Phonologic Mediation Is Fundamental to Reading: Evidence from Beginning Readers

Anna M. T. Bosman

Phonologic Mediation Is Fundamental to Reading: Evidence from Beginning Readers

Anna M. T. Bosman and Annette M. B. de Groot

University of Amsterdam, Amsterdam, The Netherlands

Three tasks were employed to investigate the role of assembled phonology in beginning readers. In two proofreading tasks, children had more trouble finding pseudohomophone misspellings (stimuli with phonology identical to that of a word) than control misspellings (stimuli that do not share their phonology with a word). In a lexical-decision task, they had more trouble deciding that pseudohomophone misspellings were non-words than deciding that control misspellings were non-words. Finally, in a semantic-categorization task, children had more trouble rejecting pseudohomophone misspellings as a member of a designated category than rejecting control misspellings. Differences between more and less advanced readers occurred, but they need not be attributed to differential use of phonology in word recognition. Instead, they were explained in terms of a difference between reader groups in spelling-verification efficiency. The results of the present studies on beginning reading parallel studies on skilled reading by Van Orden et al. (1992). The main conclusion was that assembled phonology plays an important role in word recognition in beginning readers.

The role of phonology in reading has become an issue of extensive debate in the last decade. Researchers in the field share the view that phonology plays a prominent role in reading, although the precise character of this role remains controversial. The most common view, a traditional information processing view, is expressed in dual-route theory. Baron and Strawson (1976) proposed two processes or routes that a reader can employ to pronounce printed words: an assembled route, which makes use of the relationship between letters and sounds, and a lexical route, which relies upon the specific knowledge of pronunciations stored in a mental lexicon. These two processes have

become known as non-lexical and lexical reading, or as the indirect and direct route to the lexicon (see also Coltheart, 1978; Coltheart, Curtis, Atkins, & Haller, 1993). Patterson (1982) proposed a parallel distinction between assembled and addressed phonology as alternative procedures for deriving pronunciations. Additionally, we may distinguish between procedures used to derive pronunciations and procedures used to comprehend words (Patterson & Coltheart, 1987). Taken together, these distinctions allow three possible procedures for comprehending words: (a) whole-word orthographic representations, which directly activate semantic representations, which, in turn, activate whole-word phonologic representations (i.e. direct access to semantics and postlexical access to phonology, hereafter *direct access*); (b) whole-word orthographic representations, which directly activate whole-word phonologic representations, which, in turn, activate semantic representations (i.e. addressed phonology and postlexical phonologically mediated access to semantics, hereafter *addressed phonology*); (c) subword orthographic representations, which activate subword phonologic representations, which, in turn, activate semantic representations (i.e. prelexical assembled phonology mediates access to semantics, hereafter *assembled phonology*). Both addressed and assembled phonology are forms of phonologic mediation.

The present studies all concern the question whether and to what extent assembled phonology affects word identification in beginning reading. Two opposing views concerning the development of reading behaviour have been put forward in the literature (for a review on this topic see Barron, 1986). Some research suggests that beginning readers rely on assembled phonology, but that they shift to addressed phonology and/or direct access with increasing reading experience (Backman, Bruck, Hebert & Seidenberg, 1984; Doctor & Coltheart, 1980; Reitsma, 1983a, 1984). Other research, however, suggests that beginning readers initially rely on direct and/or addressed phonology, and only with reading experience do they come to rely on assembled phonology (Coltheart & Laxon, 1990; Coltheart, Laxon, Keating & Pool, 1986; Condry, McMahon-Rideout, & Levy, 1979; Marsh, Friedman, Welch, & Desberg, 1980; Rayner & Pollatsek, 1989; Seymour & Elder, 1986). This latter hypothesis predicts that assembled-phonology effects should be absent in beginning readers, whereas the former predicts that beginning readers should show strong assembled-phonology effects.

Several studies have been conducted to test for both assembled and addressed phonologic mediation in reading comprehension of beginning and skilled readers (Coltheart, Avons, Masterson, & Laxon, 1991; Coltheart & Laxon, 1990; Coltheart et al. 1986; Coltheart, Laxon, Rickard, & Elton, 1988; Doctor & Coltheart, 1980). The conclusion that has been drawn from these studies is that addressed phonology contributes to sentence and word comprehension in both beginning and skilled readers, but that meagre evidence exists for the involvement of assembled phonology in comprehension.

The experiments that we designed also investigated the role of phonologic mediation in beginning readers, but they are restricted to the role of assembled phonology. The experiments do not test precisely how early in reading development assembled phonology emerges; instead, their goal is to see whether assembled phonology affects the performance of Dutch Grade 1 readers of varying reading ability across a variety of simple reading tasks. The tasks we used were modelled on those reported in Van Orden et al. (1992). To our knowledge, such tasks and manipulations have not previously been used with children

(except for a modified lexical-decision task described by Coltheart et al., 1986). Further-more, most of the research in beginning reading has been conducted in English. Thus, the generality of previous findings may be greatly extended by conducting the tasks with children, and in a language other than English (i.e. Dutch).

Van Orden et al. (1992) conducted proofreading, lexical-decision, and semantic-cat-egorization experiments, in which correct performance requires that subjects detect misspellings. Two types of misspellings were used—pseudohomophone misspellings [*grean*], and control misspellings [*greln*]. An assembled phonology effect is observed when subjects detect fewer pseudohomophone misspellings than control misspellings. The use of control misspellings (e.g. *greln*) insures that this effect is due to assembled phonology. Control misspellings are constructed to be as similar to the base words (e.g. *green*) in spelling as pseudohomophones are, but they are less similar to the base words in their phonology. Thus, they control for potential effects of visual and orthographic similarity due to addressed phonology and/or direct access.

In the present proofreading task, children read stories (or lists of unconnected words) and marked each misspelling that they came across. In the lexical-decision task children indicated whether presented letter strings constituted words or not, and in the semantic-categorization task they indicated whether a particular stimulus was an exemplar of a designated category. In all tasks, the experimental stimuli were pseudohomophone and control misspellings. In the proofreading task, an assembled-phonology effect would be observed if fewer pseudohomophone misspellings are detected than control misspellings. Likewise, in the lexical-decision and semantic-categorization experiments an assembled phonology effect is found when the percentage of correctly rejected pseudohomophone misspellings is smaller than that of control misspellings.

Two results obtained by Van Orden and colleagues suggested a fundamental role for assembled phonology in skilled readers. First, adult subjects generally missed more pseudohomophones than controls; second, pseudohomophones derived from low-frequency base words (*sleat* for *sleet*) were missed more often than those derived from high-frequency base words (*grean* for *green*), whereas the mean number of correctly identified errors was the same for control misspellings derived from low-frequency base words (*speet* from *sleet*) and those from high-frequency base words (*greln* from *green*). A frequency effect is taken as a signature of lexical access. Thus, if performance to pseudohomophone misspellings is differentially affected by the subjects' familiarity with the base word (the base word is never presented in the experiment), then the phonology of the pseudohomophone must have mediated this lexical effect. Van Orden et al. (1992) concluded that pseudohomophones' assembled phonology mimics that of phonologically identical lexical items, and that a common process underlies the phono-logic coding of words and non-words.

A second, closely related issue is: what is the basis for correct responses to pseudo-homophone misspellings in simple reading tasks? A widely held belief is that readers perform a spelling check, which compares knowledge of correct spellings, retrieved from lexical memory, with the spelling of the stimulus (Becker, 1976; Grossberg & Stone, 1986; Davelaar, Coltheart, Besner, & Jonasson, 1978; Paap, Newsome, McDonald, & Schvaneveldt, 1982; Rubinstein, Lewis, & Rubinstein, 1971; Van Orden, 1987, 1991).

Assuming that spellings of highly familiar base words—i.e. high-frequency words—are more available for spelling verification than those of less familiar base words, we would expect fewer false positive errors of pseudohomophones derived from high-frequency base words than of pseudohomophones derived from low-frequency base words (Jared & Seidenberg, 1991; Van Orden, 1987, 1991; Van Orden et al., 1992). Notice that this would not necessarily indicate that subjects have insufficient spelling knowledge of low-frequency base words. It only means that the availability of the base-word's spelling is more limited in the case of a less familiar word.

The verification hypothesis is related to the previously stated assembled-phonology hypothesis. It is assumed that the phonology of a pseudohomophone can retrieve the spelling of its base word, which, in turn, can be used to verify the spelling of the presented pseudohomophone. Thus, not only does a positive effect of base-word frequency on the error detection of pseudohomophone misspellings support evidence for the existence of a verification process, but it also indicates an effect of assembled phonology. Support for assembled phonology requires that an effect of base-word frequency is reduced or absent in the control stimuli.

It is possible, however, that beginning readers have not yet developed a stable, reliable verification process. We used three methods to test for effects due to spelling verification. The first test manipulated the extant base-word frequency using a frequency count. The prediction is that pseudohomophones derived from high-frequency base words are easier to detect than those derived from low-frequency base words. In another test, base-word frequency was manipulated experimentally by explicit spelling training. Children were trained more on some base words and less on others. Our final test derives from children's reading skill. We may assume that reading skill is positively correlated with effective verification (Van Orden, 1987; Van Orden, Pennington, & Stone, 1990). Even though all subjects in this study were attending Grade 1, relatively large differences in reading skill were apparent. Thus, if more advanced readers have developed a basis for more effective verification, they should have detected more pseudohomophone misspellings than less advanced readers. In three out of the four experiments (Experiment 2 being the exception) we divided the total subject group into more advanced and less advanced readers.

GENERAL METHOD

Before presenting the experiments, a few words are necessary concerning the subjects who participated in this study, Dutch orthography, the reading instruction method, and the reading test that was administered. All subjects attended first grade and had Dutch as their mother tongue. Children in the Netherlands start school at the age of 4 years, but they do not receive formal instruction in reading and spelling until the age of 6 years (in fact, 78 months on average)—that is, when they attend Grade 1 (Group 3 in the Dutch system). The school years before Grade 1 are actually kindergarten years, and pre-reading skills are trained there, but usually no real effort is made to teach the children to read, either by the teachers or by their parents.

The average age of the children who participated in the present studies was 88 months at the time of testing. That is, at the time of testing they had had about ten months of

formal reading instruction. None of the children took part in more than one experiment. All subjects (recruited from 10 different primary schools in the Netherlands) in Experiments 1 through 4 were instructed according to the same reading curriculum (*Veilig leren lezen*, which translates into "Learning to read safely"—Caeser, 1979), which has a rather rigid pre-programmed plan and imposes a strict day-by-day and week-by-week progression. (For more information on teaching early reading in the Netherlands we refer to Reitsma and Verhoeven, 1990.)

Dutch orthography has a fairly regular correspondence between spelling and sound, but not as regular as, for instance, Serbo-Croation, nor as irregular as English. Van Heuven (1980) provides a more detailed description of Dutch orthography. The most widely used reading curriculum (80% of all Dutch primary schools) is the one used by the schools that participated in our study. It starts off with the instruction of a small sight vocabulary but stresses the importance of phonics. The children are taught the grapheme–phoneme correspondence rules and are initially confronted only with regular words. After four months of instruction, the average child is capable of reading all regular Dutch words by systematic application of the grapheme–phoneme correspondence rules. Gradually less regular words are introduced.

Subjects were selected according to their scores on a standard reading test. The groups of subjects selected for participation did not include children with severe reading and spelling problems or any other type of learning disability. Because all children were instructed according to the same method, reading levels could be assessed reliably.[1] The reading test that was used (Eén-minuut-test, Caesar, 1970) is a decoding test that consists of a list of unrelated words, which have to be read aloud. The number of correctly read words in one minute is the reading score. The average is 28 ($SD = 15$) after six months in first grade.

EXPERIMENT 1
Proofreading

In Experiment 1, we manipulated type of misspelling (pseudohomophone versus control), base-word frequency (high versus low), reading level (more versus less advanced), type of reading material (story versus list), and list condition (blocked versus mixed). All of the manipulations have been discussed above, except for reading material and list condition.

Type of reading material and list condition were manipulated because they may differentially constrain the detection of spelling errors. Van Orden et al. (1992) found an effect of assembled phonology when subjects proofread stories, but Ehri and Roberts (1979) showed that more careful inspection of words may occur when subjects proofread lists of unconnected words rather than words in stories (but cf. Cohen, 1980). In other

---

[1] Scores on a non-verbal (Standard Progressive Matrices; Raven, 1958) and a verbal (vocabulary test of the RAKIT; Bleichrodt, Drenth, Zaal, & Resing, 1984) intelligence test were also assessed. In general no significant differences were found between more and less advanced readers on these variables. Scores on a spelling test (Mommers & Van Dongen, 1986) were also available, and the difference between more and less advanced readers were generally significant on this test. This is mainly due to a significant correlation ($r = .50$) that exists between reading and spelling in first grade.

words, we expect an effect of type of reading material—that is, the list condition should favour spelling verification as against a story condition such that more pseudohomophone misspellings should be detected in the list condition than in the story condition.

The list-condition manipulation is an additional test for the constraints on spelling verification. In Van Orden (1991) and Van Orden et al. (1992) subjects were always confronted with stories that contained one type of misspelling, either pseudohomophones or controls (blocked condition). In our list condition, we used both a blocked condition (lists of words mixed with one type of misspelling) and a mixed condition (lists of words with a mixture of both pseudohomophone and control misspellings). An interaction between type of misspelling and list condition would indicate strategic differences in verification. [We do not believe that either of the conditions (blocked or mixed) is free of strategic control, nor that either of the conditions provides a neutral baseline to assess proofreading behaviour—see Stone & Van Orden, 1993.]

Method

Materials

The first three reading books of the reading curriculum served as the source for the selection of the base words from which we constructed the experimental stimuli. Twelve one-syllable Dutch base words with a mean length of 3.6 letters ($SD = 0.9$) were chosen. Half of these words were of high frequency (mean count: 34.3, SD 22.3; range 13–77) and half of low frequency (mean count: 4.8, SD 3.9; range 1–11), according to their frequency count in these three books. From each selected base word a pseudohomophone misspelling and a control misspelling were created. Construction of pseudohomophones in the Dutch language is limited due to its orthography. Further restrictions had to be imposed because beginning readers served as the subject of our research. Changes in pseudohomophones and control misspellings could not be balanced with respect to type of letter and position in the word. Pseudohomophones were predominantly created by changing a vowel in the middle of the word, whereas control misspellings had to be created by changing the final consonant of the base word.

However, we used the OS-index (Orthographic Similarity) described in Van Orden (1987) as a means to guarantee similarity of orthography between pseudohomophone and control misspelling on the one hand and the base word on the other. This estimate is derived from Weber's measure of graphic similarity (Weber, 1970). The OS-index ranges from 0 to 1, where 0 indicates no orthographic similarity between the stimuli and 1 indicates identical orthography. Mean OS between high-frequency base words and pseudohomophone misspellings was .58 ($SD = .11$); between low-frequency base words and pseudohomophone misspellings it was .62 ($SD = .11$); between high-frequency base words and control misspellings it was .68 ($SD = .07$) and between low-frequency base words and control misspellings it was .67 ($SD = .09$). Note that the higher OS of the control misspellings biases our method against detecting an effect of assembled phonology. The base words, pseudohomophone and control misspellings are listed in Appendix A.

Two stories, "In the House" (149 words) and "In the Woods" (147 words), were composed from words that all appeared previously in the curriculum. Pseudohomophone (story–Psh condition) or control misspellings (story–Ctr condition) were substituted in the positions of their corresponding base words.

Additionally, four lists of 60 items each were created. All lists contained 48 correctly spelled words and 12 misspellings. List 1 consisted of the 48 words intermingled with 12 pseudohomophone misspellings (List–Psh condition). List 2 contained 12 control misspellings (List–Ctr condition)

intermingled with the 48 words. Lists 3 (Mix–A condition) and 4 (Mix–B condition) were made up of the 48 words intermixed with six pseudohomophone misspellings (three were derived from low-frequency base words and three from high-frequency base-words) and the six control misspellings derived from the remaining base words. For each child a different randomized version was created.

Procedure

A child was assigned either to the story or to the list condition. Half of the subjects in the story condition proofread stories containing only pseudohomophone misspellings; the other half proofread stories containing only control misspellings. In the list condition, a child had to proofread one of the four stimulus lists. In both the story and the list conditions each child was asked to read the page carefully and mark the misspellings (both non-existent and wrongly spelled words).

Spelling knowledge of the relevant stimuli was assessed after completion of the proofreading task by means of a forced-choice spelling test. The subjects were given an envelope with 12 strips of paper, each containing one of the base words and its derived pseudohomophone misspelling. The children were asked to mark the incorrectly spelled words.

Subjects

A sample of 120 children of Grade 1 was drawn from a population of 241 children. Forty children were assigned to the story condition (20 in the pseudohomophone and 20 in the control version) and 80 children took part in the word-list condition (20 subjects per list). Children with exceptionally high or low scores on the reading test were not included in the sample. Considerable differences still remained between the reading scores of children at the higher and the lower ends of the scale, which allowed us to form two groups: more advanced readers (32.8; $SD = 4.2$) and less advanced readers (21.6; $SD = 3.5$); $F(1, 118) = 253.28$, $p < .001$. Much care was taken to approximately match reading level across the between subject manipulations, $F(5, 108) = 0.14$, $p = .98$.

Results

Spelling-test Results.   The results of the forced-choice spelling test confirmed that the children knew the spelling of the base words; mean proportion correct was .92. Subject and item means were calculated. A 2 (reading level: more advanced versus less advanced) by 6 (condition: story–Psh versus story–Ctr versus list–Psh versus list–Ctr versus Mix–A versus Mix–B) by 2 (base-word frequency: high versus low) ANOVA was executed on the subject and item means, indicating a significant main effect of reading level, $F(1, 108) = 20.04$, $p < .001$, by subjects, and $F(1, 10) = 11.04$, $p < .01$, by items. More advanced readers (.96) had a higher proportion correct on the forced-choice task than less advanced readers (.88). Neither the main effects of condition (both F's < 1) and frequency ($p > .15$ for subjects, $F < 1$ for items), nor any interaction effect reached significant levels. The absence of main effects of between-subject conditions merely confirms the careful matching of reading level across these conditions.

Proofreading Results.   A 2 (reading level: more advanced versus less advanced) by 2 (type of reading material: story versus blocked list) by 2 (type of misspelling: pseudo-homophone versus control) by 2 (base-word frequency: high versus low) ANOVA was

conducted on subject and item means of the proportion of correctly identified mis-spellings. The children in the pseudohomophone condition detected far fewer misspell-ings (.27) than those in the control condition (.80), as indicated in a significant main effect of type of misspelling, $F(1, 72) = 223.73$, $p < .001$, by subjects, and $F(1, 20) = 121.73$, $p < .001$, by items). However, neither the main effect of base-word frequency, nor the interaction between base-word frequency and any other experimental variable was significant. Although more advanced readers were generally better at finding misspell-ings (.60) than less advanced readers (.46), as shown by the significant main effect of reading level, $F(1, 72) = 14.24$, $p < .001$, by subjects, and $F(1, 20) = 44.33$, $p < .001$, by items, this was only true for the children in the pseudohomophone misspelling condition (more advanced: .40; less advanced: .14). Performance in the control condition was identical for both groups of readers (.80). The F-values for the interaction effect between reading level and type of misspelling were $F(1, 72) = 12.59$, $p < .001$, by subjects, and $F(1, 20) = 38.96$, $p < .001$, by items.

The children in the blocked-list condition detected more misspellings (.62) than the children in the story condition (.44), $F(1, 72) = 26.78$, $p < .001$, by subjects, and $F(1, 20) = 55.47$, $p < .001$, by items. This result also needs qualification, because the children in the blocked-list condition found more pseudohomophone misspellings (.42) than those in the story condition (.12; Newman-Keuls, $p < .01$), whereas the number of correctly identified control misspellings did not vary with type of reading material (story: .76; list: .83), $F(1, 72) = 10.84$, $p < .01$, by subjects, and $F(1, 20) = 22.46$, $p < .001$, by items.

In both the story and blocked-list conditions, fewer pseudohomophone misspellings were detected than control misspellings (Newman-Keuls, $p < .01$). This result needs qualification, because the three-way interaction between reading level, type of reading material, and type of misspelling was also significant, $F(1, 72) = 11.60$, $p < .01$, by subjects, and $F(1, 20) = 28.13$, $p < .001$, by items. This interaction is depicted in Figure 1. A significant interaction between reading level and type of misspelling occurred only in the blocked-list condition, $F(1, 36) = 18.21$, $p < .001$, by subjects,
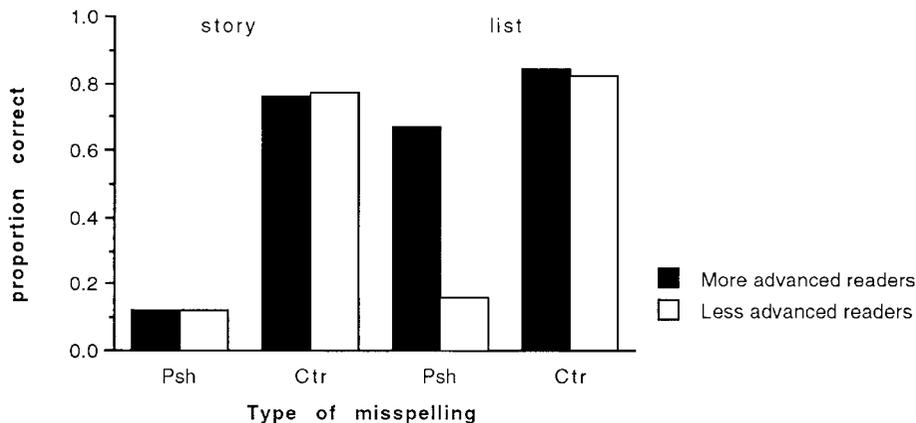

FIG. 1.   Proportion of correctly identified pseudohomophone (Psh) and control (Ctr) misspellings of the more advanced and less advanced readers in the story and blocked-list conditions of Experiment 1 (Proofreading).

and $F(1, 20) = 52.26$, $p < .001$, by items. In that condition, the less advanced readers showed a much larger difference between the number of correctly identified control misspellings (.82) and pseudohomophone misspellings (.16) than the more advanced readers (.84 and .67, respectively). The difference between the two types of misspellings in the more advanced reader group, which was in the same direction as that of the less advanced group, was significant, $F(1, 19) = 4.65$, $p < .05$, by subjects, and $F(1, 20) = 5.53$, $p < .05$, by items. In the story condition, both the more advanced and the less advanced readers detected fewer pseudohomophone misspellings (.12 for both readers groups) than control misspellings (.76 and .77, respectively). The virtually identical performance of the two groups is indicated in the non-significant interaction between reading level and type of misspelling.

It also appeared that more advanced readers were better in detecting misspellings in the blocked-list condition (.76) than in the story condition (.44; Newman-Keuls, $p < .01$), but the less advanced readers did not show this same advantage in the list condition (story: .44; list: .49). This is shown by the significant interaction between reading level and type of reading material, $F(1, 72) = 15.00$, $p < .01$, by subjects, and $F(1, 20) = 36.42$, $p < .001$, by items.

To test for change in strategic control, a 2 (type of misspelling: pseudohomophone versus control) by 2 (base-word frequency: high versus low) by 2 (list type: blocked versus mixed) by 2 (reading level: more advanced versus less advanced) ANOVA on the item means, and two separate analyses on the subject means of correctly identified pseudohomophone misspellings in the Psh–, Mix–A and Mix–B conditions and on the correctly identified control misspellings in the Ctr–, Mix–A and Mix–B conditions were conducted. Only the effects that pertain to the factor list condition are discussed, because all other effects have been addressed already.

The results of the subject analyses showed that the mean proportion of correctly identified pseudohomophones in the blocked Psh-condition (.42) was the same as in the two mixed versions (Mix–A: .46; Mix–B: .35; $F < 1$). The mean proportion of correctly identified control misspellings in the blocked condition (.83) was also equal to those in the mixed conditions (Mix–A: .79; Mix–B: 0.80; $F < 1$). The results of the item analyses paralleled those of the subject analyses. The main effect of list type was not significant ($p > .20$), nor was the interaction between list type and type of misspelling. This result needs to be qualified, because of the significant interaction between reading level, list type, and type of misspelling, $F(1, 20) = 10.18$, $p < .01$, and the one between reading level and list type, $F(1, 20) = 30.36$, $p < .001$). In the mixed condition, reader groups did not differ (more advanced readers: .62; less advanced readers: .58), but they did in the blocked condition; more advanced readers had a higher hit rate (.76) than less advanced readers (.49; Newman-Keuls, $p < .01$). More advanced and less advanced readers were equally successful in finding control misspellings in the blocked (.84 and .82, respectively) and mixed (.76 and .83, respectively) conditions, but the difference between more advanced and less advanced readers in detecting pseudohomophone misspellings was larger in the blocked condition (.68 and .16, respectively) than in the mixed condition (.48 and .33, respectively).

Discussion

Two issues pertaining to the reading behaviour of beginning readers were raised in the introduction. The first question was whether assembled phonology mediates word iden- tification in Dutch-speaking beginning readers of varying reading skill. The results of Experiment 1 support this possibility, because all readers reliably detected fewer pseudo- homophone misspellings than control misspellings. This substantial effect cannot be entirely due to insufficient spelling knowledge of the base-words' spellings, because the same children showed much better performance in the two-alternative forced-choice spelling test.

The second related issue we addressed was the question of how misspellings are detected when they *are* detected. It was suggested that word identification also includes a verification process, in which the spelling of the presented stimulus is matched with the spelling of the retrieved base-word. Two tests were used in Experiment 1 for investigating the role of spelling verification in beginning readers.

The first test was the assumed higher availability of high-frequency words compared to low-frequency words, resulting in a higher detection score for pseudohomophones derived from high-frequency base words than from low-frequency base words. This effect did not show up, however. Unlike Van Orden et al. (1992), we found no hint of a frequency effect nor an interaction between base-word frequency and type of misspell- ing. We return to this finding with Experiment 2.

The second test of the spelling-verification hypothesis was the effect of reading skill on proofreading performance. The hypothesis that more advanced readers have developed a basis for more effective verification was corroborated. More advanced readers detected more pseudohomophone misspellings than did the less advanced readers, but both reader groups were equally successful in finding control misspellings. Although the forced- choice spelling test indicated that the spelling knowledge of the less advanced readers was somewhat weaker than that of the more advanced readers, the difference between the groups (more advanced: .96; less advanced: .86) appears too small to justify the conclusion that this difference in spelling knowledge caused the difference between the subject groups in misspelling detection scores on the proofreading task. Instead, we conjecture that the difference is due to a more efficient spelling verification process in more advanced readers. In the General Discussion we discuss an alternative interpretation—namely that more advanced readers have developed a way to access the meaning of words directly.

The instability of the verification behaviour of more advanced readers is indicated by the finding that in the list condition more advanced readers detected more pseudohomo- phone misspellings than did less advanced readers, but in the story condition the per- formance of the more advanced readers dropped to the level of the group of less advanced readers.

Our hypotheses concerning the type of reading material and the list-condition manip- ulation suggested that it is spelling verification that is affected by these changes in task constraints. The blocked-list condition, the most favourable condition for spelling veri- fication, showed the highest rate of pseudohomophone detection in the group of more advanced readers. The comparison between the blocked-list condition and the (blocked) story condition favoured the list condition; more pseudohomophone misspellings were

detected in the list condition. The fact that the children found an equal number of control misspellings in the story and list conditions suggests that the difference in pseudohomophone error detection scores is not caused by a difference in length of the two types of reading materials (story: 296 words, and list: 60 words).

Finally, the absence of an overall effect of list condition (blocked versus mixed) and of an interaction between this variable and type of misspelling (pseudohomophone versus control) suggests that the effects that *do* occur are unlikely to be due to differential strategic behaviour. However, the interaction between type of reading material and reading level indicates a subtle difference in proofreading strategy between more and less advanced readers. More advanced readers detected more pseudohomophone misspellings when they occurred in lists of only one type of misspelling (blocked condition), whereas less advanced readers showed superior pseudohomophone detection performance when they occurred in lists of both types of spelling errors (mixed condition). We refrain from speculating about an explanation for this possibly interesting effect.

In sum, this proofreading experiment shows evidence for the assembled-phonology hypothesis. The effects established here are similar to those found with skilled readers in the studies of Van Orden et al. (1992). Although differences between more and less advanced readers did show up, both groups showed clear signs of phonologic mediation. The differences between groups were explained in terms of a better developed spelling verification process in more advanced readers. The most striking differences between the results from the present beginning readers and from the skilled readers of Van Orden et al. (1992) were that the former did not show an effect of base-word frequency or an interaction between base-word frequency and type of misspelling.

EXPERIMENT 2
Proofreading

In Experiment 2 we wanted not only to replicate the results of Experiment 1, but also to increase the probability of detecting an interaction between base-word frequency and type of misspelling. The base-word frequency factor was experimentally manipulated by means of spelling training prior to the proofreading task. The inclusion of a spelling training prior to the proofreading task allowed us also to test for a possible strategy effect in our design of the proofreading experiments. We compared proofreading performance of children who received spelling training prior to the proofreading task (Experiment 2) with the performance of children who did not receive spelling training (Experiment 1).

Another issue that was investigated in Experiment 2 concerns the influence of pronunciation during proofreading on the detection of misspellings. The question to be addressed is: does silent reading have a different effect on the detection of misspellings than reading aloud? According to dual-process theory, activation of the phonology of words during silent reading may be incidental, but it is obligatory when one has to read text out loud. If silent reading does not always involve phonologic activation, the "silent" and "aloud" conditions should differ. In the aloud condition, more control misspellings and/or fewer pseudohomophone misspellings should be detected than in the silent condition. The task condition (aloud versus silent) is an ideal strategy manipulation (Stone & Van Orden, 1993).

Method

Materials

The stories used in Experiment 1 were also employed in this experiment.

Procedure

Prior to the proofreading task, all children took part in spelling training. Recall that all children had encountered the base words before in their reading curriculum. In groups of 2, 3, or 4, the children were asked to copy a list of 30 base words. Six of the words (three of which occurred frequently in the curriculum and three infrequently) had to be copied four times (frequent condition) and the other six (three of which again occurred frequently in the curriculum and three infrequently) just once (infrequent condition). The assignment of words to the frequency conditions was counter-balanced. After the subjects finished the copy task, their knowledge of the words' spelling was checked with the help of a forced-choice spelling test, identical to the one in Experiment 1. They were asked to mark which of the two spellings (the base word itself or its pseudohomophone) was the wrong one. After the spelling training the subjects went back to their classroom. After a break (the minimum interval was half an hour, maximum was three hours) half of the children had to proofread the stories silently. Subjects were tested again in groups of 2, 3, or 4, and were asked to read the two stories and mark each misspelling they encountered (both non-existent and misspelled words). The other half of the children participated in the reading-out-loud condition. They were tested individually and had to read the stories to the experimenter. They were also asked to mark every misspelling they came across.

Subjects

A sample of 40 children of Grade 1 from a population of 192 children participated. The mean score of the selected group on the reading test was equal to the mean score of the total group of readers from Experiment 1 (27.1; $SD = 3.7$). Reading level was not manipulated in this experiment.

Results

Spelling-test Results.   An analysis of variance on the mean correct choices showed that the children knew the spelling of the base words they had practised only once (infrequent: .95 correct) as well as that of the ones they had practised four times (frequent: .96 correct), both F's < 1.

Proofreading Results.   A 2 (task: silently versus aloud) by 2 (type of misspelling: pseudohomophone versus control) by 2 (base-word frequency: frequent versus infrequent) ANOVA was performed on subject and item means of the proportion of correctly identified misspellings. As in Experiment 1, the children detected fewer pseudohomophone misspellings, in both the silent and the aloud condition (.41 and .52, respectively) than control misspellings (.72 and .83, respectively), as indicated by the significant main effect of type of misspelling, $F(1, 36) = 17.08$, $p < .001$, by subjects, and $F(1, 22) = 20.04$, $p < .001$, by items. The results indicate that the children in the aloud condition showed better proofreading performance (.67) than those in the silent condition (.56), $p = .14$, by

subjects, and $F(1, 22) = 6.92$, $p < .05$, by items. None of the experimental variables interacted significantly with each other. The results are depicted in Figure 2.

Misspellings derived from frequent base words (practised four times in the spelling-training task) were as often (both F's < 1) correctly identified in the proofreading task as their infrequent counterparts (practised once in the spelling-training task), .60 and .63, respectively. The factor base-word frequency did not interact with any of the other experimental variables.

Effect of a Preparatory Spelling Training

To test for the effect of spelling training a 2 (spelling training: absent versus present) by 2 (type of misspelling: pseudohomophone versus control) ANOVA was performed on the mean number of correctly identified misspellings of the subjects of the story condition in Experiment 1 (absent condition) and of the total group of subjects in Experiment 2 (present condition). The children who received a spelling training spotted more mis-spellings (.62) than did those who did not (.44) as shown by the significant main effect of spelling training, $F(1, 76) = 17.44$, $p < .001$, by subjects, and $F(1, 22) = 12.81$, $p < .01$, by items. The superior performance of the children with prior spelling training, however, was restricted to those in the pseudohomophone condition (present: .46: absent: .12). The children in the two control misspelling conditions of Experiments 1 and 2 detected an equal number of misspellings (.77 and .76, respectively). The interaction between type of misspelling and spelling training showed that this differential effect of spelling training was statistically reliable, $F(1, 76) = 15.44$, $p < .001$, by subjects, and $F(1, 22) = 20.72$, $p < .001$, by items.

The analysis on the data of the spelling test showed a marginally significant main effect of spelling training, $F(1, 78) = 3.50$, $.05 < p < .10$, by subjects, and $F(1, 12) = 3.21$, $p = .10$ by items, suggesting that the subjects who received the spelling training knew the spellings better (.96) than those who did not (.92).
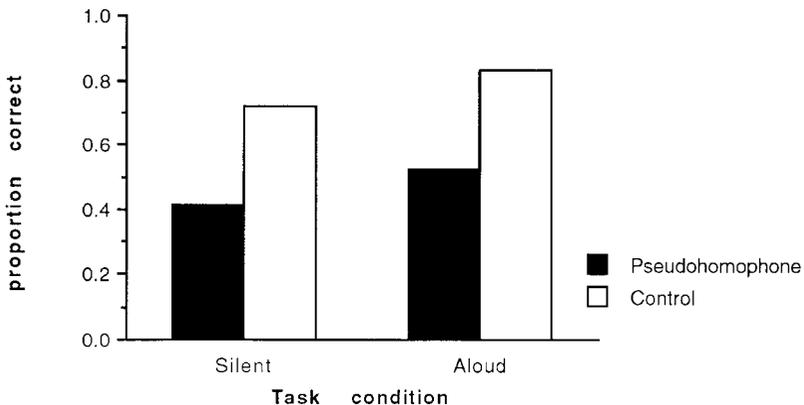

FIG. 2.   Proportion of correctly identified pseudohomophone and control misspellings in the silent and aloud conditions of Experiment 2 (Proofreading).

Discussion

The results of Experiment 2 parallel those of Experiment 1. Support for the assembled phonology hypothesis was found again, because fewer pseudohomophone misspellings were detected than control misspellings. As in Experiment 1, this effect cannot be due to insufficient spelling knowledge of the base-words' spellings, because the subjects showed near-perfect performance in the two-alternative forced-choice spelling test (more than 95% correct). Despite the preparatory spelling training, again there was no effect of base-word frequency, nor an interaction between base-word frequency and type of misspelling.

Furthermore, the results of Experiment 2 indicate that children in the aloud condition spotted more misspellings than did those in the silent condition. However, this effect was present for both the pseudohomophone and control misspellings, suggesting that the task demand did not exert a differential effect on the detection of type of misspellings. Thus, irrespective of whether or not our beginning readers have to articulate reading materials, their reading process is *always* phonologically mediated.

As shown, spelling training did improve the detection of the pseudohomophone misspellings (but not that of the control misspellings). However, one training trial was as effective as four training trials. Spelling training apparently increased the availability of the spelling of the base words, leading to more successful verification. Nevertheless, the effect of pseudohomophone misspellings again occurred, albeit reduced.

Once again no base-word frequency effect materialized. We oppose the view that our choice of frequency levels was such that this variable could not render effects. It may seem too optimistic to assume that copying a word four times constitutes a high-frequency experience, whereas copying it once counts as infrequent experience. But earlier work by Bosman (1994; Bosman & de Groot, 1992), shows that frequency effects are established with a rather limited number of presentations (0 versus 1, and 1 versus 2). However, under the present task conditions 4 versus 1 does not seem to affect verification differentially (cf. Balota & Chumbley, 1984). To establish whether the null-effect of frequency is stable over different types of tasks, the frequency factor will be investigated again in the experiments below.

EXPERIMENT 3
Lexical Decision

In this experiment we manipulated type of misspelling, base-word frequency, and reading level. The main question was again whether beginning readers show evidence of assembled phonology. This was investigated by means of a lexical-decision task in which subjects have to decide whether or not individually presented letter strings constitute words. The verification hypothesis was again tested by manipulating base-word frequency and reading level. Measuring decision latencies allows us to investigate the "getting stuck in verification" hypothesis (see further on). The same preparatory spelling training as in Experiment 2 was conducted again to try to increase the possibility of detecting an interaction between type of misspelling and base-word frequency.

Method

Materials

A frequency count for children (Kohnstamm, Schaerlaekens, de Vries, Akkerhuis, & Froonincksx, 1981) was consulted for the selection of ten semantically highly familiar one-syllable base words. Two different lists (A and B) were made up containing the base words. In List A five base words appeared four times (frequent condition) and the other five appeared only once (infrequent condition). In List B the five frequent words of list A were assigned to the infrequent condition and the five infrequent words to the frequent condition. The total number of words per list was 25.

The ten base words were changed into pseudohomophone misspellings (OS = .62; SD = .10) and control misspellings (OS = .66; SD = .05) with statistically equal OS, to serve as non-words in the lexical-decision task. The first three reading-instruction books of the reading curriculum offered 24 suitable word stimuli (12 high- and 12 low-frequency). Finally, 12 non-words were created to serve as fillers. The actual list of stimuli (56) thus consisted of 10 pseudohomophone misspellings, 10 control misspellings, 12 high-frequency and 12 low-frequency well-known words, and non-words (12). The experimental materials are listed in Appendix B.

Procedure

The experiment consisted of three parts. Subjects started with the spelling training, identical to the one applied in Experiment 2. Children were asked to copy into a note book, in groups of 2, 3 or 4, all 25 words of one of the lists. The spelling training was followed by a spelling test. The subjects received 10 strips of paper, each containing the correct spelling and the incorrect pseudohomophone misspelling of the base word, and were asked to mark the incorrectly spelled words. The final part of the experiment was the lexical-decision task. After an interval of at least half an hour and at most three hours, the subjects were asked to carry out the lexical-decision task individually using a Macintosh Classic computer. Stimuli were presented on the computer screen, and the subjects were asked to decide as quickly as they could whether the stimulus was a word or not. Two buttons (YES and NO) were connected to two separate serial ports of the computer, and subjects were asked to use their dominant hand for the YES-button. Latencies and responses were registered by the computer and the experimenter evaluated the validity of the responses and controlled the presentation of the next trial. Prior to the experiment proper, the subjects received 8 practice trials to become familiar with the task.

The practice trials were also used for instructing the subject on the task. On every trial, the experimenter asked the child which button she or he intended to press and why. Explicit feedback was supplied—that is, if the child not only correctly indicated which button to press, but also gave the correct answer as to why, the experimenter gave positive feedback. If, on the other hand, the child gave an incorrect answer, the experimenter pointed out to the child why his or her answer was wrong. The majority of the children initially erred on the pseudohomophone practice trials, but as soon as the experimenter asked the child whether he or she thought the word was spelled correctly, the child realized that that was not the case and responded by pressing the NO-button. Subjects were reminded several times always to check whether the spelling of the presented word was correct. None of the subjects appeared to have trouble understanding the instruction.

Subjects

From a population of 192 children, 20 were selected according to their scores on the reading test; 10 children were regarded as more advanced readers (mean 31.7; $SD = 4.8$) and 10 as less advanced readers (mean 19.6; $SD = 2.1$). The difference between reader groups was significant, $F(1, 19) = 53.50$, $p < .001$.

Results

Spelling-test Results.   A 2 (reading level: more advanced versus less advanced) by 2 (base-word frequency: frequent versus infrequent) ANOVA on subject and item means showed that the children generally knew the spelling of the base words (mean proportion correct .95). Neither the main effect of reading level nor the interaction between reading level and frequency was significant. But there was a main effect of base-word frequency, showing that both groups of readers had a significantly higher score on the frequent (.98) than on the infrequent words (.91), $F(1, 18) = 5.49$, $p < .05$, by subjects, and $F(1, 9) = 3.65$, $.05 < p < .10$, by items.

Lexical-decision Results.   Proportions of correctly identified errors—that is, correct non-word decisions—were calculated for pseudohomophone and control misspellings derived from frequent and infrequent base words. A 2 (reading level: more advanced versus less advanced) by 2 (type of misspelling: pseudohomophone versus control) by 2 (base-word frequency: frequent versus infrequent) ANOVA on correctly identified errors was performed on subject and item means. Pseudohomophone misspellings were less often correctly classified as non-words (.54) than control misspellings (.91), as indicated by the significant main effect of type of misspelling, $F(1, 18) = 62.79$, $p < .001$, by subjects, and $F(1, 18) = 27.17$, $p < .001$, by items). Overall, more advanced readers showed better error detection performance (.79) than less advanced readers (.66), $F(1, 18) = 8.70$, $p < .01$, by subjects, and $F(1, 18) = 12.79$, $p < .01$, by items. Although more advanced readers classified significantly more pseudohomophone misspellings correctly as non-words (.70) than did less advanced readers (.39), the two groups classified an equal number of control misspellings correctly as non-words (.89 and .93, respectively), as revealed by the significant interaction between reading level and type of misspelling, $F(1, 18) = 14.13$, $p < .01$, by subjects, and $F(1, 18) = 21.49$, $p < .001$, by items. Figure 3 presents the interaction. There was no significant main effect of base-word frequency, $F(1, 18) = 1.73$, nor a significant interaction between base-word frequency and type of misspelling ($Fs < 1$) or any of the remaining variables.

Subject's mean decision latencies of correct no-responses to pseudohomophone misspellings (54%), correct no-responses to control misspellings (91%), correct yes-responses to words (94%), and incorrect yes-responses to pseudohomophone misspellings (46%) were also computed. The mean decision latency of correct no-responses to pseudohomophone misspellings (3734 msec, $SD = 1000$) was not significantly different from that of correct no-responses to control misspellings (3886 msec, $SD = 988$), $F < 1$. The mean decision latency of correct yes-responses to words (2177 msec, $SD = 622$), however, was significantly shorter than that of incorrect yes-responses to pseudohomo-
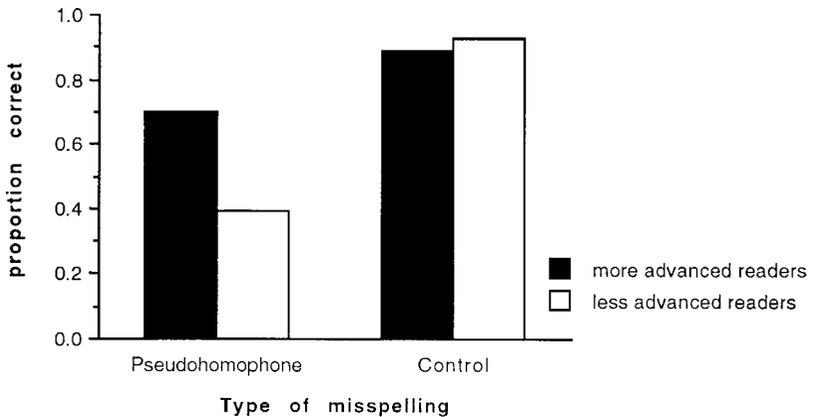
FIG. 3.   Proportion of correct no-responses to pseudohomophone and control misspellings of the more advanced and less advanced readers collapsed across the frequent and infrequent conditions of Experiment 3 (Lexical decision).

phone misspellings (3627 msec, SD = 1016), $F(1, 18)$ = 33.74, p < .001. Finally, the mean decision latency of correct no-responses to pseudohomophone misspellings (3734 msec) was not significantly different from that of incorrect yes-responses to pseudohomophone misspellings (3627 msec), F < 1.

Discussion

The results of the lexical decision task[2] parallel those of the two foregoing proofreading tasks. Positive evidence was obtained again for assembled phonology, because fewer pseudohomophone misspellings were correctly classified as non-words than control misspellings. It was also clear that all children knew the spelling of the base words rather well (an average of 95% correct). Evidence for the verification hypothesis was supplied by the interaction between type of misspelling and reading level. As in the proofreading task, more advanced readers were better at correctly classifying as non-words the pseudohomophone misspellings than the less advanced readers, but the two groups were equally good at detecting control misspellings. However, no effect of base-word frequency nor any interaction between base-word frequency and type of misspelling emerged. This result also parallels those of the two proofreading experiments.

Conclusions based on the latency data, particularly the yes-response latencies, have to be approached with caution, because the condition means are based on a rather limited number of trials. Like the skilled readers in the lexical-decision experiment of Van Orden et al. (1992), our beginning readers showed no reliable difference in time to respond correctly that pseudohomophone misspellings and control misspellings were not words. On the other hand, however, our beginning readers took considerably more time to

---

[2] The same lexical-decision task with reading training instead of spelling training as a preparatory task was performed with a similar group of beginning readers. The results of that study were nearly identical to those of the present investigation.

respond incorrectly that pseudohomophone misspellings were words than to respond correctly that words were indeed words. Van Orden, Johnston, and Hale (1988) observed the same result in a semantic-categorization task with skilled readers, but attributed this to outliers. Indeed, when they truncated the latency distributions, the difference in response times between pseudohomophone misspellings and words disappeared.

What process could underlie incorrect yes-decisions to pseudohomophones? Such a decision is clearly not easy, given the fact that it took the children nearly 1500 msec longer than to decide correctly that a word was a word. Subjects in this task appear to be caught in an impasse between the outcome of the phonology process—accepting because it *sounds* right—and the outcome of the spelling verification process—rejecting because it does not *look* right (cf. Van Orden, 1987, p. 191). The fact that the children, despite long viewing times, were still not able eventually to reject the pseudohomophone misspellings is taken as a further sign of a not fully developed verification process. Whereas the yes-response latencies (3627 msec) may show the difficulty these children had accepting pseudohomophones, the no-latencies (3734 msec) show their problems with rejecting them.

It took the children nearly as long to reject pseudohomophone misspellings as to reject non-word controls. Rejection of non-words is difficult for these subjects because they tend to accept most word-like stimuli as real words. One has to bear in mind that these beginning readers learn new words every day.

To complete this study, a semantic-categorization experiment was executed. According to many theorists, meaning evaluation is obligatory in semantic categorization, but not in lexical decision.

EXPERIMENT 4
Semantic Categorization

As in all the foregoing experiments, the main question was whether word identification in beginning readers shows signs of assembled phonology. The same variables (type of misspelling, base-word frequency, and reading level) as in the lexical-decision experiment were manipulated in a semantic-categorization task. The subject was confronted with a category name such as "flower", followed by a stimulus—the target—that was either a category exemplar (e.g. *rose*) or not (e.g. *dog*). The subjects were required to press a YES-button if the target was an exemplar of the pre-designated category and a NO-button if it was not. The verification hypothesis was also tested again by manipulating reading level and base-word frequency. Prior to the semantic-categorization task, subjects received preparatory spelling training identical to that of Experiment 3.

The same experimental stimuli as used in the lexical-decision task were employed in the semantic-categorization experiment. This enabled us to compare misspelling detection performance in the two tasks. We assumed that the task constraints of lexical decision are more favourable for spelling verification than those of semantic categorization, which should be apparent from a higher pseudohomophone misspelling detection score in the lexical-decision than in the semantic-categorization task.

Method

Materials

The lists (A and B) developed for the spelling training in the lexical-decision task (Experiment 3) were used in this experiment in the same way as materials in the spelling-training phase (see Materials Section of Experiment 3) and for the pseudohomophone and control misspellings in the categorization task. Twenty semantically highly familiar base-words were chosen from the frequency count for children (Kohnstamm et al., 1981) to serve as fillers; 10 words were actual members of the category (exemplar condition), and 10 were not (non-exemplar condition). The total list of stimuli consisted of 40 stimuli (10 pseudohomophone misspellings, 10 control misspellings, 10 exemplar, and 10 non-exemplar words). The words were instances of 10 different semantic categories. The experimental materials are listed in Appendix B.

Procedure

This experiment also consisted of three stages. The spelling training and the spelling test were identical to those of Experiment 3. After an interval of at least half an hour and at most three hours, the subjects carried out the semantic-categorization task individually on a Macintosh Classic computer. The procedure was a little different from that used in experiments with skilled adult readers (cf. Van Orden, 1987). In our experiment the experimenter named the semantic category and immediately presented the target stimulus on the screen by pushing a key on the keyboard. The child was asked to decide as quickly as possible whether the word on the screen was a member of the category just named by the experimenter. There were two reasons for conducting the experiment this way. The first was that reading the semantic category was thought to be too complicated in some cases, because these words are often difficult for beginners to read. The second reason was that the children might get confused having to change the response every other event (reading and remembering the category of the first stimulus, followed by reading and deciding on the membership of the second). The new procedure appeared to be satisfactory. Prior to the experimental session, subjects received extensive instruction and practice in the semantic-categorization task, the same way as the children in the lexical-decision task of Experiment 3. All subjects appeared to understand the task correctly.

Subjects

The beginning readers (n = 28) who participated in this experiment were all from the same classroom. They were instructed according to the curriculum that was also used by the children of all three foregoing experiments. However, the group that took part in Experiment 4 had moved to the revised edition. This meant that the initial words of instruction differed from those in the old edition, although the basic principles of the curriculum remained unchanged. Again the subjects were divided in two groups: more advanced readers (29.3, SD 4.3, n = 13) and less advanced readers (18.9, SD 2.5, n = 15), as based on the scores of the reading test, $F(1, 24) = 64.24$, $p < .001$.

Results

Spelling-test Results.   The subjects knew the spelling of the base words (mean proportion correct = .95). A 2 (reading level: more advanced versus less advanced) by 2 (base-word frequency: frequent versus infrequent) ANOVA on subject and item means showed that more advanced readers knew the spelling of the base words better (.98) than

did the less advanced readers (.92), $F(1, 26) = 5.37$, $p < .05$, by subjects, and $F(1, 9) = 9.95$, $p < .05$, by items. The main effect of base-word frequency and the interaction between reading level and base-word frequency were not significant (all F's < 1).

Semantic-categorization Results.   Analyses similar to those of Experiment 3 were performed on the data of this experiment. A 2 (reading level: more advanced versus less advanced) by 2 (type of misspelling: pseudohomophone versus control) by 2 (base-word frequency: frequent versus infrequent) ANOVA on the subjects' mean number of correct no-responses was performed. The children correctly classified pseudohomophone mis-spellings as non-exemplars (.32) less often than control misspellings (.98), $F(1, 26) = 354.99$, $p < .001$, by subjects, and $F(1, 18) = 251.89$, $p < .001$, by items. Neither the main effect of reading level (both F's < 1) nor the interaction between type of misspelling and reading level reached significant levels ($F < 1$, by subjects, and $p > .15$, by items). Figure 4 shows the results. The effect of base-word frequency was also not significant (F's < 1) and no interaction occurred between base-word frequency and type of misspelling, $F(1, 26) = 1.13$ by subjects, $F(1, 18) = 1.9$ by items, or any of the remaining experimental variables.

Mean decision latencies of correct no-responses to pseudohomophone misspellings (31%), correct no-responses to control misspellings (97%), correct yes-responses to words (93%), and incorrect yes-responses to pseudohomophone misspellings (68%) were also computed. The mean decision latency of correct no-responses to pseudo-homophone misspellings (3812 msec, $SD = 1667$) was not reliably different from that of correct no-responses to control misspellings (3937 msec, $SD = 1340$), $F < 1$. The mean decision latency of correct yes-responses to exemplar words (2339 msec, $SD = 672$), however, was significantly shorter than that of incorrect yes-responses to pseudohomo-phone misspellings (3227 msec, $SD = 1569$, $F(1, 26) = 9.31$, $p < .001$). Finally, the mean decision latency of correct no-responses to pseudohomophone misspellings (3812 msec) was not significantly different from that of incorrect yes-responses to pseudohomophone misspellings (3227 msec), $F < 1$.
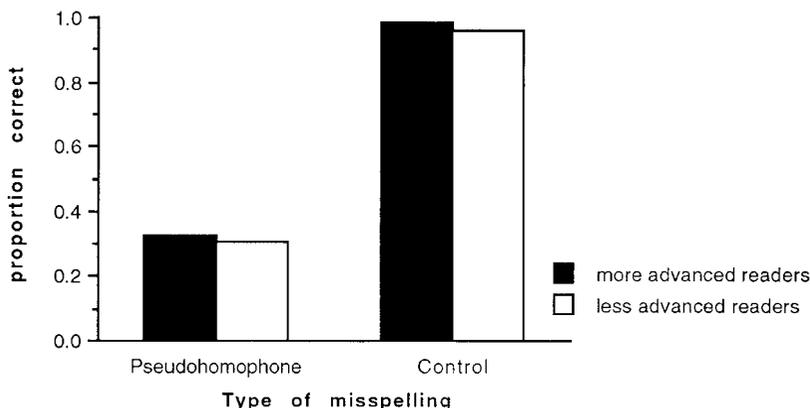

FIG. 4   Proportion of correct no-responses to pseudohomophone and control misspellings of the more advanced and less advanced readers collapsed across the frequent and infrequent conditions of Experiment 4 (Semantic categorization).

Data Analysis of Experiments 3 and 4 Combined

An analysis on the combined data of the subjects of the lexical-decision task and the semantic-categorization task was performed to test for an effect of task. A 2 (reading level: more advanced versus less advanced) by 2 (task: lexical decision versus semantic categorization) by 2 (type of misspelling: pseudohomophones versus control) ANOVA was executed on subject and item means. The discussion of the results wil be limited to those that pertain to the task variable.

The children in the lexical-decision task rejected more misspellings (.73) than those in the semantic-categorization task (.65), as shown by the significant main effect of task, $F(1, 44) = 7.04$, $p < .01$, by subjects, and $F(1, 18) = 2.44$, $p > .15$, by items. The children in the lexical-decision task rejected significantly more pseudohomophone misspellings (.54), $F(1, 44) = 17.44$, $p < .001$, by subjects, and $F(1, 18) = 11.49$, $p < .01$, by items, but significantly fewer control misspellings (.91), $F(1, 44) = 8.78$, $p < .01$, by subjects, and $F(1, 18) = 1.39$, $p > .20$, by items, than those in the semantic-categorization task (.33 and .98, respectively), which is indicated by the significant interaction between task and type of misspelling, $F(1, 44) = 26.22$, $p < .001$, by subjects, and $F(1, 18) = 10.44$, $p < .01$, by items. The interaction between reading level and task did not reach significance. The three-way interaction between reading level, task, and type of misspelling was also significant, $F(1, 44) = 9.55$, $p < .01$, by subjects, and $F(1, 18) = 5.87$, $p < 0.05$, by items). This three-way interaction reflects the presence of a two-way interaction between reading level and type of misspelling in Experiment 3 that is not apparent in Experiment 4.

To make sure that the subjects in the lexical-decision task did not differ on the variable reading level from those in the semantic-categorization task, an analysis on their reading scores was performed. There was neither a significant interaction between task and reading level ($F < 1$), nor a significant main effect of task.

Discussion

As in all foregoing experiments, a strong effect of assembled phonology showed up in the semantic-categorization task, which confirms the results obtained with adult English-speaking readers (Coltheart, Patterson, & Leahy, 1994; Van Orden et al., 1988; Van Orden et al., 1992). Both more and less advanced readers had considerably more trouble rejecting a pseudohomophone misspelling as a category exemplar than a control misspelling. Again, this difference in performance on the two types of misspellings cannot be due to insufficient spelling knowledge, because the children's performance on the spelling test was nearly perfect (proportion correct = .95).

Also, as in all of the present studies, there was no sign of a base-word frequency effect nor an interaction between base-word frequency and type of misspelling. However, in contrast to the first proofreading and the lexical-decision experiment, no effect of reading level emerged. More and less advanced readers had the same proportion of correctly identified pseudohomophone and control misspellings.

The children performing the lexical-decision task detected more pseudohomophones, but fewer control misspellings (cf. Van Orden et al., 1992, pp. 263–265). As in the story

versus list condition of Experiment 1, the advanced readers' superior performance to pseudohomophone misspelling from the lexical-decision task drops to that of less advanced readers in the semantic-categorization task. This suggest that verification in more advanced readers can be led astray when context does not favour spelling verification, as is the case in the semantic-categorization and the story-proofreading tasks.

Categorization latencies on no-responses to pseudohomophone and control misspellings were identical, whereas latencies on incorrect yes-responses to pseudohomophone misspellings were much longer than on correct yes-responses to words. This latter finding is in line with the results of our lexical-decision task (Experiment 3) and that of the semantic-categorization task of Van Orden et al. (1988). It appears again that spelling verification can be caught in an impasse between rejection and acceptance criteria.

GENERAL DISCUSSION

The general pattern of our results is clear-cut. In all tasks, all readers had more trouble detecting pseudohomophone misspellings than control misspellings. These differences are unlikely to be due to lack of spelling knowledge of the base words, because the spelling tests that were administered in all four experiments showed good spelling knowledge of the base words for all subject groups. We take these results as evidence that phonologic mediation is fundamental to word identification in Dutch-speaking beginning readers. These findings parallel those of Van Orden et al. (1992), who conducted similar experiments with skilled adult English-speaking readers.

We were also interested in the role of spelling verification in word identification. We tested the spelling-verification hypothesis in several ways. We manipulated base-word frequency (both extant and experimentally manipulated). It was hypothesized that the spellings of familiar base words are more available to verification than those of less familiar base words. If so, subjects should detect a higher proportion of pseudohomophones derived from high-frequency base words than of those corresponding to low-frequency base words. That is what Van Orden et al. (1992) found in their skilled English readers. For our beginning Dutch readers, however, all tasks and all conditions and all reader groups yielded null-effects of base-word frequency. It is not clear whether an effect of base-word frequency is genuinely absent in beginning readers or whether an unsuitable assessment of the frequency variable is responsible for this persistent null-effect. Further research is required to investigate this issue.

We also investigated the verification hypothesis by contrasting the performance of more and less advanced readers. Presumably, enhanced verification is correlated with reading skill. In most of the tasks (the semantic-categorization task and the story condition in the first proofreading experiment being the exceptions), more advanced readers outperformed less advanced readers on the pseudohomophones, but they showed equal performance on the control misspellings. Not only did we find support for the assumption that effective spelling verification is positively related to reading skill, but also that spelling verification is not yet stable in beginning readers. This is suggested by the finding that the proofreading behaviour of more advanced readers drops to the level of that of less advanced readers in some reading tasks. We will comment further on this issue.

Although we interpret the effect of reading skill in terms of better-developed verification, an alternative interpretation has been provided by Doctor and Coltheart (1980). They equate reading skill with an increased use of direct access. Their explanation was based on the result of a sentence judgement task. Meaningless sentences that were meaningful if phonologically recoded (e.g. *He ran threw the street*) generated more incorrect "yes" responses than did control meaningless sentences (e.g. *He ran sew the street*). However, the error rate to phonologically meaningful sentences was greater for the younger than for the older children. The difference scores between the two types of sentences diminished with increasing age. Doctor and Coltheart concluded that young children used phonologic recoding to a greater extent than do older children and that, conversely, older children rely more on direct access.

Because Doctor and Coltheart (1980) used *homophones* in their experiment, they could invoke the direct-access hypothesis as an explanation for the better performance of older children. In our experiment, however, we made exclusive use of *pseudohomophones*. We are not aware of any version of dual-route theory that has explained how direct access can prevent a reader from making a false-positive error to a pseudohomophone. We will explore two solutions. One possibility might be to raise the threshold of familiar "direct access" lexical entries. To cross the threshold, a more complete accumulation of stimulus features—that is, spelling and phonologic features—is required. Although this hypothesis sounds reasonable, it remains difficult to avoid errors to pseudohomophones without an emphasis on spelling features. Relying on the phonological features is deceptive, because they will match exactly those of the base word.

Another possible account of how direct access can prevent a reader from making a false positive error to a pseudohomophone would explicitly emphasize spelling features. The assembled phonology of a pseudohomophone causes errors because it exactly matches the phonology of a base word, and it activates a phonological representation of its corresponding base word (see Van Orden et al., 1992, and the present experiments). One way to reduce the error rate to pseudohomophones would be to mark each familiar "direct access" entry such that activation from assembled phonology is registered, but this activation is no longer enough to cause lexical access (as it was prior to the availability of direct access). Lexical access would then depend upon a match between spelling features of the stimulus and those of the lexical entry (Forster, 1976). This solution, however, seems indistinguishable from the spelling-verification procedure we propose here. Thus, contrary to the claims of dual process theory, direct access cannot protect a reader from making false-positive errors to pseudohomophones unless it is reformulated as verification.

At this point we comment briefly on the theoretical notion of spelling-verification as a component process. As was pointed out to us by a reviewer of an earlier version of this article, the spelling-verification hypothesis is tautological in the present case—if a spelling error is detected, verification was successful; if a spelling error is not detected, verification failed. This critique is justified, but we wonder how any contemporary hypothesis can fully escape this criticism. For example, a common argument made in favour of dual-process theory directly parallels the previous point concerning verification. If subjects make more errors to (pseudo)homophones than to controls, then phonologic mediation is present; if subjects do not make more errors to (pseudo)homophones than to controls,

phonologic mediation is absent (amongst others, Coltheart et al., 1988; Doctor & Coltheart, 1980; Jared & Seidenberg, 1991). Similarly, a recent debate over frequency-modulated lexical access was framed in terms of the presence and absence of frequency effects (Balota & Chumbley, 1984; Forster, 1992; Monsell, 1991). In that debate, Forster (1992) noted that ". . . this discussion makes sense only if we believe that there is a sharp distinction between processes that occur during access and those that take place after access has been completed. This is not an empirical issue: the distinction can only be made *vis-à-vis* some particular theoretical model" (emphasis in original, p. 423). The same point would seem to apply in discussions of any component processes of word identification, including the verification hypothesis.

Thus, verification appears an inevitable hypothesis when errors to pseudohomophones are concerned (as was argued previously), and verification gives a coherent account of additional details of our findings. A comparison of the results of the semantic-categorization and the lexical-decision tasks illustrates this. Recall that the children who partici-pated in the semantic-categorization task detected fewer pseudohomophone misspellings (33% vs. 54%), but more control misspellings (98% vs. 91%), than did those in the lexical-decision task. These are very interesting results. Why do subjects in the semantic categorization task fail to detect pseudohomophones, when they are so successful in detecting control misspellings? Also, why do the results observed in the lexical-decision task differ from those in the semantic-categorization task? Stated differently, why would subjects be more successful on pseudohomophone misspellings in one task than in the other? It appears that some constraints on performance are due to task demands.

The results of the present findings are readily explained by assuming that in all tasks phonologic activation is a primary constraint, and in some tasks context adds additional constraints that work against successful verification (see also Van Orden et al., 1992). Spelling verification in the more advanced readers appears easily destabilized by strong contextual constraints, as in the stories that were proofread and in the semantic categor-ization task.

A similar context effect was established in the study of Jared and Seidenberg (1991). Verification in their semantic-categorization task suffered in the condition where specific categories (e.g. *type of money* as category for the exemplar *cent*) were used as compared to a condition in which broad categories (e.g. *object* as category for the exemplar *cent*) were used. In the specific-category condition, subjects detected fewer homophone misspellings than control misspellings when targets were either high- or low-frequency homophones. But in the case of broad categories, the difference between high-frequency homophone targets derived from low-frequency base words and control misspellings disappeared.[3]

Jared and Seidenberg (1991), however, provided a different interpretation of their results. They argued that in the specific condition (e.g. *type of money*), priming of the

---

[3] In addition, the "true" absence of the pseudohomophone effect in a semantic-categorization task using broad categories is seriously challenged by a recent experiment of Van Orden, Aitchison, and Podgornik (submitted). They conducted a semantic-categorization experiment using broad categories and high-frequency words only. When the target stimuli were presented for a relatively long time (200 msec) no homophone effect emerged, thus replicating the effect of Jared and Seidenberg (1991). Brief presentation times (50 msec) of the target stimuli, however, yielded a clear homophone effect (see also Van Orden & Goldinger, 1995).

exemplar was responsible for the increased homophone effect.[4] The smaller or absent homophone effect in the broad condition (e.g. *object*) indicated that phonologic mediation is generally absent in the processing of high-frequency words. They concluded, therefore, that phonologic information only contributes to meaning activation of low-frequency words and not to that of high-frequency words, suggesting two different processes for high- and low-frequency words (see also Paap & Noel, 1991; Paap, Noel, & Johansen, 1992; Seidenberg, Waters, Barnes, & Tanenhaus, 1984).

We believe that it is not necessary to postulate a lexical and non-lexical process for high- and low-frequency words, respectively. The results of Jared and Seidenberg (1991) are readily interpreted in terms of changing contextual constraints between specific and broad categories. Spelling verification suffers more in a condition that supplies strong constraints in favour of the exemplar identity, as in the condition of specific categories.

In other words, we propose an interpretation of our results and those of Jared and Seidenberg (1991) in terms of a common framework in which assembled phonology underlies word recognition in all tasks, and we assume that spelling verification, a necessary requirement when pseudohomophones are to be detected, suffers in situations where stimulus identities are strongly constrained by context.

If the pseudohomophone effect, as obtained here with beginning readers and with skilled readers by Van Orden (1987, 1991; Van Orden et al., 1988; Van Orden et al., 1992), legitimizes the inference that assembled phonology underlies the reading of beginning and skilled readers alike, we have to conclude that a developmental shift of the use of phonology in beginning reading does not occur. This conclusion deviates from that in much of the earlier research on word recognition in beginning and skilled readers (see Introduction and below), where developmental differences in the occurrence of phonologic mediation are typically assumed.

The results of this study add to the growing body of evidence that phonologic mediation is obligatory in word recognition. Not only is the present new evidence obtained with beginning rather than skilled readers, but it is also obtained in Dutch rather than English. As was said before, the controversy on how reading develops still exists. There are authors who assume that initially reading is characterized by a heavy reliance on phonologic recoding and that with increased experience with written material the role of phonology diminishes or vanishes (Backman et al., 1984; Reitsma, 1983b). Conversely, there are those who hold the opposite—namely, that the developmental shift is from accessing the mental lexicon directly (without phonologic mediation) to processing words directly as well as indirectly (see for a review Rayner & Pollatsek, 1989, pp. 364–382). Our results show that none of these positions is fully tenable. It seems that proponents of the first view are right in their assumption that beginning reading is dominated by phonology, but, contrary to their belief, the role of phonology does not

----

[4] The proposed explanation by Jared and Seidenberg (1991) in terms of priming the appropriate phonology in the case of specific categories is ruled out in our experiment. Eleven children from the same population as those who participated in the present experiments were asked to write down the first thing that came to their mind when they were presented with each category name—for example, what is the first thing you think of when I say "colour". All 10 categories were presented to the children (see Appendix B). Of a total of 110 responses, only one was the exemplar used in the experiment ("bee" for ANIMAL).

vanish, nor does it seem to diminish. Support in the present data for the latter claim is that our more advanced (beginning) readers showed the same phonologic effects as did their less skilled classmates. Additional support comes from the similarities between our beginning readers and the skilled readers of Van Orden et al. (1992).

As was said in the introduction, the Dutch language has a fairly shallow orthography. One could infer from this that the present findings were expected, because a phonologic strategy suits a shallow orthography. A role for phonology may, indeed, seem more obvious in a language like Dutch than in a language with a deeper orthography, but nevertheless a survey of studies on different languages and writing systems suggests that our results with beginning readers are not exceptional. Studies of both shallow alphabetical writing systems like Serbo-Croatian (Lukatela, Carello, & Turvey, 1990; Lukatela, Turvey, & Todorovic, 1991) and Dutch (Bosman, 1994; Bosman & de Groot, 1995), and deeper alphabetical orthographies like French (Ferrand & Grainger, 1992), English (among others: Berent & Perfetti, 1995; Lesch & Pollatsek, 1993; Lukatela & Turvey, 1993, 1994; Perfetti & Bell, 1991), and Hebrew (Frost, 1995; Frost & Kampf, 1993; Navon & Shimron, 1981) all indicate a primary role for phonology (but see also for null-effects, e.g., Baluch & Besner, 1991, for Persian; Tabossi & Laghi, 1992, for Italian).

The evidence for phonology as a primary constraint on word recognition is even more compelling when it can be obtained from non-alphabetical writing systems like Chinese and Japanese. Indeed, Perfetti and Zhang (1991) showed that phonologic activation in Chinese readers occurs, and automatically so, and the work of Wydell, Patterson, and Humphreys (1993) suggests that reading logographic Japanese *Kanji* is achieved by parallel access to meaning from both orthography and phonology. They state that their results can be incorporated in a modified version of Van Orden's (1987) model on word recognition.

Given the ubiquity of phonologic effects, irrespective of the type of orthography under investigation and, as stressed in this study, the reader groups tested and the tasks used, our answer to the question of Carello, Turvey, and Lukatela (1992): "Can theories of word recognition remain stubbornly nonphonological?" is a "No" full of conviction.

REFERENCES

Backman, J., Bruck, M., Hebert, M., & Seidenberg, M.S. (1984). Acquisition and use of spelling–sound correspondences in reading. *Journal of Experimental Child Psychology*, *38*, 114–133.

Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 340–357.

Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 644–652.

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading word aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 386–393.

Barron, R.W. (1986). Word recognition in early reading: A review of the direct and indirect access hypotheses. *Cognition*, *24*, 93–119.

Becker, C.A. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance 2*, 556–566.

Berent, I., & Perfetti, C.A. (1995). A rose is a reez: The two-cycles model of phonology assembly in reading English. *Psychological Review*, *102*, 146–184.

Bleichrodt, N., Drenth, P.J.D., Zaal, J.N., & Resing, W.C.M. (1984). *Revisie Amsterdamse Kinder Intelligentie Test*, *RAKIT* [Revision Amsterdam Children Intelligence Test]. Lisse: Swets & Zeitlinger.

Bosman, A.M.T. (1994). *Reading and spelling in children and adults: Evidence for a single-route model*. Doctoral dissertation, University of Amsterdam, Dissertatie reeks 1994–2, Faculteit Psychologie.

Bosman, A.M.T., & de Groot, A.M.B. (1992). Differential effectiveness of reading and non-reading tasks in learning to spell. In F. Satow & B. Gatherer (Eds.), *Literacy without frontiers*. United Kingdom Reading Association. Cheshire: Widnes.

Bosman, A.M.T. & de Groot, A.M.B. (1995). Evidence for assembled phonology in beginning and fluent readers as assessed with the first-letter-naming task. *Journal of Experimental Child Psychology*, *59*, 234–259.

Caesar, F.B. (1970). *Eén-Minuut-Test voor de technische leesvaardigheid* [One-minute test for reading decoding]. Tilburg: Zwijsen.

Caesar, F.B. (1979). *Veilig leren lezen: Structuurmethode voor het aanvankelijk leesonderwijs.* [Learning to read safely: Structure method for beginning reading]. Tilburg: Zwijsen.

Carello, C., Turvey, M.T., & Lukatela, G. (1992). Can theories of word recognition remain stubbornly nonphonological? In R. Frost and L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 211–226). Amsterdam: Elsevier Science Publishers B.V.

Cohen, G. (1980). Reading and searching for spelling errors. In U. Frith (Ed.), *Cognitive processes in spelling*. London: Academic Press.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151–216). New York: Academic Press.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*, 589–608.

Coltheart, V., Avons, S.E., Masterson, J., & Laxon, V.J. (1991). The role of assembled phonology in reading comprehension. *Memory & Cognition*, *19*, 387–400.

Coltheart, V., & Laxon, V.J. (1990). The development of phonological mediation in reading comprehension. *Language and Cognitive Processes*, *5*, 81–104.

Coltheart, V., Laxon, V.J., Keating, G.C., & Pool, M.M. (1986). Direct access and phonological encoding processes in children's reading: Effects of word characteristics. *British Journal of Educational Psychology*, *56*, 255–270.

Coltheart, V., Laxon, V., Rickard, M., & Elton, C. (1988). Phonological recoding in reading for meaning by adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 387–397.

Coltheart, V., Patterson, K., & Leahy, J. (1994). When a rows is a rose: Phonological effects in written word comprehension. *The Quarterly Journal of Experimental Psychology*, *47a*, 917–955.

Condry, S.M., McMahon-Rideout, M., & Levy, A.A. (1979). A developmental investigation of selective attention to graphic, phonetic, and semantic information in words. *Perception & Psychophysics*, *27*, 88–94.

Davelaar, E., Coltheart, M., Besner, D., & Jonasson, J.T. (1978). Phonological recoding and lexical access. *Memory & Cognition*, *6*, 391–402.

Doctor, E.A., & Coltheart, M. (1980). Children's use of phonological encoding when reading for meaning. *Memory & Cognition*, *8*, 195–209.

Ehri, L.C., & Roberts, K.T. (1979). Do beginners learn printed words better in contexts or in isolation? *Child Development*, *50*, 675–685.

Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked non-word priming. *The Quarterly Journal of Experimental Psychology*, *45A*, 353–372.

Forster, K.I. (1976). Accessing the mental lexicon. In R.J. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 257–287). Amsterdam: North-Holland.

Forster, K.I. (1992). Memory-addressing mechanisms and lexical access. In R. Frost and L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 413–434). Amsterdam: Elsevier Science Publishers B.V.

Frost, R. (1995). Phonological computation and missing vowels: Mapping lexical involvement in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 398–408.

Frost, R., & Kampf, M. (1993). Phonetic recoding of phonologically ambiguous printed words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 23–33.

Grossberg, S., & Stone, G.O. (1986). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review*, *93*, 46–74.

Jared, D., & Seidenberg, M.S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General*, *120*, 358–394.

Kohnstamm, G.A., Schaerlaekens, A.M., de Vries, A.K., Akkerhuis, G.W., & Froonincksx, M. (1981). *Nieuwe streeflijst woordenschat voor 6-jarigen*. [New target-vocabulary list for 6-year-olds]. Lisse: Swets & Zeitlinger.

Lesch, M.F., & Pollatsek, A. (1993). Automatic access of semantic information by phonological codes in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 285–294.

Lukatela, G., Carello, C., & Turvey, M.T. (1990). Phonemic priming with words and pseudowords. *The European Journal of Cognitive Psychology*, *2*, 375–394.

Lukatela, G., Turvey, M.T., & Todorovic, D. (1991). Is alphabet biasing in bialphabetical word perception automatic and prelexical? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 653–663.

Lukatela, G., & Turvey, M.T. (1993). Similar attentional, frequency, and associative effects for pseudo-homophones and words. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 166–178.

Lukatela, G., & Turvey, M.T. (1994). Visual lexical access is initially phonological: 1. Evidence from associative priming by words, homophones and pseudohomophones. *Journal of Experimental Psychology: General*, *123*, 107–128.

Marsh, G., Friedman, M., Welch, V., & Desberg, P. (1980). A cognitive–developmental theory of reading acquisition. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 339–353). New York: Academic Press.

Mommers, M.J.C., & van Dongen, D. (1986). *Zorgverbreding bij het leren lezen*. Tilburg: Zwijsen.

Monsell. S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G.W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 148–197). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Navon, D., & Shimron, J. (1981). Does word naming involve grapheme-to-phoneme translation? Evidence from Hebrew. *Journal of Verbal Learning and Verbal Behavior*, *20*, 97–109.

Paap, K.R., Newsome, S.L., McDonald, J.E., & Schvaneveldt, R.W. (1982). An activation-verification model for letter and word recognition: The word superiority effect. *Psychological Review*, *89*, 573–594.

Paap, K.R., & Noel, R.W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research*, *53*, 13–24.

Paap, K.R., Noel, R.W., & Johansen, L.S. (1992). Dual-route models of print to sound: Red herrings and real horses. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 293–318). Amsterdam: Elsevier Science Publishers B.V.

Patterson, K.E. (1982). The relation between reading and phonological coding: Further neuropsychological observations. In A.W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 77–111). London: Academic Press.

Patterson, K., & Coltheart, V (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 421–447). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Perfetti, C.A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language*, *30*, 473–485.

Perfetti, C., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 633–643.

Raven, J.C. (1958). *Standard Progressive Matrices*. London: H.K Lewis.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice-Hall.

Reitsma, P. (1983a). Word-specific knowledge in beginning reading. *Journal of Research in Reading*, *6*, 41–56.

Reitsma, P. (1983b). *Phonemic and graphemic codes in learning to read*. Unpublished doctoral dissertation. Free University of Amsterdam, Amsterdam, The Netherlands.

Reitsma, P. (1984). Sound priming in beginning readers. *Child Development*, *55*, 406–423.

Reitsma, P., & Verhoeven, L. (1990). Acquisition of written Dutch: An introduction. In P. Reitsma & L. Verhoeven (Eds.), *Acquisition of reading in Dutch* (pp. 1–13). Dordrecht, The Netherlands: Foris Publications.

Rubinstein, H., Lewis, S.S., & Rubinstein, M.A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, *10*, 645–657.

Seidenberg, M.S., Waters, G.S., Barnes, M.A., & Tanenhaus, M.K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior*, *23*, 383–404.

Seymour, P.H.K., & Elder, L. (1986). Reading without phonology. *Cognitive Neuropsychology*, *3*, 1–40.

Stone, G.O., & Van Orden, G.C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 744–774.

Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory & Cognition*, *20*, 303–313.

van Heuven, V.J. (1980). Aspects of Dutch orthography and reading. In J.F. Kavanagh & R.L. Venezky (Eds.), *Orthography, reading, and dyslexia* (pp. 57–73). Baltimore, MD: University Park Press.

Van Orden, G.C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, *15*, 181–198.

Van Orden, G.C. (1991). Phonologic mediation is fundamental to reading. In D. Besner & G.W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 77–103). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Van Orden, G.C., Aitchison, C.S., & Podgornik, M.N. (submitted). When a rows is not a rose: Null effects and the absence of mental structure.

Van Orden, G.C., & Goldinger, S.D. (1995). Phonologic mediation in skilled and dyslexic reading. In C. Chase, G. Rosen & G. Sherman (Eds.), *Developmental dyslexia: Neuro, cognitive and genetic mechanisms*. Timonium, MD: York Press.

Van Orden, G.C., Johnston, J.C., & Hale, B.L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 371–386.

Van Orden, G.C., Pennington, B.F., & Stone, G.O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, *97*, 488–522.

Van Orden, G.C., Stone, G.O., Garlington, K.L., Markson, L.R., Pinnt, G.S., Simonfy, C.M., & Brichetto, T. (1992). "Assembled" phonology and reading: A case study in how theoretical perspective shapes empirical investigation. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 249–292). Amsterdam: Elsevier Science Publishers B.V.

Weber, R.M. (1970). A linguistic analysis of first-grade reading errors. *Reading Research Quarterly*, *5*, 427–451.

Wydell, T.N., Patterson, K.E., & Humphreys, G.W. (1993). Phonologically mediated access to meaning for Kanji: Is a Rows still a Rose in Japanese Kanji? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 491–514.

APPENDIX A

Experimental Materials Used in Experiments 1 and 2

English translations in brackets.

Base-word	Pseudohomophone	Control
bij [near/bee]	bei	bijn
kijkt [looks]	keikt	kijnt
zijn [to be]	zein	zijm
weg [road]	wech	wag
zout [salt]	zaut	zolt
klein [small]	klijn	krein
voet [foot]	foet	voot
bijl [ax]	beil	bijk
pijp [pipe]	peip	pijg
hout [wood]	haut	huut
bijt [bites]	beit	bijf
geeft [gives]	geevt	geekt

APPENDIX B

Experimental Materials Used in Experiments 3 and 4

English translations in brackets

Base-word	Pseudohomophone	Control	Category
wijn [wine]	wein	wijg	drinken [drink]
zout [salt]	zaut	zoum	eten [food]
tand [tooth]	tant	tald	lichaamsdeel [body part]
acht [eight]	agt	aft	getal [number]
bij [bee/near]	bei	bijf	insekt [insect]
krijt [chalk]	kreit	krijl	schrijfgerei [writing material]
kind [child]	kint	kild	mens [human being]
geit [goat]	gijt	geim	dier [animal]
kous [stocking]	kaus	koum	kledingstuk [piece of clothing]
grijs [grey]	greis	grijk	kleur [colour]