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

Cremer, M.

Citation for published version (APA):
Chapter 5

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

The findings of Chapter 4 identified speed of access as a factor in children’s reading performance. In this chapter\textsuperscript{10}, we will again address availability of semantic word knowledge (research question 3) and speed of access (research question 4). Speed of access will be investigated in terms of lexical decision speed and semantic classification speed. In addition, we will go one step further and investigate the activation of abstract, semantic word knowledge as reflected by semantic priming (research question 5). Finally, the contributions of semantic word knowledge, speed and priming to reading comprehension will be analysed (research question 6). The chapter starts with a discussion of relevant previous research (section 5.1), followed by a discussion of the research methodology used (section 5.2). The results are presented in section 5.3 and discussed in section 5.4.

5.1 Background

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

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

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

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

5.1.1 Semantic representations and processing

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

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

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

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

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

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

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

5.1.2 Monolingual and bilingual children

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

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

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

5.2 Method

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

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

Reading comprehension

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

Semantic word knowledge

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

---

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

**Word decoding**

To control for differences between children in word decoding skill the ‘Drie Minuten Test’ was administered, as in Chapter 4. It has a good reliability and validity (Verhoeven, 1992). As in Chapter 4, children were asked to read two word lists out loud. The lists consist of regular, non-related words increasing in decoding difficulty. The first list consists of 150 monosyllabic words; the second list consists of 120 multisyllabic words: 60 two-, 30 three-, and 30 four-syllable words. Children were given a break between the two lists. For each list the score is the number of words read correctly in one minute. In this study, the scores for the two word decoding lists are strongly correlated (Pearson’s $r = .822$) and show a high split-half reliability of $r_{1,2} = .90$; therefore, the two list scores are combined (averaged) to derive a single score for word decoding which is used in further analyses. Mean word decoding scores are normally distributed: skewness is -.156 (se .212) and kurtosis -.054 (se .420).
Lexical decision

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

The critical stimuli consisted of 24 semantically related, non-associated prime-target pairs (see Appendix B and C for the Dutch stimuli and an English translation). Ten were coordinately related (e.g., *nose* – *ear*) and fourteen were subordinately related (e.g., *knife* – *cutlery*). For the construction of test pairs, the word association norms collected in the study reported on in Chapter 3 were used. The 24 prime words were stimulus words from the word association study; the 24 targets had never occurred as response. Thus, association strength for prime-target pairs was close to zero. To ensure children’s familiarity with the stimuli, all stimuli used were high-frequent nouns checked in the child school-language corpus 12 (Schrooten & Vermeer, 1994). Word frequency and word length in terms of numbers of letters were balanced across conditions and versions (see Table 5.1). Care was taken to avoid pairings in which prime and target were orthographically or phonologically related.

---

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

<table>
<thead>
<tr>
<th></th>
<th>PRIMED</th>
<th>CONTROL</th>
<th>Non-words (51)</th>
<th>Fillers (27)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Primes (12)</td>
<td>Targets (12)</td>
<td>Primes (12)</td>
<td>Targets (12)</td>
</tr>
<tr>
<td></td>
<td>fre</td>
<td>lett</td>
<td>fre</td>
<td>lett</td>
</tr>
<tr>
<td>version 1</td>
<td>310</td>
<td>4.5</td>
<td>266</td>
<td>5.3</td>
</tr>
<tr>
<td>version 2</td>
<td>300</td>
<td>5.5</td>
<td>156</td>
<td>6.6</td>
</tr>
</tbody>
</table>

|        | fre  | lett | fre  | lett | fre  | lett | fre  | lett | fre  | lett |
|        | 5.9  | 124  | 6.2  |  |

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

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

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

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

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

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

**Semantic classification**

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

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

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

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

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

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

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

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

**Language background**

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

**5.2.3 Procedure**

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

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

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

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

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

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

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

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

<table>
<thead>
<tr>
<th>Measures</th>
<th>Total (N=130)</th>
<th>Monolingual (N=83)</th>
<th>Bilingual (N=47)</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>23.3</td>
<td>4.6</td>
<td>24.9</td>
<td>3.8</td>
</tr>
<tr>
<td>WAT</td>
<td>20.7</td>
<td>4.9</td>
<td>22.3</td>
<td>3.7</td>
</tr>
<tr>
<td>Word Decoding</td>
<td>86.7</td>
<td>15.2</td>
<td>86.2</td>
<td>14.7</td>
</tr>
<tr>
<td>Lexical decision speed</td>
<td>828</td>
<td>134</td>
<td>805</td>
<td>114</td>
</tr>
<tr>
<td>Semantic classification speed</td>
<td>862</td>
<td>150</td>
<td>825</td>
<td>130</td>
</tr>
<tr>
<td>Priming(_{LDT})</td>
<td>39.8</td>
<td>88.2</td>
<td>38.7</td>
<td>80.6</td>
</tr>
<tr>
<td>Priming(_{SCT})</td>
<td>70.7</td>
<td>142.2</td>
<td>59.7</td>
<td>136.6</td>
</tr>
</tbody>
</table>

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

5.3.2 Semantic priming

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

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

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

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

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

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

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

**Semantic classification task**

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

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

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

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

### 5.3.3 Monolingual and bilingual children

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

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

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

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

5.3.5 Contributions to reading comprehension

**Lexical decision task**

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

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

Note: * significant at the 0.05 level.

**Semantic classification task**

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

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

Note: * significant at the 0.05 level

Language background
To examine to what extent differences between monolingual and bilingual minority children are accounted for by differences in speed and priming scores, we conducted additional regression analyses of reading comprehension. When entered into the regression equation after decoding, semantic word knowledge, semantic classification speed and priming, the effect of language background (i.e., monolingual or bilingual) was still significant (see Table 5.7) but considerably reduced. On its own, language background accounted for 21.7% of explained variance ($\Delta R^2 = .217$; $F_{change}(1,128)=35.58; p = .000$).
A regression analysis with interaction terms was carried out to determine whether the contribution of speed and priming to reading comprehension is different for monolingual and bilingual minority children (i.e., whether speed and priming interact with language background). Inclusion of the interaction term between language background and speed or priming, either with the lexical decision or semantic classification variables, in the final step of the model did not account for additional variance, showing that the effect of speed or priming is not significantly different for the language groups.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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