drop	*chocolate*	*liquorice*
egel	mol	*hedgehog*	*mole*	trui	broek	*sweater*	*pants*
vlinder	wesp	*butterfly*	*wasp*	gitaar	piano	*guitar*	*piano*
zebra	ezel	*zebra*	*donkey*	auto	vliegtuig	*car*	*air plane*
eend	gans	*duck*	*goose*	rok	bloes	*skirt*	*blouse*
hamster	cavia	*hamster*	*guinea pig*	trompet	fluit	*trumpet*	*flute*
zwaan	meeuw	*swan*	*sea gull*	potlood	viltstift	*pencil*	*felt-pen*
krekel	spin	*cricket*	*spider*	fiets	brommer	*bike*	*moped*
uil	duif	*owl*	*dove*	regen	sneeuw	*rain*	*snow*
hond	wolf	*dog*	*wolf*	kaas	worst	*cheese*	*sausage*
Context-dependent pairs							
Location - animate				Person - inanimate			
bos	eekhoorn	*forest*	*squirrel*	visser	rivier	*fisherman*	*river*
zee	walvis	*sea*	*whale*	kapper	borstel	*hair dresser*	*brush*
bakkerij	klant	*bakery*	*customer*	juf	lokaal	*teacher (female)*	*class room*
woestijn	slang	*desert*	*snake*	baby	melk	*baby*	*milk*
camping	toerist	*camp site*	*tourist*	chauffeur	stuur	*driver*	*steering wheel*
stal	boerin	*stable*	*farmer (female)*	dokter	pleister	*doctor*	*band-aid*
station	conducteur	*train station*	*conductor*	prinses	koets	*princess*	*carriage*
oceaan	dolfijn	*ocean*	*dolphin*	dief	tas	*thief*	*bag*
tuin	mus	*garden*	*sparrow*	kapitein	haven	*captain*	*harbour*
markt	vis	*market*	*fish*	kok	oven	*chef*	*oven*

Appendix C: Stimulus characteristics by subsets and halved subsets in the semantic priming experiment (Chapter 3)

Table C.1

Mean pair relatedness, frequencies and durations for primes and targets by subsets and halved subsets

		All items (N = 10)		Half 1 (N = 5)		Half 2 (N = 5)		Comparison 1 & 2	
		M	*SD*	*M*	*SD*	*M*	*SD*	t (df = 8)	*p*
Animal coordinates	Prime frequency	247.70	325.17	243.20	241.50	252.20	423.70	0.04	0.97
	Target frequency	128.10	89.78	130.00	82.10	126.20	106.71	0.06	0.95
	Prime duration	505.80	99.48	494.40	95.81	517.20	112.97	-0.34	0.74
	Target duration	580.80	95.93	581.20	131.09	580.40	59.36	0.12	0.99
	Relatedness	3.93	0.38	3.97	0.32	3.88	0.46	0.35	0.73
Object coordinates	Prime frequency	247.90	296.82	256.00	368.02	239.80	250.24	0.08	0.94
	Target frequency	126.90	98.73	135.80	94.44	118.00	113.20	0.27	0.79
	Prime duration	582.10	96.50	545.00	119.33	619.20	57.19	-1.25	0.25
	Target duration	629.00	115.28	610.60	122.90	647.40	118.11	-0.48	0.64
	Relatedness	4.31	0.15	4.24	0.18	4.37	0.10	1.37	0.21
Location - animate	Prime frequency	235.20	263.84	267.60	317.90	202.80	230.10	0.37	0.72
	Target frequency	127.40	136.50	118.20	74.57	136.60	190.14	-0.20	0.85
	Prime duration	617.40	108.57	625.20	95.24	609.60	131.52	0.22	0.84
	Target duration	641.70	106.57	629.80	108.29	653.60	116.07	-0.34	0.75
	Relatedness	4.07	0.36	4.05	0.44	4.09	0.30	0.18	0.86
Person - inanimate	Prime frequency	274.70	325.74	305.00	431.51	244.40	224.15	0.28	0.79
	Target frequency	121.00	116.15	132.60	124.83	109.40	120.16	0.30	0.77
	Prime duration	600.80	136.29	553.20	82.43	648.40	171.27	-1.12	0.30
	Target duration	611.90	67.82	615.60	51.97	608.20	87.26	0.16	0.88
	Relatedness	4.15	0.29	4.28	0.20	4.01	0.32	1.58	0.15

Note: frequency data are from Schrooten and Vermeer (1994). Durations of spoken words measured in milliseconds. Relatedness represents mean ratings of semantic relatedness of the pairs on a five-point Likert scale by 33 adult native speakers.

Appendix D: Scatter plots with trend line including and excluding bivariate outliers (Chapter 3)

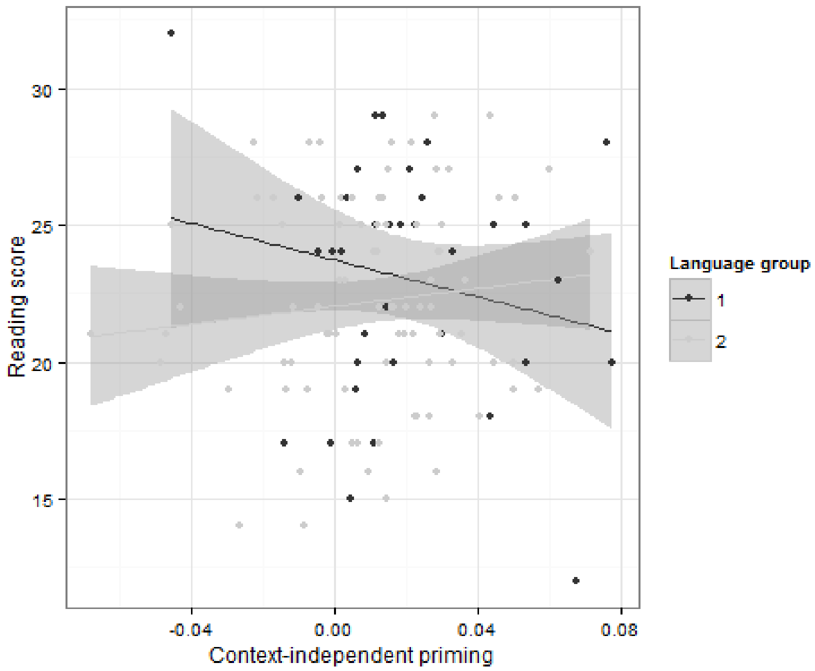

Figure D.1. Scatter plot with trend line and 95% confidence interval (grey area) of context-independent priming and reading scores by language group, including bivariate outliers (N = 122)

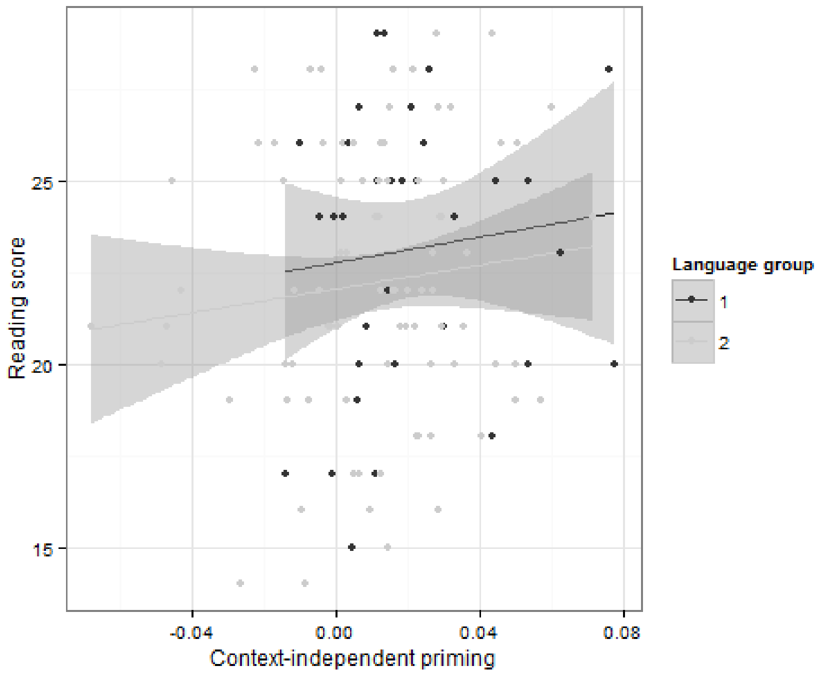

Figure D.2. Scatter plot with trend line and 95% confidence interval (grey area) of context-independent priming and reading scores by language group, excluding bivariate outliers (N = 120)

Appendix E: Related word pairs in the self-paced reading experiment (Chapter 4)

Table E.1

Word pairs including English translation

Context-independent pairs								
Non-associated					Associated			
chocola	drop	*chocolate*	*liquorice*	*	auto	vervoer	*car*	*transport*
eend	gans	*duck*	*goose*	*	banaan	fruit	*banana*	*fruit*
egel	mol	*hedgehog*	*mole*	*	bruid	vrouw	*bride*	*woman*
fiets	brommer	*bike*	*moped*	*	gitaar	instrument	*guitar*	*instrument*
hamer	boor	*hammer*	*drill*		hamster	huisdier	*hamster*	*pet*
hond	wolf	*dog*	*wolf*	*	kaas	melk	*cheese*	*milk*
jam	honing	*marmalade*	*honey*		komkommer	groente	*cucumber*	*vegetable*
jas	muts	*coat*	*cap*		koning	baas	*king*	*boss*
kat	tijger	*cat*	*tiger*	*	krekel	insect	*cricket*	*insect*
regen	sneeuw	*rain*	*snow*	*	oceaan	zee	*ocean*	*sea*
trompet	fluit	*trumpet*	*flute*	*	pop	speelgoed	*doll*	*toy*
trui	broek	*sweater*	*pants*	*	rok	kleding	*skirt*	*clothing*
zebra	ezel	*zebra*	*donkey*	*	uil	vogel	*owl*	*bird*
zwaan	meeuw	*swan*	*sea gull*	*	vlinder	rups	*butterfly*	*caterpillar*

The non-associated pairs with * were also used in the single-word priming study in Chapter 3.

Table E.1 (continued)

Word pairs including English translation

Context-dependent pairs								
Non-associated					Associated			
bakkerij	klant	*bakery*	*customer*	•	baby	luier	*baby*	*diaper*
bos	eekhoorn	*forest*	*squirrel*	•	boer	vee	*farmer*	*cattle*
chauffeur	stuur	*driver*	*steering wheel*	•	camping	tent	*camp site*	*tent*
dief	tas	*thief*	*bag*	•	heks	hoed	*witch*	*hat*
dokter	pleister	*doctor*	*band-aid*	•	indiaan	veren	*indian*	*feathers*
juf	lokaal	*teacher (female)*	*classroom*	•	jager	geweer	*hunter*	*rifle*
kapitein	haven	*captain*	*harbor*	•	kast	spullen	*closet*	*stuff*
kapper	borstel	*hair dresser*	*brush*	•	kasteel	ridder	*castle*	*knight*
kok	oven	*chef*	*oven*	•	markt	kleren	*market*	*clothes*
prinses	koets	*princess*	*carriage*	•	piloot	vliegtuig	*pilot*	*air plane*
station	conducteur	*train station*	*conductor*	•	schilder	verf	*painter*	*paint*
tuin	mus	*garden*	*sparrow*	•	stal	hooi	*stable*	*hay*
visser	rivier	*fisherman*	*river*	•	toerist	vakantie	*tourist*	*holiday*
woestijn	slang	*desert*	*snake*	•	winkel	geld	*store*	*money*

The non-associated pairs with • were also used in the single-word priming study in Chapter 3.

Appendix F: Length and frequency information for primes and targets in the self-paced reading experiment (Chapter 4)

Table F.1

Length (in letters) and frequency information for primes and targets

	Context-independent, non-associated		Context-independent, associated		Context-dependent, non-associated		Context-dependent, associated	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Length related primes	4.6	1.3	5.4	1.8	5.9	2.1	5.6	1.5
Length control primes	5.0	1.5	5.2	1.1	6.1	2.0	5.3	1.4
Frequency related primes	270.9	297.6	173.6	242.6	281.6	325.7	181.8	128.5
Frequency control primes	250.0	292.5	165.8	185.4	272.6	311.1	181.3	131.2
Length targets	4.8	1.1	6.0	2.0	5.7	1.9	5.4	1.7
Frequency targets	143.1	96.2	225.5	255.0	102.4	91.8	191.3	164.4

Frequency data from Schrooten and Vermeer (1994).

Appendix G: Test sentences in the self-paced reading experiment (Chapter 4)

Table G.1

Experimental sentence pairs (prime-target pairs in boldface, related items always followed by unrelated items) by semantic type and plausibility

Context-independent, non-associated Plausible items
Kevin mocht **chocola** en **drop** meenemen toen we op kamp gingen.
Kevin mocht **limonade** en **drop** meenemen toen we op kamp gingen.
Ik kan alleen de **hamer** of de **boor** meenemen want allebei wordt te veel.
Ik kan alleen de **plank** of de **boor** meenemen want allebei wordt te veel.
Ik zag bij die **egel** een **mol** zitten en dat is toch wel bijzonder.
Ik zag bij die **struik** een **mol** zitten en dat is toch wel bijzonder.
Amir speelt **trompet** en **fluit** en kan ook heel goed zingen.
Amir speelt **voetbal** en **fluit** en kan ook heel goed zingen.
Het nieuws ging over de **regen** en de **sneeuw** waar iedereen last van heeft op de weg.
Het nieuws ging over de **drukte** en de **sneeuw** waar iedereen last van heeft op de weg.
Het was een film over een **hond** en een **wolf** die heel goede vrienden werden.
Het was een film over een **meisje** en een **wolf** die heel goede vrienden werden.
Het plaatje met de **eend** en de **gans** vind ik niet zo mooi.
Het plaatje met de **hut** en de **gans** vind ik niet zo mooi.
Implausible items
Dina tekent een **kat** naast de **tijger** op de kleine zachte kauwgom.
Dina tekent een **zon** naast de **tijger** op de kleine zachte kauwgom.
Onderweg zag ik een **fiets** bij een **brommer** langs de kant staan printen.
Onderweg zag ik een **paard** bij een **brommer** langs de kant staan printen.
Op het schilderij zag je een **zwaan** naast een **meeuw** met grote vleugels scholen eten.
Op het schilderij zag je een **hek** naast een **meeuw** met grote vleugels scholen eten.
Is hij nou zijn **trui** en zijn **broek** weer vergeten mee te bakken?
Is hij nou zijn **deken** en zijn **broek** weer vergeten mee te bakken?
Het verhaal ging over een **zebra** en een **ezel** die samen avonturen in de vloer wassen.
Het verhaal ging over een **rover** en een **ezel** die samen avonturen in de vloer wassen.
Hij pakte zijn **jas** en zijn **muts** om mee te kunnen hoesten in het krijtje.
Hij pakte zijn **bril** en zijn **muts** om mee te kunnen hoesten in het krijtje.
Zet je de **jam** en de **honing** ook terug in de vulkaan?
Zet je de **olie** en de **honing** ook terug in de vulkaan?

Table G.1 (continued)

Experimental sentence pairs (prime-target pairs in boldface, related items always followed by unrelated items) by semantic type and plausibility

Context-dependent, non-associated Plausible items
Als je in dit **bos** een **eekhoorn** zou vinden, is dat niet gek.*
Als je in dit **schip** een **eekhoorn** zou vinden, is dat wel gek.*
Mijn **kapper** is altijd haar **borstel** kwijt als ze hem nodig heeft.
Mijn **dochter** is altijd haar **borstel** kwijt als ze hem nodig heeft.
Hij had in de **woestijn** een **slang** gezien waardoor hij erg was geschrokken.
Hij had in de **speeltuin** een **slang** gezien waardoor hij erg was geschrokken.
Toen zijn **juf** het **lokaal** in liep was het een grote bende.
Toen zijn **oma** het **lokaal** in liep was het een grote bende.
Op het **station** zocht de **conducteur** naar de man die hem had geroepen.
Op het **kantoor** zocht de **conducteur** naar de man die hem had geroepen.
De oude **kapitein** kijkt naar de **haven** terwijl de grote boot wegvaart.
De oude **mevrouw** kijkt naar de **haven** terwijl de grote boot wegvaart.
Mijn **kok** heeft alvast de **oven** aangezet voor zijn speciale koekjes.
Mijn **neef** heeft alvast de **oven** aangezet voor zijn speciale koekjes.
Implausible items
In de **bakkerij** heeft de **klant** een praatje gemaakt met de snoeischaar.
In de **woonkamer** heeft de **klant** een praatje gemaakt met de snoeischaar.
De jonge **visser** kijkt naar de **rivier** terwijl de maan gras maait.
De jonge **keizer** kijkt naar de **rivier** terwijl de maan gras maait.
De **chauffeur** pakte het **stuur** wat steviger vast in de blauwe sla.
De **directeur** pakte het **stuur** wat steviger vast in de blauwe sla.
In een grote **tuin** kan een **mus** heel ver vliegen met tafels.
In een grote **stad** kan een **mus** heel ver vliegen met tafels.
De aardige **dokter** pakte een **pleister** uit de grote televisie met lepels.
De aardige **dame** pakte een **pleister** uit de grote televisie met lepels.
Ik wil niet dat die **dief** mijn **tas** mee kan nemen uit de pindakaas.
Ik wil niet dat die **kerel** mijn **tas** mee kan nemen uit de pindakaas.
De **prinses** heeft de **koets** al horen aankomen op de kroket.
De **soldaat** heeft de **koets** al horen aankomen op de kroket.

* This pair differs in the penultimate word to make sure both sentences are plausible, but since this position is outside of the spillover region, this difference does not affect our measurements of interest.

Table G.1 (continued)

Experimental sentence pairs (prime-target pairs in boldface, related items always followed by unrelated items) by semantic type and plausibility

Context-independent, associated
Plausible items
Cas houdt van **banaan** en rood **fruit** in zijn yoghurt bij het ontbijt.
Cas houdt van **suiker** en rood **fruit** in zijn yoghurt bij het ontbijt.
Zara wil een **gitaar** en een **instrument** met van die belletjes, een tamboerijn.
Zara wil een **puzzel** en een **instrument** met van die belletjes, een tamboerijn.
Op een **krekel** lijkt dit **insect** wel een beetje, vind ik.
Op een **schelp** lijkt dit **insect** wel een beetje, vind ik.
Emma maakt een tekening van een **vlinder** en een **rups** met bloemen en gras erbij.
Emma maakt een tekening van een **vijver** en een **rups** met bloemen en gras erbij.
De gemene **koning** is de **baas** van het hele land in het sprookje.
De gemene **leeuw** is de **baas** van het hele land in het sprookje.
Die oude **pop** als **speelgoed** gebruiken is geen goed idee, leg hem maar terug.
Die oude **foto** als **speelgoed** gebruiken is geen goed idee, leg hem maar terug.
Mamma wil nog een **rok** en warme **kleding** voor Maaike kopen deze week.
Mamma wil nog een **vaas** en warme **kleding** voor Maaike kopen deze week.
Implausible items
Milan houdt van de **oceaan** en de **zee** en wandelen in de scherpe papiertjes.
Milan houdt van de **lente** en de **zee** en wandelen in de scherpe papiertjes.
Daan zag de **bruid** bij de **vrouw** in de kleine kruimel lopen.
Daan zag de **wesp** bij de **vrouw** in de kleine kruimel lopen.
Door de **auto** ging het **vervoer** van het grote oerwoud naar de tranen wel goed.
Door de **lucht** ging het **vervoer** van het grote oerwoud naar de tranen wel goed.
Ik wil geen **hamster** als **huisdier** maar een grote zonnige stoeptegel.
Ik wil geen **garnaal** als **huisdier** maar een grote zonnige stoeptegel.
De **uil** en de kleine **vogel** zitten stil in de nietmachine.
De **zoon** en de kleine **vogel** zitten stil in de nietmachine.
Pappa moest nog **kaas** en **melk** halen in de groene bijl.
Pappa moest nog **zeep** en **melk** halen in de groene bijl.
Neem je nog een **komkommer** en **groente** uit de diepvries mee van de verwarming?
Neem je nog een **pudding** en **groente** uit de diepvries mee van de verwarming?

Table G.1 (continued)

Experimental sentence pairs (prime-target pairs in boldface, related items always followed by unrelated items) by semantic type and plausibility

Context-dependent, associated
Plausible items
Gelukkig heeft de oude **boer** nog **vee** en ook veel land waar hij op werkt.
Gelukkig heeft de oude **heer** nog **vee** en ook veel land waar hij op werkt.
Ik had in de **winkel** mijn **geld** laten liggen, echt stom van me.
Ik had in de **trein** mijn **geld** laten liggen, echt stom van me.
Bij een **kasteel** was de **ridder** gestopt om uit te rusten.
Bij een **heuvel** was de **ridder** gestopt om uit te rusten.
De **indiaan** heeft de **veren** gevonden in het hoge gras.
De **knecht** heeft de **veren** gevonden in het hoge gras.
In het verhaal raakte de **heks** haar **hoed** kwijt in het donkere woud.
In het verhaal raakte de **zus** haar **hoed** kwijt in het donkere woud.
De aardige **schilder** heeft de **verf** nog even uit zijn busje gehaald.
De aardige **buurman** heeft de **verf** nog even uit zijn busje gehaald.
De bange **jager** liet het **geweer** per ongeluk op de grond vallen.
De bange **prins** liet het **geweer** per ongeluk op de grond vallen.
Implausible items
Er stond een **piloot** naast het **vliegtuig** op de landingsbaan van trommels.
Er stond een **ladder** naast het **vliegtuig** op de landingsbaan van trommels.
Voor de **camping** stond een **tent** met veel groene slagroom en sleutels.
Voor de **garage** stond een **tent** met veel groene slagroom en sleutels.
Zijn kleine **baby** wilde geen **luier** aan omdat hij de straat moest typen.
Zijn kleine **broer** wilde geen **luier** aan omdat hij de straat moest typen.
Op de **markt** zijn mijn **kleren** helemaal vies geworden van het nadenken.
Op de **kermis** zijn mijn **kleren** helemaal vies geworden van het nadenken.
Bij de **stal** lag het **hooi** met alle mooie telefoons erdoorheen.
Bij de **put** lag het **hooi** met alle mooie telefoons erdoorheen.
De **toerist** vond zijn **vakantie** heel erg leuk en eetbaar.
De **postbode** vond zijn **vakantie** heel erg leuk en eetbaar.
Daar achter de **kast** vind je de **spullen** waarmee je de beker kunt oefenen.
Daar achter de **muur** vind je de **spullen** waarmee je de beker kunt oefenen.

Summary in English

Developing semantic networks
Individual differences in Dutch monolingual and bilingual children's semantic knowledge and reading comprehension

Vocabulary knowledge is a fundamental requirement for school success, not least because of its importance for the acquisition of literacy skills, which in turn becomes the basis for acquiring content knowledge. However, the development of vocabulary knowledge involves more than extending the number of words in one's mental lexicon. Additionally, the richness or so-called depth of semantic information associated with lexical entries, and the speed and facility with which this information can be accessed, need to be developed. This dissertation examines access to semantic information and a specific instantiation of vocabulary depth, namely, the semantic network, in monolingual and bilingual minority children, and assesses the importance of these vocabulary aspects for individual differences in reading comprehension. All three studies were conducted in the Netherlands, in Dutch. The participating children were fifth-graders (*groep 7* in the Dutch system), between 10-11 years old and spoke either exclusively Dutch (monolinguals), or were raised with a minority language, additionally acquiring Dutch at home and/or at school (bilingual minority children).

The mental lexicon is hypothesized to be organized as a network, in which words are connected through semantic relations of various types. These semantic relations are activated via spreading activation during language processing, meaning that a lexical representation for a word such as *squirrel* can activate related concepts such as *animal*, via a taxonomic relation; *fur*, via a feature relation; *forest*, via a situational relation; and *cute*, via a subjective relation. These semantic relations can be characterized as occupying spaces on a continuum from context-independent to context-dependent knowledge: *squirrel* and *animal* share

many intrinsic features, and are therefore related independent of context, while *squirrel* and *forest* are related through co-occurrence in context, and the pair *squirrel* – *cute* is subjectively related and therefore specific to one's personal context. Additionally, a distinction can be made between associative and non-associative relations: these are empirically determined by means of word association tasks. An associative relation holds between words that often occur as stimulus – response pairs in word association tasks: a common example of a word pair with a strong associative relation is *blood* – *red*. Word pairs that are not common stimulus – response pairs in such a task may still hold any type of semantic relation, such as for example *desert* – *snake* (situational) or *dog* – *wolf* (taxonomic).

Typically, during language acquisition, children first focus mostly on context-dependent relations and gradually add more context-independent knowledge, thus abstracting from episodic experience in order to develop decontextualized knowledge. The presence and automatic activation (so-called semantic priming) of these relations in children's semantic networks is studied from various perspectives in this dissertation. Both the contribution of this semantic knowledge to reading comprehension and differences between monolingual and bilingual minority children are examined.

Bilingual minority children are often found to lag behind their monolingual peers in the acquisition of vocabulary and reading comprehension. For knowledge of semantic relations specifically, evidence from declarative tasks requiring the identification of certain semantic relations suggests that bilinguals' knowledge of especially context-independent relations is less advanced. However, word association data and priming experiments so far fail to uniformly confirm these patterns, leading to the question whether semantic network organization and activation are indeed so different between these two language groups.

Furthermore, theories of reading comprehension suggest that vocabulary knowledge is central to the complex process of understanding texts, and that the quality and accessibility of word knowledge is pivotal. Fluent access to a word's semantic representation may free up processing capacity for higher-order comprehension processes. Furthermore, the activation of semantically related words, i.e. semantic priming, is hypothesized to aid the establishment of coherence during reading, by linking related concepts in the mental representation of a text. Again, studies targeting declarative knowledge of context-independent relations find that it contributes to comprehension, and additionally, group studies of semantic priming suggest that poor comprehenders show less activation of especially context-independent relations. However, it is unclear whether individual differences in semantic activation also relate to individual differences in comprehension in normal readers: evidence for this is currently suggesting that it may not be. The present dissertation attempted to contribute to answering these questions by assessing the influence of individual differences in associative preferences, semantic access and both single-word and sentence-level priming to reading comprehension in monolingual and bilingual minority children.

A multiple word association task, requiring three responses to each stimulus word, and a reading comprehension task were administered to 207 children for the study described in Chapter 2. A detailed classification system for the semantic characteristics of the responses was devised based on previous word association and language acquisition research, including four semantic categories: taxonomic, feature-based, situational and subjective relations. The data showed limited differences between the monolingual and bilingual children. Although bilingual children scored lower on the reading comprehension test and produced more null responses on the word association task, suggesting that their vocabularies were smaller, there were no significant differences between the two groups within each of the semantic categories. The organization of their semantic knowledge thus appears to be similar to that of monolinguals in this sample, despite other delays in language proficiency. The fact that the stimuli used were highly frequent words in Dutch may be responsible for this discrepancy: because differences between monolinguals and bilinguals may be largely due to differences in exposure, highly frequent words may be less susceptible to delays in bilinguals.

In contrast, contributions of semantic preferences to reading comprehension were found. Children producing more taxonomic associations such as *squirrel – animal*, both as initial and as later responses, were found to perform better on the reading comprehension test as well.

Conversely, participants who produced more feature associations such as *squirrel – fur* as first responses, showed worse reading comprehension scores. This is contrary to evidence from many studies that assessed the contribution of vocabulary depth to comprehension, which indicate that declarative knowledge of intrinsic features is positively related to comprehension. The knowledge measured in the relatively spontaneous word association task may thus relate differently to comprehension compared to knowledge of semantic relations which is explicitly elicited. Additionally, compared to adults, producing more features as first rather than later responses in an association task is atypical, and it was argued that this may therefore be a sign of a less developed semantic network, which is apparently also related to a disadvantage in comprehension.

The second study, reported in Chapter 3, assessed speed of access to semantic information and priming of non-associated context-dependent and context-independent semantic relations in an auditory semantic classification task. In this task, children heard single words and were asked to decide for each word whether it denoted an animate or inanimate being. Reaction times were recorded to determine whether responses were speeded when a target word such as *wolf* was preceded by a related prime word, such as *dog*, compared to an unrelated prime word, such as *captain*. 122 children were administered this task and additionally performed tasks assessing reading comprehension, vocabulary size, decoding speed and processing speed. In this sample, the monolinguals and bilinguals did not show significant differences on any of the language measures, neither the standardized tasks assessing comprehension, size, decoding and processing speed, nor on the priming and access measures from the classification task. This may be a sign that the tasks were not sensitive enough to detect the perhaps small differences between the language groups. On the other hand, an absence of differences is in line with recent evidence from large-scale assessment reports in the Netherlands, which suggest that differences between monolinguals and bilingual minority children may be disappearing.

The analyses assessing the influence of individual differences in semantic access and semantic priming on comprehension revealed no significant contributions of either of the vocabulary aspects. Although the choice for an auditory task was deliberate, so that any connection to comprehension would be located at the semantic level, this may also be the reason why, contrary to previous studies, semantic access was not related to comprehension. The difference between access to semantic information from written stimuli rather than spoken stimuli may be crucial for an effect to arise. Furthermore, although group-level priming of the context-independent pairs was observed, differences in this activation on the individual level were not found to be related to differences in comprehension. This is contrary to evidence from group studies, but it does coincide with findings from a similar individual differences study performed in the same population. The data suggest that within a group of normal readers, individual differences in priming may be too small to be predictive of comprehension. Additionally however, the instability and noisiness of semantic priming effects on the individual level may complicate relating this measure of semantic knowledge to comprehension.

In order to check whether the subtlety of differences in semantic activation is indeed to blame for the lack of contribution to comprehension in Chapter 3, Chapter 4 targeted activation during sentence reading. In this way, the role of semantic activation during the comprehension process could be approximated more closely. Both associated and non-associated context-dependent and context-independent word pairs were tested in a self-paced reading paradigm, which was administered to 137 children. In this task, children were presented with sentences that were initially masked with dashes. Each word could be revealed and read by the child by pressing a button, and was masked again to reveal the next word at the next button press. In this way, reading times at each word could be measured, and facilitation in reading time due to a previously occurring related word could be measured. Related and unrelated word pairs similar to those used in the single-word priming experiment were embedded in the sentences. Additionally, the participants performed tasks measuring reading comprehension, vocabulary size, declarative knowledge of context-independent semantic relations and decoding. Again, differences between monolinguals and bilinguals were limited: significant differences were found for the context-independent knowledge task and speed in the self-paced reading task. No differences were found on any of the other tasks, including semantic activation during sentence reading.

In line with the findings from Chapter 3, a clear group effect of non-associated context-independent priming was identified: reading times at the second word following a related target were reduced compared to an unrelated target. Additionally, a smaller joint effect of non-associated context-dependent and context-independent pairs was observed at the first word following related targets. However, again individual scores of these priming effects were not found to be related to reading comprehension, corroborating the findings from previous research and Chapter 3. Even when measured during reading, facilitation of context-independent knowledge appears not to be predictive of individual differences in comprehension. The consistency with previous individual differences research suggests that the group effects of priming in poor and normal comprehenders may not translate to the individual level. Whether this is because the connection is too weak in normal readers or because priming effects are too variable within individuals, remains an open question.

Combined, the three studies showed few significant differences between monolinguals and bilingual minority children. The fact that the majority of the bilingual participants spoke both Dutch and another language at home and that all were schooled and became literate in Dutch, is likely responsible for these findings. The contributions of individual differences in semantic network organization, semantic access and semantic priming to comprehension were also limited. The relative lack of contributions to comprehension from access and priming may be partly due to methodological issues, such as the use of highly frequent stimuli, or the inherent noisiness of priming effects. However, the lack of contributions to comprehension in individual differences studies targeting semantic priming is consistent so far, suggesting that this vocabulary aspect may not be the most fruitful direction for future research and interventions. The three studies in this dissertation mainly confirm the importance of vocabulary size and knowledge of context-independent semantic relations as main contributors to reading comprehension.

Nederlandse samenvatting

Semantische netwerken in ontwikkeling: Individuele verschillen in semantische kennis en begrijpend lezen bij Nederlandse eentalige en meertalige kinderen

Een goed ontwikkelde woordenschat is een fundamentele voorwaarde voor schoolsucces, niet in het minst vanwege het belang van woordkennis voor het ontwikkelen van een goede leesvaardigheid. Het verwerven van woordenschat heeft echter meer om het lijf dan alleen het uitbreiden van het *aantal* gekende woorden: ook de kwaliteit van de kennis die we hebben van deze woorden speelt een rol. Hieronder valt bijvoorbeeld kennis van de betekenisrelaties tussen het woord en andere woorden, en de snelheid waarmee de woordbetekenis en -relaties kunnen worden opgehaald tijdens het gebruik van taal. In dit proefschrift wordt het niveau van deze woordenschataspecten bestudeerd bij Nederlandse kinderen van 10-11 jaar oud (groep 7). Zowel de woordenschat van eentalig opgroeiende kinderen, als van meertalige kinderen die naast het Nederlands een minderheidstaal spreken, is onderzocht. Daarnaast werd bekeken hoe deze woordenschataspecten samenhangen met leesvaardigheid: zijn de kinderen bij wie deze woordenschataspecten beter ontwikkeld zijn, ook beter in begrijpend lezen?

In deze samenvatting wordt eerst meer toelichting gegeven op de belangrijkste begrippen en conclusies uit eerder onderzoek, waarna de opzet en de uitkomsten van de verschillende deelstudies uit dit proefschrift worden besproken.

Semantische toegang, semantisch netwerk en semantische priming

De woorden die we kennen zijn opgeslagen in het zogenoemde *mentale lexicon*. In dit proefschrift wordt ingegaan op drie aspecten van het mentale lexicon: de snelheid van de *semantische toegang* tot de betekenis van de opgeslagen woorden, de *structuur* van het semantisch netwerk en de mate van *priming* binnen dat netwerk. Het eerste aspect, de *semantische toegang*, wordt in dit proefschrift gedefinieerd als de snelheid waarmee een taalgebruiker de betekenis van gekende woorden kan ophalen. Bij het zoeken naar woorden voor bijvoorbeeld woordherkenning worden woorden in het mentale lexicon geactiveerd, en boven een bepaalde activatiedrempel komt de betekenis die bij een (schriftelijke) woordvorm hoort beschikbaar: dit is de semantische toegang. Hoe meer dit proces geautomatiseerd is, hoe sneller de woordbetekenis kan worden opgehaald. Deze automatisering bespaart cognitieve capaciteit die gebruikt kan worden voor andere deelprocessen tijdens het verwerken van de taal.

De *structuur* van het semantisch netwerk en de *priming* binnen dit netwerk hangen samen met de manier waarop woorden opgeslagen zijn in het mentale lexicon. We kunnen deze opslag voorstellen als een netwerk, waarin alle woorden met elkaar verbonden zijn via betekenisrelaties, die verschillende soorten verbanden in kunnen houden. Zo heeft het woord *eekhoorn* taxonomische relaties met de woorden *dier* en *rat*: *dier* is het superordinaat van *eekhoorn*, en *rat* is een coördinaat van *eekhoorn*. *Eekhoorn* is ook verbonden met het woord *staart*, via een kenmerk-relatie; met het woord *bos* via een situationele relatie; en met het woord *schattig*, via een subjectieve relatie. Deze vier relatietypen kunnen we zien als punten op een continuüm van context-onafhankelijke naar context-afhankelijke verbanden. Taxonomische relaties, het meest context-onafhankelijke type, zijn gebaseerd op gedeelde kenmerken: *eekhoorn* en *dier* delen veel intrinsieke kenmerken en zijn daarom verbonden onafhankelijk van context. Bij kenmerkrelaties wordt één van deze intrinsieke kenmerken uitgelicht, zoals bij *eekhoorn* – *staart*. Bij een situationele relatie als *eekhoorn* – *bos* daarentegen is er geen sprake van gedeelde intrinsieke kenmerken: *bos* vormt een context waarin *eekhoorns* vaak voorkomen, maar het concept *bos* is niet inherent aan het concept *eekhoorn* zoals het concept *staart* dat wel is. Subjectieve relaties op hun beurt zijn voornamelijk afhankelijk van iemands persoonlijke context: niet iedereen zal van mening zijn dat een *eekhoorn schattig* is. Het onderscheid tussen context-onafhankelijke en context-afhankelijke verbanden speelt een rol in de taalverwerving. Tijdens het verwerven van de moedertaal focussen kinderen over het algemeen in eerste instantie meer op context-afhankelijke relaties, omdat deze direct te observeren zijn in de omgeving. Van hieruit worden context-onafhankelijke relaties ontwikkeld, door geleidelijk te abstraheren van de directe ervaring en te generaliseren naar context-onafhankelijke kennis.

Een ander belangrijk onderscheid dat gemaakt kan worden met betrekking tot de relaties tussen woorden in een taal en daarmee in de semantische netwerken van sprekers, is tussen geassocieerde en niet-geassocieerde relaties. In woordassociatie-onderzoek krijgen proefpersonen woorden als stimulus voorgelegd, met de vraag om aan te geven aan welke woorden zij moeten denken in reactie op deze stimuluswoorden. Hieruit kunnen we afleiden welke woordparen *geassocieerd* zijn in een taal en welke niet: als veel sprekers een bepaald woord geven als reactie op een specifiek stimuluswoord, beschouwen we dit als een geassocieerd woordpaar. Het woordpaar *hond* – *kat* is bijvoorbeeld sterk geassocieerd in het Nederlands, maar *hond* – *wolf* niet, terwijl beide paren een duidelijke semantische relatie hebben. Geassocieerde woordparen kunnen alle hierboven genoemde typen woordrelaties behelzen – van taxonomisch tot subjectief – en omgekeerd kunnen alle semantische typen ook niet-geassocieerd zijn. Zowel geassocieerde als niet-geassocieerde relaties hebben een plaats in het semantisch netwerk van een individu. De structuur van het semantisch netwerk van een individu wordt gekenmerkt door de verbanden die in het netwerk zijn opgenomen: niet iedereen zal precies dezelfde verbanden in zijn of haar netwerk hebben, en bij de één zullen bepaalde betekenisrelaties prominenter zijn dan bij de ander, wat bijvoorbeeld kan blijken uit een voorkeur voor bepaalde relaties in woordassociatieonderzoek.

Daarnaast vindt in het semantisch netwerk *spreiding van activatie* plaats: tijdens het taalgebruik activeren we woorden, die vervolgens omliggende woorden in het semantisch netwerk zullen activeren. Als een spreker van het Nederlands bijvoorbeeld net *eekhoorn* gehoord heeft, zal dit woord waarschijnlijk ook *dier* en *bos* activeren, waardoor deze woorden tijdelijk makkelijker herkend of opgehaald kunnen worden. We noemen dit semantische *priming* van een *target* (*bos*) door een *prime* (*eekhoorn*), en onderzoek laat zien dat de mate waarin deze priming-effecten optreden varieert van persoon tot persoon: verschillende mensen kunnen meer of minder priming laten zien voor de verschillende typen betekenisrelaties.

Doel van het onderzoek

De drie woordenschataspecten die in dit proefschrift onder de loep zijn genomen – de structuur van en mate van priming binnen het semantisch netwerk en de semantische toegang – behelzen dus alle drie iets anders dan de omvang van de woordenschat. Zowel de relatie tussen deze woordenschataspecten en begrijpend lezen, als verschillen tussen eentalige en meertalige kinderen op het gebied van deze woordenschataspecten worden onderzocht.

Kinderen die meertalig opgroeien met een minderheidstaal vertonen vaak achterstanden in vergelijking met hun eentalige leeftijdsgenootjes op het gebied van woordenschat en leesvaardigheid. Deze achterstanden zijn niet alleen aangetoond voor de woordenschatgrootte, maar ook voor de declaratieve kennis van betekenisrelaties: in taken waarin kinderen wordt gevraagd specifieke betekenisrelaties aan te duiden of te selecteren, presteren meertalige kinderen over het algemeen minder goed. Met name de ontwikkeling van context-onafhankelijke relaties lijkt bij deze groep achter te blijven. Dit doet vermoeden dat er ook verschillen zullen zijn op het gebied van woordassociaties en semantische priming, maar daar is minder duidelijk bewijs voor. Er is nog niet eenduidig aangetoond dat eentalige en meertalige kinderen bijvoorbeeld een voorkeur hebben voor andere semantische relaties wanneer hun gevraagd wordt woordassociaties te produceren, en het is nog niet duidelijk of meertalige kinderen andere patronen van semantische priming vertonen dan hun eentalige leeftijdsgenoten.

Voor wat betreft het verband tussen woordenschat en begrijpend lezen, zijn er verschillende theorieën die suggereren dat woordenschatkennis in al haar facetten een centrale rol speelt bij begrijpend lezen: niet alleen de woordenschatgrootte, maar ook de kwaliteit van de semantische informatie die bij een woord hoort en de toegang tot deze informatie zou cruciaal zijn. Snelle en automatische toegang tot de semantische representatie van een woord, dus een goed ontwikkelde semantische toegang, kan helpen bij het begrijpend lezen: als het ophalen van de woordbetekenis relatief weinig moeite kost, blijft er meer verwerkingscapaciteit over voor de meer complexe leesprocessen zoals het leggen van verbanden in de tekst en het scheiden van hoofd- en bijzaken. Daarnaast wordt verondersteld dat een goed ontwikkeld semantisch netwerk bijdraagt aan de leesvaardigheid. Het activeren van semantisch gerelateerde woorden kan bijvoorbeeld helpen bij het opbouwen van een coherente tekstrepresentatie, wat het tekstbegrip bevordert: doordat gerelateerde woorden automatisch geactiveerd worden tijdens het lezen, kunnen de geactiveerde verbanden makkelijker worden opgenomen in het zogenoemde mentale model van de tekst. Er is bewijs dat individuele verschillen in declaratieve kennis van met name context-onafhankelijke betekenisrelaties bijdragen aan tekstbegrip, en groepsstudies tonen aan dat lezers die minder semantische priming van met name niet-geassocieerde context-onafhankelijke betekenisrelaties vertonen, ook minder goed scoren op begrijpend lezen. Hoewel we op basis van dergelijke resultaten zouden verwachten dat verschillen in semantische priming ook zouden bijdragen aan *individuele* verschillen in begrijpend lezen, lijkt eerder onderzoek dit niet uit te wijzen. Daarom is in dit proefschrift nader onderzoek gedaan naar de relatie tussen individuele verschillen in begrijpend lezen en de drie focale woordenschataspecten: de structuur van het semantisch netwerk, de mate van semantische priming en de snelheid van de semantische toegang.

Studie 1: structuur van het semantisch netwerk

De structuur van het semantisch netwerk is onderzocht in de studie die beschreven wordt in Hoofdstuk 2. Hiervoor kregen 207 kinderen een toets begrijpend lezen en een meervoudige woordassociatietaak voorgelegd. In een meervoudige woordassociatietaak wordt, in tegenstelling tot in de enkelvoudige woordassociatietaak, van de deelnemers gevraagd om niet één maar drie woordassociaties te geven in reactie op de stimuluswoorden, wat een gedetailleerder beeld oplevert van de verbanden in hun mentale lexicon. De 11.725 semantische woordassociaties die de kinderen hebben gegeven zijn geanalyseerd met behulp van een uitgebreid classificatiesysteem, dat gebaseerd is op eerder woordassociatie- en taalverwervingsonderzoek. De vier semantische typen die eerder besproken zijn – taxonomische, kenmerk-, situationele en subjectieve relaties – waren opgenomen in dit systeem.

De dataset liet slechts beperkte verschillen zien tussen de eentalige en de meertalige kinderen. Hoewel de meertalige kinderen lager scoorden op de leestaak en vaker géén associatie gaven in reactie op de stimuli, en dus mogelijk een kleiner vocabulaire hadden, waren er geen significante verschillen tussen de taalgroepen binnen de vier semantische categorieën. De organisatie van het semantisch netwerk lijkt dus vergelijkbaar te zijn tussen de twee groepen, ondanks andere verschillen in taalvaardigheid. Het feit dat de stimuluswoorden allemaal hoogfrequente zelfstandige naamwoorden waren, kan hierbij een rol hebben gespeeld: omdat verschillen in woordenschat tussen eentalige en meertalige kinderen waarschijnlijk vooral veroorzaakt worden door verschillen in blootstelling aan de taal, kan het zijn dat hoogfrequente woorden minder verschillen laten zien tussen eentalige en meertalige kinderen.

Een verband tussen individuele verschillen in semantische organisatie en leesvaardigheid kon echter wel worden aangetoond. Kinderen die meer taxonomische associaties produceerden, zoals *eekhoorn – dier*, zowel als eerste antwoord als als tweede of derde antwoord, presteerden ook beter op begrijpend lezen. Aan de andere kant werd gevonden dat kinderen die veel kenmerkrelaties zoals *eekhoorn – staart* produceerden als eerste antwoord, juist minder goede leesscores behaalden. Het produceren van veel kenmerkrelaties als tweede en derde antwoorden vertoonde geen significante samenhang met de leesvaardigheid. Eerder onderzoek laat zien dat declaratieve kennis van taxonomische relaties én intrinsieke kenmerken van woorden een positieve relatie heeft met begrijpend lezen: voor de taxonomische relaties werd dit dus ook gevonden in dit woordassociatie-onderzoek, maar voor kenmerkrelaties kwam een tegenovergesteld effect naar voren. Blijkbaar hangt de kennis van intrinsieke kenmerken zoals gemeten in de relatief spontane associatietaak anders samen met de leesvaardigheid dan kennis die in een meer expliciete, gestuurde taak wordt gemeten. Daarnaast is het mogelijk veelzeggend dat juist de kinderen die vaak een kenmerkrelatie als *eerste* antwoord gaven, minder goed presteerden op de leestaak. Bij volwassenen is het namelijk atypisch om vaak een kenmerk als eerste antwoord te geven in een meervoudige associatietaak: zij geven doorgaans de voorkeur aan taxonomische relaties als eerste antwoord. Mogelijk is het vaak produceren van kenmerkrelaties als eerste associatie dan ook een teken van een minder ver ontwikkeld semantisch netwerk, wat mogelijk negatief samenhangt met de leesvaardigheid.

Studie 2: semantische toegang en priming op woordniveau

In de tweede empirische studie, gerapporteerd in Hoofdstuk 3, werden de semantische toegang en semantische priming onder de loep genomen. De priming van niet-geassocieerde context-onafhankelijke en context-afhankelijke semantische relaties werd gemeten in een priming-experiment, met behulp van een auditieve semantische classificatietaak. In dit experiment hoorden de deelnemers losse woorden, en werd hun gevraagd voor elk woord zo snel mogelijk te beslissen of het een levend wezen of een ding betrof. Ze gaven deze beslissing aan door een knop in te drukken, waardoor naast de accuratesse de snelheid van hun reactie in milliseconden kon worden gemeten. De reactietijden op woorden werden vergeleken in een gerelateerde en ongerelateerde conditie: zo werd bekeken of kinderen sneller het woord *wolf* konden beoordelen wanneer dit op het woord *hond* volgde, versus wanneer dit op het woord *kapitein* volgde. Indien de reactie versneld is na het gerelateerde woord, kan worden geconcludeerd dat het kind priming van deze context-onafhankelijke relatie vertoont. Daarnaast kon de snelheid van de semantische toegang gemeten worden door de reactietijden op niet-experimentele woorden, zogenoemde *fillers*, te middelen.

De taak werd uitgevoerd door 122 kinderen, die naast het experiment met aparte taken getoetst werden op woordenschatgrootte, technische leesvaardigheid en cognitieve verwerkingssnelheid. De eentalige en de meertalige kinderen in deze groep lieten op geen enkele taak significante verschillen zien. Dit kan een teken zijn dat de taken niet gevoelig genoeg waren om de mogelijk kleine verschillen tussen de taalgroepen te meten. Aan de andere kant is de afwezigheid van verschillen in overeenstemming met aanwijzingen uit verschillende recente nationale onderzoeken dat taalvaardigheidsverschillen tussen eentalige en meertalige kinderen in Nederland langzaam verdwijnen.

Er werd ook geen significante bijdrage aan de leesscores gevonden van individuele verschillen in de snelheid van de semantische toegang of semantische priming. Het gebrek aan een significante relatie tussen de semantische toegang en het begrijpend lezen heeft mogelijk te maken met de keuze om een auditieve taak te gebruiken: eerdere studies die wel een relatie vonden tussen leesvaardigheid en semantische toegang, gebruikten geschreven stimuli. In dit onderzoek werd een auditieve taak gebruikt om vast te kunnen stellen of een eventueel effect van semantische toegang en semantische priming daadwerkelijk op het betekenisniveau plaatsvond, en niet bijvoorbeeld (gedeeltelijk) veroorzaakt werd door beter ontwikkelde technische leesvaardigheid. Mogelijk is het echter voor de relatie tussen semantische toegang en begrijpend lezen noodzakelijk dat de toegang tot het concept verloopt via het geschreven woord.

Ook het verwachte verband tussen priming van semantische relaties en begrijpend lezen kon niet worden aangetoond. Hoewel de kinderen als groep een priming-effect op de context-onafhankelijke paren lieten zien (een versnelde herkenning van *wolf* na *hond*), was het niet zo dat kinderen die sterkere activatie van deze relatie lieten zien, ook beter scoorden op begrijpend lezen. Dit is in tegenspraak met de bevindingen uit eerdere groepsstudies, die juist lieten zien dat normale lezers als groep meer priming van niet-geassocieerde, context-onafhankelijke woordparen vertoonden dan slechte lezers. Het komt echter wel overeen met een eerdere studie in dezelfde populatie, die ook geen effect van individuele verschillen in priming op begrijpend lezen kon aantonen. De resultaten suggereren dat de individuele verschillen in semantische priming binnen een groep normale lezers mogelijk te klein zijn om het begrijpend lezen te kunnen voorspellen. Daarnaast is aangetoond dat semantische priming-effecten, die doorgaans een grootte hebben van enkele tientallen milliseconden, onderhevig zijn aan ruis in de experimentele opzet en relatief instabiel zijn binnen individuen. Dit kan een reden zijn waarom een verband met andere complexe taalvaardigheden op individueel niveau moeilijk aan te tonen is.

Studie 3: priming op zinsniveau

Om na te gaan of de veronderstelling dat binnen de groep normale lezers de verschillen in semantische priming te subtiel waren om een verband met leesvaardigheid aan te tonen, werd voor Hoofdstuk 4 een tweede experiment ontworpen en voorgelegd aan 137 kinderen. In dit experiment werden woordparen niet als losse woorden gepresenteerd, maar in een zinscontext, zodat de priming tijdens het lezen kon worden gemeten. Zowel geassocieerde als niet-geassocieerde woordparen met zowel context-afhankelijke als context-onafhankelijke semantische relaties werden verwerkt in een *self-paced reading* experiment, waarbij deelnemers in hun eigen tempo, woord voor woord zinnen lezen. De zinnen werden eerst gepresenteerd in gemaskeerde vorm, d.w.z. dat alle letters waren vervangen door streepjes. Door op een toets te drukken, kon het kind het eerste woord zichtbaar maken en lezen. Door nogmaals op een toets te drukken, werd het eerste woord weer gemaskeerd en werd het tweede woord leesbaar, etc. Op deze manier werd voor elk woord de leestijd in milliseconden geregistreerd. Gerelateerde en ongerelateerde woordparen werden ingebed in zinnen, zodat gemeten kon worden of de leestijd van een woord versnelde wanneer het in een zin voorafgegaan werd door een gerelateerd woord, versus wanneer het in precies dezelfde zin voorafgegaan werd door een niet-gerelateerd woord. Verder werden taken voor begrijpend lezen, woordenschatgrootte, declaratieve kennis van context-onafhankelijke semantische relaties en technische leesvaardigheid afgenomen.

Ook in deze studie waren de verschillen tussen de eentalige en de meertalige kinderen beperkt. Alleen op declaratieve kennis van context-onafhankelijke relaties lieten de meertalige kinderen een significant lagere score zien, en een klein maar significant verschil in het voordeel van de meertalige kinderen trad op voor het technisch lezen en de leessnelheid in het experiment. Afgezien van deze controlematen presteerden de meertalige kinderen dus op een vergelijkbaar niveau met hun eentalige klasgenoten.

Net als in Hoofdstuk 3 werd een duidelijk groepseffect van niet-geassocieerde, context-onafhankelijke relaties gevonden: op het tweede woord na targetwoorden die voorafgegaan werden door context-onafhankelijk gerelateerde primewoorden waren leestijden significant korter. Daarnaast werd een kleiner effect gevonden van alle niet-geassocieerde woordparen, zowel context-afhankelijk als context-onafhankelijk gerelateerd, op de woorden die direct op de target volgden. Op individueel niveau hingen deze priming-effecten echter net als in het vorige experiment niet significant samen met de individuele leesscores. Zelfs wanneer semantische priming tijdens het lezen werd gemeten, kon niet worden aangetoond dat deze een verband had met het begrijpend lezen op individueel niveau. De consistentie tussen Hoofdstuk 3 en 4 en eerder onderzoek suggereert dat het eerder aangetoonde verband op groepsniveau tussen begrijpend lezen en semantische priming zich niet zonder meer laat vertalen naar een verband op individueel niveau. Of dit komt doordat het eventuele verband te zwak is binnen een groep normale lezers of doordat semantische priming-effecten te variabel zijn, blijft nog de vraag.

Discussie

De drie studies lieten samen weinig significante verschillen tussen de eentalige en de meertalige kinderen zien. Het feit dat de meerderheid van de meertalige kinderen thuis opgroeide met zowel Nederlands als een minderheidstaal en dat ze allemaal geschoold en gealfabetiseerd waren in het Nederlands, speelt hierin waarschijnlijk een rol. De bijdragen van individuele verschillen in de structuur van het semantisch netwerk, semantische toegang en semantische priming aan begrijpend lezen waren ook beperkt. Het ontbreken van een significant verband tussen begrijpend lezen, semantische toegang en semantische priming in de gerapporteerde studies is mogelijk (gedeeltelijk) toe te schrijven aan methodologische kwesties, zoals het gebruik van hoogfrequente stimuli of de inherente instabiliteit van semantische priming-effecten. Echter, de afwezigheid van significante verbanden tussen semantische priming en begrijpend lezen op individueel niveau is tot nu toe consistent, wat suggereert dat dit woordenschataspect mogelijk niet de meest vruchtbare weg is voor toekomstig onderzoek en interventies. De drie studies in dit proefschrift bevestigen voornamelijk het belang van woordenschatgrootte en kennis van context-onafhankelijke relaties voor begrijpend lezen.

Dankwoord

Het is een cliché, maar: hoewel een promotie vaak een solo-project bij uitstek is, is het uiteindelijk, direct of indirect, een gezamenlijke inspanning. Dit proefschrift was er niet gekomen zonder de hulp van een heleboel mensen, en ik ben iedereen ontzettend dankbaar voor alle steun, advies, gezelligheid en liefde die ik onderweg ontvangen heb. Een aantal mensen wil ik graag persoonlijk noemen.

Allereerst ben ik veel dank verschuldigd aan mijn promotoren, Rob Schoonen en Jan Hulstijn. Zonder jullie betrokkenheid, zowel op academisch als persoonlijk vlak, was dit proefschrift er zeker niet gekomen. Rob, naast onze vele gezellige en leerzame discussies over stimuli, classificaties, analyses en nulresultaten, waardeer ik boven alles je oprecht betrokken, stimulerende en waar nodig pragmatische houding: ik streef ernaar die zelf ook toe te passen. Jan, een mailtje aan jou met een nieuw plan of een vers geschreven stuk kon altijd rekenen op heerlijk positief commentaar; bedankt dat je zo motiverend en begaan was.

Ik heb veel baat gehad bij het advies van zowel naaste als iets verdere collega's: in het bijzonder Sible Andringa, Judith Rispens, Paul Boersma, Huub van den Bergh, Simon de Deyne; de leden van CASLA, die meermaals hebben meegedacht over het kiezen tussen honderd experimentele kwaden; en Dirk-Jan Vet, die altijd klaarstond om te helpen bij het inprogrammeren van het uiteindelijk gekozen minste kwaad.

Taalontwikkelingsonderzoek kan niet van de grond komen zonder de hulp van schooldirecties, docenten en kinderen. Zonder hun bereidheid om tijd in te ruimen in drukke lesprogramma's of wéér een toets te maken, had ik niet de data kunnen verzamelen die het hart van dit proefschrift vormen. Vooral de vrolijkheid, nieuwsgierigheid en behulpzaamheid van de kinderen trof mij steeds weer. Maar, hoe leuk het testen ook kon zijn, ik ben óók dankbaar dat ik niet alle data alleen heb hoeven verzamelen en verwerken: gelukkig had ik hulp van Iris, Jasmijn, Dorothee, Mandy, en Floor.

In de afgelopen paar jaar heb ik me soms verbaasd over het grote aantal ontzettend leuke, lieve, fijne mensen dat ik heb leren kennen onder mijn collega-promovendi in het Bungehuis en het PCH. Zonder voorbij te willen gaan aan ieders persoonlijke kwaliteiten, durf ik wel te zeggen dat taalwetenschappers een bijzonder slag mensen zijn! Veel dank aan mijn PCH roomies – Tiffany, Imme, Merel, Iris D. en Iris B. – voor onze knusse kamer, de gezelligheid, en de ruimte om het soms even niet zo leuk te vinden; aan onze buren op de 6^{e}, in het bijzonder Klaas, Marloes, Vadim, Ulrika en Marieke, voor de lunches, koffieruns en de mogelijkheid om altijd binnen te lopen; en aan nog meer fijne (oud-)collega's, voor steun en gezelligheid in het verleden en het nu – Margreet, Sterre, Marjolein, Akke, Sanne, Marlou, Camille en Sanna. Een extra woord van dank gaat uit naar Tiffany en Klaas, mijn lieve paranimfen! Zonder jullie had ik het niet gered: dank voor jullie vriendschap, alle plezier, koffie, wijn, dansen, weekends, etentjes, advies, knuffels, soms het huilen, maar vooral het lachen!

Al het denkwerk dat bij een proefschrift komt kijken, heeft veel afleiding nodig als tegenhanger. Ik ben heel blij met mijn lieve vriendinnen bij wie ik niet alleen gezelligheid, maar ook veel peptalks en knuffels kon vinden, naast etentjes, films, concerten, nacho's met guacamole, vakanties, naailessen, terrasjes, Bossche bollen, etc. In willekeurige volgorde: dankjewel Maartje, Linda, Joyce, Trynke, Aletta en Denise, Anke en Renee, Lily, en Sofie. Twee extra bedankjes zijn ook hier op hun plaats. Joyce, de stoffenmarkt, de weekendjes, de spontane etentjes en taartjes waren altijd precies wat ik nodig had! Anke, Michiel en Ollie, dank jullie wel dat ik altijd bij jullie op de bank kan ploffen, jullie hebben me er doorheen gesleept de laatste maanden! En lieve Anke, dankjewel dat je zo dichtbij staat en dat we alles kunnen delen.

Last but not least: mijn familie. Familiebanden zijn iets heel bijzonders, en dat heb ik de afgelopen jaren extra duidelijk kunnen ervaren. Veel dank aan familie Spätgens voor de gezelligheid, hilariteit, en oprechte belangstelling. Mijn lieve, stoere Oma Sylvie, jouw optimisme, schrijfwerk en dapperheid zijn voor mij grote inspiratiebronnen; *laat ze het ons maar nadoen*! Lydia en Rutger, dank jullie wel dat jullie zo lief en betrokken zijn, en vaak een tweede thuis voor me waren. Taco en Marieke, dank dat jullie er waren toen het het hardst nodig was. En vooral, lieve pappa en mamma: als ik ga opsommen wat jullie allemaal voor mij hebben gedaan, zijn we zo vijf pagina's verder, dus ik ga voor kort maar krachtig. Jullie zijn de allerliefsten, en ik kan jullie nooit genoeg bedanken.

About the author

Tessa Marie Spätgens (Roermond, the Netherlands, 1987) pursued her BA and MA degrees in Linguistics at Radboud University Nijmegen, graduating cum laude from both programs. During her studies, she spent a semester in Nice, France, worked as a student assistant, and did internships at Radboud University Nijmegen and the Expertisecentrum Nederlands. With a grant from the Huygens Scholarship Program, she then continued her studies at the University of Edinburgh, Scotland, obtaining her MSc degree in Evolution of Language and Cognition with distinction.

Following her interests in language acquisition, multilingualism and experimental psycholinguistics, Tessa started her PhD research at the Amsterdam Center for Language and Communication (ACLC) under the supervision of Rob Schoonen and Jan Hulstijn, resulting in the present dissertation. During her PhD project, she presented her work at multiple international conferences, taught guest lectures and courses at the BA and MA level and was active as representative of the PhD researchers within the ACLC advisory board and as a member of the Werkverband Amsterdamse Psycholinguïsten.