

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/182098>

Please be advised that this information was generated on 2021-03-09 and may be subject to change.

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

TESSA SPÄTGENS

*University of Amsterdam*

ROB SCHOONEN

*Radboud University Nijmegen*

Received: July 16, 2016

Accepted for publication: July 11, 2017

## ADDRESS FOR CORRESPONDENCE

Tessa Spätgens, University of Amsterdam, Spuistraat 134, 1012 VB Amsterdam, The Netherlands.

E-mail: [t.m.spatgens@uva.nl](mailto:t.m.spatgens@uva.nl)

## 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 mental processing speed were tested by means of standardized tasks. Mixed-effects modeling 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 show that the monolingual and bilingual children perform similarly on all measures, suggesting that in the current Dutch context, language status may not be highly predictive of vocabulary knowledge and reading comprehension skill.

Keywords: lexical access; monolingual and bilingual children; reading comprehension; semantic network; semantic priming; vocabulary knowledge

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<sup>1</sup> 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, 2007; Perfetti & Hart, 2002) and confirmed by a number of empirical studies investigating lexical fluency and the semantic network (Qian, 1999; Tannenbaum, Torgesen, & Wagner, 2006; Verhallen & Schoonen, 1993). This could also be a source of differences in reading comprehension between monolingual and bilingual children, as young bilinguals have been found to lag behind their monolingual peers in terms of both knowledge of semantic relations (e.g., Cremer, Dingshoff, de Beer, & Schoonen, 2011; Verhallen & Schoonen, 1993) and reading comprehension (e.g., Centraal Bureau voor de Statistiek [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, 2003) 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 that occur together in context, the latter are more intrinsically motivated, existing between words that are related independent of context and that 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, as children need to abstract from direct experience to more generalized, decontextualized knowledge (Elbers, van Loon-Vervoorn, & van Helden-Lankhaar, 1993; Lin & Murphy, 2001; Nelson, 1977, 1982, 1985, 1991, 2007; Petrey, 1977). Bilinguals have been found to have generally less extensive semantic knowledge in their second language (L2) compared to monolinguals in their native first language (L1), for example, providing fewer semantically oriented word associations (Cremer et al., 2011). 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 et al. (2006) found that in monolinguals aged 9–10, the ability to provide synonyms and

multiple attributes such as category, function, and location for nouns, that is, 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- to 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 of 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 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.

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 that 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 the same 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 stimulus onset asynchrony, 800 ms in their study, compared to an interstimulus

interval (ISI) of 500 ms 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? Larkin, Woltz, Reynolds, and Clark (1996) used a semantic priming experiment involving a synonym judgment 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 also category coordinates.<sup>2</sup> Therefore, while the study by Larkin et al. 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 cause the different findings.

In this study, our first aim is to partially replicate and extend Cremer's (2013) and Nation and Snowling's (1999) findings 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 better 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.

The 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, Perfetti, & McKeown, 1982). We will use the term *semantic access*, or *access* for short, because *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, 2007; Perfetti & Hart, 2002) 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 to be devoted to higher level comprehension processes. Cremer (2013) found that semantic access as measured by response times in a semantic classification task could explain a small amount of variance in the reading comprehension scores of monolingual and bilingual readers, namely, 2%, in addition to the variance already explained by vocabulary size 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 (Cremer, 2013; Heesters, van Berkel, van der Schoot, & Hemker, 2007; Sijtsma, van der Schoot, & Hemker, 2002; Smits & Aarnoutse, 1997; van Berkel, van der Schoot, Engelen, & Maris, 2002). These consistent delays are found even though most primary 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 (Heesters et al., 2007; Sijtsma et al., 2002; van Berkel et al., 2002).

The weaker links hypothesis (Gollan, Montoya, Cera, & Sandoval, 2008; Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Gollan & Silverberg, 2001; 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 negatively affected in the development of context-independent semantic knowledge.

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. van Berkel et al., 2002; Verhallen & Schoonen, 1993) and may access semantic information more slowly (Cremer, 2013).<sup>3</sup> These findings, combined with the other findings that especially context-independent knowledge may be particularly important for reading comprehension (cf. Nation & Snowling, 1999), lead us to expect that these lower level 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.

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

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). In addition, 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, that is, 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 toward 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, Jones, Thomson, Kelly, and McRae (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.

Table 1. *Age and gender in monolingual and bilingual groups*

	<i>N</i>	Girls	Boys	Age (years; months)	
				<i>M</i>	<i>SD</i>
Monolingual	36	19	17	11;1	0;6
Bilingual	86	45	41	11;3	0;6
Total	122	64	58	11;3	0;6

## METHOD

### *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. Socioeconomic status (SES) in these neighborhoods was average to low (Sociaal en Cultureel Planbureau, 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 attention-deficit/hyperactivity disorder. One child had been diagnosed with both attention-deficit/hyperactivity disorder and dyslexia, a further 12 children had been diagnosed with dyslexia. The data from these children were removed. No other cases were reported. A further 9 children were not able to participate in all tasks or had missing data on some of the tasks. Finally, 7 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. Thirty-six 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 years, 3 months (11;3;  $SD = 0;6$ ), ranging from 10;4 to 12;6. Table 1 below shows the age and gender distributions across the monolingual and bilingual groups.

### *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 technical reading 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.

*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 this decision, 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.

*Priming experiment.* As was discussed earlier, an auditory semantic decision task was used to measure context-dependent and context-independent word knowledge. Forty prime-target pairs were made for the experiment, 20 for each semantic relation. These can be found in Table A.1 in Appendix A.

*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 *type of 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.

Two fully counterbalanced versions of the experiment were made, in which one half of the targets appeared in the related condition, the other half in the unrelated condition, and vice versa. Each participant thus encountered each target once. Unrelated control pairs were formed by repairing 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 2, with primes marked typographically to clarify the repairing 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 208 children from the same

Table 2. *Stimulus pairings per condition*

Condition	Context-Independent		Context-Dependent		
	Prime Target Type	Animate Coordinates	Inanimate Coordinates	Location Animate	Person Inanimate
Semantic decision		<i>creature–creature</i>	<i>thing–thing</i>	<i>thing–creature</i>	<i>creature–thing</i>
Related		<i>duck–goose</i>	<b>guitar</b> –piano	<i>forest–squirrel</i>	<b>thief</b> –purse
Control		<b>thief</b> –goose	<i>forest–piano</i>	<b>guitar</b> –squirrel	<i>duck–purse</i>

target population (Spätgens & Schoonen, 2017). Eighty 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 allows us to control for associations that are maybe not as immediate but still prevalent.

To form the prime-target pairs for the present experiment, the stimulus words from the association task were used as primes. The related targets never occurred as first associations in the data set, 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 5-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 or unrelated conditions and none of the critical pairs form 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 primary 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 that 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 Table B.1 in Appendix B.

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 an extra 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, and locations), was roughly equal. Finally, the ISI was 1000 ms. After a response, there was a blank screen for 500 ms, followed by a screen with a fixation point (+) for 500 ms, 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. RTs 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.

*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.

*Word decoding.* Technical reading skills were 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 1 min, 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 ( $r = .821$ ,  $p = .000$ ) and were therefore combined into one measure by averaging the scores for each child.

*Cognitive processing speed.* Cognitive processing speed was measured using the Rapid Automatized Naming Test (RAN; Denckla & Rudel, 1974) and Rapid

Alternating Stimulus Test (RAS; Wolf, 1986). 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 and the number of errors made is thus not incorporated. The test makers consider more than five errors or self-corrections to be “excessive” (e.g., over 10%; Wolf & Denckla, 2005), and a potential reason for retesting 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.

*Language interview.* To establish language dominance, a short questionnaire on linguistic background was done with each participant. The children were asked whether they were born in The Netherlands, from what age onward 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.

### *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 1 or 2 school days, depending on group size. In the morning on the first day, testing began with the reading task, which lasted about 35 min including instruction, followed by the vocabulary task, which took roughly 25 min 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: the semantic decision task, then word decoding, RAN, RAS, and finally the language interview. In all, the individual sessions took around 25 min. Before starting the experiment, the participants received a verbal instruction that 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 12 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.

Table 3. *Skewness and kurtosis values for main measures*

	Skewness		Kurtosis	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
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
Rapid automatized naming (ms)	1.201	0.219	2.050	0.435
Rapid alternating stimulus (ms)	0.715	0.219	0.329	0.435
Access (mean response time to fillers)	0.347	0.219	-0.164	0.435

After the practice items, they could ask more questions if needed, and then the experiment began. During the experiment, no feedback was provided.

### *Data handling and analysis*

RTs for inaccurate responses were set to missing (1,360 items, 5.6% of data). Then, the average RT for each child was calculated. RTs over 2.5 *SD* from the mean (the individual's means and standard deviations were used) were defined as outliers, and removed (614 items, 0.03% of data). RTs under 350 ms were removed so that only real responses and not accidental taps were recorded (15 items, <0.01% of data). We used 350 ms instead of the commonly used 250 ms (i.e., Betjemann & Keenan, 2008; Cremer, 2013) as the cutoff 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 100 ms.

Three children with accuracy scores under 85% and four children with mean RTs over 1700 ms 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, Mächler, Bolker, & Walker, 2015).

## RESULTS

### *Descriptives*

Skewness and kurtosis values for all main measures are reported in Table 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  $\alpha = 0.634$ ). Finally, the internal consistency of the vocabulary task was satisfactory, with Cronbach  $\alpha = 0.709$ .

Table 4. *Descriptives for task scores in the monolingual and bilingual groups*

	Total ( <i>N</i> = 122)		Monolingual ( <i>N</i> = 36)		Bilingual ( <i>N</i> = 86)		<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
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
Rapid automatized naming (ms)	22.05	3.65	21.96	3.14	22.10	3.86	0.04
Rapid alternating stimulus (ms)	30.24	5.63	29.74	5.79	30.44	5.59	0.12
Access (mean response time to fillers)	1149.7	160.67	1117.73	164.31	1163.09	158.17	0.28

Table 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  $d > 0.20$ , a small effect size (Cohen, 1969). However, none of these differences were found to be significant.

#### *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 RTs to the target items. Since the RTs to the target items were skewed and peaked (skewness = 1.722,  $SE = 0.037$ ; kurtosis = 4.91,  $SE = .074$ ), 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  $2 \times 2 \times 2$  design: Relatedness (0 related, 1 unrelated)  $\times$  Relationship Type (0 context-independent, 1 context-dependent)  $\times$  Animacy (0 inanimate, 1 animate). 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 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 RTs to the individual targets. Since this variable was used as a control variable, it will not be discussed any further here.

Table 5. *Fixed and random effects estimates for the overall priming model (4,415 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)	<i>t</i>
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

*Note:* Absolute *t* values over 2 are considered significant.

The results indicate that there are two positive main effects among the three dichotomous predictors. First, there is a main effect of relatedness. This indicates that overall, RTs 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.

Second, there is a positive main effect of animacy. Here, the inanimate items yield a higher, and therefore longer, RT. In other words, identifying objects took participants longer than identifying animates, even though the instruction for the experiment was focused on making the semantic decision for both types of items as similar as possible. Perhaps this difference occurs because the set of inanimate items to search through in the mental lexicon is larger, or because participants 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?” Nevertheless, 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 RTs. 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 very similar.

Of the four interactions that were tested, only the interaction between relatedness and relationship type was significant. [Figure 1](#) shows that the overall priming effect is due to a large priming effect for the context-independent items, while the

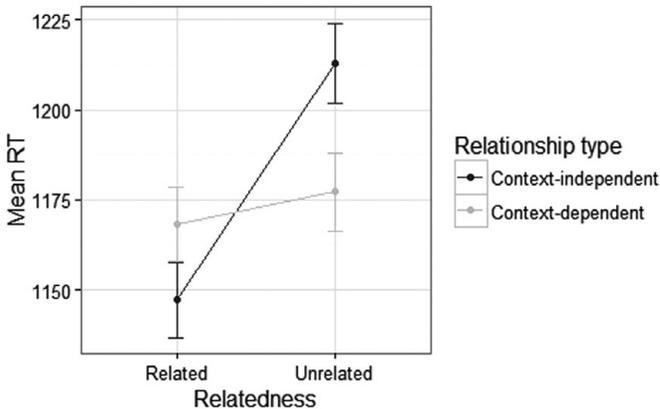


Figure 1. Mean response times by relatedness and relationship type (error bars represent standard error).

difference between unrelated and related items in the context-dependent condition is much smaller. 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.

### *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-dependently related primes were analyzed separately.<sup>4</sup> Again, the log transformed RT was modeled with random intercepts for classes, subjects, and items. As fixed effects, language group and relatedness and their interaction were included, and mean

Table 6. *Fixed and random effects estimates for the monolingual and bilingual priming models (N = 122)*

Random Effects	Context-Independent Priming (2,136 Items)	Context-Dependent Priming (2,279 Items)
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)	<i>t</i>	Estimate (SE)	<i>t</i>
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

*Note:* Absolute *t* values over 2 are considered significant.

RT was entered as a covariate. Table 6 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 condition 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.

#### *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 items. 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 RTs 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 RTs on the unrelated items. By estimating the priming scores in this way, rather than subtracting mean RTs on the related items from mean RTs 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,

Table 7. Summary of individual priming scores on the context-independent items, with and without transformation ( $N = 122$ )

Priming Based on	First Quartile	Mean ( <i>SD</i> )	Third Quartile
Log transformed RTs	-0.00024	0.01359 (0.02596)	0.02773
Untransformed RTs	-1.61171	6.86388 (14.35792)	15.78807

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 RT on the related items) and the individual slopes (i.e., how much the average RT to the unrelated items deviates from the average RT to the related items). This was evidenced by the model collapsing onto perfectly correlated random intercepts and slopes. This is potentially because 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 7. Note that the numbers are very small due to the log transformation of the RTs. Table 7 also shows a summary of the individual scores when calculated by the same model but with untransformed RTs, as an illustration of what the individual priming scores would be in that case. However, due to the skewness and peakedness of the RTs, these numbers should be interpreted with caution.

#### *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 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 1,000, respectively, to make sure the values of all variables were on comparable scales. Two children were removed from the data set 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.<sup>5</sup> 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.<sup>6</sup>

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, and processing speed) were entered as fixed effects. As we can see in Table 8, only the vocabulary size measure is associated with the reading scores in this case, with children who scored 1 point above the mean

Table 8. *Fixed and random effects estimates for the reading comprehension models (N = 120)*

Random Effects	1		2		3	
Variance between classes	0.5145		0.5649		0.4048	
Residual variance	10.0317		10.1424		10.4018	
Fixed Effects	Estimate (SE)	<i>t</i>	Estimate (SE)	<i>t</i>	Estimate (SE)	<i>t</i>
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
Rapid automatized naming	-0.138 (0.104)	-1.3	-0.136 (0.104)	-1.3	-0.13 (0.107)	-1.2
Rapid alternating stimulus	-0.023 (0.071)	-0.3	-0.024 (0.075)	-0.3	-0.041 (0.078)	-0.5
Access (mean response time 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 ( <i>ns</i> )		1.86 ( <i>ns</i> )	
Difference <i>df</i>			2		5	
Compared to model			1		1	

*Note:* Absolute *t* values over 2 are considered significant.

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 was significantly associated with the reading scores. Therefore, it is not the case that there is an effect of access or priming that is filtered out by the control tasks. Finally, to compare our results to Nation and Snowling's (1999), we divided the group into 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 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.

## DISCUSSION

### *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 (for an overview, see Lucas, 2000), 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 nonassociative context-independent semantic priming.

That no overall context-dependent priming effect was found is likely to be because we controlled very strictly for association strength to make sure that we were tapping into purely semantic connections. The word association data that were utilized (Spätgens & Schoonen, 2017) 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 that 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 (e.g., Nation & Snowling, 1999, but also Hare et al., 2009, who tested very similar context-dependent pairs, including *location-person/animal*).

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 (2008). 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 context-dependent relations being supported by the co-occurrence of concepts in experience.

A methodological point that may be of use for future studies on semantic priming is 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, that is, 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.

#### *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 et al., 2005, 2008; Gollan & Silverberg, 2001; Michael & Gollan, 2005), suggesting that the hypothesis may 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 hard 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

primary 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 because all bilingual children in this study had gone to Dutch schools from age 4 onward, and the vast majority was born in The Netherlands. In addition, 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 education report (Kuhlemeier et al., 2014) shows that when SES is taken into account, differences between monolinguals and bilinguals disappear, in line with our findings. 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. Large-scale national studies examining reading comprehension and vocabulary of Dutch primary school children have found that 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 Centraal Bureau voor de Statistiek, 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.

### *Reading comprehension and the influence of access and priming*

The analyses of the reading comprehension scores were done in three steps: looking at 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, 2012). The cognitive processing tasks (RAN and RAS) were mainly included as 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, Necochea, & 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 that 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.

Our priming results do corroborate Cremer's (2013), who used a stimulus onset asynchrony 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 (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 modeling 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 priming scores reflecting automatic processing may be too noisy for use in individual differences studies. 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, because 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%, 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 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 the text 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. The relation may be bidirectional to some degree. However, as we have argued, 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.

## APPENDIX A

Table A.1. *Critical stimuli in the semantic priming experiment*

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

APPENDIX B

Table B.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									
249	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

Table B.1 (*cont.*)

	All Items ( $n = 10$ )		Half 1 ( $n = 5$ )		Half 2 ( $n = 5$ )		Comparison 1 & 2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	$t (df = 8)$	$p$
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 are measured in milliseconds. Relatedness represents mean ratings of semantic relatedness of the pairs on a 5-point Likert scale by 33 adult native speakers.

APPENDIX C

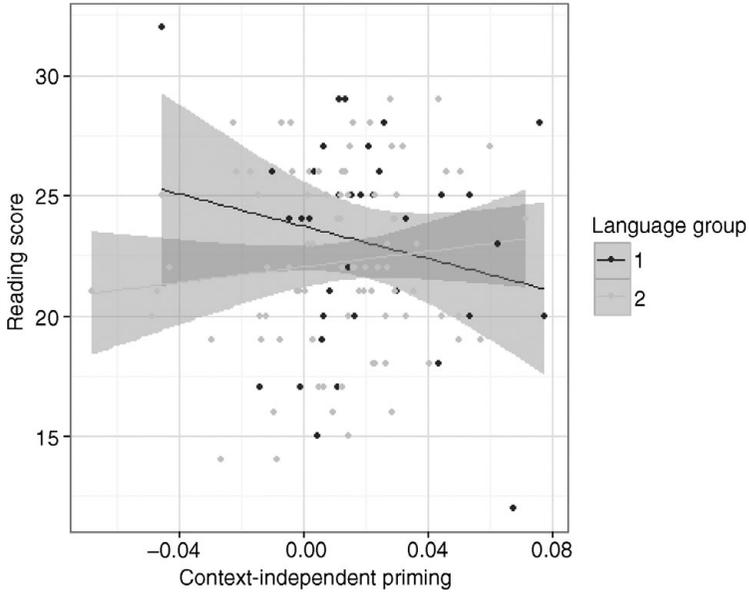


Figure C.1. Scatter plot with trend line and 95% confidence interval (gray area) of context-independent priming and reading scores by language group, including bivariate outliers ( $N = 122$ ).

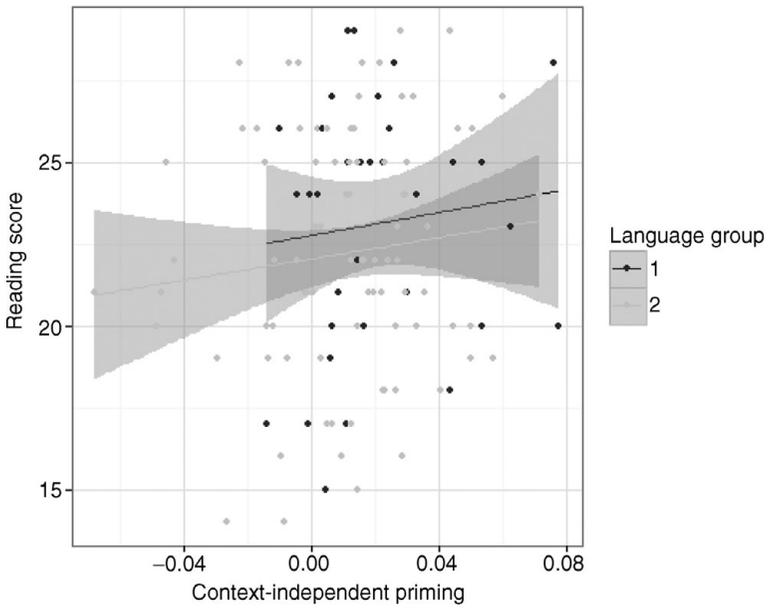


Figure C.2. Scatter plot with trend line and 95% confidence interval (gray area) of context-independent priming and reading scores by language group, excluding bivariate outliers ( $N = 120$ ).

## ACKNOWLEDGMENTS

This paper benefited from the help of many who we thank here. This research would not have been possible without the contribution of time and effort by the children, teachers, and directors from the participating schools. Furthermore, Dorothee Bliem and Mandy Luiten collected and processed the bulk of the data as part of their internships. Valuable advice on statistics was provided by Huub van den Bergh and Paul Boersma, and the research benefited from many discussions with Jan Hulstijn. Finally, the comments provided by the editor and three anonymous reviewers helped to improve and clarify the paper.

## NOTES

1. The bilingual minority children in our study can be characterized as successive bilinguals, learning a minority language as their L1 and Dutch as their L2. We will refer to them as bilinguals throughout this paper.
2. Nobre and 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 setup 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.
3. Gollan et al. (2005) found similar semantic classification times for pictures in bilingual adults in their dominant language and monolinguals, 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.
4. 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.
5. 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 Figures C.1 and C.2 in Appendix C.
6. 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.

## REFERENCES

- Aarnoutse, C., & Kapinga, T. (2006). *Begrijpend lezen 678* [Reading comprehension grades 456]. Ridderkerk, The Netherlands: 678 Onderwijsadvisering.
- Aitchison, J. (2003). *Words in the mind: An introduction to the mental lexicon* (3rd ed.). Malden, MA: Blackwell.
- Alderson, J. C. (2005). *Assessing reading*. Cambridge: Cambridge University Press.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. doi:10.18637/jss.v067.i01

- 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*, 506–521.
- Betjemann, R. S., & Keenan, J. M. (2008). Phonological and semantic priming in children with reading disability. *Child Development, 79*, 1086–1102.
- 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*, 309–328.
- Centraal Bureau voor de Statistiek. (2014). *Jaarrapport integratie 2014* [Year report integration 2014]. Den Haag/Heerlen, The Netherlands: Author.
- Centraal Bureau voor de Statistiek. (2016). *Jaarrapport integratie 2016* [Year report integration 2016]. Den Haag/Heerlen/Bonaire, The Netherlands: Author.
- 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*, 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*, 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*, 1195–1217.
- De Deyne, S., & Storms, G. (2008). Word associations: Network and semantic properties. *Behavior Research Methods, 40*, 213–231.
- Denckla, M. B., & Rudel, R. (1974). Rapid “automatized” naming of pictured objects, colors, letters and numbers by normal children. *Cortex, 10*, 186–202.
- Di Filippo, G., Brizzolaro, 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*, 349–361.
- Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly, 38*, 78–103.
- 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 Press.
- Erdos, C., Genesee, F., Savage, R., & Haigh, C. A. (2011). Individual differences in second language reading outcomes. *International Journal of Bilingualism, 15*, 3–25.
- 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*, 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*, 1220–1234.
- Gollan, T. H., & Silverberg, N. B. (2001). Tip-of-the-tongue states in Hebrew–English bilinguals. *Bilingualism: Language and Cognition, 4*, 63–83.
- 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*, 151–167.
- Heesters, K., van Berkel, S., van der Schoot, F., & Hemker, B. (2007). *Balans van het leesonderwijs aan het einde van de basisschool 4* [Balance of reading education at the end of primary school 4] (No. 33 PPON series). Arnhem, The Netherlands: Cito.
- Kliegl, R., Masson, M. E., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition*, *18*, 655–681.
- 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 primary and special education 2] (No. 54, PPON series). Arnhem, The Netherlands: Cito.
- Larkin, A. A., Woltz, D. J., Reynolds, R. E., & Clark, E. (1996). Conceptual priming differences and reading ability. *Contemporary Educational Psychology*, *21*, 279–303.
- Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General*, *130*, 3–28.
- Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review*, *7*, 618–630.
- McNamara, T. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York: 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.
- Michael, E. B., & Gollan, T. H. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In J. R. Kroll & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic Approaches* (pp. 389–407). New York: Oxford University Press.
- 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.
- Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: A review of research and theory. *Psychological Bulletin*, *84*, 93–116.
- Nelson, K. (1982). The syntagmatics and paradigmatics of conceptual development. In S. Kuczaj (Ed.), *Language development: Vol. 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: 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*, 753–770.
- Olkonen, 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*, 554–566.
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*, 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: John Benjamins.
- Petrey, S. (1977). Word associations and the development of lexical memory. *Cognition*, *5*, 57–71.

- 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, 282–307.
- R Core Team. (2015). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
- 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.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime reference guide*. Pittsburgh, PA: Psychology Software Tools Inc.
- Schoonen, R., & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing*, 25, 211–236.
- Schrooten, W., & Vermeer, A. (1994). *Woorden in het basisonderwijs. 15.000 woorden aangeboden aan leerlingen* [Words in primary education. 15,000 words presented to pupils]. Tilburg: Tilburg University Press.
- Sociaal en Cultureel Planbureau. (2015). *SCP statusscores* [SCP SES scores]. Retrieved from [http://www.scp.nl/Onderzoek/Lopend\\_onderzoek/A\\_Z\\_alle\\_lopende\\_onderzoeken/Statusscores](http://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores)
- Sijstra, J. M., van der Schoot, F., & Hemker, B. T. (2002). *Balans van het taalonderwijs aan het einde van de basisschool 3* [Balance of language education at the end of primary school 3] (No. 19, PPO series). Arnhem, The Netherlands: 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 autochthonous and minority children]. *Nederlands Tijdschrift Voor Opvoeding, Vorming En Onderwijs*, 13, 33–52.
- Spätgens, T., & Schoonen, R. (2017). *The structure of developing semantic networks: Evidence from single and multiple word associations in monolingual and bilingual readers*. Manuscript submitted for publication.
- Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, 36, 139–152.
- Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: 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, 284–336.
- Swanson, H. L., Trainin, G., Necochea, D. M., & Hammill, D. D. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlation evidence. *Review of Educational Research*, 73, 407–440.
- 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, 381–398.
- van Berkel, S., van der Schoot, F., Engelen, R., & Maris, G. (2002). *Balans van het taalonderwijs halverwege de basisschool 3* [Balance of language education halfway through primary school 3] (No. 20, PPO series). Arnhem, The Netherlands: Cito.
- Verhallen, M., & Schoonen, R. (1993). Lexical knowledge of monolingual and bilingual children. *Applied Linguistics*, 14, 344–363.
- Verhoeven, L. (1992). *Drie-minuten-toets: Handleiding* [Three-Minute-Test]. Arnhem, The Netherlands: Cito.
- Verhoeven, L., & van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology*, 22, 407–423.
- Verhoeven, L., & van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Reading and Writing*, 25, 1805–1818.

- Verhoeven, L., & Vermeer, A. (1995). *Leerlingvolgsysteem 5678: Leeswoordenschat* [Monitoring and Evaluation System Grades 3456: Reading vocabulary]. Arnhem, The Netherlands: Cito.
- 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, TX: Pro-Ed.
- 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, 375–386.
- 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: From methods to insights* (pp. 203–226). New York: Psychology Press.