SCIENTIFIC STUDIES OF READING, *9*(2), 167–188

Learning to Read Words: Theory, Findings, and Issues

Linnea C. Ehri
City University of New York

Reading words may take several forms. Readers may utilize decoding, analogizing, or predicting to read unfamiliar words. Readers read familiar words by accessing them in memory, called sight word reading. With practice, all words come to be read automatically by sight, which is the most efficient, unobtrusive way to read words in text. The process of learning sight words involves forming connections between graphemes and phonemes to bond spellings of the words to their pronunciations and meanings in memory. The process is enabled by phonemic awareness and by knowledge of the alphabetic system, which functions as a powerful mnemonic to secure spellings in memory. Recent studies show that alphabetic knowledge enhances children's learning of new vocabulary words, and it influences their memory for doubled letters in words. Four phases characterize the course of development of sight word learning. The phases are distinguished according to the type of alphabetic knowledge used to form connections: pre-alphabetic, partial, full, and consolidated alphabetic phases. These processes appear to portray sight word learning in transparent as well as opaque writing systems.

Life is indeed exciting but demanding these days for researchers who study reading. Because many educators are seeking evidence as the basis for decisions about reading instruction, there is great interest in scientific studies of reading processes and instruction. My studies over the years have focused on how beginners learn to read words. My plan is to review what I think we know about learning to read words, particularly sight words; to present some new findings that involve children's vocabulary learning and memory for orthographic structure; and to point out some issues that linger. An issue of special interest is whether this research in English is relevant for more transparent orthographies.

Request for reprints should be sent to Linnea C. Ehri, Program in Educational Psychology, Graduate Center of the City University of New York, 365 Fifth Avenue, New York, NY 10016. E-mail: lehri@gc.cuny.edu

Most advances in research proceed in small steps and depend on the contributions of a community of researchers. My work is no exception. There are many colleagues who have contributed to the picture of reading acquisition that I discuss. Regrettably, I lack the space to acknowledge all of them.

One of the great mysteries that has challenged researchers is how people learn to read and comprehend text rapidly with ease. When people read text, the print fills their minds with ideas. The route to these ideas begins with individual printed words. Eye movement studies show that when readers read a text, their eyes land on practically every word (Rayner & Pollatsek, 1989). Because words are always spelled the same way, this makes them reliable units for readers' eyes to process. In contrast, grapheme–phoneme correspondences may vary. The same phoneme may be spelled more than one way, and the same letter may stand for more than one phoneme. Moreover, written words activate meanings whereas single graphemes do not. Thus, words are the basic units that readers' eyes pick up and process to construct meaning out of print. The key to understanding how reading skill develops is understanding how beginners learn to recognize written words accurately and automatically.

We can distinguish four different ways to read words (Ehri, 1991). The first three ways help us read unfamiliar words. The fourth way explains how we read words we have read before. One way is by decoding, also called phonological recoding. We can either sound out and blend graphemes into phonemes, or we can work with larger chunks of letters to blend syllabic units into recognizable words. Another way is by analogizing (Goswami, 1986). This involves using words we already know to read new words—for example, using the known word *bottle* to read *throttle*. Another way is by prediction (Goodman, 1970; Tunmer & Chapman, 1998). This involves using context and letter clues to guess unfamiliar words. The fourth way of reading words is by memory or sight. This applies to words we have read before. We can just look at the words and our brain recognizes them.

SIGHT WORD READING

Let's take a closer look at sight word reading. When readers eyes alight on a word known by sight, the word's identity is triggered in memory very rapidly (Ehri, 1992). When sight words are known well enough, readers can recognize their pronunciations and meanings automatically without any attention or effort at sounding out letters (LaBerge & Samuels, 1974). Researchers have studied automatic word reading using Stroop tasks. In these tasks, readers are shown drawings of familiar objects with words printed on them, such as a horse with *cow* written on it, or a banana with *apple* written on it. Or they are shown written words naming colors with the letters appearing in a different color, such as *red* colored blue, or *yellow* colored green. Readers are told to name aloud the drawings or colors rapidly and to ignore the written words. Despite their intention to ignore the words, reader's brains can't help but recognize them, and this slows them down in naming the competing pictures or colors. Children as young as those finishing first grade show evidence of automaticity in these tasks (Guttentag & Haith, 1978).

Another important property of sight word reading is that words come to be read as single units with no pauses between word parts, referred to as *unitization*. In one experiment (Ehri & Wilce, 1983), we had students read familiar object words (i.e., *book*, *man*, *car*, *tree*), read consonant–vowel–consonant (CVC) nonwords (e.g., *baf*, *jad*, *nel*, *des*), and name single digits (4, 6, 3, 9). We measured their latencies to read each type of stimulus. We studied younger skilled and older less skilled readers who were reading at the second- and fourth-grade equivalent levels. As evident in Figure 1, both groups read the familiar words much faster than the unfamiliar nonwords. This shows the advantage of reading familiar words from memory over decoding unfamiliar words. Of importance, the skilled readers at both grade levels read the words as quickly as they named the single digits. This indicates that the words were read as single, whole units rather than as letters processed sequentially. In contrast, less skilled readers did not show unitization until fourth grade. This is consistent with other findings indicating that poor readers have difficulty with sight word reading (Ehri & Saltmarsh, 1995).

Some people limit the term *sight word* to refer only to high-frequency words or to irregularly spelled words. However, this is not accurate. Any word that is read sufficiently often becomes a sight word that is read from memory.

Another misconception is to consider sight word reading as a strategy for reading words. However, being strategic involves choosing procedures to optimize outcomes. Readers are strategic when they figure out unknown words by decoding,

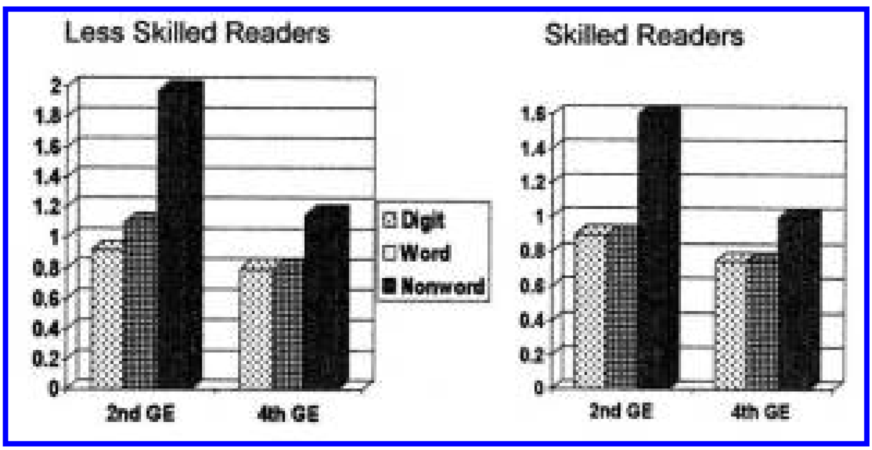

FIGURE 1 Mean latency to name digits and to read words and nonwords by skilled and less skilled readers in second and fourth grades (Ehri & Wilce, 1983).

analogizing, or predicting. But they are not behaving strategically when they read words by sight, which happens automatically and is not a matter of choice.

Given that there are multiple ways to read words, consider which way makes text reading most efficient. If readers know words by sight and can recognize them automatically as they read text, then word reading operates unconsciously. In contrast, each of the other ways of reading words requires conscious attention. If readers attempt to decode words, to analogize, or to predict words, their attention is shifted from the text to the word itself to identify it, and this disrupts comprehension, at least momentarily. It is clear that being able to read words automatically from memory is the most efficient, unobtrusive way to read words in text. Hence, building a sight vocabulary is essential for achieving text-reading skill.

SIGHT WORD LEARNING

How do children learn to read words by sight? The process at the heart of sight word learning is a *connection-forming* process. Connections are formed that link spellings of written words to their pronunciations and meanings in memory (Ehri, 1980, 1992; Perfetti, 1992; Rack, Hulme, Snowling, & Wightman, 1994).

What kinds of connections are formed to remember sight words? People used to believe that readers memorized associations between visual features such as the shapes of words and their meanings. This was one justification for the look–say, whole-word method of teaching reading. However, the visual explanation is inadequate, because visual–semantic connections lack sufficient mnemonic power to explain the facts. That is, they do not explain how the spellings of words are capable of being encoded in memory easily with very little practice. They do not explain how skilled readers are able to recognize many thousands of words in an instant with high accuracy. If meanings were the anchors for words in memory, we would expect many more synonymous readings, for example, misreading the word *pupil* as *student*. In actuality, semantic errors are rare.

Based on our findings, we have proposed that readers learn sight words by forming connections between letters in spellings and sounds in pronunciations of the words (Ehri, 1992, 1998). The connections are formed out of readers' knowledge of the alphabetic system. This includes knowledge of grapheme–phoneme relations and phonemic awareness, that is, knowing how to distinguish the separate phonemes in pronunciations of words. This also includes knowledge of spelling patterns that recur in different words. When readers learn a sight word, they look at the spelling, they pronounce the word, they distinguish separate phonemes in the pronunciation, and they recognize how the graphemes match up to phonemes in that word. Reading the word a few times secures its connections in memory.

In Figure 2, I have depicted how readers might form connections to learn several sight words. Capital letters designate the spellings of words, spacing be-

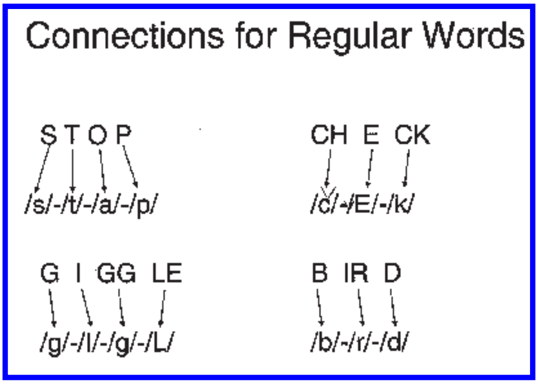

FIGURE 2 Connections between graphemes (capital letter groups) in spellings and phonemes (phonetic symbols between hyphens) in pronunciations to retain regularly spelled words in memory.

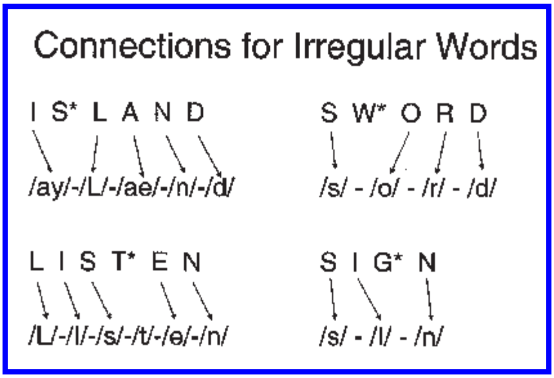

FIGURE 3 Connections between graphemes in spellings and phonemes in pronunciations to retain irregularly spelled words in memory.

tween letters distinguishes constituent graphemes that may involve single letters or digraphs, phonetic symbols between slashes indicate phonemes, and lines linking graphemes to phonemes indicate connections. A property of regularly spelled words is that all of the graphemes can be connected to phonemes in pronunciations.

This connection-forming process also depicts memory for irregularly spelled words. It turns out that most letters in irregularly spelled words conform to grapheme–phoneme conventions, for example, all but the letters with asterisks shown in Figure 3. Thus, exception words are only exceptional when someone tries to read them by applying a decoding strategy. When they are learned as sight words, they are secured in memory by the same connections as regularly spelled words, with only the exceptional letters unsecured.

Whether all the letters in spellings become secured to pronunciations in memory depends not only on irregularities in spellings but also on the readers' knowledge of the alphabetic system. If readers do not know short vowel spellings, or they do not know that *ph* symbolizes /f/, then when they encounter these letters in particular words, the letters will not become bonded to their phonemes in memory. Knowledge of these graphophonemic relations must be learned through either explicit instruction or implicit learning and practice before bonding can occur.

As readers learn about spelling patterns that recur in different words, these larger units are used to form connections to remember words (Bhattacharya & Ehri, 2004). These chunks include spellings of common little words appearing in larger words, spellings of common rimes, and spellings of morphemes and syllables.

It is important to understand how learners apply their general alphabetic knowledge to retain specific words in memory. Schema theory offers a way of portraying how this works (Anderson, Reynolds, Schallert, & Goetz, 1977). Learners possess schemata in the form of alphabetic knowledge about many spelling–sound relations and patterns. This provides them with general expectations about how any written word might be pronounced and how any pronounced word might be written. These schemata may entail more than one possibility, for example, the two ways to spell /z/ as *s* or *z*, or the two ways to spell /ayt/ as *-ite* or *-ight*. When readers see and pronounce a particular word, the relevant spelling–sound relations are instantiated and they secure that word's spelling in memory. Alternative unseen relations are not even considered.

When readers acquire sufficient knowledge of the alphabetic system, they are able to learn sight words quickly and to remember them long term. Reitsma (1983) taught Dutch first graders to read a set of words and found that a minimum of four practice trials was sufficient to enable students to read the words from memory. More recently Share (2004) reported that even one exposure to words enabled Israeli third graders to retain information about the spellings of specific words in memory, and this memory persisted a month later. Sight word learning this rapid and lasting is possible only because readers possess a powerful mnemonic system in the form of alphabetic knowledge that is activated when words are read.

To summarize, readers learn to process spellings of words as phonemic maps that lay out elements of their pronunciations visually. Beginners become skilled at computing these mapping relations spontaneously when they read new words. This is the critical event for sight word learning. Grapheme–phoneme connections provide a powerful mnemonic system. They provide the glue that bonds letters in written words to their pronunciations in memory along with meanings. Once the alphabetic mapping system is known, readers can build a vocabulary of sight words easily. Unfortunately, some children have difficulty with the automatic mapping between print and speech, and they may require much more practice to achieve a normal level of sight word learning (Ehri & Saltmarsh, 1995; Reitsma, 1983).

PHASES OF DEVELOPMENT

Now that we have considered the mnemonic system for sight word learning, let us examine the course of development. I distinguished four phases to identify significant advances that occur as children learn to read words by sight. The phases are labeled to reflect the type of alphabetic knowledge that predominates in the connections that are formed. The four phases are pre-alphabetic, partial alphabetic, full alphabetic, and consolidated alphabetic (see Ehri, 1999, in press; Ehri & McCormick, 1998, for a more complete portrayal of phase theory and evidence).

The pre-alphabetic phase characterizes sight word learning at the earliest period. Because children know little about the alphabetic system, they do not form letter–sound connections to read words. If they read words at all, they do it by remembering selected visual features. For example, they might remember *look* by the two eyeballs in the middle, or *dog* by the tail at the end, or *camel* by the humps in the middle (Gough, Juel, & Griffith, 1992).

Environmental print is read from contextual cues such as golden arches, not from letters (Mason, 1980). In one study (Masonheimer, Drum, & Ehri, 1984), we altered letters in familiar labels that the children could read, for example, showing *PEPSI* as *XEPSI*. Pre-alphabetic readers did not notice the change, even when they were cautioned about a possible mistake. Others have shown that preschoolers pay closer attention to letters in their own names, but the letters are not connected to sounds in the names (Bloodgood, 1999; Share & Gur, 1999; Treiman & Broderick, 1998).

Because most written words do not contain easily remembered cues, children in this phase are essentially nonreaders. Of course, they can pretend read stories they have heard many times, and they can guess words from pictures. However, all of their feats of reading are performed by using cues that do not involve the alphabetic system.

Children progress to the partial alphabetic phase when they learn the names or sounds of alphabet letters and use these to remember how to read words. However, they form connections between only some of the letters and sounds in words, often only the first and final letter sounds, which are easier to detect, for example, the letters *s* and *n* to read *spoon* (Savage, Stuart, & Hill, 2001). They may confuse similarly spelled words such as *spoon* and *skin* having the same boundary letters.

They are limited to forming partial connections because they are unable to segment the word's pronunciation into all of its phonemes. Also they lack full knowledge of the alphabetic system, especially vowels. Because of this, partial phase readers have much difficulty decoding unfamiliar words. They invent partial spellings of words by writing only the more salient sounds and leaving out medial letters.

To show the difference between the pre-alphabetic and partial alphabetic phases, we tested kindergartners and separated them into the two phases (Ehri & Wilce, 1985). We gave them several practice trials to learn to read two kinds of spellings. One type involved visual spellings with varied shapes but no relationship to sounds, for example, *mask* spelled uHo. The other type involved phonetic spellings whose letters represented some sounds in the words, for example, *mask* spelled MSK. As evident in Figure 4, pre-alphabetic phase readers learned to read visual spellings more easily than phonetic spellings, confirming our idea that they must depend on visual cues because they lack knowledge of letters. In contrast, partial alphabetic readers showed the opposite pattern and were able to use letter–sound cues to remember the words.

Others have replicated and extended our findings (de Abreu & Cardoso-Martins, 1998; Rack et al., 1994; Roberts, 2003; Scott & Ehri, 1989; Treiman & Broderick, 1998; Treiman & Rodriguez, 1999). Roberts manipulated preschoolers' knowledge of letter names experimentally and showed that children who were taught the names of letters learned phonetic spellings better whereas children not taught letters learned visual spellings better. De Abreu and Cardoso-Martins obtained the same results with Brazilian children reading in Portuguese. These findings combine to show that when children learn the names of alphabet letters, they have available a more effective mnemonic system that improves their sight word learning.

Not only novice beginning readers but also older children with a reading disability qualify as partial alphabetic phase readers. They too retain in memory only partial representations of words, with medial letters poorly bonded to pronunciations (Ehri & Saltmarsh, 1995).

Children become full alphabetic phase readers when they can learn sight words by forming complete connections between letters in spellings and phonemes in

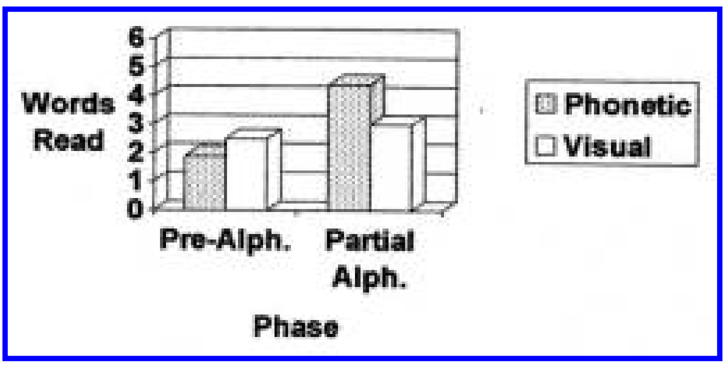

FIGURE 4 Mean phonetically and visually spelled words that were read correctly per trial over five trials by readers at the pre-alphabetic and partial alphabetic phases of development (Ehri & Wilce, 1985).

pronunciations (see Figures 2 and 3). This is possible because they know the major grapheme–phoneme correspondences (Venezky, 1970, 1999) and they can segment pronunciations into phonemes that match up to the graphemes they see (Ehri et al., 2001; Liberman, Shankweiler, Fischer, & Carter, 1974; Share, Jorm, Maclean, & Matthews, 1984). From the perspective of schema theory, those alphabetic schemata that fit the graphophonemic relations they encounter in particular words are the ones that are activated spontaneously to connect letters to sounds in memory. When spellings are instantiated in this way, that is, when they are recognized as an instance of the reader's background alphabetic knowledge, they become fully bonded to pronunciations in memory.

One advantage of representing sight words completely in memory is that word reading becomes much more accurate, and similarly spelled words are seldom confused. At this phase, readers are able to decode unfamiliar words, they can invent spellings that represent all the phonemes, and they can remember correct spellings of words better than partial phase readers.

We conducted a study to show the difference in sight word learning between full and partial phase readers (Ehri & Wilce, 1987). We selected kindergartners in the partial alphabetic phase. We randomly assigned them to a treatment or a control group. The treatment group received training to make them full phase readers. They practiced reading similarly spelled words that required processing all the grapheme–phoneme relations in the words to read them accurately. The control group remained partial phase readers. They practiced isolated grapheme–phoneme relations. Then we gave both groups practice learning to read a set of 15 words over several trials. The words had similar spellings that made them harder to learn by remembering partial cues (e.g., *spin, stab, stamp, stand*). None of the children could read more than 2 of these words prior to training.

Figure 5 shows the mean percentage of words read correctly by each group. A huge difference is apparent. Whereas full phase readers learned to read most of the words in three trials, partial phase readers never reached this level of learning. One reason for their difficulty is that they confused similarly spelled words. This shows the great advantage to readers when they can form full connections to retain sight words in memory.

The consolidated phase emerges as full phase readers retain increasingly more sight words in memory. As they become familiar with letter patterns that recur in different words, the grapheme–phoneme connections in these words become consolidated into larger units. These include spellings of rimes, syllables, morphemes, and whole words that have become unitized. Knowing letter chunks is valuable for remembering how to read multisyllabic words. Readers who know the relevant chunks can learn a word such as *interesting* more easily, because fewer connections are required to secure the word in memory. The number is reduced from 10 grapheme–phonemes to four syllabic chunks.

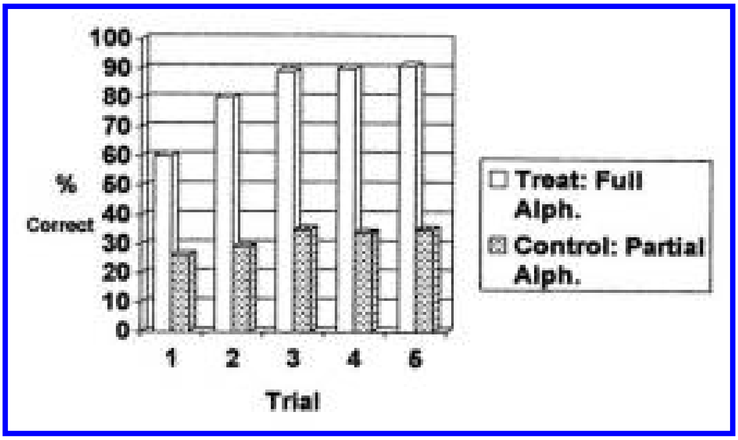

FIGURE 5 Mean percentage of similarly spelled words (i.e., *spin, stab, stamp, stand*) read correctly over trials by children receiving the full alphabetic phase treatment and by control children receiving the partial alphabetic phase treatment (Ehri & Wilce, 1987).

To describe the course of development, phases rather than stages have been proposed to avoid stringent criteria that a stage theory must meet. Phase theory provides a looser view of the properties that portray the course of acquisition. The evidence also does not support a stage view of acquisition (Ehri, in press; Frith, 1985; Stuart & Coltheart, 1988).

Stage theory requires each stage to be a prerequisite for the next stage. However, in the case of phase theory, word reading at the pre-alphabetic phase does not contribute to word reading during later phases. Children who select visual cues like golden arches to read words do this only by default because they have not yet acquired alphabetic knowledge. Knowing how to read golden arches does not help them learn to read words alphabetically.

Once children begin to use alphabetic processes, the phases emerge successively. However, unlike stages that are qualitatively distinct, children may use connections from more than one phase to learn sight words. Phases simply characterize the predominant types of alphabetic knowledge used. For example, readers in the full phase use mainly grapheme–phoneme connections to learn words but they may resort to partial connections for longer words. Also they may process well-known words such as *at, in,* and *up* as consolidated units when they appear in longer words.

VOCABULARY LEARNING

The alphabetic system functions as a very powerful mnemonic device to enable sight word learning. According to our theory, readers store visual spellings of words in memory by analyzing how graphemes symbolize phonemes in pronunciations. We reasoned that if this is true, then spellings should improve memory for spoken words. A common example of this is when you ask someone to spell an unusual name so you can remember it better (Ehri, 1984).

To test the mnemonic power of spellings, we used first and second graders (Ehri & Wilce, 1979). They were taught to pronounce on command a set of four nonwords such as *mav, rel, kip,* and *guz,* each paired with a number. A paired associate learning task was used. During study periods, children practiced saying the nonwords. During test trials, each number was shown, and children recalled the nonword. In one condition, spellings were shown with the numbers during study periods as depicted in Figure 6. The spellings simply appeared beneath the numbers, and no attention was drawn to them. In the control conditions, children either repeated the nonword extra times or saw misspellings of the nonwords but never correct spellings.

Children were given several trials to learn the nonwords. We found that those who saw correct spellings learned the nonwords faster than children who did not see spellings. Looking at misspellings made it especially hard to learn the nonwords. Our explanation is that spellings improved memory for sounds because they were retained as visual symbols preserving the sounds in memory.

More recently, we have examined the mnemonic power of spellings for vocabulary learning (Rosenthal & Ehri, 2005). We wondered whether seeing spellings of new words would enhance learners' memory for pronunciations and meanings of the words. The students were second graders. They were taught two sets of six low-frequency nouns and their meanings, for example, *gam* meaning a family of

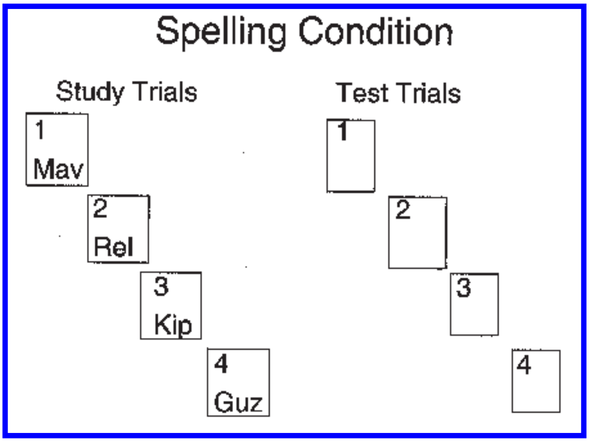

FIGURE 6 Depiction of the stimuli and procedures to teach students to remember spoken nonwords in the condition where they saw spellings during study trials but not during test trials (Ehri & Wilce, 1979).

whales, *yag* meaning fake jewelry, *sod* meaning wet grassy ground, and *fet* meaning a big fun party.

Children learned both sets of words by hearing and repeating them. In one condition, they also saw the words' spellings during study periods, but no attention was drawn to their presence. In the other condition, they did not see spellings. To make up for the possibility that spellings prompted students to rehearse the words an extra time, students in the control condition spoke the words an extra time. Students also saw the words pictured and heard sentences that defined the words in both conditions. Students received two types of test trials that were interleaved. On one trial, they saw meanings depicted and they recalled the words. On another trial, they heard the words and recalled their meanings. Incorrect responses were corrected.

Children's memory for words and meanings during test trials was scored. Figure 7 reveals that students learned the words much better when they were exposed to spellings than when they were not. In other words, seeing spellings during study periods improved their ability to recall and pronounce the new vocabulary words when spellings were not present. Figure 8 shows that students who saw spellings learned the meanings of the words better than students who did not see spellings. Remembering how to say the words was much harder than remembering their meanings, indicating that spellings were especially helpful with the harder part of vocabulary learning.

From these findings we conclude that the alphabetic system provides a mnemonic that helps students secure new vocabulary words in memory, both their pronunciations and their meanings. This constitutes one more reason why beginners need a strong alphabetic foundation when they learn to read. It helps them acquire new vocabulary words.

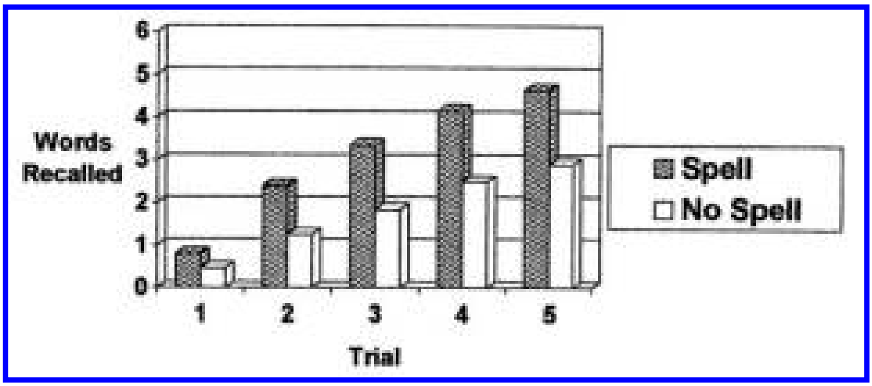

FIGURE 7 Mean vocabulary words recalled during test trials by students who saw spellings of the words during study trials and by students who did not see spellings but rehearsed the words an extra time.

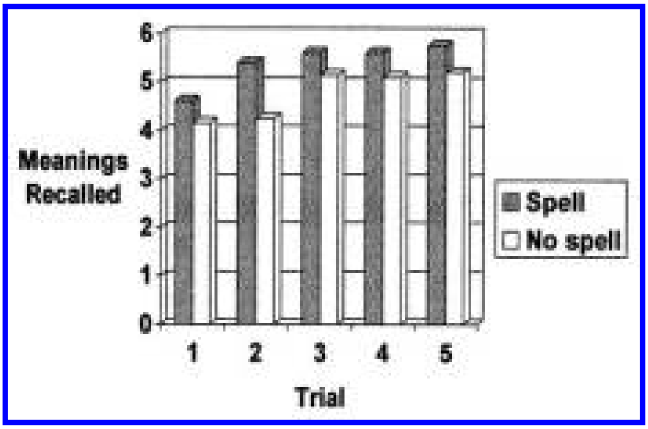

FIGURE 8 Mean number of meanings of vocabulary words recalled during test trials by students who saw spellings during study trials and by students who did not see spellings.

These findings bear on Share's (1995, 1999, 2004) self-teaching mechanism. They raise questions about the claim that learners need to apply a decoding procedure to retain sight words in memory when the words are read. In our studies, learners did not decode the words taught with spellings. Rather they heard the words pronounced by someone else. In fact, spellings appeared but were never even mentioned. Nevertheless students retained specific information about the spellings in memory. This suggests that it is not the conscious application of a decoding procedure (i.e., sounding out letters and blending them) that is critical but the implicit, spontaneous activation of alphabetic knowledge that connects graphemes to phonemes to secure the spellings of specific words in memory. Use of a decoding strategy may help students apply this knowledge, but it is the knowledge rather than the act of decoding that is critical. Alternatively, it may be that some of the children decoded the spellings subvocally when they saw them. This is an issue for future study.

MEMORY FOR ORTHOGRAPHIC STRUCTURE

Another issue we have pursued is the question of when children acquire knowledge about orthographic structure when they learn to read words. According to phase theory, readers use mainly grapheme–phoneme knowledge to remember spellings of words during the partial and full phases. Knowledge of spelling patterns is later to emerge. However, Cassar and Treiman (1997) suggested that orthographic knowledge emerges earlier than might be expected by phase theory. They performed a study with kindergartners and first graders. The students were given a two-choice judgment task. They were shown pairs of nonwords, one with a legal doubled letter and one with an illegal doublet (e.g., *baff* vs. *bbaf* and *pess* vs. *ppes*). Results indicated that children were sensitive earlier than expected. Even the kindergartners could select which of the two doubled consonants was orthographically legal, beyond the level of chance.

Note that doubling in English differs from that in other languages. In Finnish, doubled letters contrast with their single mates in symbolizing different phonemes, with doublets representing longer phonemes than single letters. In English, doubled letters are just like their single mates in representing the same phoneme (e.g., the medial consonant in *later* and *latter*).

We conducted a word-learning study to see whether beginners would remember doubled consonants in words they were taught to read (Wright & Ehri, 2005). We tested and classified kindergartners and first graders as partial or full phase readers. We taught them to read 12 CVC words. Four words contained final doublets, either *T* or *D* as in *jet* spelled JETT, and *mud* spelled MUDD. Four words contained initial doublets, either *L* or *R*, as in *rug* spelled RRUG, and *luck* spelled LLUK. Four words contained single consonants, for example, *FAN* and *TUB*. All letters but the initial doublets were legal patterns in English. Children practiced reading the words for several trials. They took somewhat longer learning words with initial doublets, indicating that the illegal letters slowed down learning.

Practice continued until students learned to read all the words. Then they were asked to spell the words from memory. We scored their memory for initial and final consonants. Figure 9 shows that both partial and full phase readers remembered the spellings of all but one type of consonant very well. They remembered final doubled

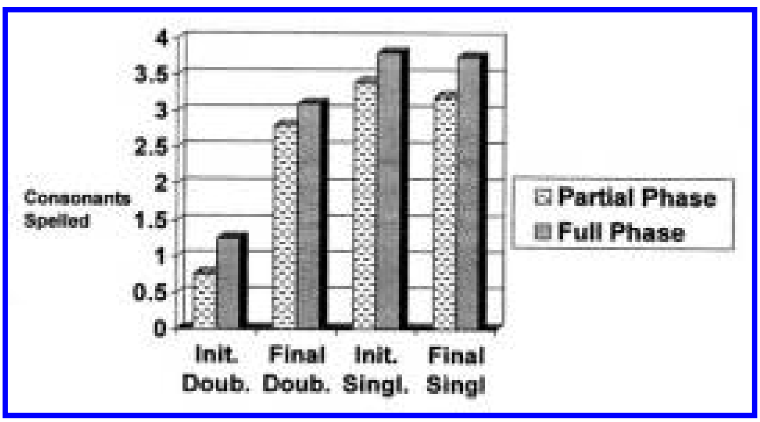

FIGURE 9 Mean number of initial and final single and double consonants recalled in the spelling test by students in the partial and full alphabetic phases of development.

consonants and initial and final single consonants. However, they did not remember the initial doubled consonants, that is, the doubled *R* and *L* that were illegal.

Present findings pose a challenge. In phase theory, the sound mapping function of letters is emphasized and their visual form is slighted in explaining how beginners in the partial and full alphabetic phases form connections to store words in memory when they read them. However, even partial alphabetic readers remembered which final letters were doubled and which were not, despite the fact that both letter types were graphophonemic and represented phonemes in the words. This shows that beginners' memory for words includes a visual component.

In light of their good memory for final doubled letters, children's poor memory for initial doubled letters was surprising. We know that beginners pay special attention to initial letters in words when they learn to read them, so these letters should have been among the easiest to remember (Bowman & Treiman, 2002). Apparently, children's general alphabetic knowledge regarding legal and illegal spellings influenced their word memory more than surface features such as a salient letter position.

Children made many errors spelling words with initial doublets, so we examined how they misspelled the words. In most cases, they wrote the *R* or *L* as single consonants. However, sometimes in addition they doubled the final consonant. Some examples are shown in Figure 10. This mistake was limited to the illegal words, as they rarely doubled final letters in the single-consonant words. The pattern was more common among full phase readers than among partial phase readers.

One explanation is that readers remembered that something was doubled in these words. This caused them to double a consonant but to do it in a legal position rather than an illegal position. These findings reveal the impact of general alphabetic knowledge on specific word learning. When word spellings are inconsistent with general knowledge, they are harder to remember and they may get regularized.

Training Word Seen	Conventional Spelling (Not Seen)	Misspelling of Initial Doublets
LLIM	(limb)	LEMM, LAMM, LUMM
LLUK	(luck)	LUKK, LAKK, KAKK
RRUG	(rug)	RUGG, ROGG
RRIP	(rip)	RIPP, REPP

FIGURE 10 Misspellings of words containing initial doublets produced by students in the spelling test.

RELEVANCE TO TRANSPARENT ORTHOGRAPHIES

Sight word learning processes and phase theory have interested researchers studying languages other than English. One question that arises is whether theory and findings pertain only to reading acquisition in English? Analyses of the structure of alphabetic writing systems gives some reason for doubt. Seymour, Aro, and Erskine (2003) recently categorized several European alphabetic languages using two dimensions, orthographic depth, and syllabic complexity. Shallow languages exhibit consistent mappings between letters and phonemes whereas deep orthographies contain inconsistencies. Languages with simple syllable structures have open CV syllables with few consonant clusters. Complex languages have numerous closed CVC syllables and complex consonant clusters. The English writing system stands apart from the other languages in being both deep and complex. This suggests that English might very well be a unique case.

Seymour et al. (2003) compared first graders learning to read in different languages. Students' errors reading words and nonwords were examined toward the end of the school year. The English sample from Scotland performed much worse than students reading in all of the other languages. In fact, their rate of development was more than twice as slow as in the shallow orthographies. Perhaps phase theory and findings based on English are not generalizable. Alternatively, perhaps the same processes operate across languages but take more time to acquire when writing systems are more complex.

One issue is whether sight word reading occurs in more transparent orthographies, which can be read by transforming letters into sounds, blending the sounds, and then recognizing the meaning of the blend. Because decoding is a viable strategy that works to read words, sight word reading may not be necessary. However, according to our theory, if readers have alphabetic knowledge that they apply to read words by forming graphophonemic connections, the words should become secured in memory and read by sight. Defior, Cary, and Martos (2002) and Wimmer and Goswami (1994) provided relevant data. They compared German, Spanish, and Portuguese students' latencies to name digits and to read number words and nonwords. Their data mirror the unitization study previously discussed and shown in Figure 1.

Results of the Defior et al. (2002) study with Spanish students are presented in Figure 11. Students from first through fourth grades read the familiar words faster than the unfamiliar but decodable pseudowords, showing that the words were read from memory rather than decoded. Unitization was evident among second through fourth graders who read the words as fast as they named the digits. These findings indicate that sight word reading processes are relevant in transparent orthographies. Students do not persist in decoding words once the words are practiced and retained in memory. However, it is important to note that even when sight words are unitized and read from memory, the process of accessing them in memory is

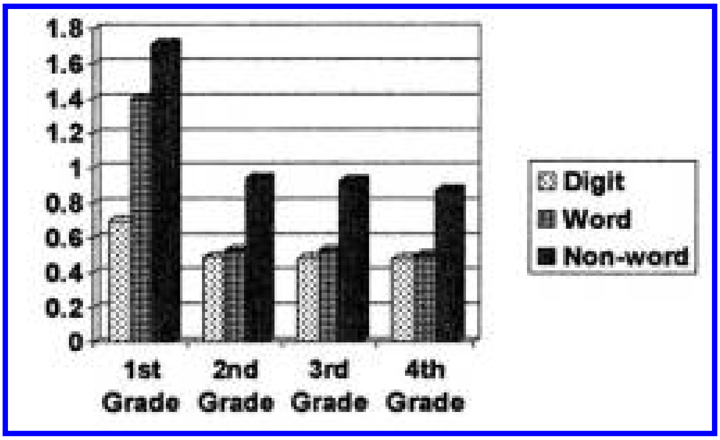

FIGURE 11 Mean latency to name digits and to read words and nonwords by Spanish-reading students in first through fourth grades (Defior et al., 2002).

still phonological in that graphophonemic connections are rapidly activated to retrieve pronunciations and meanings in memory.

Another issue is whether phase theory applies to languages other than English. Because the phases are structured around knowledge of an alphabetic writing system, one might expect the theory to apply fully to transparent systems. Alternatively, because the systems are so transparent, the partial alphabetic phase might not be relevant. Wimmer and Hummer (1990) found that students began reading German using full alphabetic cues and showed little evidence of partial cue use in reading words. They suggested that the partial phase may not apply in transparent orthographies.

To examine this possibility, Cardoso-Martins (2001) studied beginners reading in Portuguese, another transparent system. She compared kindergartners who were receiving instruction with a whole-word method to beginners taught with a phonics method after about 3 months of reading instruction. Students knew at least some letters and sounds. She found that the whole-word group did exhibit the partial phase in their reading and spelling. In contrast, beginners who were taught with the phonics method did not show evidence of the partial phase but started out reading by decoding words. Her findings indicate that the instructional method influences how long beginners show evidence of the partial phase and how quickly they acquire use of full graphophonemic connections in more transparent orthographies.

The reason why Wimmer and Hummer (1990) did not observe partial phase reading in their study is that their beginning readers had received 8 months of reading instruction with a phonics method and had learned to decode words. So these students had already moved beyond this phase.

Another issue raised about phase theory involves the full alphabetic phase in transparent writing systems: What are the skills that enable beginners to move from the partial to the full phase? Recall that in the full phase, readers bond spellings fully to pronunciations to store words in memory. In our studies in English, knowledge of the major grapheme–phoneme relations plus decoding skill has distinguished full from partial phase readers. However, Share (2004) reported a study with Hebrew-reading first graders that appears to show otherwise.

Share (2004) tested first graders who had received intense phonics instruction in a transparent orthography. They practiced reading several target words embedded in stories, for example, *ketem*, *kaskasim*, *atalef*. (The words were written in Hebrew, not English.) Then 7 days later their memory for specific letters in the words was assessed. Despite eight trials of practice, these beginners did not remember the letters. Their performance contrasted markedly to third graders who clearly remembered the specific letters, even after reading the words only once, even after a month had passed.

Share's (2004) findings suggest that the emergence of fully bonded sight words in memory may be delayed in beginners learning to decode a transparent orthography. However, there are some alternative explanations to be considered. Share's first graders knew grapheme–phoneme relations well enough, but the decoding strategy they applied to read words involved sounding out individual letters but not blending them to pronounce words as identifiable wholes (Share, personal communication, June 2004). According to our theory, storing sight words in memory requires securing grapheme–phoneme connections to pronunciations and meanings of word units in memory. Share's students said the sounds of letters but did not connect them to whole words. This may be why they did not remember letters in the spellings of the words. Alternatively, perhaps Share's multiletter words were too hard for first graders to remember, or they needed more practice, or the 8-day delay between learning and recall was too great. These possibilities await further study.

CONCLUSION

To summarize, my intent was to present a coherent picture of the processes involved in sight word reading and its development. I described how sight words are learned, through a connection-forming process that bonds spellings to pronunciations in memory. The connections are secured with glue consisting of readers' knowledge of the alphabetic system that is activated when specific words are seen and pronounced. Knowledge of the alphabetic system provides readers with a powerful mnemonic that enables sight word learning. It enhances beginners' memory for vocabulary words. It influences beginners' memory for doubled letters in words.

Sight word learning develops in four phases labeled to reflect the alphabetic knowledge that is used to connect the sight words in memory. These phases portray development not only in opaque writing systems such as English but also in transparent orthographies. However, some properties of the phases may differ. The partial alphabetic phase may have a shorter life in transparent systems because decoding skill emerges sooner. However, decoding skill may not be sufficient to move readers to the full phase if it is not practiced as a tool for building a sight vocabulary but is simply applied as a strategy for sounding out the letters in words. Although our studies and others have revealed much about how sight word reading develops, there is more to be learned from research.

To conclude, I offer some words of inspiration for researchers in their quest to advance our knowledge about literacy. I draw these words from the New York Yankees baseball legend and semantic wizard, Yogi Berra (1998).

- First: "Never give up, because it ain't over 'till its over."
- Second: "During the years ahead, when you come to a fork in the road, take it."
- Third: "Don't always follow the crowd, because nobody goes there anymore. It's too crowded."
- Fourth: "You've got to be careful if you don't know where you're going, 'cause you might not get there."
- Fifth: In conducting your experiments, "remember that you can observe a lot by watching."
- Sixth: Replicating your findings is important. "It's deja vu all over again."
- Seventh and last: "Remember that whatever you do in life, 90 percent of it is half mental."

In closing, I wish all of you great success in your research. I am extremely grateful to all my colleagues who have supplied ideas and findings to fuel my studies and thinking. As Yogi Berra would say, "Thank you for making this day necessary."

ACKNOWLEDGMENT

This article is adapted from Linnea Ehri's Distinguished Scientific Contributions Award address at the annual meeting of the Society for the Scientific Study of Reading held in Amsterdam, The Netherlands, June 2004.

REFERENCES

Anderson, R., Reynolds, R., Schallert, D., & Goetz, E. (1977). Frameworks for comprehending discourse. *American Educational Research Journal*, *14*, 367–382.

Berra, Y. (1998). *The Yogi book.* New York: Workman.

Bhattacharya, A., & Ehri, L. (2004). Graphosyllabic analysis helps adolescent struggling readers read and spell words. *Journal of Learning Disabilities*, *37,* 331–348.

Bloodgood, J. (1999). What's in a name? Children's name writing and name acquisition. *Reading Research Quarterly*, *34*, 342–367.

Bowman, M., & Treiman, R. (2002). Relating print and speech: The effects of letter names and word position on reading and spelling performance. *Journal of Experimental Child Psychology*, *82*, 305–340.

Cardoso-Martins, C. (2001). The reading abilities of beginning readers of Brazilian Portuguese: Implications for a theory of reading acquisition. *Scientific Studies of Reading*, *5*, 289–317.

Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology*, *89*, 631–644.

de Abreu, M., & Cardoso-Martins, C. (1998). Alphabetic access route in beginning reading acquisition in Portuguese: The role of letter-name knowledge. *Reading and Writing: An Interdisciplinary Journal*, *10*, 85–104.

Defior, S., Cary, L., & Martos, F. (2002). Differences in reading acquisition development in two shallow orthographies: Portuguese and Spanish. *Applied Psycholinguistics*, *23*, 135–148.

Ehri, L. (1980). The development of orthographic images. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 311–338). San Diego, CA: Academic.

Ehri, L. (1984). How orthography alters spoken language competencies in children learning to read and spell. In J. Downing & R. Valtin (Eds.), *Language awareness and learning to read* (pp. 119–147). New York: Springer Verlag.

Ehri, L. (1991). Development of the ability to read words. In R. Barr, M. Kamil, P. Mosenthal, & P. Pearson (Eds.), *Handbook of reading research Volume II* (pp. 383–417). New York: Longman.

Ehri, L. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107–143). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ehri, L. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. Metsala & L. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 3–40). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Ehri, L. (1999). Phases of development in learning to read words. In J. Oakhill & R. Beard (Eds.), *Reading development and the teaching of reading: A psychological perspective* (pp. 79–108). Oxford, UK: Blackwell.

Ehri, L. (in press). Development of sight word reading: Phases and findings. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook*. Oxford, UK: Blackwell.

Ehri, L., & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading and Writing Quarterly*, *14*, 135–163.

Ehri, L., Nunes, S., Willows, D., Schuster, B., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, *36*, 250–287.

Ehri, L. & Saltmarsh, J. (1995). Beginning readers outperform older disabled readers in learning to read words by sight. *Reading and Writing: An Interdisciplinary Journal*, *7*, 295–326.

Ehri, L. & Wilce, L. (1979). The mnemonic value of orthography among beginning readers. *Journal of Educational Psychology*, *71*, 26–40.

Ehri, L. C., & Wilce, L. S. (1983). Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology*, *75*, 3–18.

Ehri, L. C., & Wilce, L. S. (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly*, *20*, 163–179.

Ehri, L., & Wilce, L. (1987). Cipher versus cue reading: An experiment in decoding acquisition. *Journal of Educational Psychology*, *79*, 3–13.

Frith, V. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 301–330). London: Lawrence Erlbaum Associates, Inc.

Goodman, K. (1970). Behind the eye: What happens in reading. In K. Goodman & O. Niles (Eds.), *Reading: Process and program* (pp. 3–38). Urbana, IL: National Council of Teachers of English.

Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology*, *42*, 73–83.

Gough, P., Juel, C., & Griffith, P. (1992). Reading, spelling and the orthographic cipher. In P. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 35–48). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Guttentag, R., & Haith, M. (1978). Automatic processing as a function of age and reading ability. *Child Develoment*, *49*, 707–716.

LaBerge, D., & Samuels, J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293–323.

Liberman, I., Shankweiler, D., Fischer, F., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, *18*, 201–212.

Mason, J. (1980). When *do* children learn to read: An exploration of four-year old children's letter and word reading competencies. *Reading Research Quarterly*, *15*, 202–227.

Masonheimer, P., Drum, P., & Ehri, L. (1984). Does environmental print identification lead children into word reading? *Journal of Reading Behavior*, *16*, 257–272.

Perfetti, C. (1992). The representation problem in reading acquisition. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Rack, J., Hulme, C., Snowling, M., & Wightman, J. (1994). The role of phonology in young children learning to read words: The direct-mapping hypothesis. *Journal of Experimental Child Psychology*, *57*, 42–71.

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

Reitsma, P. (1983). Printed word learning in beginning readers. *Journal of Experimental Child Psychology*, *75*, 321–339.

Roberts, T. (2003). Effects of alphabet letter instruction on young children's word recognition. *Journal of Educational Psychology*, *95*, 41–51.

Rosenthal, J., & Ehri, L. (2005). *Spellings enhance readers' memory for new vocabulary words and definitions.* Unpublished manuscript, Graduate Center of the City University of New York.

Savage, R., Stuart, M., & Hill, V. (2001). The role of scaffolding errors in reading development: Evidence from a longitudinal and a correlational study. *British Journal of Educational Psychology*, *71*, 1–13.

Scott, J., & Ehri, L. (1989). Sight word reading in prereaders: Use of logographic vs. alphabetic access routes. *Journal of Reading Behavior*, *22*, 149–166.

Seymour, P., Aro, M., & Erskine, J. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–174.

Share, D. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*, 151–218.

Share, D. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*, 95–129.

Share, D. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology*, *87,* 267–298.

Share, D., & Gur, T. (1999). How reading begins: A study of preschoolers' print identification strategies. *Cognition and Instruction*, *17*, 177–213.

Share, D., Jorm, A., Maclean, R., & Matthews, R. (1984). Sources of individual diferences in reading achievement. *Journal of Educational Psychology*, *76*, 1309–1324.

Stuart, M., & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition*, *30*, 139–181.

Treiman, R., & Broderick, V. (1998). What's in a name: Children's knowledge about the letters in their own names. *Journal of Experimental Child Psychology*, *70*, 97–116.

Treiman, R., & Rodriguez, K. (1999). Young children use letter names in learning to read words. *Psychological Science*, *10*, 334–338.

Tunmer, W., & Chapman, J. (1998). Language prediction skill, phonological recoding ability, and beginning reading. In C. Hulme & R. Joshi (Eds.), *Reading and spelling: Development and disorders* (pp. 33–67). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Venezky, R. (1970). *The structure of English orthography.* The Hague, The Netherlands: Mouton.

Venezky, R. (1999). *The American way of spelling*. New York: Guilford.

Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development: Word recognition in English and German children. *Cognition*, *51*, 91–103.

Wimmer, H., & Hummer, P. (1990). How German-speaking first graders read and spell: Doubts on the importance of the logographic stage. *Applied Psycholinguistics*, *11*, 349–368.

Wright, D., & Ehri, L. (2005). *The case of doubled letters: Do beginners remember orthography when they learn to read words?* Manuscript submitted for publication.

Manuscript received June 2004
Accepted September 2004