Children's use of phonological encoding when reading for meaning

ESTELLE ANN DOCTOR and MAX COLTHEART
Birkbeck College, University of London, London WC1E 7HX, England

The role of phonological recoding in children's reading was investigated by means of a task requiring comprehension of sentence meaning: The child's task was to decide whether a sequence of printed letter strings was a meaningful sentence or not. Meaningless sentences that are meaningful when phonologically recoded (e.g., "He ran threw the street") produced more incorrect responses than did meaningless sentences that remain meaningless when phonologically recoded (e.g., "He ran sew the street"). The difference in error rates between the two sentence types diminished as a function of age. Control experiments showed that these results were not due to visual similarity effects, nor to imperfect ability to spell homophones. It was concluded that very young readers rely extensively on phonological recoding when reading for meaning; as they grow older, reliance on visual encoding becomes progressively more important.

This paper is concerned with the strategies that young children use to encode words when reading continuous text for meaning. A variety of theoretical views can be taken about this, the extremes being represented, on the one hand, by the view that children treat words as visual wholes (and ignore the fact that words contain components, letters, which are common to many words and which permit, in principle, the conversion of a visual stimulus into a phonological representation) and, on the other hand, by the view that children profit from the fact that they can already understand speech when they are trying to learn to read and so make use of the alphabetic nature of English script by converting letter sequences into phonological representations, which are then comprehended.

This theoretical contrast, belonging to the psychology of reading, is paralleled by that between whole-word and phonic methods of teaching reading, a contrast that belongs to the pedagogy of reading. A distinction needs to be drawn between the two contrasts, however, since even if we knew precisely how children read for meaning, we would still not necessarily know how best to teach them. For example, if it were convincingly demonstrated that young readers did read by carrying out conversions from printed words to phonological representations, it would remain for us to discover whether they could best be taught this by explicitly teaching conversion procedures (as in phonics instruction) or by providing many examples of words and heir pronunciations (as in whole-word instructional methods) and relying on the children to learn implicitly how print is converted to sound. Our concern in this paper is with the psychological question rather than the pedagogical question.

The relative importance of visual vs. phonological encoding for young readers has been investigated in a variety of ways. Perhaps the most frequent approach has been to contrast groups of good and poor readers. Such research has often been prompted by the reasoning that, if one can discover some basic cognitive task that good readers are good at and bad readers are bad at, then one can deduce that learning to read must rely at least in part on this cognitive task. It has often been pointed out, however, that this reasoning confuses correlation with causation. For example, if one shows that good readers are faster and more accurate at naming single letters than are bad readers (which is the case; Chall, 1967, reviews this work), one cannot therefore conclude that knowing letter names helps children to learn to read (and hence that learning to read must rely in part on using knowledge of letter names to convert a printed representation to a phonological representation). It is possible, even plausible, to argue that the direction of causality is opposite, that learning about letter names is a consequence of learning to read well, not a prerequisite, or that poor progress in reading discourages children and hence makes them perform more poorly in any task involving printed letters. Moreover, the possibility also exists that the relationship between reading ability and letter-naming ability is not causal at all and that the apparent correlation is due to the influence of some third variable (intelligence, socioeconomic class, ability to attend selectively; there are many possible candidates) that influences both how well a child learns to read and how well he learns to name letters.

This kind of objection can be raised against any investigation that seeks, by comparing good and bad readers, to discover something about the causation of reading deficiencies. Thus attempts to deduce something about how children read by comparing good and bad readers, which rely on such causal inferences, will scarcely be satisfactory. This objection applies even to such

findings as those of Rozin, Bressmann, and Taft (1974), who showed that poor readers were so unaware of the relationships between letters and phonemes that they could not correctly pair up the spoken words "mow" and "motorcycle" with their printed equivalents. This inability is striking, but we do not know whether it is a cause or a consequence or poor reading. The same point can be made concerning demonstrations that good readers have superior knowledge of spelling-sound correspondences (e.g., Calfee, Venezky, & Chapman, Note 1). Can we be sure that they are good at reading as a consequence of this superior knowledge, rather than having acquired this superior knowledge as a consequence of being good readers?

A second difficulty that arises in some basic research on the nature of reading in young children is the choice of a measure of reading ability, in particular, the choice between an oral test of reading and a silent test. Firth (1972) showed that good readers differed to an extraordinary degree from bad readers in the ability to convert letters to sounds. His measure of the latter ability was the number of pronounceable nonsense words a child could pronounce accurately. A group of 50 good readers averaged 128 out of 150 right; a group of 50 poor readers averaged 29 out of 150. However, the tests used by Firth to measure reading ability were the Schonell R1 and the Neale accuracy tests. Both are oral tests in which success on an item is credited when the item is pronounced correctly. It is not especially surprising that there is a strong correlation between how accurately a child can pronounce words and how accurately he can pronounce nonwords. If one takes the view that the essence of reading is understanding what a printed word means, rather than being able to pronounce it aloud, then it is not clear what Firth's finding tells us about how children encode words when they are reading. This would not be a difficulty if understanding what a word meant was always accompanied by being able to pronounce it and vice versa, but this is not always so. The two abilities must be at least partially separable, since damage to the left hemisphere can produce a disorder in which printed words are given completely incorrect pronunciations, yet at least some knowledge of their meanings remain (e.g., a patient reads "GNOME" as "pixie" and "THUNDER" as "storm"). This syndrome is known as "deep dyslexia" (Marshall & Newcombe, 1973) or "phonemic dyslexia" (Shallice & Warrington, 1975). The same point is made in another way by considering the findings of Thackray (1971), who showed that after 3 years of tuition in either i.t.a.[1] or traditional orthography (t.o.), children who had been taught by i.t.a. were superior to t.o. children in accuracy of oral reading (Schonell and Neale accuracy) but not in comprehension (Neale comprehension). Thackray (1971, p. 215) commented thus: "The author, on a number of occasions, asked for the meaning of a word a child had given correctly, to find the child did not know what the word meant. It would seem then that in the author's sample there was some evidence to suggest that, in the i.t.a. reading, accuracy was outstripping understanding of the words read." Neale (1967) found that girls were superior to boys at oral reading, but not in reading comprehension. Viitaniemi (1965, cited in Downing & Thackray, 1971, p. 19) found, with Finnish children, that boys were actually superior to girls in reading comprehension, but girls showed a small superiority over boys in oral reading. Sebesta (1964) found that oral reading was facilitated if children learned to read with an i.t.a.-like artificial orthography, but there was no effect on reading comprehension. These various examples demonstrate the dissociability of the ability to pronounce a printed word aloud correctly from the ability to understand what a printed word means.

Provided that one is interested in reading as reading for meaning, then, research attempting to show the importance of phonological encoding for oral reading is unsatisfactory; it is essential to use a reading task that necessarily requires access to meaning while not necessarily requiring the conversion of print to sound.

This methodological problem is, if anything, more acute when one considers research on skilled reading. Here the experimental tasks most frequently used (for example, the lexical decision task, reading aloud, same-different judgment, tachistoscopic report, and so on) bear no obvious relationship to the task of reading continuous prose for meaning, and hence, they may not be capable of providing much enlightenment about how such reading is performed. In particular, these experimental tasks do not require the subject to gain access to the meanings of the words with which the experimenter confronts him. Such considerations led Baron (1973) to devise a task that would require his subjects, skilled readers, to gain access to the meaning of printed text. The task was to classify such phrases as "Tie the not," "He is know," or "He is ill" as meaningful or meaningless. To the extent to which access to the meanings of printed words is accomplished by a conversion of print to phonology, the subject will have difficulty in giving a correct "no" response to a phrase like "Tie the not," which becomes meaningful once it has been converted from a visual to a phonological representation.

The results of this experiment were unclear (phrases that sounded meaningful did not produce slower "no" responses than phrases that sounded meaningless, but they did produce more than twice as many errors, and this difference in error rates was significant), but the task seems a promising one with children from 6 to 10 years of age. Briefly, each child was given 16 short printed sentences, one at a time, and was asked to say whether each sentence made sense or not. Eight of the sentences were meaningful; eight were nonsensical. Of the eight meaningless sentences, four became meaningful when converted to a phonological representation (e.g., "She blue up the balloon") and four remained meaning-

less (e.g., "She know up the balloon"). Comparisons of responses to the former kinds of sentences with responses to the latter kinds will provide information about the use of letter-sound conversion when young children read continuous text for meaning.

EXPERIMENT 1

Method

Subjects. The subjects were 60 girls and 60 boys randomly selected from two London primary schools. They had been taught to read by eclectic ("mixed") methods. There were 12 boys and 12 girls of each of the following mean-age groups: 6.2 years, 7.3 years, 8.35 years, 9 years, and 10.5 years. None of the children had any reading problems, and all spoke English at home.

Materials. Twelve separate blocks of sentences were constructed. Each block consisted of 16 different sentences. Eight of the sentences in each block were meaningful and required a "yes" response; the remaining eight were meaningless and required a "no" response.

The eight meaningless sentences in a block were of four types, two sentences per type. (1) All word, sounds correct: An example is "I have know time." When phonologically recoded, this sounds meaningful. (2) All word, sounds wrong: An example is "I have blue time." This remains meaningless when phonologically recoded. (3) Nonword, sounds correct: An example is "I have noe time." This is like Type 1 above, except that it contains a nonword. (4) Nonword, sounds wrong: An example is "I have bloo time." This is like Type 2 above, except that it contains a nonword.

The eight meaningful sentences in a block consisted of two sentences containing homophones (e.g., "I have no time") and two nonhomophone sentences (e.g., "I have the time"). The remaining four meaningful sentences were merely fillers, to equate the number of "yes" and "no" responses, and they will not be discussed further.

Although each block included two of each of the six different sentence types, no two sentences in a block were phonologically identical. For example, if one of the sentences in the block was "I know him," the sentence for that block that was all word and sounded correct could not be "I no him," nor could the sentence in which the nonword sounded correct be "I noe him." Thus the 16 sentences seen by any one subject were all phonologically distinct from each other. The 144 sentences, which were formed into 12 blocks of six types, with two exemplars per type, are listed in Appendix A, Table I. Prior to the experiment, the sentences had been given to a pilot group of children aged between 5 and 8 years, and the final blocks of sentences were such that all the words used in them were known by all of these children.

Procedure. Testing took place over 10 days. Sessions were held in a quiet room in the school. Each child was tested individually in sessions lasting approximately 10 min each.

The child was seated behind a table, facing the experimenter. He was shown a practice sentence, with the initial letter and proper names, if any, written in large uppercase letters, and the remainder written in lowercase letters. Half of the boys and half of the girls at each age level were requested to read the sentence aloud and the remainder were asked to read it silently before answering "yes" if the sentence made sense and "no" if it did not. A distinction was drawn between the truthfulness of the sentence (e.g., a boy having to read the sentence "I have a green dress") and its meaningfulness. Other practice sentences were given until the child had been shown a couple of examples of each sentence type and his errors had been corrected and explained.

Subjects were randomly assigned to the 12 blocks of sentences, with two children of each age group receiving the same sentence, and their responses were recorded. A week before this experiment took place, every child had been shown each homophone and nonword that was to appear in his block of sentences. He was asked to pronounce the word and to make a sentence with it to ascertain whether he knew its meaning in isolation. The results of these pretests are discussed later.

Results

The means for the percentage of correct responses to all six sentence types are given in Table 1.

Meaningful sentences containing no homophones vs. meaningful sentences containing homophones. A 5 (age levels) by 2 (sex) by 2 (sentence type) by 2 (reading aloud vs. reading silently) analysis of variance was carried out on these data. There was no significant effect of sentence type, which indicates that children do not find homophones more difficult to comprehend than nonhomophones.[2] The only significant effect in this analysis was that reading aloud produced superior performance to reading silently [$F(1,100) = 4.29$, $p < .05$]. Since this effect was not significant with the other four sentence types, we will not consider it further. No other effects were significant, although there was a consistent trend for sentences containing homophones to be more difficult. This effect may well be genuine and concealed in these data by a ceiling effect. Arcsin transformations were applied to the scores, and the analysis of variance results were similar to the above.

Effects of homophony on all-word meaningless sentences. Analysis of variance was used with the data from meaningless sentences containing only words to determine whether incorrect "yes" responses are more likely to occur when such sentences are phonologically meaningful. The relevant means are in Rows 3 and 4 of

Table 1
Percentage Correct Responses to Each of the Six Sentence Types and Percentage of Substitutions

Sentence Type	Correct Response	Age In Years				
		6	7	8	9	10
Meaningful sentences containing no homophones	Yes	89.6	97.9	95.8	95.8	95.8
Meaningful sentences containing homophones	Yes	85.4	93.8	87.5	95.8	89.6
Meaningless all-word sentences that sound correct	No	29.2	56.3	68.8	81.3	79.2
Meaningless all-word sentences that sound wrong	No	91.7	95.8	95.8	97.9	97.9
Meaningless sentences containing nonwords that sound correct	No	43.7	72.9	75.0	95.8	95.8
Meaningless sentences containing nonwords that sound wrong	No	89.6	93.8	95.8	97.9	100.0

Table 1. Performance was significantly worse with sentences that become meaningful when translated from a visual to a phonological code than with sentences that remain meaningless even in their phonological form $[F(1,100) = 108.73, \ p < .001]$. Hence, children are more likely to say that a meaningless sentence makes sense if its phonological form is meaningful; in other words, they make considerable use of letter-to-sound translation in reading printed sentences for meaning.

Performance also improved significantly with age $[F(4,100) = 8.71, \ p < .001]$, and there was a significant interaction between sentence type and age $[F(4,100) = 7.12, \ p < .001]$: The size of the performance difference between the two types of sentences decreased with age. This interaction may be a spurious consequence of a ceiling effect, since even with the youngest children, performance was above 90% with the phonologically meaningless sentences. Nevertheless, with phonologically meaningful sentences, performance increased steeply with increasing age, indicating that the older a child is, the less use he or she makes of a phonological reading strategy. No other effects were significant.

The same main effects and interactions were significant when an arcsin transformation was applied to the data.

Effects of homophony on meaningless sentences containing a nonword. The relevant means here are in Rows 5 and 6 of Table 1. Once again, performance was significantly worse with sentences whose phonological form is meaningful than with those whose phonological form is meaningless $[F(1,100) = 42.09, \ p < .001]$, again indicating considerable use of phonological encoding in these readers. The reliance on such encoding decreases as age increases, since performance improves significantly with age $[F(4,100) = 14.45, \ p < .001]$. Once again, there was an Age by Sentence Type interaction $[F(4,100) = 8.12, \ p < .001]$, but this, too, may be due to a ceiling effect, in view of the uniformly high level of performance on phonologically meaningless sentences.

The age of the subject interacted significantly both with sex $[F(4,100) = 2.70, \ p < .05]$ and with reading aloud vs. reading silently $[F(4,100) = 2.60, \ p < .05]$. This was due to rather poor performance by 6-year-old girls and by 6-year-old oral readers, but since this effect was not evident at other age levels, we will not discuss it further.

No other effects were significant.

Once again, reanalysis of the data after arcsin transformation yielded the same pattern of results.

Effects of the presence of a nonword. A comparison of the means in Row 3 of Table 1 with the means in Row 5 provides information about effects of nonwords. Both types of sentence here are meaningless but sound meaningful. The difference is that the sentences represented by Row 3 consist entirely of words, whereas those represented by Row 5 contain a nonword. Thus

it would in principle be possible to respond "no" to the latter type of sentence simply by detecting the presence of a nonword. Performance was in fact significantly better with sentences that contained nonwords than with all-word sentences $[F(1,100) = 16.65, \ p < .001]$. No other effects were significant. Performance also improved with age $[F(4,100) = 16.13, \ p < .001]$.

Discussion

The many errors the children made in dealing with meaningless sentences that were phonologically meaningful suggests that phonological coding was taking place to a great extent while they were reading for meaning. Reading was not entirely phonologically mediated, however, even for 6-year-olds, since they did not invariably respond "yes" to nonsense sentences that were phonologically acceptable and since performance was improved when nonwords were present, even though these nonwords were homophonic with real words and thus could only have been detected as nonwords by using a visual code. The difficulties engendered by meaningless sentences that sounded correct diminished as the age of the reader increased. We suggest, then, that in the evaluation of the meaning of these printed sentences, the younger children relied more on a phonological code, while children of 9 years and older relied largely on a visual code.

There were no significant differences between the sexes in this comprehension task, although sex differences have been found in reading. Thompson's (1975) review of the literature shows that girls are often superior to boys until the age of 10 years and that more boys than girls are found in remedial reading classes. Other studies have found differences in favor of girls in oral reading (Viitaneimi, 1965, cited in Downing & Thackray, 1971, p. 19) and in reading rate, but not in accuracy or comprehension (Neale, 1967). Sex differences may occur in oral reading tasks that differ from the comprehension test used in this study. Some sex differences appear to be cultural, with American studies showing more frequent sex differences in favor of girls. Overall, the evidence is not conclusive.

Coltheart (1978) and Forster and Chambers (1973) and others have noted that a word's phonological code can be derived from its printed form in at least two ways: a nonlexical method (applying letter-sound rules) and a lexical method (accessing a word's entry in the internal lexicon and retrieving from this entry the word's phonological representation). Although often neglected, this is an important distinction, and it applies to the results of this experiment. The phonological code that causes children to have difficulty in rejecting meaningless sentences containing homophones that are phonologically appropriate could be obtained prelexically (by applying letter-to-sound rules to the printed form and using the resulting phonological code for purposes of lexical access) or postlexically (by accessing the

lexical entry corresponding to the individual printed word, retrieving the word's phonological representation from this entry, and using the resulting phonological code during the attempt to comprehend the sentence as a whole). Either of these two possibilities would lead to the difficulties with meaningless sentences containing phonologically appropriate homophones that the children, especially the younger children, encountered. It is possible that even if the children did not use the nonlexical method, they may have been forced into using the lexical method because of the confusion caused by the visually incorrect word being phonologically appropriate to the syntactic and semantic context of the sentence. The older children may have been less tempted to yield to the wrong information provided by this code because they had more reading experience and were more familiar with the written forms of the words. Only the prelexical method, of course, is available for nonwords, so that the difficulties experienced with the meaningless sentences that contained a nonword that sounded correct in the context must represent a prelexical encoding effect.

There are several possible objections to the interpretation of this experiment. One may be to the design of the study, another may be inherent in the subjects, and a third in the materials.

First, it could be argued that the critical stimuli, the homophones, were unknown to the children and that they could not distinguish between the members of the homophone pairs. Next, it could be argued that the children were poor spellers and that this accounts for their inability to make semantic judgments based on visual features. Finally, it may have been the case that an incorrect word in the meaningless sentences that contained phonologically appropriate homophones and nonwords also looked more similar to the correct word than was the case with meaningless sentences that were both phonologically and visually wrong. For example, in the sentence "I will sea him today," the word "sea" looks like the correct word "see," whereas in the corresponding sentence "I will for him today," no such visual similarity exists. Perhaps children have difficulty with the former type of sentence because of this visual similarity effect, rather than because these sentences make sense phonologically.

Each of these objections will be dealt with in turn. The next experiment was an attempt to observe how the children pronounced and used the homophones and nonwords that were included in the sentences used in the above study when these items were presented in isolation and not in the context of a sentence.

EXPERIMENT 2

This experiment has to do with the objection that a child says "yes" to the printed sentence "The wind blue" because he believes the past tense of "to blow" is

spelled "blue." If this were so, the effects observed in Experiment 1 would not represent evidence of phonological recoding; they would instead demonstrate imperfect knowledge of spelling. We therefore investigated how well a child could understand the homophonic words used in Experiment 1 and the extent to which he confused a single printed homophone with the other member of the homophonic pair. We refer to such a confusion as a "substitution." We also investigated the pronunciation of each word, to ensure that the homophonic words actually were homophonic in the child's speech.

Method

Subjects. The same subjects participated as in the previous study.

Materials. Each child was tested on a set of eight different words, six of which were the homophones and four the nonwords used in Experiment 1. Each stimulus was written with a thick black felt-tipped pen on a 6 x 4 in. record card. Each child saw the individual words, nonwords, and homophones that were presented to him in the context of a sentence in Experiment 1.

Procedure. Testing took place over 2 weeks, and sessions were held in a quiet room in the school. Each child was tested individually. Sessions lasted approximately 30 min for the younger children and 10 min for the older children.

The child was seated at a table, facing the experimenter. The cards were presented in random order, and after the first card of the deck had been shown, the child was asked to pronounce the word written on the card and then to explain its meaning either by using it in a sentence or as best he could. By way of warning the child that there were nonwords present in the deck, the experimenter explained that he might never have seen some of the words before, but that he should deal with them as best he could.

Results and Discussion

Table 2 summarizes the data and gives the total percentages, averaged over all the words, for correct responses and errors, substitutions (i.e., using a homophone as its mate), and the percentage these substitutions were of the total errors in usage. This table reveals that there were relatively few pronunciation errors and that pronunciation improved with age. All the children made more usage errors than pronunciation errors, but these also decreased with age.

Although substitutions of one homophone for another, its mate, did occur, especially with the 6- and 7-year-old children, they were much too infrequent to explain the effects obtained in Experiment 1; a comparison between Table 1 and Table 2 indicates this. To take the worst case, 6-year-old children confused 24% of the homophones by using their mates when the homophone was presented as a single printed word. In Experiment 1 this would reduce percent correct on meaningless but phonologically meaningful sentences to 76%; but in Experiment 1 we observed performance far lower than this (29.2%). Applying this argument consistently across Table 1 shows that substitutions cannot explain the effects evident there. This conclusion is confirmed, using a different technique, by the results of Experiment 3.

Table 2
Pronunciation (P) and Usage (U) of Homophones

| | Age in Years | | | | | | | | | |
| | 6 | | 7 | | 8 | | 9 | | 10 | |
	P	U	P	U	P	U	P	U	P	U
Percentage correct	86.9	67.8	97.4	87.00	96.9	91.1	98.4	97.4	100	96.90
Percentage wrong	13.1	32.2	2.6	13.00	3.1	8.9	1.6	2.6	0	3.15
Percentage of all responses that were substitutions		24.0		10.40		5.7		1.6		3.15
Substitutions as percentage of total errors		74.0		75.35		67.0		66.0		100.00

Inspection of the responses to individual homophones revealed that more errors were made on one member of some pairs than on others. For example, "blew," "four," "knew," "know," "won," "sew," "their," and "wear" tended to result in more substitution errors than did their homophonic mates. In order to ascertain whether children were substituting more within pairs that were highly similar in graphic similarity than within pairs that were less similar, the graphic similarities of the words were calculated and correlated with the children's errors in substituting one homophone for its pair. Weber's (1970b) formula for calculating graphic similarity was used. She has described it as follows: "The features taken into account are the number of letters shared by the stimulus-response pair, the position of shared letters within the words, the position of shared letters relative to each other, the average length of the words, and the difference in lengths between the stimulus and response words."

The formula is: $GS = 10[(50F + 30V + 10C)/A + 5T + 27B + 18E]$, where F = the number of pairs of adjacent letters in the same order shared by S and R; V = the number of pairs of adjacent letters in reverse order shared by S and R; C = the number of single letters shared by S and R; A = average number of letters in S and R; T = ratio of numbers of letters in the shorter word to the number in the longer; $B = 1$ if the first letter in the response is the same as the first letter in the stimulus, otherwise $B = 0$; and $E = 1$ if the last letter in the response is the same as the last letter in the stimulus, otherwise $E = 0$. The percentage of incorrect substitutions and the graphic similarity index for each pair of homophones is presented in Appendix B, Table I, Columns 3 and 4, respectively. The correlation was not significant ($r = .22$), indicating that homophone pairs with high graphic similarity were not more likely to produce the wrong homophone in response to its mate than were low similarity pairs.

A further possibility was that if the homophones differed in frequency the child might use the more frequent member of the homophone pair more often and so incur errors when the less frequent member of the pair was presented. Word frequency tables such as that of Kučera and Francis (1970) had not been consulted when the homophone pairs were selected because their tables are derived from an adult population, and no comparable tables exist that have been standardized on British school children. However, the correlation between word frequency based on the Kučera and Francis ratings and the percentage of incorrect substitutions was calculated (see Appendix B, Table I, Columns 5 and 3, respectively) and was found to be negligible ($r = -.04$), indicating that usage of the more frequent members of the pair was not related to errors.

Nonwords. The nonword data were inspected in a manner similar to the above homophone data, but there was no significant correlation between graphic similarity of a nonword to a homophone and usage of the nonword as that homophone (see Appendix B, Table II, Columns 4 and 3, respectively) or between frequency of the graphically similar homophone and usage of the nonword as that homophone (Columns 5 and 3, respectively).

Many of the children, and particularly the older ones, recognized that the nonwords were not real words. This is indicated in Table 3 by the increase of "no" responses to nonwords and by the children's comments about not having encountered those words in print before. They could not have been quite sure that the nonwords were not simply unfamiliar real words, as the tendency to use "soe" as the less familiar "sow" (less familiar in comparison with "so" and "sew") reveals. However, the older subjects refused to use more than half the nonwords. Those nonwords that were used were usually given in the sense of similar-sounding homophones rather than as some other word.

The main purpose of this experiment was to show that the children knew how to pronounce and use homophones and to investigate how they would deal with nonwords when these items were presented to them in isolation.

A comparison of children's substitution of one homophone by its mate in usage and their responses to sentences in which a meaningless but phonologically appropriate homophone was incorporated in a sentence is shown in Table 4. Children of all ages made more errors in reading sentences than in the pronunciation or usage of isolated words, and there was no correlation between performance on these two tasks ($r = -.09$) (see Appendix C, Table I).

Comparisons of children's usage of nonwords with their responses to sentences containing nonwords that

Table 3
Pronunciation (P) and Usage (U) of Nonwords

	Age in Years									
	6		7		8		9		10	
	P	U	P	U	P	U	P	U	P	U
As a similar-sounding real homophone	68.8	63.5	69.8	40.7	44.8	36.5	55.2	23.95	51.1	15.6
As some other real word	15.6	15.6	6.2	4.1	12.5	7.2	2.1	7.30	1.1	2.1
No response	15.6	20.9	24.0	55.2	42.7	56.3	42.7	68.75	48.0	82.3
Percentage used as a similar-sounding real word*		80.2		90.8		83.5		76.64		88.0

*Percentage used as a similar-sounding real word of total nonwords used as real words.

Table 4
Comparison Among the Percentage of Incorrect Substitutions in Usage, Incorrect Acceptances, and Incorrect Substitution in Spelling

	Age in Years				
	6	7	8	9	10
Percentage of substitution of one homophone by its mate in usage**	24.0	10.4	5.7	1.60	3.15
Percentage of incorrect acceptances of meaningless all-word sentences that sound correct*	70.8	43.7	31.2	18.70	20.80
Percentage of nonwords used as phonologically similar real words**	63.5	40.7	36.5	23.95	15.60
Percentage of incorrect acceptances of meaningless nonword sentences that sound correct*	56.3	27.1	25.0	4.20	4.20
Percentage of substitution in spelling†	17.0	9.0	6.0	3.00	1.50

*Experiment 1.   **Experiment 2.   †Experiment 3.

Table 5
Spelling Responses

	Age in Years				
	6	7	8	9	10
Percentage correct	31	42	86	90	94.5
Percentage substitution of one homophone by its mate	17	9	6	3	1.5
Substitutions as percentage of errors	25	28	40	29	27.0

were phonologically meaningful in the context of the sentence is also shown in Table 5. Fewer nonwords were incorrectly accepted in sentences than were incorrectly used as real words, and there was no correlation between performance on these two tasks (r = .27) (see Appendix C, Table II).

If the meaning of a word is accessed visually, then it should be relatively easy to access the meaning of homophones, because their similar sounds should not interfere with the meanings associated with them. Relatively few homophones were used as their mates by the children, indicating that they might well have been using a visual route to the lexicon for this task. This might lead one to expect that children should not accept the wrong form of a homophone when it appears in the context of a sentence in which it is phonologically, but not visually, appropriate. However, the results of the pronunciation and usage of nonwords experiment confound this result. If lexical access is purely visual, none of the nonwords should have been used as they were, and, more particularly, none of them should have been used as a similar-sounding homophone, yet many were. These results may mean that the child first tries

to gain access to the lexicon visually. If the entry is not found, he may then try gaining access by encoding the item phonologically. This would allow him to make correct responses to homophones and to use nonwords as real words, and the effect might be increased when other semantic and syntactic cues are present, as in sentences. If the task demands differ in pronunciation and usage of isolated words compared with reading for meaning, the results might also differ. Neither words nor nonwords need to be recoded phonologically during reading for meaning. Without this demand on them, subjects might be encouraged to rely more on the visual features of the stimuli, which would enable them to reject nonwords more easily. This could apply to homophones as well, but the wrong homophones used in a phonologically suitable context might be more difficult to reject than a corresponding nonword: Children may reject unfamiliar words and nonwords because they do not recognize them as being genuine words. This strategy would help children of all ages to reject nonwords but would hinder younger children, who tend to make more substitution errors in rejecting phonologically appropriate but visually inappropriate homophones

when both are equally familiar to them, especially if recoding to a phonological code takes place after the occurrence of access to individual word entries in the internal lexicon, but before semantic judgment of the sentence as a whole is made. Once again, this would not affect nonwords that are not represented visually in the lexicon unless they are phonologically encoded prior to lexical access, as they are in a pronunciation task, but as they need not be in a task involving semantic judgment.

EXPERIMENT 3

The aim of this study was to demonstrate that children were sufficiently proficient at spelling the homophones used in Experiment 1 to rule out any explanation of the results of that experiment in terms of inability to distinguish the spellings of homophones.

Method

Subjects. Where possible, the same children were used who had participated in the previous experiment. However, the spelling test was administered 1 year later, so that a new sample of 6-year-olds was selected. In addition, some children had left school, and it was not possible to obtain equal numbers of girls and boys in each age group. Twelve 6-year-old girls and 11 6-year-old boys, 10 7-year-old girls and 12 7-year-old boys, 12 8-year-old girls and 8 8-year-old boys, 12 9-year-old girls and 12 9-year-old boys, and 12 10-year-old girls and 6 10-year-old boys served as subjects.

Materials. The 12 homophone pairs used in the previous experiments were the stimuli that had to be spelled.

Procedure. The children in a particular age group were seated at desks in their classrooms. A sheet of paper was distributed to each child. They were requested to furnish biographical data and were told that they were to be given a spelling test. The class teacher dictated 12 homophones, one at a time, pausing between each word for the children to write their answers. After this, each of the homophones was dictated in the context of a meaningful sentence, and the children were again asked to write down each homophone.

Results and Discussion

Table 5 gives the percentage of correct responses averaged over all the homophones, as well as the percentage of substitutions averaged over all the homophones and the percentage of errors that substitutions formed. The main reason for carrying out an analysis of spelling errors on the homophones was for purposes of comparison with the children's responses to homophones embedded in the context of meaningful sentences. It could be argued that the high rate of errors on meaningless sentences that nevertheless sounded correct when encoded phonologically (e.g., "He went to by the ball") occurred simply because the children did not know how to spell such homophones as BY and BUY. The results of this spelling study refute such an argument. Children who were presented aurally with the homophones in context and asked to spell them rarely produced the wrong member of the homophone pair, whereas they commonly accepted the wrong member

when it was presented in print. A comparison of errors is presented in Table 4, where it is clear, for example, that an incorrect "yes" to the printed sentence "He went to by the ball" is far more common than is spelling "buy" as "by" in response to the spoken sentence "He went to buy the ball." Thus the results of Experiment 1 cannot be explained in terms of poor ability to spell homophones.

EXPERIMENT 4

It might be argued that the homophones and nonwords in the meaningless sentences that sounded correct when translated into a phonological code looked more similar to the correct word that should have been there than the corresponding words in the meaningless sentences, in which the homophones and nonwords neither looked nor sounded correct. On this argument, the effects observed in Experiment 1 are visual, not phonological. If the argument is correct, then the degree to which a child tends to respond with an incorrect "yes" to meaningless but phonologically correct sentences should depend on the degree of similarity between the incorrect word or nonword in the sentence and the word that should be there. Therefore, for each of these sentences, the graphic similarity index of Weber (1970a, 1970b) was used to measure graphic similarity of the incorrect word in the sentence to the word that should have been there. The results are shown in Appendix B, Table I for words and Appendix B, Table II for nonwords. Then the correlation between the proportion of error responses and the degree of graphic similarity was calculated across sentences. This correlation was not significant for the meaningless sentences that contained homophones that sounded correct ($r = .01$) (Columns 6 and 3, respectively, both tables), indicating that those with high graphic similarity were no more likely to produce an incorrect "no" response than those with low graphic similarity. Thus the difficulty with these sentences is a phonological effect, not a visual effect.

The correlation, however, was significant for the meaningless sentences containing phonologically appropriate nonwords ($r = -.46$, $p < .02$). It is thus possible that the difficulty here is a visual one and is not due to phonological encoding. In order to clarify this, a final experiment was carried out.

The point of this experiment was to compare performance on nonword sentences that were meaningless but sounded meaningful with performance on nonword sentences that neither looked nor sounded meaningful, when the two types of sentence were matched on graphic similarity. For example, the nonwords "bie" and "bak" have equal graphic similarity to the word "by," according to Weber's (1970a, 1970b) formula. Therefore, if children make more incorrect responses to "He went bie the town" than to "He went bak the

town," this cannot be due to differential graphic similarity of the nonwords to the correct word "by"; it must be a phonological effect.

Method

Subjects. All the boys and girls in the second, third, and fourth grades of one of the London primary schools used in the previous experiment were tested. There were 15 girls and 12 boys whose average age was 6 years, 12 girls and 22 boys whose average age was 7 years, and 17 girls and 9 boys whose average age was 8 years. No older children were included in the sample because the previous experiments had shown that they made few errors.

Materials. In addition to some of the nonwords used in the previous experiments, several new nonwords were created, each of which had the same graphic similarity score as a particular nonword in the previous experiment, according to Weber's (1970a, 1970b) index. Sentence frames were built around all of these nonwords, and there were four different sets of sentences, as shown in Appendix A, Table II. Sentences in the first set contained a nonword that sounded like an English homophone, and the sentence made sense when translated into a phonological code. The sentences in the second set were matched to these sentences for graphic similarity, but the nonword that sounded like an English word did not fit the phonological context of the sentence, so that the complete sentence did not make sense when translated into a phonological code. The third set of sentences contained the same nonwords that appeared in the second set of sentences, but all of the sentences made sense when translated into a phonological form. The fourth set of sentences were graphically similar to the third set, although the nonwords in them, which sounded like English homophones, did not fit the phonological context of the sentence, and the whole sentence was meaningless when translated into a phonological form. The graphic similarity between the nonwords and the English homophones and words they replaced in the sentences was calculated and is given alongside the subjects' errors in Appendix D.

The different sentence types were randomly assigned to four blocks, so that each block contained five sentences of each of the four types, to all of which the correct answer was "no." Each block also contained 20 filler items to which the correct answer was "yes." The same filler items were used in each block, but a different set of 20 test items was selected from the four sentence types and used in each block. Different children in each age group saw the different blocks.

Because of the difficulty of constructing the stimuli, only eight of the original nonwords from the first experiment could be used to derive the new nonwords. Two additional nonwords that sounded identical to English homophones were added, each of which had a corresponding homophone listed by Edwards and Gibbon (1973) as a word likely to be used by 6-year-olds, as were all the real words corresponding to the newly created nonwords.

Procedure. Testing took place at school during the morning. The children were tested in groups in their classrooms. Each session lasted approximately 30 min. Within each grade, the children were randomly assigned to one of the four blocks of sentences. Questionnaires were distributed, and the children were asked to fill in personal details such as name, age, sex, and grade. The answering technique differed from Experiment 1 in that the children were asked to mark their responses to the questionnaires. Eight practice trials were presented to explain the nature of the task. The first four of these were already correctly filled in on the questionnaire to demonstrate to the children that a tick (check mark) meant that the sentence made sense, whereas a cross indicated that it did not. The next four practice examples were completed by the children and the experimenter together. When it was clear that the children understood the procedure, they were requested to turn to the second page of the questionnaire and to complete the remaining 40 items in the same way.

Results

Effects of graphic similarity. The percentage of correct "no" responses to the four types of sentence is shown in Table 6.

(1) Nonword (like English homophone): Sentence sounds correct vs. Nonword (like English word): Sentence sounds wrong. Performance was significantly worse for the former set of sentences than for the latter set, indicating that phonological encoding was taking place [$F(1,81) = 17.61$, $p < .001$]. Performance improved with age [$F(2,81) = 9.79$, $p < .001$], and there was a sex difference [$F(1,81) = 6.54$, $p < .05$]. Six- and 7-year-old boys were more accurate than girls of those ages, whereas 8-year-old girls were marginally more accurate than 8-year-old boys, hence the Sex by Age interaction [$F(2,81) = 3.24$, $p < .05$]. Ceiling effects may be present in the high level of performance of the 7- and 8-year-olds.

(2) Nonword (like English homophone): Sentence sounds wrong vs. Nonword (like English word): Sentence sounds correct. These two types of sentence were equated for graphic similarity. Performance was significantly worse on sentences in which the nonword sounded like a real word that would have been meaningful in the context [$F(1,81) = 28.99$, $p < .001$], again indicating phonological encoding. Performance improved with age [$F(2,81) = 12.3$, $p < .001$], and there was a Sentence by Age interaction, possibly due to ceiling effects, as all the children were accurate in rejecting the

Table 6
Percentage Correct "No" Responses Made by Children to Sentences in Experiment 4

	Age in Years		
	6	7	8
Phonologically meaningful sentences in which the nonword sounds like an English homophone	77.80	92.35	93.85
Phonologically meaningless sentences in which the nonword sounds like an English word	89.63	98.24	96.92
Phonologically meaningless sentences in which the nonword sounds like an English homophone	91.11	97.65	97.69
Phonologically meaningful sentences in which the nonword sounds like an English word	69.63	86.47	95.38

Note—The first two and the last two sentence types listed are graphically similar to each other.

phonologically meaningless sentences, and the 7-year-olds did well on both sentence types [$F(2,81) = 5.14$, $p < .05$].

The correlations between graphic similarity and amount of errors calculated from the data in Appendix D, Table I were not significant. Therefore, although nonwords in the one set of sentences differed in graphic similarity from nonwords in the other set of sentences with respect to the correct word they replaced (e.g., "I noe on the door" vs. "I nok on the door"), this cannot be the explanation of the difference in error rates, since graphic similarity does not correlate with error rates; therefore, this difference, too, is evidence for phonological encoding.

GENERAL DISCUSSION AND CONCLUSIONS

The first experiment in this series suggested that, when children perform a task requiring silent reading of text for meaning, phonological recoding of the printed stimuli is employed to a considerable degree by young children. The older the child, the less evident phonological recoding is. Subsequent experiments provide reasons for rejecting claims that the effects observed in the first experiment were due to imperfect ability to spell homophones or to effects of visual similarity.

The extent to which phonological recoding is used by skilled adult readers has been a subject of debate throughout this century. Two major problems that have complicated the issue are (1) that many pronouncements for or against the phonological recoding hypothesis have not been accompanied by any evidence and (2) that when evidence is adduced, it is often of an inappropriate form. These problems have been discussed by Coltheart (1980). His conclusion was that, when one considers only evidence collected in methodologically appropriate ways, there is no clear evidence that adults reading single words for meaning use phonological recoding.

Few investigations of the role of phonological recoding in reading continuous text have been reported; very few of these have used appropriate techniques, and for this reason we know virtually nothing about phonological recoding during the reading of text. At least for single-word reading, however, it appears that phonological recoding is not a significant process for skilled reading.

Our suggestion is that when children begin to read, they rely to a great degree on phonological recoding. Over the next 4 or 5 years, the contributions of phonological recoding decrease, and reading for meaning relies more and more heavily on visual representations of words. There are several possibilities as to what might cause this shift. One is that it may be related to tuition in spelling, which is not usually extensive until a child has been reading for some time. A second possibility is that the words a child is given to read in early years are often deliberately selected; they are words with

regular correspondences between grapheme and phoneme, that is, words that can correctly be recoded into phonological form by grapheme-phoneme correspondence rules. Only later are "irregular" or "exception" words introduced (Mathews, 1966); these cannot be recoded by standard grapheme-phoneme rules, and so require "visual" reading.

Learning to read by phonological recoding has its advantages and disadvantages. Among its advantages are that it allows reading to be parasitic upon speech comprehension (a process well developed in normal 5-year-olds), that what is learned generalizes to completely novel words, and that the method makes relatively modest demands upon rote learning capacities. Among its disadvantages are that learning of grapheme-phoneme correspondence is an abstract task and one whose relevance to understanding print may not be obvious to a 5-year-old; reading by phonological recoding cannot cope with exception words, nor with the task of distinguishing between homophones.

The traditional alternative method of learning to read, treating words as visual wholes, has no problems with exception words or homophones and allows immediate word comprehension, but it makes massive demands upon rote learning and does not permit generalization to visually novel words.

Since we have no a priori way of deciding which manner of learning to read has the superior balance of advantages over disadvantages, what is needed is evidence concerning which of the advantages are important and which of the disadvantages severe. Very much more evidence of this kind is needed than is yet available, but some relevant points can be made. For example, Japanese is written in a mixture of two scripts, Kanji (similar to the Chinese script; it is ideographic, so it can only be learned on a whole-word basis) and Kana (a syllabic script ideally designed for learning by phonological recoding). The practice in Japan is for children to learn to read Kana first (most children have mastered it before even coming to school) and then gradually, over a period of several years, to learn to read Kanji. In this sense, 6-year-old Japanese children rely almost exclusively on phonological recoding, whereas 10-year-old Japanese children make substantial use of visual recoding. Our results suggest that this is true of English children, too.

REFERENCE NOTE

1. Calfee, R. C., Venezky, R. L., & Chapman, R. S. *Pronunciation of synthetic words with predictable and unpredictable letter-sound correspondences* (Tech. Rep. 71). Wisconsin Research and Development Center for Cognitive Learning, 1969.

REFERENCES

BARON, J. Phonemic stage not necessary for reading. *Quarterly Journal of Experimental Psychology*, 1973, **25**, 241-246.
CHALL, J. *Learning to read: The great debate.* New York: McGraw-Hill, 1967.

CLARK, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 335-359.

COLTHEART, M. Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies in information processing*. London: Academic Press, 1978.

COLTHEART, M. Reading, phonological encoding and deep dyslexia. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia*. London: Routledge & Kegan Paul, 1980.

DOWNING, J., & THACKRAY, D. *Reading readiness*. London: University of London Press, 1971.

EDWARDS, R. P. A., & GIBBON, V. *Words your children use*. London: Burke Books, 1964.

FIRTH, I. *Components of reading disability*. Unpublished doctoral dissertation, University of New South Wales, 1972.

FORSTER, K. I., & CHAMBERS, S. H. Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 627-635.

KUCERA, H., & FRANCIS, W. N. *Computational analysis of present-day American English*. Providence, R.I: Brown University Press, 1970.

MARSHALL, J. C., & NEWCOMBE, F. Patterns of paralexia. *Journal of Psycholinguistic Research*, 1973, **2**, 175-199.

MATHEWS, M. M. *Teaching to read*. Chicago: University of Chicago Press, 1966.

NEALE, M. *Neale analysis of readability*. Houndsmills, Basingstoke, Hants., England: Macmillan Educational, 1967.

ROZIN, P., BRESSMAN, B., & TAFT, M. Do children understand the basic relationship between speech and writing? The mow-motorcycle test. *Journal of Reading Behaviour*, 1974, **6**, 327-334.

SCHONELL, F. *Schonell graded word reading test*. Edinburgh: Oliver & Boyd, 1945.

SEBESTA, S. L. Artificial orthography as a transitional device in first grade reading instruction. *Journal of Educational Psychology*, 1964, **55**, 253-257.

SHALLICE, T., & WARRINGTON, E. K. Word recognition in a phonemic dyslexic patient. *Quarterly Journal of Experimental Psychology*, 1975, **27**, 187-199.

THACKRAY, D. V. *Reading for meaning with i.t.a. and t.o.* London: Geoffrey Chapman, 1971.

THOMPSON, G. B. Sex differences in reading attainments. *Educational Research*, 1975, **18**, 16-23.

WEBER, R. M. First-graders' use of grammatical context in reading. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, 1970. (a)

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

WIKE, E. L., & CHURCH, J. D. Comments on Clark's "The language-as-fixed-effect fallacy." *Journal of Verbal Learning and Verbal Behavior*, 1976, **15**, 249-255.

NOTES

1. Initial teaching alphabet (i.t.a.) is a simplification of English orthography in which the mapping of letters onto phonemes is highly regular (see, e.g., Thackray, 1971, Chapter 3). It was used quite widely for teaching reading in England in the 1960s.

2. Further analyses of these data were carried out using the min F' method advocated by Clark (1973), which treats both subjects and sentences as random effects. It is controversial whether this is necessary, or even justifiable (Wike & Church, 1976). Fortunately, we have no need to take sides in this controversy, since all of the effects that are significant in the analyses we report remain significant when min F' is used.

APPENDIX A

Table I
List of 144 Sentences Used in Experiment 1

WOOD	WOULD	BLUE	BLEW
We walk in the street	Jane will come with us	The sky is big	She picks up the balloon
We walk in the wood	Jane would come with us	The sky is blue	She blew up the balloon
We walk in the would	Jane wood come with us	The sky is blew	She blue up the balloon
We walk in the woud	Jane woud come with us	The sky is bloo	She bloo up the balloon
We walk in the wun	Jane wun come with us	The sky is noe	She noe up the balloon
We walk in the won	Jane one come with us	The sky is no	She know up the balloon

SO	SEW	NEW	KNEW
This house is very big	She will make a dress	Tell us a funny story	I saw that boy
This house is so big	She will sew a dress	Tell us a new story	I knew that boy
This house is sew big	She will so a dress	Tell us a knew story	I new that boy
This house is soe big	She will soe a dress	Tell us a nue story	I nue that boy
This house is throo big	She will throo a dress	Tell us a wair story	I wair that boy
This house is threw big	She will through a dress	Tell us a where story	I wear that boy

SEA	SEE	WON	ONE
We swim in the water	I will call him today	That boy has a toy	All of us will come
We swim in the sea	I will see him today	That boy won a toy	One of us will come
We swim in the see	I will sea him today	That boy one a toy	Won of us will come
We swim in the cea	I will cea him today	That boy wun a toy	Wun of us will come
We swim in the faw	I will faw him today	That boy woud a toy	Woud of us will come
We swim in the four	I will for him today	That boy wood a toy	Would of us will come

BY	BUY	THREW	THROUGH
He went into the town	She will give him a toy	He hit the ball	He ran in the street
He went by the town	She will buy him a toy	He threw the ball	He ran through the street
He went buy the town	She will by him a toy	He through the ball	He ran threw the street
He went bie the town	She will bie him a toy	He throo the ball	He ran throo the street
He went thair the town	She will thair him a toy	He soe the ball	He ran soe the street
He went their the town	She will there him a toy	He so the ball	He ran sew the street

Table I Continued

FOUR	FOR	NO	KNOW
I have five dogs	The toy is with John	I have the time	I like your name
I have four dogs	The toy is for John	I have no time	I know your name
I have for dogs	The toy is four John	I have know time	I no your name
I have faw dogs	The toy is faw John	I have noe time	I noe your name
I have cea dogs	The toy is cea John	I have bloo time	I bloo your name
I have sea dogs	The toy is see John	I have blue time	I blew your name

THEIR	THERE	WHERE	WEAR
We will stay at my house	That is a leaf on the tree	Tell me why he went	I will have a green dress
We will stay at their house	There is a leaf on the tree	Tell me where he went	I will wear a green dress
We will stay at there house	Their is a leaf on the tree	Tell me wear he went	I will where a green dress
We will stay at thair house	Thair is a leaf on the tree	Tell me wair he went	I will wair a green dress
We will stay at bie house	Bie is a leaf on the tree	Tell me nue he went	I will nue a green dress
We will stay at by house	Buy is a leaf on the tree	Tell me new he went	I will knew a green dress

Table II
List of 80 Sentences Used in Experiment 4

(BUY, BIE, and BEL are graphically similar)	(BY, BIE, and BAK are graphically similar)	(WHERE, WAIR, and WOAR are graphically similar)	(WEAR, WAIR, and WAUR are graphically similar)
She will bie him a toy	He went bie the town	Tell me wair he went	I will wair a green dress
I ring the bie	Let's go bie to the house	They fight in the wair	He fights in the wair
I ring the bel	Let's go bak to the house	They fight in the woar	He fights in the waur
She will bel him a toy	He went bak the town	Tell me woar he went	I will waur a green dress

(SEA, CEA, and BEA are graphically similar)	(SEE, CEA, and KEA are graphically similar)	(NO, NOE, and NOK are graphically similar)	(KNOW, NOE, and NOK are graphically similar)
We swim in the cea	I will cea him today	I have noe time	I noe him
A cea can fly	I have a cea for the door	I noe on the door	Noe is white
A bea can fly	I have a kea for the door	I nok on the door	Sno is white
We swim in the bea	I will kea him today	I have nok time	I sno him

(NEW, NUE, and NOE are graphically similar)	(KNEW, NUE, and WUN are graphically similar)	(ATE, AIT, and ATT are graphically similar)	(EIGHT, AIT, and ITT are graphically similar)
Tell us a nue story	I nue what to tell him	Tom ait his apple	He is ait years old
She said nue to him	He is nue of my friends	My dog is ait home	Ait is hot today
She said noe to him	He is wun of my friends	My dog is att home	Itt is hot today
Tell us a noe story	I wun what to tell him	Tom att his apple	He is itt years old

(ONE, WUN, and KAN are graphically similar)	(WON, WUN, and WEN are graphically similar)	(BE, BEA, and BEL are graphically similar)	(BEE, BEA, and BEL are graphically similar)
Wun of us will come	That boy wun a toy	He wants to bea a doctor	A bea can fly
I wun sing a song	Wun are you going home?	The bea is ringing	She rings the bea
I kan sing a song	Wen are you going home?	The bel is ringing	She rings the bel
Kan of us will come	That boy wen a toy	He wants to bel a doctor	A bel can fly

(FOUR, FAW, and FEL are graphically similar)	(FOR, FAW, and FIL are graphically similar)	(BLUE, BLOO, and BLAK are graphically similar)	(BLEW, BLOO, and BLAK are graphically similar)
I have faw dogs	The toy is faw John	The sun shines in the bloo sky	She bloo up the balloon
I faw down	He will faw the box with sand	That man is bloo	That girl is bloo
I fel down	He will fil the box with sand	That man is blak	That girl is blak
I have fel dogs	The toy is fil John	The sun shines in the blak sky	She blak up the balloon

APPENDIX B

Table I
Percentage of Substitutions (PS), Graphic Similarity (GS), Word Frequency (WF), and Errors for Homophones

Presented as	Used as Meaning	PS	GS	WF	Errors
NO	KNOW	7.5	258.33	2201	2
KNOW	NO	2.5	258.33	683	5

Table I Continued

Presented as	Used as Meaning	PS	GS	WF	Errors
ONE	WON	.0	283.33	3292	5
WON	ONE	5.0	283.33	68	2
SEW	SO	7.5	343.50	6	2
SO	SEW	.0	343.50	1984	2
WHERE	WEAR	7.5	376.67	938	8
WEAR	WHERE	10.0	376.67	36	3
BLEW	BLUE	7.5	520.00	12	3
BLUE	BLEW	.0	520.00	143	3
THROUGH	THREW	30.0	522.17	969	7
THREW	THROUGH	12.5	522.17	46	3
SEE	SEA	.0	553.33	772	1
SEA	SEE	.0	553.33	95	2
BY	BUY	12.5	563.50	5305	8
BUY	BY	10.0	563.50	70	3
NEW	KNEW	5.0	588.93	1635	8
KNEW	NEW	25.0	588.93	395	3
THERE	THEIR	.0	600.00	2724	8
THEIR	THERE	42.5	600.00	2670	3
WOULD	WOOD	7.5	607.78	2714	0
WOOD	WOULD	7.5	607.78	55	3
FOR	FOUR	2.5	716.07	9489	3
FOUR	FOR	7.5	716.07	359	1

Note—Word frequencies given are those for the presented homophones.

Table II
Percentage Substitutions (PS), Graphic Similarity (GC) Scores, Word Frequency (WF) Scores, and Errors for Nonwords

Presented as	Used as Meaning	PS	GS	WF	Errors
WUN	ONE	5.0	83.30	3292	1
CEA	SEE	15.0	83.30	772	1
NUE	KNEW	12.5	94.60	395	1
NOE	KNOW	12.5	237.50	683	1
FAW	FOUR	2.5	336.07	359	2
BIE	BY	5.0	343.50	5305	0
BIE	BUY	30.0	353.33	70	1
FAW	FOR	10.0	353.33	9489	2
WAIR	WHERE	17.5	354.44	938	4
SOE	SEW	22.5	386.67	6	2
NUE	NEW	20.0	386.67	1635	2
CEA	SEA	10.0	463.30	95	3
THAIR	THERE	40.0	480.00	2724	4
BLOO	BLUE	5.0	495.00	143	6
BLOO	BLEW	32.5	495.00	12	4
WUN	WON	4.75	566.67	68	5
WAIR	WEAR	32.5	575.00	36	4
THROO	THREW	22.5	580.00	46	3
NOE	NO	12.5	583.50	2201	1
THROO	THROUGH	27.5	583.50	969	2
WOUD	WOOD	7.5	700.00	55	1
THAIR	THEIR	7.5	780.00	2670	5
SOE	SO	7.5	783.50	1984	2
WOUD	WOULD	35.0	801.11	2714	6

APPENDIX C

Table I
Percentage Correct "No" Responses (PC) to All-Word Sentences Containing Meaningless but Phonologically Appropriate Homophones and Percentage Incorrect Usage of Homophones by Substitution (PI) Given for Children Aged 6-10 Years

Presented as	Used as Meaning	PC	PI
NO	KNOW	80	7.5
KNOW	NO	50	2.5
ONE	WON	50	.0
WON	ONE	80	5.0
SEW	SO	80	7.5
SO	SEW	80	.0
WHERE	WEAR	20	7.5
WEAR	WHERE	70	10.0
BLEW	BLUE	70	7.5
BLUE	BLEW	70	.0
THROUGH	THREW	30	30.0
THREW	THROUGH	70	12.5
SEE	SEA	90	.0
SEA	SEE	80	.0
BY	BUY	20	12.5
BUY	BY	70	10.0
NEW	KNEW	20	5.0
KNEW	NEW	70	25.0
THERE	THEIR	20	.0
THEIR	THERE	70	42.5
WOULD	WOOD	100	7.5
WOOD	WOULD	70	7.5
FOR	FOUR	70	2.5
FOUR	FOR	90	7.5

Table II
Percentage Correct "No" Responses (PC) to Meaningless Sentences in Which the Nonword Sounded Correct and Percentage Incorrect Usage of Nonwords as Homophones (PI) by Children Aged 6-10 Years

Presented as	Used as Meaning	PC	PI
WUN	ONE	90	5.0
CEA	SEE	90	15.0
NUE	KNEW	90	12.5
NOE	KNOW	90	12.5
FAW	FOUR	80	2.5
BIE	BY	100	5.0
BIE	BUY	90	30.0
FAW	FOUR	80	10.0
WAIR	WHERE	60	17.5
SOE	SEW	80	22.5
NUE	NEW	80	20.0
CEA	SEA	70	10.0
THAIR	THERE	60	40.0
BLOO	BLUE	40	5.0
BLOO	BLEW	60	32.5
WUN	WON	50	4.75
WAIR	WEAR	60	32.5
THROO	THREW	70	22.5
NOE	NO	90	12.5
THROO	THROUGH	80	27.5
WOUD	WOOD	90	7.5
THAIR	THEIR	50	7.5
SOE	SO	80	7.5
WOUD	WOULD	40	35.0

APPENDIX D

Table I
Graphic Similarity (GS) Scores for the Different Types of Items and Corresponding Errors for Experiment 4

Presented as		Meaning	GS	Errors		Presented as		Meaning	GS		Errors	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
WUN	KAN	ONE	83.30	2	0	KAN	WUN	CAN	463.33	263.30	3	2
CEA	KEA	SEE	83.30	1	1	KEA	CEA	KEY	533.33	83.30	5	1
NUE	WUN	KNEW	94.64	2	0	WUN	NUE	ONE	83.30	296.70	2	1
NOE	SNO	KNOW	237.50	2	0	SNO	NOE	SNOW	746.79	237.50	0	0
FAW	FEL	FOUR	336.07	0	2	FEL	FAW	FELL	858.93	336.10	7	1
BIE	BAK	BY	343.50	1	0	BAK	BIE	BACK	716.07	336.10	4	0
BIE	BEL	BUY	353.33	2	0	BEL	BIE	BELL	858.93	354.60	4	4
FAW	FIL	FOR	353.33	2	0	FIL	FAW	FILL	858.93	336.10	5	0
WAIR	WOAR	WHERE	354.44	6	2	WOAR	WAIR	WAR	716.07	716.07	2	1
NUE	NOE	NEW	386.67	4	0	NOE	NUE	NO	583.50	343.50	0	0
CEA	BEA	SEA	463.30	2	0	BEA	CEA	BEE	553.33	83.30	7	0
BLOO	BLAK	BLUE	495.00	1	2	BLAK	BLOO	BLACK	801.10	465.60	6	0
BLOO	BLAK	BLEW	495.00	5	1	BLAK	BLOO	BLACK	801.10	465.60	3	2
WUN	WEN	WON	566.67	2	2	WEN	WUN	WHEN	716.07	554.60	3	2
WAIR	WAUR	WEAR	575.00	8	0	WAUR	WAIR	WAR	716.07	716.07	3	2
NOE	NOK	NO	583.50	3	0	NOK	NOE	KNOCK	410.00	205.00	4	2
AIT	ITT	EIGHT	260.00	0	1	ITT	AIT	IT	763.50	493.50	1	1
AIT	ATT	ATE	386.70	2	3	ATT	AIT	AT	763.50	563.50	2	1
BEA	BEL	BEE	553.30	3	1	BEL	BEA	BELL	858.93	507.50	4	1
BEA	BEL	BE	583.50	3	2	BEL	BEA	BELL	858.93	507.50	6	0

Note—Column 4 refers to the GS scores of the corresponding words and nonwords in Columns 1, 2, and 3. All three have the same GS score except for the entries on Line 18: ATE and ATT have a greater GS of 513.30. Column 5 refers to errors made by accepting nonwords in Column 1 as meaning the same as homophones in Column 3. Column 6 refers to errors made by accepting nonwords

in Column 2 as meaning the same as the homophones in Column 3, from which they differ phonologically even though they are graphically similar. Column 10 refers to the GS between the nonword in Column 7 and the corresponding word in Column 9. Column 11 refers to the GS between the nonword in Column 8 and the corresponding word in Column 9. Column 12 refers to errors made in accepting the nonword in Column 7 as meaning the same as the phonologically similar word in Column 9. Column 13 refers to errors made in accepting the nonword in Column 8 as meaning the same as the word in Column 9 although the two are phonologically different.