8. THE PSYCHOLINGUISTICS OF BASIC LITERACY

Charles A. Perfetti, Julie Van Dyke, and Lesley Hart

We review major issues in research on reading, including theories of word reading, cross-writing system comparisons, comprehension, reading difficulties, learning how to read, and cognitive neuroscience studies of reading. Each of these topics has psycholinguistic components that reflect the language foundations of reading. These foundations lie in two facts: (1) a writing system connects to a linguistic system at one or more levels, meaning that word reading is partially a psycholinguistic process; and (2): reading comprehension shares processes (e.g., parsing) with general language comprehension. One trend of recent research is the development of models of word identification that rely on single rather than dual mechanisms and their extension to explain reading difficulties. Another is the conclusion that phonology plays a role in reading that cuts across writing systems. Reading comprehension research continues to reflect two different traditions, sentence parsing and text comprehension. Both show increasing influence of general cognitive explanations, as opposed to strictly linguistic explanations, for comprehension phenomena. Studies of brain activation bring converging evidence on reading processes and provide neuroanatomical correlates of reading problems. In one area, the acquisition and teaching of reading, advances come from increasing consolidation and practical use of previous research gains.

Psycholinguistic processes are central to reading, which is the exclusive focus of this review. Our definition of reading is focused: the conversion of written forms into linguistic messages. In this definition, the study of reading is, in part, the study of language processes, including comprehension. What distinguishes reading most clearly from spoken language processes is the conversion process, or decoding. Beyond decoding, reading shares some linguistic and general cognitive processes with spoken language in the processes of comprehension. Given this perspective on the penetration of language processes
throughout reading, the following review sections examine skilled reading, reading difficulties, and the acquisition of literacy. In presenting a general framework for the review, we attempt to characterize what we see as important developments in the psycholinguistics of reading.

Language Processes in Skilled Reading

Reading uses both the language of the reader and the writing system that encodes that language. Specifically important are (a) the identification of words and (b) the engagement of language and general cognitive mechanisms that assemble these words into messages. Because the process begins with a visual input that is part of a conventionalized writing system, an analysis of reading includes the role of the writing system and the units it provides to map onto the language system.

Reading and Writing Systems

Because writing systems are the result of human invention and refinement, they vary in their deep design principles (which define the writing system itself), their implementation in a particular language (the orthography), and their appearance (the script). The significance of writing system analysis for literacy lies in the importance of comparative studies for questions of skilled word identification and reading acquisition. The conventional distinction among alphabetic, syllabic, and logographic systems (Gelb, 1952) groups English, Italian, Russian, and Korean as alphabetic writing systems, in which graphic units associate with phonemes. Arabic, Hebrew, and Persian represent modified alphabetic systems, in which vowels can be omitted. Japanese Kana is a syllabary system, in which graphic units correspond to spoken language syllables. The logographic system is, at least among current systems, uniquely represented by Chinese (along with adaptations of this system into the Japanese Kanji and Korean Hanja). In such a system, the graphic units associate primarily with whole words or lexical morphemes.

Although Chinese approximates this logographic concept, its partial representation of single syllable-single morpheme pronunciation suggests a classification as morpho-syllabic (DeFrancis, 1989). Thus in Chinese, the compound character 竹 (pronounced “yang2” and meaning ocean) has two components; the left component signifies that the character meaning has to do with water, while the right component is a character that when it stands alone is also pronounced as “yang2” (羊, meaning sheep.) This combination of meaning and pronunciation information inside a character is very far from perfect, but it is frequent enough to challenge the assumption that it is a pure logographic system. Indeed, pure logographic systems seem not to exist among current writing systems.
Written Word Identification

Visual word identification is the process most distinctive to reading. In an alphabetic writing system, perceptual processing of a letter string results in the activation of the grapheme units (individual and multiple letters) of words. In traditional models, these words are represented in the reader’s lexicon, a mental representation of word forms and meanings. Successful word reading is a match between the input letter string and a word representation. As part of this process, phonological units, including individual phonemes associated with individual letters, are also activated. Indeed, the research now points to the conclusion that it is the joint contributions of graphemic and phonological units that yield the identification of a word.

It is common to refer to the phonological contribution to this process as “phonological mediation.” In the framework of traditional representational models, there are two pathways, one from graphemic units to meaning directly, and one from graphemic units to phonological units, and then to meaning (the mediation pathway). In this Dual Route Model (Coltheart, 1978), the direct pathway can be used for words that become highly familiar and must be used to read so called “exception words” (e.g., café) for which an indirect phonological route would fail. And the phonological route must be used to read pseudowords (e.g., nufe) for which there is no lexical representation to access. These sorts of mediation and one vs. two routes are central points of contrast between traditional representational models and more recent alternative theoretical models.

These alternative models share the idea that words are not represented in a mental lexicon, but rather emerge from processing activity. One class of models assumes that words emerge from patterns of parallel and distributed activation (Plaut, McClelland, Seidenberg, & Patterson, 1996). In resonance models, word identification results from the stabilization of dynamic patterns that are continuously modified by interactions among inputs and various dynamic states resulting from prior experience (Van Orden & Goldinger, 1994). An interesting feature of this model is that patterns of graphic-phonological activation stabilize more rapidly than do patterns of graphic-semantic activation. In effect, a word form becomes identified primarily through the convergence of orthography and phonology. Meaning is slower to exert an influence on the identification process.

These contrasting approaches to cognitive architectures have been central in word reading research in recent years. The empirical results in this area have accumulated enough so that there is strong consensus about the facts of alphabetic word reading. Although there appear to be relevant data for deciding between representational and nonrepresentational models (e.g., Besner, Twilley, McCann, & Seergobin, 1990; Coltheart, Atkins, & Haller, 1993; Seidenberg & McClelland, 1990), both classes of models can account for many of the same results.
Models of word identification have not generally been sensitive to the linguistic structure of words. In both representational and nonrepresentational models, a letter string has usually been treated like any other stimulus. And its spoken language counterpart, to which the graphic stimulus must be connected, has been treated like any other association. One theory, however, specifies a functional internal linguistic structure during the word identification process. In this two-cycles model (Berent & Perfetti, 1995), vowels and consonant phonemes are assembled separately from a graphic input. Consonants are assembled more quickly than vowels. This separation of vowels and consonants can arise either from fundamental phonological considerations or from the fact that, in English, the grapheme-phoneme mapping is more reliable for consonants than for vowels.

The case of Chinese provides maximal contrast with alphabetic reading. However, recent research has overturned the idea that reading Chinese avoids phonology by mapping characters to meaning. For example, when Chinese readers make judgments about whether two characters have the same meaning, interference results when the two characters happen to have the same pronunciation (Perfetti & Zhang, 1995). This remarkable role of phonology where the writing system does not require it suggests that reading includes highly general processes that must refer to spoken language representations. A strong form of this generalization is the Universal Phonological Principle (Perfetti, Zhang, & Berent, 1992), which is that reading in all writing systems engages those phonological units provided by the writing system at the earliest opportunity. The Constituency Model (Perfetti & Tan, 1999) incorporates the basic idea that Chinese characters, like English words, are interconnected networks of graphic, phonological, and semantic units with immediate activation of all components from visual input.

**Skill Differences in Word Identification**

Individual differences in word reading skill are wide ranging, observable among select groups of college students as well as children. Individuals of extremely low skill are said to have a specific reading disability, or dyslexia, provided they also show a discrepancy between their achievement in reading and their achievement in other domains. However, there are reasons to blur the distinction between specific and nonspecific reading problems for some purposes. The processes that go wrong in a specific disability may not be much different from those that go wrong for an individual who also has a problem in some other area (Stanovich & Siegel, 1994). For both groups, their problems must be understood in terms of the processes of reading that can go wrong.

Readers of low skill show problems in reading words and pseudowords. Theoretical models of reading processes provide distinctive explanations for these problems. Dual Route Models allow two different sources of word reading difficulties: Either the direct (print-to-meaning) route or the indirect (print-to-phonology-to-meaning) route can be impaired (Coltheart et al. 1993). These models provide an explanation for both developmental and acquired dyslexia.
acquired dyslexia, surface dyslexics have selective damage to the direct route, whereas phonological dyslexics have selective damage to the phonological route. In developmental dyslexia, children may have a phonological deficit or an "orthographic" (direct route) deficit. In contrast, the newer single mechanism models with learning procedures give an alternative account of developmental dyslexia: Only phonological dyslexia is the result of a processing defect. Surface or orthographic dyslexia becomes a delay in the acquisition of word-specific knowledge (Harm & Seidenberg, 1999), which comes through experience. The successful application of both single-route and dual route models to reading problems illustrates that understanding reading problems depends on the theoretical understanding of reading processes.

Reading Comprehension

Whether reading uses just those processes that serve spoken language or requires something more turns out to be a difficult question. Clearly, reading comprehension depends on spoken language comprehension. For adults, correlations between spoken language and reading comprehension are in the range of $r = .90$ for adult samples (Gernsbacher, 1990), leaving little unique variance to reading. For children, the correlations are lower and increase during schooling (Curtis, 1980; Sticht & James, 1984). The interpretation of these correlations is that reading comprehension skill approaches listening comprehension skill as printed word identification is mastered. However, such correlations are based on the use of formally equivalent texts across listening and reading. Although this equivalence is a methodological necessity, it masks the possibility that the correlations include a text component; the texts are literate texts, characteristic of printed language rather than spoken language. However, what we read constitutes a sample of language different along formal, semantic, and pragmatic dimensions from what we speak. It is even possible that with literacy come some significant changes in the way people process spoken language (Olson, 1977), a development that would also boost the correlation of listening and reading comprehension. The conclusion thus has to be carefully cast: Reading comprehension strongly depends on listening comprehension, but whether that is all there is to the relationship remains an open question.

Components of Comprehension

Reading comprehension entails processes beyond visual word identification: the selection of contextually appropriate word meaning, the parsing of sentences, and the construction of an integrated understanding of the text across sentences.

Parsing. Most of the research on parsing has been carried out in reading experiments, giving an ironic twist to the assumption among linguists that spoken language is primary, with written language secondary. This makes parsing a more appropriate topic for literacy than might be supposed.
The main question has been how the language processor determines the linguistic constituent structure of an input string. For example, a sentence that begins with *When the boys strike the horse* encourages the comprehender to construct a transitive representation, with *the horse* as the object of the verb *strike*. However, when the sentence continues with the verb *kicks*, the comprehender now must restructure the parse so that *the horse* is the subject of the main clause verb (*kicks*) rather than the subordinate clause object of *strike*. The main question for psycholinguistic research has been what factors influence the original decision. A secondary question has been how the correct restructuring is achieved.

A central fact to be explained is the processing difficulty that occurs at a critical word that signals a parsing error (*kicks* in the above example). Increased error rates (Ferreira & Henderson, 1991), slower reading times (Frazier & Rayner, 1982; Sturt, Pickering, & Crocker, 1999), longer eye fixations (Frazier & Rayner, 1982), and increased amplitudes in wave-forms registered in ERPs (Event Related Potentials) (Friederici & Meckinger, 1996; Osterhout, Holcomb, & Swinney, 1994) are associated with the appearance of this critical word. Most of the explanations center on the parser’s earlier decision on the phrase *strike the horse*, the source of the parsing error.

Frazier and colleagues proposed two highly general structural principles to control initial parsing decisions, Minimal Attachment and Late Closure (Frazier 1987; Frazier & Fodor, 1978; Frazier & Rayner, 1982). On Minimal Attachment, the parser prefers attachments that require fewer nonterminal nodes; Late Closure keeps a constituent open as long as possible, closing it only if the current item cannot be attached to it. In most cases, these two principles create a preference for attaching items low in the parse tree, leading to a preference to attach the NP *the horse* to the open VP constituent (*V: strike / NP: __*), and making it difficult to raise the NP from the lower object position to the subject position of the following clause. This is because the parser applies its structure-building principles without regard for the lexical properties of the items in the word string. Mitchell (1994) provides a still valuable review of the research on these attachment issues.

In recent years, alternative proposals have emphasized a broader range of influences on initial parsing, sacrificing the simplicity of a small number of principles for multiple constraint satisfaction, focusing especially on lexical information (MacDonald, Pearlmuter, & Seidenberg, 1994; Trueswell & Tannenhaus, 1994). For example, the processing difficulty in the above example (*When the boys strike the horse kicks*) comes from a frequency-based preference for the transitive reading of *strike* over its intransitive reading. Because the problem sentence requires the less frequent intransitive reading, a parsing error occurs. Although these models include structural information, they emphasize *lexical* over *structural* operations in parsing.

A recent development is the differentiation of constituent structures in terms of parsing influences. For example, restructuring constituents that contain
arguments, such as the example above (strike-horse; kicks-horse) is highly susceptible to lexical variables, such as subcategorization preferences, frequency of competing structures, and plausibility of arguments (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Sturt et al., 1999; Trueswell & Tannenhaus, 1994). In contrast, restructuring of nonobligatory constituents such as modifying adjuncts appears to be less affected by lexical factors and more influenced by contextual factors (Britt, Perfetti, Garrod, & Rayner, 1992). This has led Frazier and Clifton (1996) to propose a bifurcated parsing mechanism with distinct attachment mechanisms for arguments and adjuncts. The effect of such a proposal is to make processing principles less global—more connected to both structural specifications and cognitive mechanisms.

This development may signal a move in psycholinguistic research towards a cognitive resource-oriented approach, within which the role of memory becomes critical. Recent proposals have suggested that restructuring, as well as initial parsing preferences, may be limited by the availability of items required during restructuring (Gibson & Thomas, 1999; Lewis, 1998). Thus, preferences for low attachments are also preferences for recent (more available) attachment sites, and lexical preferences reflect "resting activation" levels that are determined by relative frequencies. Even the argument/nonargument distinction can be recast in terms of resource limitations, insofar as argument reanalyses require the retrieval of decayed alternatives, whereas nonargument adjuncts have no ready-made alternatives to retrieve. This approach for understanding the nature of language comprehension appears to bring psycholinguistics into closer contact with research on memory interference (e.g., Fernandes & Moscovitch, 2000; Lewis, 1996) and lexical retrieval (e.g., MacDonald, Perlmutter, & Seidenberg, 1994; Twilley & Dixon, 2000). Finally, another recent development is research on the role of prosodic and intonational information in parsing (Bader, 1998; Kjelgaard & Speer, 1999; Schafer, Speer, Warren, & White, 2000), which can help test the assumption that results from reading generalize to spoken language.

**Text comprehension.** Research on how readers understand texts beyond single sentences has little concern with syntax, assuming that the sentences to be understood and integrated into a coherent representation have already been parsed. What counts here is how the reader comes to construct mental models of the text and the situations described in the text.

Two classes of mental models are needed, a model of what the text says (the text base) and a model of what the text is about (the situation model; Kintsch, 1998; van Dijk & Kintsch, 1983). The text base is a mental representation of the propositions of the text, as extracted from the reading of successive sentences, supplemented only by inferences necessary to make the text coherent. The reader builds a situation model from the text base by combining knowledge sources through additional inference processes. Thus, a text base is essentially linguistic, consisting of propositions derived from sentences, whereas a situation model is essentially agnostic in its form of representation.
Some research has sought to demonstrate that situational models are, however, fundamentally nonlinguistic. One approach has been to study texts that encourage spatial representations. A variety of experimental tasks has been employed to demonstrate that the spatial information readers represent cannot be reconstructed solely on the basis of linguistic propositions. In particular, these representations preserve unstated but inferrable spatial relations, more readily conceived as spatial analogs rather than propositions (Glenberg, Krueley, & Langston, 1994; Haenggi, Kintsch, & Gernsbacher, 1995).

More recently, there has been a growing interest in narrative situations, where time, rather than space may organize the reader's model. For example, Zwaan (1996) has demonstrated that readers use time phrases in a narrative—phrases such as an hour later or a moment later—to build temporal models. Zwaan, Langston, and Graesser (1995) argue that readers construct representations of stories along five dimensions—time, space, protagonist, causality, and intentionality. Their event-indexing model assumes that events and the intentional actions of characters are the focal points of situation models, which are updated during reading along some or all of these dimensions. Another trend has been the study of multiple texts, which, when they deal with a single shared situation, more strongly force a distinction between texts and situations (Perfetti, Rouet, & Britt, 1999).

Inferences. Because texts are never fully explicit, text research has had an enduring interest in how readers make inferences that add to a text's explicit content. Indeed, the main difference between a text base and a situation model is assumed to be one of inferences, with text bases inferentially poor and situation models inferentially rich.

Inferences that are needed to maintain text coherence may be made routinely by linking anaphora with their antecedents. This is the main device seen in an influential theory of comprehension, the Construction-Integration Model (Kintsch, 1988; 1998). Readers seek to make what they are reading referentially coherent, so either a pronoun or a noun without a clear referent triggers a process to establish co-reference with something available from the text representation. These co-referential inferences are considered minimal—needed merely to make reading a text something different from reading a list of unrelated sentences.

More controversial is the occurrence of inferences beyond these minimal ones. The idea that many inferences are made during reading "on-line" (at the first moment they are allowed by the text) has been attractive for theories that emphasize knowledge use over linguistic input. However, the evidence on how likely readers are to go beyond the minimum is mixed. Although readers may make elaborative inferences under the right conditions (O'Brien, Shank, Myers, & Rayner, 1988), the bulk of the evidence suggests they do not do so routinely (Corbett & Doshier, 1978; McKoon & Ratcliff, 1992). Inferences that are intrinsic to the causal structure of the text may be more likely to be made than other kinds
of elaborative inferences (Trabasso & Suh, 1993). The major theoretical issue in inferences can be captured by the contrast between the minimalist hypothesis (McKoon & Ratcliff, 1992) and the constructionist hypothesis (Graesser, Singer, & Trabasso, 1994), which allows a variety of inferences—those that address the reader’s comprehension goals, those that explain why things occur, and those that establish global as well as local coherence.

We can summarize the state of affairs this way: Readers make some inferences in the service of coherence, perhaps those that aid the coherence of the narrative structure (a situation model for a story) as well as that of the text. Other inferences, e.g., predictive inferences that anticipate some event or some consequence of an action not required by considerations of either text or causal coherence, are less probable in many circumstances. They require effort and are subject to error. Many readers can become more active readers, making more inferences and other elaborations, under the right circumstances. But under the mundane demands of typical experiments, there is little motivation to become an inference generator.

Skill Differences in Reading Comprehension

Any component of comprehension is a candidate for “breakdown” that causes low reading comprehension achievement—parsing, sentence integration, inference making, and other components may be involved. However, because the higher levels of processing rely on output from lower levels, an observed problem in text comprehension can also result from lower level processes, including word identification.

Structure building problems. One important set of ideas on individual differences comes from the Structure Building Framework (Gernsbacher, 1990), which frames comprehension skill around the assumption that readers, in constructing a coherent framework for a text, must activate and enhance relevant concepts while suppressing irrelevant concepts. The suppression hypothesis is that less skilled readers have deficient suppression mechanisms. To illustrate, in the sentence, He dug with the spade, the final word has two meanings, but only one fits the context of the sentence. However, when adult readers are immediately asked to decide whether a following word is related to the meaning of the sentence, their decisions are initially slow for the word ace (related to the inappropriate meaning of spade). Both appropriate and inappropriate meanings may be activated at first. With more time before the appearance of ace, skilled readers show no delay in rejecting it; i.e., they “suppress” the irrelevant meaning; however, less skilled readers continue to react slowly to ace, as if they have not completely suppressed the irrelevant meaning of spade. A failure to use context is not quite what is involved here. Research on children’s word identification found that less skilled readers use context in word identification at least as much and perhaps more than do skilled readers (Perfetti, 1985; Stanovich 1980). Gernsbacher’s research with adults shows a similar result.
A new direction in this problem is the attempt to account for comprehension skill differences in terms of differences in lexical representations. The Lexical Quality Hypothesis (Perfetti & Hart, in press) argues that variations in the quality of lexical representations—defined in terms of interlocking knowledge of phonology, orthography, and meaning for specific words—lead to variations in observed comprehension. In this account, suppression failure is not an independent contributor to comprehension problems.

**Skill differences in syntactic processes.** Less skilled readers often show a wide range of problems in syntax and morphology. The question is whether such problems arise from a deficit in processing syntax or from some other source, especially working memory or lexical processing limitations, that affects performance on syntactic tasks.

To illustrate one class of syntactic problems, consider two sentences with relative clauses below. (1) is the easier subject relative; (2) is the more difficult object relative.

(1) The girl that the boy believed understood the problem.
(2) The girl that believed the boy understood the problem.

The greater difficulty of (2) compared with (1) can arise from different degrees of interference they produce in the attempt to assigning a subject for “understood,” which is greater in (2) where both “boy” and “girl” might be considered. Research with both children (Crain & Shankweiler, 1988) and adults (Carpenter, Miyake, & Just, 1994) suggests that syntactic problems of this kind can arise from processing limitations rather than structural deficits. King and Just (1991) found that readers with low working memory spans have problems with object-relative sentences such as (2). Moreover, these problems were most severe where the processing load was hypothesized to be the greatest—at the second verb in the object relative, i.e., “understood” in (2). Comprehension difficulties may be localized at points of high processing demands—whether from syntax or something else. If this analysis is correct, then the problem is not intrinsic deficits in syntax, but the processing capacity to handle complexity.

Another perspective on this issue is the opportunity for practice. Because some syntactic structures are more typical of written language than spoken language, the opportunity for practice is limited by the ability to read. Thus, continuing development of reading skill as a result of initial success at reading—and the parallel increasing failure as a result of initial failure—is undoubtedly a major contributor to individual differences in reading.

**Higher level skill differences in text comprehension.** The overall complexity of text comprehension implies several possibilities for processing failure, beyond those considered above. Problems in any component can result in a less integrated or less coherent representation of the text.
Problems in inference making have been the target of considerable research. Oakhill and Garnham (1988) summarize evidence suggesting that less skilled readers fail to make a range of inferences in comprehension. When skill differences in inferences are observed, their cause—occurring in the absence of lexical, working memory, or general language processes—is seldom clearly demonstrated. However, there has been some success in identifying a relatively small percentage of children whose problems can be considered at least comprehension-specific, although highly general across reading and spoken language (Stothard & Hulme, 1996).

Another example is comprehension monitoring, the reader’s implicit attempts to assure a consistent and meaningful understanding. Skilled readers can use the detection of a comprehension breakdown (e.g., an apparent inconsistency) as a signal for re-reading and repair. Less skilled readers may fail to engage this monitoring process (Baker, 1984; Garner, 1980). However, such differences may not be independent of the reader’s ability to construct a simple understanding of the text (Otero & Kintsch, 1992). The general interpretive problem here is that comprehension monitoring, like inference making, both contributes to and results from the reader’s text representation. This makes it difficult to attribute comprehension problems uniquely to failures to monitor comprehension, as opposed to more basic comprehension failures.

For comprehension to succeed, readers must import knowledge from outside the text. Thus, a powerful source of comprehension skill differences is the reader’s access to knowledge needed for a given text (Anderson, Reynolds, Shallert, & Goetz, 1977). However, readers of high skill can compensate for lack of knowledge to some extent (Adams, Bell, & Perfetti, 1995). It is the reader who lacks both knowledge and reading skill who is assured failure. The deleterious effect of low reading skill (and its motivational consequences) on learning through reading creates readers who lack knowledge of all sorts.

**Literacy Acquisition**

In learning an alphabetic writing system, a child comes to learn that letters and strings of letters correspond to speech segments. In principle, this learning might occur implicitly (through the extraction of print-speech correspondences in text) or explicitly (through direct instruction). How much success is to be expected from strictly implicit learning is a key question for literacy education.

Complicating the picture for English is the much lamented problem of inconsistent English orthography at the letter-phoneme level. Only a little of this inconsistency translates to gain in morphological transparency. The problem in English is not the contrast between *nation* and *national*, but that between *choir* and *chore* and *head* and *bead*. European languages, including Romance, Germanic, Slavonic, and Finno-Ugrian languages, tend to be coded by orthographies that consistently map graphemes to phonemes. In contrast, the child learning English
must learn that the letter string ow is read as /ow/ in the words how and cow and as /o/ in low and tow. An important recent development in the study of orthographies has been attention to mappings of larger sub-syllabic units. The rime unit, the vowel plus the syllable ending together, turns out to be very reliably pronounced in English (Treiman, Mullenax, Bijeljac-Babic, & Richmond-Welty, 1995). Consistent with this discovery is research suggesting that learning by rime-based analogy can be effective for English speaking children at the beginning (Goswami, 1993).

**Phonological Sensitivity**

A major link to discovering the alphabetic principle is the child’s phonemic awareness, the understanding (more-or-less explicit) that the speech stream can be segmented into a set of meaningless units (phonemes). Phonemic awareness shows a strong correlation with early reading success, and training studies with explicit instruction in phoneme segmentation produce gains in reading. The causal relationship between reading and phonemic awareness appears to be reciprocal, however. First, adults whose cultural context does not include formal schooling may fail to show phonemic awareness (Morais, Cary, Alegría, & Bertelson, 1979). Second, Chinese children who learned to read Chinese without first learning the Chinese alphabetic (pin-yin) system fail to show phonemic awareness (Read, Zhang, Nie, & Ding, 1986). Third, longitudinal studies of first grade children learning to read English find a reciprocal relationship, such that initial gains in simple phonemic awareness precede gains in reading, which in turn lead to further gains in phonemic awareness (Perfetti, Beck, Bell, & Hughes, 1987). For a summary of the evidence on the role of phonemic awareness and other factors in learning to read, see the National Research Council Report (1998). Among the most important recent developments for practice is the increased acceptance of phonological sensitivity as an important support for learning to read (see also Geva and Wang, this volume).

**Theories of Learning to Read**

Progress in acquiring reading skill typically has been viewed as a series of stages (e.g., Ehri, 1991; Frith, 1985; Gough & Hillinger, 1980). In such theories, the earliest stage of reading may be characterized by an attempt to learn associations between visual features of graphic forms (not complete orthographic word forms) and spoken words. A subsequent stage of graphic-phonological decoding brings on a truly productive capability in reading. The use of letter names as a bridge to phonology is an important beginning step (Ehri, 1991).

Alternative theoretical accounts emphasize the incremental acquisition of decodable lexical representations, rather than discrete stages (Perfetti, 1992; Share, 1995). Important in these theories is the role of phonology in helping to establish word-specific orthographic representations, a proposal that has come to be known as the “bootstrapping hypothesis” (Share, 1995).
Spelling

The initial expression of the alphabetic principle appears more often in spelling than in reading (Frith, 1985). Indeed, children's attempts at spelling prior to formal reading instruction typically reveal a real understanding that the sounds of a word are to be found in its letters (Chomsky, 1970; Read, 1971). This early realization can form the basis for later alphabetic reading. Eventually, in nearly all languages, the learner confronts an important fact about spelling: Typically, the mapping from pronunciation to spelling is less consistent than the mapping from spelling to pronunciation. Reading is more reliable than spelling.

An important recent idea in spelling comes from research on word identification. The more ways a sequence of phonemes can be spelled, the longer it takes to read a word that contains that sequence. For example, the word "shelf" is more efficiently read than the word "sneer" because its rime unit /elf/ is always spelled elf, whereas the rime unit /er/ is spelled variously as eer, ere, ier, and er. Notice that this is not a question of consistency in the direction of orthography to phonology: eer is always pronounced /er/. Stone, Vanhoy, and Van Orden (1997) reported the first demonstration of this backward consistency effect. An interesting theoretical implication is that reading words involves a feedback mechanism from phonology to orthography. It's not enough to convert a written input into a phonological representation, but a reader also verifies that the phonological representation can be spelled in the way presented. A more general implication is that spelling and reading are intimately related. The fact that spelling is typically more difficult than reading is not because these tasks use different representations, but because spelling requires additional processes that benefit from specific practice at spelling (Bosman & Van Orden, 1997).

There is a growing interest in theoretical aspects of spelling and in comparisons across languages and writing systems. For a collection of theoretical and empirical papers on spelling in different languages, see Perfetti, Rieben, and Fayol (1997).

Teaching Reading

The instructional goals for alphabetic systems would seem clear: Children need to learn that the letters of their alphabets map onto speech segments of their language. However, in English speaking countries especially, well known controversies abound over how to teach reading. Rather than focus on letter-sound correspondences, the dominant instructional approaches have emphasized instead meaning-focused instruction built around story reading, exposure to print, and enhanced language environments.

For over 30 years, the research has tended to support the effectiveness of methods that are based on direct instruction in decoding. A comprehensive ground breaking study by Chall (1967) was confirmed by later studies (see Adams, 1990)
and even more recent studies (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1999). However, the practice of reading instruction remained out of touch with the research, emphasizing a variety of language activities but excluding teaching of grapheme-phoneme relationships. Much of the debate was fueled by philosophical stances and professional advocacies that have little to do with the research basis for effective teaching. (See Stanovich, 2000, for a researcher’s perspective on this problem.)

However, recent indications are that a stronger consensus is emerging in favor of research-guided practice. The U.S. National Research Council (the research arm of the National Academy of Sciences) revisited this issue in its 1998 report, *Preventing Reading Difficulties in Young Children*. The NRC report reviewed available research and concluded that beginning reading “depends critically on mapping the letters and the spellings of words onto the sounds and speech units that they represent.” Directly counter to the idea that somehow comprehension can proceed on its own, the report adds that “Failure to master word recognition impedes text comprehension” (p. 321).

The Cognitive Neuroscience of Reading

Finally, we take brief notice of recent developments in the cognitive neuroscience of reading. Neuroscience methods have allowed both tests of existing theories and information not available with current behavioral techniques.

One method is Event-Related Potentials (ERPs), collections of brain electrical activity recorded from electrodes pasted on the scalp. ERP is very time-sensitive, detecting events at the millisecond scale, e.g., the access of meaning from a visual word presentation. It is not, however, sensitive to spatial information, providing only approximate brain localization information. For information on brain regions, Functional Magnetic Resonance Imaging (fMRI) is proving to be a useful method, joining more established imaging methods such as PET (position emission tomography) in providing spatial information to a centimeter or even millimeter scale, but with little time sensitivity. Between them, ERP and fMRI can provide fine grained temporal and spatial information about brain activity that accompanies reading.

In skilled readers, studies of reading find lateralized function, with greater activity in the left hemisphere for frontal, temporal, parietal, and occipital sites (Crosdon et al., 1999; Simos, Basile, & Papanicolaou, 1997; Small, Noll, Perfetti, & Hlustik, 1996; Wilding & Rugg, 1997). Precise fMRI locations (Crosdon, et al., 1999; Pugh, Shaywitz, Shaywitz, & Shankweiler, 1997) and ERP latencies (Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999; Martin-Loeches, Hinojosa, Gomez-Jarabo, & Rubia, 1999) may partially distinguish orthographic, phonological, and semantic processes. Imaging studies may also identify distinct verbal working memory locations (Crosdon et al., 1999) and networks of regions that support oral reading (Small, Noll, Perfetti, & Hlustik, 1996). Such studies also
suggest that reading Chinese involves the same brain circuitry as found in English (Tan et al., 2000), converging with the picture emerging from behavioral studies.

In studies of differences in reading skill, imaging studies have found higher skill associated with lateralization to the left hemisphere (Pugh, Shawitz, Shawitz, & Shankweiler, 1997; Segalowitz, Wagner, & Menn, 1992). In ERP experiments, skilled readers show shorter latencies of an early ERP component (P200) that indicates perception of a recognizable word (Rudell & Hua, 1997). Dyslexic subjects can be differentiated from nondyslexic subjects on the basis of brain activation, but the differential pattern that distinguishes dyslexics from nondyslexics appears not to be the same one that distinguishes nondyslexics of varying skill from each other. Dyslexia, whether developmental or acquired, is associated with brain circuitry responsible for attention (Pugh et al., 1997) and phonological processing (Georgiewa et al., 1999; Small, Flores, & Noll, 1998). Segalowitz et al. (1992) suggest that different predictors are required depending on whether the reader’s ability is below or above some threshold. In addition, some structural differences have been found to distinguish dyslexics from controls (Semrud-Clikeman, Hooper, Hynd, & Hern, 1996).

One of the most promising uses of neuroscience methods concerns the effects of training on brain activation and circuitry. Not only does repeated practice increase brain efficiency (Rudell & Hua, 1997), but it can actually influence brain development and neuronal connectivity. A recent fMRI study found a thicker band of callosal connective fibers between parietal lobes for literate than for illiterate subjects (Castro-Caldas et al., 1999). Another study designed to shift an acquired dyslexic subject from a whole word reading strategy to a phonological reading strategy found that the activation patterns change from pre- to post-therapy (Small et al., 1998). Both papers describe the shift as an alteration in brain circuitry. Neuronal connectivity appears plastic into adulthood, and intervention can actually change the way the brain is structured and wired for reading.

Cognitive neuroscience methods promise to be useful across many issues in the study of reading—the components of word identification, cross-language comparisons, the acquisition of skill, and even comprehension. There remains much to learn about the specific brain regions that support specific processes. However, rather than the mapping of single regions to single processes, the results demonstrate a complex network of connections broadly distributed across brain regions in support of reading and language that emerge in response to task requirements and are subject to reorganization during learning.

ANNOTATED BIBLIOGRAPHY

This volume includes a collection of chapters on psycholinguistic topics, including speech perception, parsing, and reading. Some chapters present cognitive analysis of language and research reviews, while others review neuroimaging and ERP methods of language study. A chapter by Perfetti (Comprehending written language: A blueprint of the reader, pp. 167–208) contains a review of the research in reading.


This book presents a detailed summary of evidence, including cross-linguistic data, motivating parsing principles discussed here, including experimental data supporting the differential treatment of argument and adjuncts during grammatical restructuring.


This article supplies a recent example of single-mechanism approaches to word reading. A connectionist learning model is applied to traditional classes of dyslexia, demonstrating how a "lesioned" language system prior to reading leads to phonological dyslexia.


An integration of 20 years of research on knowledge representation and language comprehension, this book presents an excellent, in-depth treatment of data on comprehension at the word-level, sentence-level, and text-level. The Construction-Integration Model is described, together with numerous examples of its application as a computational model of phenomena at all of these levels.


This report reviews a wealth of research on how children learn to read, with broad coverage of pre-school language factors, and school and social factors, as well as cognitive and psycholinguistic research. It includes practical research-based recommendations.
UNANNOTATED BIBLIOGRAPHY


Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia:*)


