Semantic Processing and the Development of Word-Recognition Skills: Evidence from Children with Reading Comprehension Difficulties

Kate Nation and Margaret J. Snowling
University of York, York, United Kingdom

Children with specific reading comprehension difficulties were compared with control children on tests of language skill. The two groups performed at a similar level on tests requiring predominantly phonological skills, but the poor comprehenders performed less well on tests tapping semantic ability. Although the two groups were matched for decoding ability (as assessed by nonword reading), the poor comprehenders were worse at reading words with irregular spelling patterns and low-frequency words. These results show that despite having adequate phonological decoding skills, poor comprehenders have problems reading words that are typically read with support from semantics. These findings are related to connectionist models of reading development in which phonological and semantic processes interact.

Traditional accounts of reading development have stressed the importance of phonological awareness to the acquisition of reading skill (Goswami & Bryant, 1990; Gough, Ehri, & Treiman, 1992). Such models have been silent as to the role of semantic skills and other language resources in learning to read, although a number of studies have reported that dyslexic children have primary deficits in phonological but not semantic processing (Shankweiler et al., 1995; Share & Stanovich, 1995).

However, not all poor readers have difficulties at the level of word recognition. Gough and Tunmer (1986) proposed that reading comprehension could be described as the product of two skills, decoding ability and linguistic comprehension. Both of these skills were necessary, but neither alone was sufficient for successful reading. Gough and Tunmer’s argument implies that children can have difficulties with either of these fundamental aspects of reading. While dyslexic children often comprehend in advance of their ability to decode (Frith & Snowling, 1983), children who have specific difficulties with reading comprehension have normal word-recognition skills (Oakhill, 1982, 1984; Stothard & Hulme, 1992, 1995). In this paper, we investigate the language and reading skills of children with reading comprehension difficulties. Just as studying dyslexic children has increased theoretical understanding of the role of phonological representations in reading development (e.g., Snowling & Hulme, 1994), so studying children with comprehension impairments should elucidate the role of semantic processes in reading development.

The majority of studies of children with reading comprehension problems have focused on written text-level processes. Children with comprehension problems are poorer than controls matched for reading accuracy on tasks tapping anaphoric reference, sentence integration, and inference making skills. They also show a range of metacognitive difficulties including problems with comprehension mon-
itoring (for a detailed review, see Yuill & Oakhill, 1991). However, Stothard and Hulme (1992, 1995) found that poor comprehenders have poorer listening comprehension, lower verbal ability and weaker receptive language skills than skilled comprehenders matched for chronological age and decoding ability. These findings point to a more general language processing impairment.

If children with comprehension difficulties have poor verbal skills, it is important to consider how their weaknesses will impact upon reading development. Vellutino and colleagues (Vellutino & Scanlon, 1985; Vellutino, Scanlon, & Spearing, 1995) have highlighted the importance of semantic coding in early reading development. In paired associate learning for example, children more readily learn to associate ideographic characters with verbal labels that are high in referential meaning. It may follow that children with verbal-semantic impairments will experience difficulty on such tasks. Thus, it is plausible that poor semantic knowledge will constrain the early development of a sight vocabulary. Consistent with this hypothesis, Perfetti and Hogaboam (1975) observed that children with comprehension difficulties were slower at reading single words, especially low-frequency words and nonwords, than normal readers. Similarly, poor comprehenders also had difficulty accessing the meanings of words in a categorization task (Perfetti, Hogaboam, and Bell, cited in Perfetti & Lesgold, 1979). However, these findings are difficult to interpret because neither study included an assessment of the decoding ability of the good and poor comprehenders. This leaves open the possibility that the poor comprehenders were generally poor readers, i.e., that they had impaired decoding and comprehension. To address this issue more clearly, we investigated the speed and accuracy of single-word processing in children with specific comprehension difficulties in whom poor decoding skills as a factor affecting performance could be ruled out.

The investigation of semantic skills in children who vary in reading proficiency raises the important question of how semantic skills interact with critical aspects of phonological processing during the course of reading development. One framework in which to consider this issue is offered by parallel distributed processing (PDP) models. In their connectionist model of word recognition, Seidenberg and McClelland (1989) demonstrated that single-word pronunciation can be reasonably captured by a system that learns to map between orthographic and phonological representations via a set of hidden units. Following training, the implemented model was able to compute the pronunciation of nonwords, consistent words, and some high-frequency exception words without the need for explicit rules or word-level representations. However, as acknowledged by Seidenberg and McClelland, such a system is incomplete because knowledge of word meanings is essential for skilled reading. For example, in order to read homographs such as bass or lead, semantic as well as phonological knowledge is required. Moreover, the performance of the model when reading nonwords was significantly less good than that of normal skilled readers (Besner, Twilley, McCann, & Seergobin, 1990).

In a more recent connectionist account, Plaut, McClelland, Seidenberg, and Patterson (1996) hypothesized that the interaction between phonology and semantics may be best characterized as a “division of labour” between a phonological process which deals with mappings between orthographic and phonological representations and an interacting semantic process which deals with mappings between semantic, phonological, and orthographic representations. Consistent with Plaut et al., we refer to these two processes as the phonological pathway and the semantic pathway, respectively. However, the use of the term pathway does not imply separate routes analogous to those described in dual route models employing rules and stored lexical knowledge (see Plaut et al. (1996) for an account of the differences between their model and dual route accounts).

As evidence for this interaction between phonological and semantic pathways, Plaut et al. demonstrated that a phonological network
trained with additional support from a semantic pathway resembled skilled reading more closely than a network trained without semantics. Their simulations show that, as learning continues, there is a redistribution of labor between phonological and semantic processes. Initially, the combined model learns to pronounce words and nonwords in much the same way as the phonology-alone model does. However, the accuracy of the combined model accelerates considerably as the contribution from semantics increases. In turn, the model comes to depend less on orthographic-phonological mappings for the pronunciation of words, especially exception words, as these are more economically processed via the semantic input. This causes the weights on the connections in the phonological pathway to become smaller (as the amount of error is reduced due to the effects of the semantic pathway). As a consequence, the phonological pathway becomes more adept at pronouncing consistent spelling-sound correspondences and less accurate at pronouncing exceptional or irregular spelling-sound patterns. Thus, with extended reading experience, the phonology-alone model and the phonological pathway in the combined model were shown to operate in different ways; in the combined model, support from semantics allows the phonological pathway to become increasingly specialized for the pronunciation of consistent spelling-sound correspondences. This does not occur in the phonology-alone model as the weights on the connections have to stay large in order that exception words are pronounced correctly. The pronunciation of novel words suffers as a consequence.

Perhaps the clearest demonstration of the way in which the development of the phonological pathway is influenced by the semantic pathway arises when the network is lesioned in order to model acquired surface dyslexia. Patients with surface dyslexia typically show intact regular word and nonword reading, but impaired exception word reading (Patterson, Marshall, & Coltheart, 1985). Interestingly, some patients with semantic dementia show a surface dyslexic reading profile (e.g., Graham, Hodges, & Patterson, 1994). In earlier models containing only orthography-phonology connections (e.g., Seidenberg & McClelland, 1989), lesioning this pathway did not result in severely and selectively impaired exception word reading. Plaut et al. (1996) simulated acquired surface dyslexia by lesioning the semantic pathway, thus leaving the phonological pathway to operate in isolation. This simulation did cause exception word reading to become selectively impaired. They reasoned that this pattern of reading performance reflects the “normal operation of a phonological pathway that is not fully competent on its own because it learned to rely on support from the semantic pathway (which is subsequently impaired by brain damage)” (Plaut et al., 1996, p. 100).

To summarize, Plaut et al.’s model shows how semantics and phonology interact and how the nature of this interaction may change as a function of learning and reading experience. Their simulations suggest that, during normal development, the semantic pathway causes the phonological pathway to operate differently to the way that same pathway would operate in a model trained without semantics. With this model as a backdrop, an important question to ask is whether word recognition develops atypically in children who have poor semantic processing skills. If, in the normal course of reading development, the semantic pathway supplements the operation of the phonological pathway, then children whose semantic representations are in some way under-specified or degraded may show a division of labor that is more heavily reliant on the phonological pathway than children who have well-specified or rich semantic representations. According to this view, individual differences in semantic processing skill will influence word recognition. In particular, we would expect to see differences in the reading of those words that are ordinarily read with heavy support from the semantic pathway, namely, low-frequency exception words.

In this paper, we set out to test two major predictions. First, we sought evidence that
children with specific reading comprehension difficulties do indeed show impairments on tasks that require semantic processing. In contrast, we predicted they would show normal-range performance on tasks which require phonological processing (cf. Stothard & Hulme, 1995). Clearly, adequately specified semantic representations at the single-word level are crucial to reading comprehension. In addition, we predicted that individual differences in semantic processing ability would influence word recognition. As poor comprehenders have good decoding skills, they should have little difficulty reading regular words or high-frequency words. However, if they have semantic weaknesses, poor comprehenders will have difficulty reading orthographically inconsistent words that are not easily identified via simple decoding. Thus, a prediction tested in Experiment 3 is that poor comprehenders will be less accurate at pronouncing words that place greater demands on semantic skills, namely low-frequency exception words.

SAMPLE CHARACTERISTICS

Selection Procedures

We used a reading-age matched design to compare children who have specific reading comprehension difficulties with a group of skilled comprehenders matched for decoding ability, chronological age, and nonverbal ability. This design allows us to discount concomitant problems with decoding as a potential explanation of group differences in performance (cf. Perfetti, 1985). The initial phase of this research involved assessing the reading skills of 172 children (ages between 8 years, 6 months and 9 years, 6 months). All of the children attended the same school serving a socially mixed catchment area in the city of York, with the majority of the children being from lower middle class families. This screening phase involved assessments of basic decoding and text reading skills and an assessment of nonverbal ability.

Decoding ability. This was assessed using the Graded Nonword Reading Test (GNWRT; Snowling, Stothard, & McLean, 1996) as, arguably, reading novel words is the purest measure of decoding skill. This test is graded and provides norms for children of ages between 6 and 11 years.

Text reading accuracy and comprehension. In the Neale Analysis of Reading Ability—Revised (Neale, 1989), children read aloud short passages of text and are then asked a number of questions, some of which could be answered correctly using verbatim memory while others required inferences to be made. Two reading-age equivalents can be calculated: a reading accuracy age and a reading comprehension age.

Nonverbal skills. Nonverbal reasoning skills were assessed using the Matrix Analogies Test (short form, Naglieri, 1985). This has a multiple-choice format and assesses pattern completion, reasoning by analogy, serial reasoning, and spatial visualization skills.

Children were recruited into the study according to the following criteria. All of the normal readers had at least average-for-age nonword reading, reading accuracy, and reading comprehension. Children in the poor comprehender group were carefully matched with normal readers for nonword reading and nonverbal ability. However, their reading comprehension was at least one year below the expected level.

In previous studies of poor comprehenders by Oakhill and colleagues (for review see Yuill & Oakhill, 1991), children with specific reading comprehension difficulties were equated with normal readers for Neale Analysis reading accuracy age. However, Nation and Snowling (1997), following an investigation of the performance of 184 children of ages between 7 and 9 years on the Neale Analysis of Reading Ability, reported that text reading accuracy as measured by this test is influenced by comprehension skills. Thus, children with severe comprehension impairments will necessarily show poorer text reading than would be predicted on the basis of their basic-level decoding (which was largely independent of comprehension). In the present study, therefore, it was decided to equate the
two groups for decoding skill (nonword reading) instead of text reading accuracy. However, to make sure that no "garden-variety" type poor readers were selected, all children included in this study obtained at least age-appropriate reading accuracy as measured by the Neale Analysis.

To confirm that children with specific reading comprehension difficulties have impairments in semantic knowledge, we administered the Test of Word Knowledge (TOWK; Wiig & Secord, 1992). This is a standardized test of receptive and expressive vocabulary valid for children of ages between 5 and 17 years. Four core subtests from the TOWK were used, two receptive and two expressive. Using these four tests allowed composite receptive, expressive, and total language standard scores to be calculated for each child.

Receptive vocabulary: Synonyms. In this test, children see and hear a target word presented alongside four distracters which consist of antonyms of the target, members of the same semantic class, or words related by association. The child’s task is to select the synonym of the target word, for example, REAL (alternative choices: sturdy; actual; amazing; mere).

Receptive vocabulary: Figurative usage. This test requires children to match figurative expressions to meanings. It has a multiple-choice format, for example, "What does pat on the back mean?" (alternative choices: bother someone; be sneaky; turn around; give a compliment).

Expressive vocabulary: Word definitions. A word is presented to the child in both written and spoken format and they must provide a definition of the word. The definitions are scored to measure the child’s understanding of category membership and semantic feature knowledge.

Expressive vocabulary: Multiple contexts. This tests children’s ability to provide two meanings or contexts for a given target word. For example, the target BAT has two meanings (the animal and the object one hits a ball with).

### TABLE 1

<table>
<thead>
<tr>
<th>Performance of Poor Comprehenders and Normal Readers on Selection Tests and Tests of Word Knowledge (TOWK)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Poor comprehenders</strong></td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td><strong>n = 16</strong></td>
</tr>
</tbody>
</table>
| Chronological age
  \(M\) | 9.33 | 9.22 |
| SD | 0.68 | 0.66 |
| Nonword reading
  \(M\) | 18.50 | 18.13 |
| SD | 1.71 | 1.67 |
| Reading comprehension
  Age equivalent
  \(M\) | 7.18 | 10.59 |
| SD | 0.96 | 1.01 |
| Standard score
  \(M\) | 82.88 | 104.44 |
| SD | 6.64 | 5.05 |
| Reading accuracy
  Age equivalent
  \(M\) | 10.40 | 11.29 |
| SD | 1.26 | 1.36 |
| Standard score
  \(M\) | 105.69 | 112.06 |
| SD | 6.09 | 6.34 |
| Matrix analogy
  test \(M\) | 103.38 | 105.69 |
| SD | 14.60 | 15.23 |
| TOWK receptive
  \(M\) | 90.13 | 110.75 |
| SD | 11.47 | 9.87 |
| TOWK expressive
  \(M\) | 93.00 | 116.94 |
| SD | 10.85 | 6.16 |
| TOWK total language
  \(M\) | 90.88 | 115.13 |
| SD | 10.46 | 6.98 |

 dared to measure the child’s understanding of category membership and semantic feature knowledge.

Participants

This procedure resulted in the selection of 16 normal readers and 16 poor comprehenders, matched for decoding skill, chronological age and nonverbal ability (see Table 1); analyses of variance confirmed that the two groups did not differ on these measures (all \(F\)’s < 1). In contrast, the poor comprehenders’ reading comprehension was over three years lower than that of the control children \((F(1,30) = 100.85, \text{MSE} = 92.99, p < .001)\). Their Neale Analysis reading accuracy age was some 10 months lower than controls’, a difference that approached statistical significance \((F(1,30) = 3.74, \text{MSE} = 6.44, p = .06)\). This is consistent with our finding that comprehension skill in-
fluences text reading accuracy (Nation & Snowling, 1997). The implications of this will be discussed later in this paper.1

As is clear from Table 1, the mean score of both groups of children on the GNWRT appears to be approaching the ceiling of the test. Thus, it could be argued that they are not adequately matched for nonword reading. However, the GNWRT is a graded test and as such, each item discriminates between readers. Both groups obtained age-equivalent non-word reading scores well in advance of their chronological ages (\( M = 10.93 \) years for both groups). However, to confirm that the two groups of children have comparable nonword reading skills, the children read an additional list of 40 nonwords and we measured naming response latencies. The nonwords contained various spelling-sound patterns and while some shared rime unit correspondences with many words, others contained relatively less-common correspondences. The two groups made a similar number of errors (poor comprehenders 7% and normal readers 6.4%); this difference was not significant (\( F(1,30) = 0.33, MSE = 0.13, p > .80; F(1,39) = 2.33, MSE = 1.80, p > .10 \)). There was considerable variation in the children’s naming latencies, although overall, the poor comprehenders were slower than the normal readers (mean RTs (standard deviation) 1255.74 ms (316.10) and 1069.59 ms (457.79) for the poor comprehenders and the normal readers, respectively). However, this difference was not significant, either when the data were analyzed across subjects (\( F(1,30) = 1.81, MSE = 277205.21, p > .10 \)) or items (\( F(1,39) = 3.59, MSE = 18824.94, p > .07 \)). Thus, we are confident that the poor comprehenders and normal readers have comparable nonword reading skills.

Table 1 shows that the poor comprehenders achieved lower receptive (\( F(1,30) = 29.71, MSE = 3403.13, p < .001 \)), expressive (\( F(1,30) = 58.89, MSE = 4584.03, p < .001 \)) and total (\( F(1,30) = 59.51, MSE = 4704.50, p < .001 \)) language scores than the normal readers.

In order to validate these findings, a second analysis was performed in which children were selected as having high word knowledge scores (total language score greater than 100; \( M = 115.93 \)) or low word knowledge (total language score less than 100; \( M = 87.33 \)). Although the two groups did not differ in terms of nonword reading accuracy (\( F(1,30) = 1.11, MSE = 3.33 \)), the high language group had significantly higher reading comprehension scores than the low language group (\( M = 10.10 \) years vs \( M = 7.17 \) years, respectively; \( F(1,30) = 52.30, MSE = 70.23, p < .001 \)). Thus, there is a clear and consistent relationship between impoverished vocabulary knowledge and reading comprehension difficulties. Moreover, the poor comprehenders’ inferior performance across all subtests of the TOWK suggests that they have difficulty with both basic receptive and expressive vocabulary as well as with higher level aspects of word knowledge such as multiple meanings or figurative usage.

**EXPERIMENT 1**

Poor comprehenders have weaker vocabulary skills than control children and this suggests that their difficulties are not limited to comprehending written text but instead may encompass more general language problems. Vocabulary knowledge, however, is just one aspect of language skill. In this experiment, we used two judgment tasks to investigate good and poor comprehenders’ ability to ac-
cess semantic and phonological information. In synonym judgment, the children must decide whether or not two spoken words have similar meanings. In rhyme judgment, the children decide whether or not two spoken words rhyme. Arguably, in order to decide whether two words are synonymous, it is necessary to access the meanings of the words. In contrast, rhyme judgment taps phonological processing but does not require any consideration of word meaning.

In the synonym task, we compared the performance of the two groups of children on items that differ in imageability. We anticipated that low-imageability words will be harder than high-imageability words, especially for those children with reading comprehension difficulties. In the rhyme task, we used word pairs that are either orthographically similar or dissimilar. It has been found that children (Rack, 1985) and adults (Seidenberg & Tanenhaus, 1979) are faster and more accurate at deciding whether or not two words rhyme if they share a similar spelling pattern (e.g., ROSE–NOSE) than if they are spelled differently (e.g., ROSE–GOES). We predicted that poor comprehenders and normal readers would perform in a qualitatively similar manner across stimulus type.

**Method**

**Materials**

*Synonym judgment.* Twenty pairs of synonyms were selected. Imageability was assessed using the norms provided by Paivio, Yuille, and Madigan (1968). Ten pairs consisted of high-imageability items (e.g., BOAT–SHIP; RUG–MAT). These were matched for frequency of occurrence (Carroll, Davies, & Richman, 1971), letter length, and syllable length with ten low imageability pairs (e.g., FAST–QUICK; CRY–SOB). The pairs were randomly reordered to produce 20 pairs of nonsynonymous items. In total therefore, the test contained 40 pairs of words.

*Rhyme judgment.* Eighteen rhyming and 18 nonrhyming pairs of words, matched for frequency of occurrence in children’s reading materials (Carroll et al., 1971) and letter length were chosen. Half of the rhyming words also shared similar spellings (e.g., ROPE–HOPE; JOKE–COKE) and half were spelled differently (e.g., ROPE–SOAP; JOKE–SOAK). Likewise, half of the nonrhyming words shared similar spellings (e.g., WORD–CORD; CASH–WASH) and half were spelled differently (e.g., WORD–DROP; CASH–SHOP).

**Procedure**

The order of presentation of the synonym task and the rhyme task was counterbalanced across participants in each group. In all cases, a period of at least one day intervened between sessions.

The stimuli were recorded by a female speaker onto DAT tape and then digitized. The children heard two spoken words, separated by a 500 ms interval, via headphones. The children were instructed to decide whether the words had similar meanings (synonym judgment) or rhymed (rhyme judgment) and they were asked to respond by making a key press on a designated “yes” key or a designated “no” key. Following each response, there was a 1000 ms interval and then the next pair of items was heard. Following 10 practice trials with corrective feedback, the experiment was split into two blocks of trials, each consisting of an equal number of yes- and no-pairs. A few minutes break was allowed between blocks. The yes-pairs and the no-pairs were randomly presented with the proviso that any individual item could not appear in a yes-pair and a no-pair in the same block. The children were encouraged to attempt each trial and to respond as quickly and accurately as possible.

**Results and Discussion**

In both tasks, only those trials which elicited a correct response were included in the RT analysis. As the spread and variability in RTs was large, it was very difficult to detect true outliers and the danger of ignoring extreme but valid RTs seemed great. In an analysis of RT studies, Ulrich and Miller (1994)
TABLE 2
Mean Error Rate and RTs of the Poor Comprehenders and Normal Readers on the Synonym Judgement Task (Experiment 1)

<table>
<thead>
<tr>
<th></th>
<th>“Yes” pairs</th>
<th></th>
<th>“No” pairs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High image</td>
<td>Low image</td>
<td>High image</td>
<td>Low image</td>
</tr>
<tr>
<td>Poor comprehenders</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td>1.00</td>
<td>0.82</td>
<td>2.13</td>
<td>1.96</td>
</tr>
<tr>
<td>RT</td>
<td>788.58</td>
<td>421.37</td>
<td>1019.67</td>
<td>541.02</td>
</tr>
<tr>
<td>Normal readers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td>0.63</td>
<td>0.72</td>
<td>0.94</td>
<td>0.99</td>
</tr>
<tr>
<td>RT</td>
<td>600.17</td>
<td>212.25</td>
<td>634.85</td>
<td>174.80</td>
</tr>
</tbody>
</table>

*Maximum number of errors is 10 for high and low imageability; 20 for “no” pairs.

Recommended that extreme RTs should not be removed, as truncating data in this way may introduce considerable bias. For this reason, we report analyses based on all correct RTs. As each child was tested individually, we are confident that all extreme responses were indeed genuine and not due to equipment failure. In the synonym judgment task, 19 trials (2.97%) were discarded as a result of equipment failure; 22 trials (3.82%) were discarded in the rhyme judgment task.

**Synonym Judgment**

Table 2 summarizes the children’s performance on the synonym judgment task. Two 2 (imageability: high vs low) × 2 (reader group) analyses of variance were computed, one taking number of errors as the dependent variable and one using mean RTs as the dependent measure. The nonsynonymous “no” items could not be included as imageability was not manipulated in these pairs. However, the speed and accuracy of rejecting the nonsynonymous pairs was examined in separate analyses.

In the RT analysis, the main effect of group was significant ($F(1,30) = 5.44$, $MSE = 1314374.21$, $p < .02$; $F(1,18) = 23.99$, $MSE = 1130396.70$, $p < .001$) showing that poor comprehenders were slower to respond than the normal readers. The main effect of imageability was significant by subjects only ($F(1,30) = 8.98$, $MSE = 282524.63$, $p < .001$; $F(1,18) = 2.50$, $MSE = 203863.59$, ns). Similarly, the interaction between group and imageability was significant by subjects ($F(1,30) = 4.90$, $MSE = 154303.77$, $p < .05$; $F(1,18) = 2.21$), but that the poor comprehenders were slower on the low imageability items ($F(1,30) = 7.33$, $MSE = 283991.47$, ns) but that the poor comprehenders were slower to respond than the normal readers. The main effect of imageability was significant by subjects only ($F(1,30) = 8.98$, $MSE = 282524.63$, $p < .001$; $F(1,18) = 2.50$, $MSE = 203863.59$, ns). Similarly, the interaction between group and imageability was significant by subjects ($F(1,30) = 4.90$, $MSE = 154303.77$, $p < .05$; $F(1,18) = 2.21$, $MSE = 104238.80$, ns). Follow-up analyses confirmed that the two groups did not differ on the high imageability items ($F(1,30) = 2.55$, $MSE = 283991.47$, ns) but that the poor comprehenders were slower on the low imageability items ($F(1,30) = 7.33$, $MSE = 1184686.50$, $p < .01$).

Essentially, the pattern of results in the accuracy analysis was consistent with the results of the RT analysis. The main effect of imageability by items in either the accuracy or RT analyses, despite strong effects across subjects. Of the high-imageability items, EARTH–SOIL induced a high error rate and slow reaction times. In contrast, the low-imageability items, ANGRY–CROSS, FAST–QUICK, and LOVELY–NICE induced few errors and fairly fast reaction times. Arguably, these items are particularly familiar to young children and this may explain the lack of a main effect of imageability across items.
readers \((F1(1,30) = 5.34, \text{MSE} = 9.77, p < .02; F2(1,18) = 7.08, \text{MSE} = 19.60, p < .02)\). The main effect of imageability, was significant across subjects \((F(1,30) = 9.69, \text{MSE} = 8.27, p < .01; F(1,18) = 2.76, \text{MSE} = 19.60, \text{ns})\). However, the interaction between group and imageability was not significant \((F(1,30) = 0.22, \text{MSE} = 12.291.62, p > .05)\) although the group factor did not interact significantly with any other factor.

The results of the rhyme judgment task need to be interpreted cautiously. Although the two groups did not differ significantly in terms of accuracy, they was a tendency for the poor comprehenders to be a little slower than controls, but this was only significant in the by items analysis. In summary, although it may be argued that their rhyme judgment skills are a little weaker than controls, the poor comprehenders showed markedly impaired performance on the synonym judgment task. We suggest that these results point to a specific deficit in semantic but not phonological processing in children who decode well but have problems with reading comprehension.

**EXPERIMENT 2**

In this experiment, we again examined the nature of the children’s semantic and phonological skills. Here, however, we use verbal fluency tasks to investigate how efficiently semantic and phonological information can be accessed and retrieved.

**Method**

**Materials and Procedure**

*Semantic fluency.* This was measured using the Word Association subtest from the *Clinical Evaluation of Language Fundamentals—Revised* (CELF-R; Semel, Wiig, & Secord, 1987). In this test, children are given three spoken category names (animals, ways of getting from one place to another, and kinds of work that people do) and are asked to generate as many examples of category members as possible in 60 s.

*Rhyme fluency.* An analogous rhyme production task was devised in which children were instructed to generate as many rhymes to three spoken words (plate, fright, and chair) as possible in 60 s.
TABLE 3
Mean Error Rate and RTs of the Poor Comprehenders and Normal Readers on the Rhyme Judgement Task (Experiment 1)

<table>
<thead>
<tr>
<th>Orthographically similar</th>
<th>Orthographically dissimilar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhyming</td>
<td>Nonrhyming</td>
</tr>
<tr>
<td>Poor comprehenders</td>
<td>M  SD</td>
</tr>
<tr>
<td>Errors</td>
<td>0.06 0.25</td>
</tr>
<tr>
<td>RTs</td>
<td>476.16 154.71</td>
</tr>
<tr>
<td>Normal readers</td>
<td>M  SD</td>
</tr>
<tr>
<td>Errors</td>
<td>0.19 0.40</td>
</tr>
<tr>
<td>RTs</td>
<td>420.76 133.48</td>
</tr>
</tbody>
</table>

* Maximum number of errors is 9.
* In milliseconds.

Results and Discussion

Table 4 shows the performance of the two groups on these tasks. A 2 (reader group) × 2 (task: semantic vs rhyme fluency) analysis of variance revealed a main effect of group indicating that poor comprehenders generated fewer words ($F(1,30) = 12.36$, $MSE = 915.06, p < .01$) than the normal readers. The main effect of task was also significant showing that more words were provided in semantic fluency than rhyme fluency ($F(1,30) = 14.45$, $MSE = 900.0, p < .001$). Importantly, the interaction between group and task was significant ($F(1,30) = 10.86$, $MSE = 676.0, p < .01$). Analyses of simple main effects confirmed that the two groups did not differ for rhyme fluency ($F(1 < 1.0)$ but that the poor comprehenders produced significantly fewer words in semantic fluency ($F(1,30) = 28.46$, $MSE = 1582.03, p < .001$).

Consistent with the results of Experiment 1, these data show that poor comprehenders have difficulties with semantic processing, here revealed as problems in accessing and retrieving semantic information. In contrast, when predominantly phonological skills were required, the poor comprehenders’ performance was more comparable to that of the normal readers. Thus, the semantic knowledge of poor comprehenders appears to be less well developed than their linguistic competence in the phonological domain.

EXPERIMENT 3

Experiments 1 and 2 established that poor comprehenders have semantic processing weaknesses. Our selection criteria were such that they all experienced significant problems with reading comprehension. It is plausible that their difficulties at the level of single-word semantics contributes to their problems with comprehension; slowness or inability to access word meanings must inevitably affect processing at higher levels, for example the sentence or text level (Perfetti, 1985). What is less obvious is whether semantic processing difficulties will also compromise word recog-
nition. Theories that emphasize the critical role of phonological skills to the development of automatic word recognition (e.g., Shankweiler et al., 1995; Share & Stanovich, 1995) downplay the contribution of semantics. Yet, from a very early stage, children learn concrete words more readily than abstract items (McFalls, Schwanenflugel, & Stahl, 1996; Vellutino & Scanlon, 1985; Vellutino et al., 1995) and there is the implicit assumption that the phonetic cues provided by print are used to access stored semantic knowledge about words (Ehri, 1992; Pring & Snowling, 1986; Rack, Hulme, Snowling, & Wightman, 1994). It follows, therefore, that children with specific comprehension difficulties may have problems developing automatic word recognition, even if they possess proficient (phonological) decoding skills.

It is important here to make explicit the differences between decoding and word recognition. Whereas good decoding skills will support the reading of words with regular spelling-sound correspondences, other skills may be required in order to read words which have more unusual spelling patterns. Indeed, Plaut et al. (1996) demonstrated that exception words are learned more easily by a model that embodies both phonological and semantic representations than by a model that lacks semantic representations. This model predicts that children who have poor semantic skills (i.e., poor comprehenders) will have specific difficulty reading words that are typically read with support from semantics, namely low-frequency exception words.

It will be recalled that poor comprehenders and normal readers were matched for decoding skill as measured by a nonword reading test. The prediction that, nonetheless, the poor comprehenders would be poorer in word recognition receives initial support from the finding that they achieved lower Neale reading accuracy age scores than the normal readers. Despite this difference, it is important to stress that the poor comprehenders all achieved at least adequate for age reading accuracy. Thus, they are not simply ‘‘garden-variety’’ poor readers.

In this experiment, we asked poor comprehenders and control children to read words varying in frequency and regularity. If the hypothesis that the poor comprehenders have poorer word-recognition skills than the good readers is correct, the poor comprehenders should demonstrate specific difficulties reading irregular and low-frequency words accurately and efficiently.

**Method**

**Materials**

Fifty-six words, varying in frequency and regularity, were used in this experiment. Fourteen items of each of the following word types were used: regular/high frequency (e.g., *mouth, dark*), regular/low frequency (e.g., *mince, ditch*), exception/high frequency (e.g., *month, door*) and exception/low frequency (e.g., *mould, dread*). The words were selected from a larger set used by Patterson and Hodges (1992) and were matched for length and initial phoneme. Words were designated as regular or exception on the basis of vowel pronunciation in the context of the terminal consonant/consonant cluster (see Patterson & Hodges, 1992, for further details).

**Procedure**

The words were presented in the center of a Macintosh SE/30 screen in lowercase type. A fixation point appeared on the screen for 1000 ms and immediately at the offset of the fixation, a stimulus word appeared on the screen, where it remained until a response was initiated. A voice-activated relay interfaced with the computer-timed naming response latencies (in milliseconds) from the appearance of the stimulus to the onset of the child’s response. The children were tested individually and were instructed to read each item as quickly as possible. All errors and equipment failures were noted. The words were split randomly into two lists such that an equal number of words of each type occurred in each list. The lists were presented in separate experi-
TABLE 5
Mean Error Rate and Naming Latencies for Words Varying in Frequency and Regularity (Experiment 3)

<table>
<thead>
<tr>
<th></th>
<th>High frequency</th>
<th>Low frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Regular</td>
<td>Exception</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Poor comprehenders</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td>0.50</td>
<td>0.73</td>
</tr>
<tr>
<td>RTs</td>
<td>767.52</td>
<td>114.05</td>
</tr>
<tr>
<td>Normal readers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Errors</td>
<td>0.38</td>
<td>0.62</td>
</tr>
<tr>
<td>RTs</td>
<td>746.18</td>
<td>186.55</td>
</tr>
</tbody>
</table>

a Maximum number of errors is 14.
b In milliseconds.

mental sessions and order of presentation was counterbalanced across the two groups of children.

Results

Mean RTs and number of errors are presented in Table 5. As in Experiment 1, the spread and variability in RTs was large, but to avoid any bias due to truncating data, we report analyses based on all correct RTs. Twenty-nine trials (1.62%) were voice key errors and these were not included in the analyses.

Two 2 (frequency: high vs low) × 2 (regularity: regular vs exception) × 2 (reader group: normal readers vs poor comprehenders) analyses of variance were conducted, one with number of errors as the dependent variable and one using mean RT. For accuracy, there was a main effect of group (F1(1, 30) = 107.48, MSe = 162.56, p < .001) and regularity (F1(1, 30) = 19.18, MSe = 17.29, p < .001) showing that the poor comprehenders made more errors than the normal readers. The main effects of frequency (F1(1, 30) = 211.16, MSe = 341.26, p < .001; F2(1, 13) = 14.76, MSe = 371.57, p < .001) and regularity (F1(1, 30) = 77.63, MSe = 79.70, p < .001; F2(1, 13) = 7.76, MSe = 85.75, p = .02) were also significant showing that fewer errors were made in the high-frequency and regular word conditions.

The interaction between reader group and frequency was significant (F1(1, 30) = 4.65, MSe = 7.51, p < .04; F2(1, 13) = 30.82, MSe = 9.14, p < .001). A test of simple main effects confirmed that the poor comprehenders read fewer low-frequency words than the normal readers (F1(1, 30) = 8.20, MSe = 27.56, p < .01; F2(1, 13) = 86.91, MSe = 25.79, p < .001), whereas the two groups did not differ for high-frequency words. The interaction between group and regularity was not significant (F1(1, 30) = 2.83; F2(1, 13) = 0.43). The interaction between regularity and frequency was also significant (F1(1, 30) = 96.91, MSe = 96.26, p < .001; F2(1, 13) = 8.98, MSe = 108.04, p < .01) and tests of simple main effects confirmed that the effect of regularity was significant for low-frequency words (F1(1, 30) = 107.48, MSe = 162.56, p < .001; F2(1, 13) = 36.57, MSe = 440.16, p < .001) only. The three-way interaction between reader group, frequency, and regularity was not significant in either analysis (both Fs < 1.0).

3 Although the interaction between regularity and group was not significant, exploratory simple main effects showed that whereas the two groups of children read regular words equally well (F1(1, 30) = 2.67, MSe = 4.0, ns), the poor comprehenders were significantly less accurate at reading exception words (F1(1, 30) = 13.10, MSe = 19.14, p < .01).
In the RT analysis, the main effect of group was not significant by subjects ($F1(1,30) = 2.13$) although it was by items ($F2(1,13) = 4.31$, $MSE = 317117.90$, $p < .001$). There were also significant effects of frequency ($F1(1,30) = 29.10$, $MSE = 1115753.21$, $p < .001$; $F2(1,13) = 64.49$, $MSE = 1124524.74$, $p < .001$) and regularity ($F1(1,30) = 7.56$, $MSE = 177070.0$, $p < .01$; $F2(1,13) < 1.0$, ns).

There were significant interactions between group and frequency ($F1(1,30) = 4.31$, $MSE = 165128.37$, $p < .05$; $F2(1,13) = 12.19$, $MSE = 142067.43$, $p < .01$) and between group and regularity ($F1(1,30) = 4.41$, $MSE = 103189.50$, $p < .05$; $F2(1,13) = 4.33$, $MSE = 51800.76$, $p < .001$). Simple main effects on these interactions demonstrated that the poor comprehenders read high-frequency and regular words as fast as control children. However, they were marginally slower than controls at reading both low-frequency words ($F1(1,30) = 3.09$, $MSE = 478530.48$, $p = .08$; $F2(1,13) = 37.90$, $MSE = 441847.53$, $p < .001$) and exception words ($F1(1,30) = 3.08$, $MSE = 565434.88$, $p < .05$; $F2(1,13) = 26.14$, $MSE = 312627.0$, $p < .001$). Neither the interaction between frequency and regularity nor the three-way interaction between group, frequency, and regularity was significant in the RT analyses (all $F$'s $< 1.2$).

Finally, it is interesting to note that when the children are split into two groups according to vocabulary skills as measured by the TOWK (high word knowledge vs low word knowledge), a similar pattern of results emerges (although not all of the interactions reached significance in the items analyses). Thus, vocabulary group interacted with both frequency (accuracy: $F1(1,30) = 6.36$, $MSE = 9.79$, $p < .02$; $F2(1,53) = 13.93$, $MSE = 0.04$, $p < .001$; RT: $F1(1,30) = 3.98$, $MSE = 154260.15$, $p = .05$; $F2(1,53) = 1.83$, $MSE = 70411.97$, ns) and regularity (accuracy: $F(1,30) = 3.86$, $MSE = 3.71$, $p = .05$; $F2(1,53) = 5.74$, $MSE = 0.18$, $p < .03$; RT: $F1(1,30) = 4.32$, $MSE = 101427.69$, $p < .05$; $F2(1,53) = 0.04$, $MSE = 1430.78$, ns).

**Discussion**

The results show that poor comprehenders have word-recognition weaknesses relative to normal readers to whom they are matched for decoding (nonword reading) ability. When reading high-frequency and regular words, poor comprehenders’ reading accuracy and speed was equivalent to that of control children. However, the poor comprehenders had specific difficulties recognizing low-frequency and exception words. As predicted, therefore, word recognition was most compromised when reading words that require a significant level of semantic support.

According to the Plaut et al. model, poor comprehenders should experience particular difficulty reading words that are both low-frequency and irregular and hence one should find a significant three-way interaction between reader group, frequency, and regularity. One reason for the lack of a significant three-way interaction in our data may be that our participants were children and were yet to develop a fully competent level of word-recognition skill. Although the two-way interaction between regularity and frequency was present in the accuracy analysis, there was no significant interaction in the RT data. Previous investigations examining frequency and regularity effects in children have found that a reliable frequency by regularity interaction does not emerge until later in development. For example, Waters, Seidenberg, and Bruck (1984) found that while fifth-grade children showed a regularity effect for low-frequency words only, younger and less-skilled readers also showed regularity effects for high-frequency words. It is possible that in the present study, group interactions with frequency by regularity are masked by such developmental factors. Moreover, as the words were defined as high- and low-frequency according to an adult frequency count (Kucera & Francis, 1967) it is possible that some of the high-frequency words were, in reality, rather low in frequency for the children. If older children with greater word-recognition skills were compared, we predict that the three-way interaction between
reader group, frequency, and regularity would be significant.

**GENERAL DISCUSSION**

The studies reported in this paper compared children with reading comprehension difficulties to normal readers matched for decoding skills and nonverbal ability. In spite of being closely matched in nonverbal ability and possessing phonological processing skills that fall within normal range, children with reading comprehension difficulties have weak semantic processing abilities. Moreover, although they had good decoding skills, the poor comprehenders experienced greater difficulty reading low-frequency and exception words.

Plausibly, individual differences in semantic processing may underlie individual differences in reading comprehension. The performance of the poor comprehenders at synonym judgment indicates that they have difficulty processing word meanings and their tendency for poorer performance on the low-imageability words suggests that their semantic knowledge of abstract items is especially weak. In addition, their word knowledge was poorer and they generated fewer exemplars of semantic categories in the fluency tasks. These data converge with those of Stothard and Hulme (1992) who found that children with specific reading comprehension problems have weak receptive language skills and poor listening comprehension. In short, these findings indicate a more general verbal–semantic impairment, rather than a circumscribed problem with reading comprehension.

Although this study does not provide a framework in which to test causal mechanisms, it is possible to speculate on the nature of the relationship between semantic skills and comprehension. A simple view might be that the presence or absence of word knowledge causes or hampers reading comprehension (Anderson & Freebody, 1981). Thus, text will be comprehended if the meaning of each of the individual words is familiar. However, attempts at improving comprehension through vocabulary instruction do not improve comprehension, even for passages containing the taught vocabulary (Stahl & Fairbanks, 1986). Alternatively, the relationship between semantic skills and reading comprehension might be best explained by the speed or efficiency of semantic access. Given limited cognitive resources, if semantic access is slow or effortful then less capacity will be available for comprehension. Our findings offer some support for this view as the poor comprehenders were slower than controls at synonym judgment and semantic fluency (as well as at reading low-frequency and exception words, which arguably, are read with support from semantics) even though decoding skill was carefully controlled (cf. Perfetti, 1985). In addition, it is also important to consider the influence reading comprehension impairments may have on language development. Stanovich (1986) has argued that children with word-recognition problems (dyslexia) become more disadvantaged as they get older as they benefit less from the “riches” of reading experience. A similar argument may be constructed for children with comprehension problems. As these children are less sensitive to contextual information, their reading experiences are less likely to facilitate improvements in vocabulary or general knowledge. In turn, this will fail to fuel further increases in reading comprehension ability.

An important finding of this study was that poor comprehenders have weaker word-recognition skills for exception and low-frequency words than predicted from their basic decoding skill. This implies that underlying semantic skills may constrain not only comprehension, but also the development of skilled word recognition. Indeed, the finding that the poor comprehenders performed at a lower level than the normal readers may be considered surprising given our stringent selection criteria demanding that all children show at least age-appropriate reading accuracy. By adopting this criterion, we may well have excluded those poor comprehenders with more severe semantic impairments. As such, our results are likely to underestimate the extent to which poor reading comprehension
compromises the development of word recognition.

Developmental models of word recognition (e.g., Ehri, 1992; Frith, 1985; Gough & Hillinger, 1980) have tended to focus on the contribution of decoding skills and the importance of underlying phonological skills. According to Share (1995; Share & Stanovich, 1995), decoding skills are central to the acquisition of skilled word recognition. He argues that decoding operates as a "self-teaching device" as it offers a unique opportunity to build wordspecific representations (e.g., Ehri, 1992). Our observation that poor comprehenders have at least average (and in many cases very good) decoding skills yet fail to develop fully competent word recognition suggests that this view has limitations. The principles of the self-teaching hypothesis could be extended to accommodate the finding that individual differences in semantic processing skill also account for variations in word-recognition ability. Connectionist models, in particular the "three-cornered" model of Plaut et al. (1996) provide an appropriate framework in which to consider the complex nature of the relationships between phonological skills, semantic skills, and word-recognition development.

If it is assumed that the major goal facing the beginning reader is to develop decoding skills, the quality of semantic input might be less important than phonological skills at this stage of reading development. Consistent with this is the observation that the combined semantic plus phonological model implemented by Plaut et al. did not learn any faster than the phonology-alone model during the early stages of training. However, well-specified semantic representations may become more important as the range and number of words to be read increases. In a connectionist model without semantics, exception words are pronounced correctly only if they are high in frequency (and therefore they have been encountered many times by the network). By training the phonological pathway with support from semantics, the combined model required fewer training trials than the phonology-alone model to learn the pronunciations of exception words. Plaut et al. concluded that the development of word recognition is best described in terms of a division of labor between interacting phonological and semantic pathways. Plausibly, therefore, individual differences in either phonological or semantic processing will be related to individual differences in reading ability.

We have argued previously (Hulme, Quinlan, Bolt, & Snowling, 1995; Snowling, Hulme, & Goulandris, 1994) that the mechanisms underlying dyslexic children’s reading difficulties can be well described within a connectionist framework. Briefly, dyslexic children have poorly specified phonological representations and their ability to set up mappings between orthography and phonology is thus severely compromised. More recently, we have found that dyslexic children are more likely to use contextual support to facilitate word recognition than reading-age matched children (Nation & Snowling, in press). These findings led us to argue that dyslexic children may compensate for their difficulties with orthographic-phonological mappings by developing a reading system with a division of labor that is more heavily dependent on semantic input than is usually the case. This provides some explanation for how dyslexic children can and do learn to read, despite persistent difficulties with phonological processing and nonword reading (e.g., Bruck, 1990; Snowling et al., 1994).

The difficulties exhibited by poor comprehenders are also well accommodated within this framework. A direct prediction from the model is that poor comprehenders will have greater difficulty reading words that are typically read with support from semantics. Our finding that poor comprehenders were relatively poor at reading exception and low-frequency words is consistent with this prediction. Poor comprehenders also show less contextual facilitation of word recognition than reading-age matched children (Nation & Snowling, in press). When faced with an unfamiliar word that cannot be decoded easily, they are less adept at combining contextual information with information gleaned from
partial decoding. In contrast to dyslexic children, therefore, poor comprehenders rely less on the semantic pathway.

In summary, this study has shown that poor comprehenders have weaker semantic skills than children with normally developing reading abilities. We speculate that these weaknesses may be causally related to their reading comprehension difficulties, but this possibility requires empirical validation. We also propose that these same semantic deficits may constrain the development of word recognition. In support of this proposal, we found that poor comprehenders have greater difficulty reading words that are typically read with support from semantics. Although there is an intimate relationship among phonological skills, decoding, and word recognition, our results demonstrate that variations in semantic skill also contribute to the development of skilled word recognition.

REFERENCES


(Received September 3, 1997)
(Revision received December 31, 1997)