28 Linguistics and Reading

REBECCA TREIMAN

Linguists are primarily concerned with the structure and processing of spoken language. In this chapter, the focus changes to written language. The goal of the chapter is to review what is known about the processes involved in reading and in learning to read. Topics to be discussed include the controversies about the best way to teach children to read and the reasons why some apparently normal children have great difficulty mastering this skill. Another question is whether knowledge of written language changes people's basic intellectual or linguistic abilities. Researchers from a variety of disciplines, including cognitive psychology, developmental psychology, and education, have been active in research on reading. This mix reflects the fact that the study of reading is both theoretically interesting and practically important. Reading is a domain in which experimental psychologists study fundamental questions such as how knowledge and experience affect perception. Reading is also a domain in which research findings have implications for important social issues, such as the education of children. It is no wonder, then, that a large amount of research has been carried out on reading. The discussion of this research begins with a consideration of the cognitive processes that are involved in skilled reading.

1 Bottom-Up and Top-Down Processing in Reading

In the case of reading, as with other cognitive processes, psychologists have distinguished between two kinds of processing. *Bottom-up processes* are those that take in stimuli from the outside world – letters and words, for reading – and deal with that information with little recourse to higher-level knowledge. With *top-down* processes, on the other hand, the uptake of information is guided by an individual's prior knowledge and expectations. In most situations,

bottom-up and top-down processes work together to ensure the accurate and rapid processing of information. However, theories about the cognitive processes involved in reading differ in the emphasis that they place on the two approaches. Theories that stress bottom-up processing focus on how readers extract information from the printed page, claiming that readers deal with letters and words in a relatively complete and systematic fashion (e.g., Gough 1972). Theories that stress top-down processing hold that readers form hypotheses about which words they will encounter and take in only just enough visual information to test their hypotheses (e.g., Goodman 1967, Smith 1971). In the words of Goodman, reading is a "psycholinguistic guessing game."

An example may help to clarify the distinction between theories that stress bottom-up processing and those that stress top-down processing. Suppose that a reader has just read, "Daylight savings time ends tomorrow, and so people should remember to change their . . ." According to the top-down view, the reader guesses that the next word in the sentence will be "clocks." The reader checks that the word begins with a "c" and, because the hypothesis has been supported, does not take in the remaining letters of the word. Theories of reading that stress bottom-up processing claim that the reader processes all of the letters in the last word of the sentence, regardless of its predictability.

Studies of readers' eye movements provide some insight into the roles of bottom-up and top-down processes in reading. Research has shown that the eye does not sweep across a line of text in a continuous fashion. Rather, the eye comes to rest for somewhere around a quarter of a second, in what is called a *fixation*, and then makes a rapid jump (a *saccade*) to the next fixation. It is during the fixation that visual stimulation is taken in; little or no useful information is extracted during a saccade. Researchers have found that skilled readers fixate at least once on the majority of words in a text. They do not skip a large number of words, as the top-down view predicts, but instead process the letters and words rather thoroughly. Readers do this, in part, because their span of useful vision is fairly small. For example, a reader who fixates on the "a" of "daylight" will be able to see all of the letters in this word. The reader may or may not be able to see enough to identify the next word, "savings," but will be unable to identify "time." Thus, the eye movement data portray reading as more of a bottom-up process than a top-down process. (See Rayner and Pollatsek 1989 for a review of the research.)

Comparisons of good and poor readers further support the claim that bottom-up processes play an important role in reading. If reading were a linguistically guided guessing game, as top-down theorists maintain, one would expect guessing ability to discriminate between good and poor readers. In this view, good readers are highly sensitive to context and use it to guide their uptake of print, whereas poor readers have trouble predicting the upcoming words in a sentence. However, research has shown that poor and unskilled readers use context at least as much as good readers (e.g., Perfetti et al. 1979). Skilled readers' perceptually based recognition skills are so accurate and automatic that they do not usually need to guess.

The statement that bottom-up processes play a central role in reading does not necessarily mean that top-down processes are completely unimportant. Studies have shown that words that are predictable from context are fixated for shorter periods of time and are skipped more often than words that are less predictable, although the effects are relatively modest (see Rayner and Pollatsek 1989). One interpretation of these results is that readers sometimes use their higher-order thinking skills to predict the upcoming words in a sentence. However, the results may alternatively reflect low-level associative processes within the *lexicon* (mental dictionary) itself. For example, readers may spend less time on "cake" in the sentence "The guests ate the wedding cake" than in the sentence "The guests ate the large cake" because the activation of "wedding" automatically sends some activation to "cake." Whatever the mechanism responsible for context effects, we must keep in mind that most words are not predictable or only minimally predictable from context. After "the," for example, almost any adjective or noun could occur. Therefore, bottom-up processing is often essential for reading.

2 Word Recognition

Many of the processes that are involved in understanding what we read are similar to the processes involved in comprehension of spoken language. In both cases, we must often use our knowledge of the world to make sense of and elaborate on the information. When reading about a wedding, for example, it is helpful to know about the kinds of activities that usually take place on such occasions. The grammatical knowledge that is necessary to understand a sentence is similar, too, whether the words are read or heard. What distinguishes reading from speech is the need to identify words by eye. Readers must recognize printed words accurately and automatically, linking them to representations stored in the mental lexicon. This process of *word recognition* has been a central focus of reading research.

To understand the processes that are involved in the recognition of words, one needs to consider the way in which printed words map onto speech. Although writing systems differ from one another in many ways, all full writing systems are based on speech (DeFrancis 1989, see also chapter 3, Writing Systems). For example, each syllable (roughly speaking) in spoken Japanese has its own symbol in the writing system called kana and so this system maps onto speech at the level of syllables. In alphabetic languages, in contrast, the link between print and speech is at the level of individual sounds or phonemes. Some alphabetic writing systems, such as Italian and Finnish, exemplify the alphabetic principle almost perfectly, with each letter representing one and only one phoneme. English is not a pure alphabetic writing system, which has led to widespread criticism of the system and many calls for spelling reform. Some English sounds have more than one possible spelling, as

when /k/ is alternatively spelled as "c" ("cat"), "k" ("kit"), or "ck" ("pack"). Moreover, some letters have more than one possible pronunciation. For example, "c" can correspond to /k/ as in "cat" or /s/ as in "city." Although such complications make the English writing system more complex than some other writing systems, they do not negate the usefulness of the alphabetic principle. "Gove" could be pronounced to rhyme with "cove" or "love," for example, but skilled readers would never pronounce it as "mab." Certain deviations from the alphabetic principle are themselves principled, reflecting the tendency of English to spell morphemes (units of meaning) in a consistent fashion. For example, the past tense ending is variously pronounced as /t/ (as in "jumped"), /d/ (as in "hemmed") or /əd/ (as in "wanted"), but it is generally spelled as "ed." As another example, the "a" in "health," which makes the word an exception from an alphabetic standpoint, reveals the relationship in meaning to "heal."

Just as the printed forms of words reflect their linguistic forms, so the processing of printed words involves the recovery of the words' linguistic forms. In many cases, readers access the phonological (or sound) forms of words as part of the recognition process. This phonological activation is covert, for skilled readers who are reading silently, but psychologists have devised clever ways to detect it. In one technique, people are presented with a category name (e.g., "type of food") and must then rapidly decide whether various printed words belong to the category. College students sometimes misclassify words that sound like category members (e.g., "meet") as members of the category, even when they know the words' correct spellings. Participants make fewer errors on words that look equally like a member of the category but that do not sound like one (e.g., "melt") (Van Orden 1987). (See Frost 1998 for a review.)

There is some debate about exactly how readers derive the phonological forms of words from their spellings. Do skilled readers use explicit rules of the kind taught in phonics lessons ("b" corresponds to /b/, "m" to /m/, and so on), or do they rely on a network of implicit connections? Are the links between spellings and sounds based on individual graphemes, or letters and letter groups that correspond to single sounds (e.g., "b," "sh")? Alternatively, do readers sometimes rely on larger units, linking units such as "ead" and "ine" to their pronunciations? These units have been called *orthographic rimes*; they correspond to the *phonological rimes* (vowel + final consonant units) of spoken syllables (Bernstein and Treiman, in press). To investigate questions such as those described above, researchers are devising explicit models of the spelling-to-sound translation process and are testing the predictions of such models (Coltheart et al. 1993, Plaut et al. 1996, Seidenberg and McClelland 1989). These tests are no longer restricted to small-scale experiments but often involve assessing readers' performance on large samples of words (Spieler and Balota 1997, Treiman et al. 1995). Although areas of disagreement remain, it is widely believed that rapid, automatic word recognition is critical to reading success and that such recognition often involves activation of words' spoken forms.

3 Learning to Read

Much of the research on learning to read has focussed on the acquisition of alphabetic writing systems, especially English. In the United States and other English-speaking countries, there has been a debate between advocates of two different approaches to learning to read. (See Adams and Bruck 1995, Bergeron 1990, Liberman and Liberman 1992 for discussion.) The first of these, the *whole language* approach, is based on the idea that top-down processing plays an important role in reading. If fluent readers use context to predict the upcoming words in a sentence, only processing the print to the degree necessary to confirm their expectations, then children should do the same. Children should focus on the meaning of what they read rather than laboriously sounding out the individual words. Just as children will master spoken language if they are spoken to by others and given the opportunity to respond, so children will become literate if they are placed in an environment that is rich in print and are encouraged to explore it. Whole language teachers thus focus on the meaning and purpose of printed language rather than on individual letters and sounds. Activities may include reading stories to children and helping children use the pictures or the context to figure out the words. Sounding out an unknown word is typically considered a strategy of last resort, and children are given little guidance on how to do this. Whole language teachers also encourage the integration of reading and writing, expecting children to write independently from an early age and offering little or no instruction in conventional spelling.

The second class of approaches to literacy instruction, known as *phonics*, places primary stress on the bottom-up processing of letters and words. In this view, learning to read is quite different from learning to talk. Spoken language is deeply rooted in biological evolution and is as old as the human species itself. All normal members of the species learn to speak and understand without explicit tuition, provided only that they are exposed to a spoken language. However, the situation is quite different for written language. Writing is a cultural achievement dating back no more than 5,000 years; it is found among some groups of people but not others. Learning to read, phonics advocates claim, usually requires explicit instruction. Children must learn to convert the unfamiliar printed words into their familiar spoken forms by learning that "b" is pronounced as /b/, that "c" can be pronounced as /k/ or /s/, and so on. This sounding out process is slow and laborious at first, but becomes fast and automatic with practice. The phonics approach thus focusses on individual letters and sounds, repetition, and practice. Content and interest are not the only criteria for choosing reading materials; the words must also be easy to decode. For example, a story about a *bug* that can *fish* would be preferred to a story about a *worm* that can *talk*, as *o* does not have its typical pronunciation in *worm* and *a* and *l* do not have their typical pronunciations in *talk*.

In practice, many programs include a combination of whole language and phonics methods. For example, children who are receiving phonics instruction can learn about the meaning and function of print by reading (or being read) interesting stories and using written language for meaningful purposes. As another example, writing can be emphasized in phonics classrooms as well as in whole language classrooms. The central question is whether early reading instruction should include instruction in phonics. The answer to this question, most researchers now agree, is "yes." Across a large number of studies, programs that include explicit, intensive attention to phonics generally yield better results than programs that do not. (See Adams 1990 for a review.)

Still, dissatisfaction with conventional phonics remains. Part of the reason is that some children have trouble grasping phonics instruction and leave first grade able to read only a few words. Researchers have thus begun to examine the factors that make it difficult for some children to benefit from phonics instruction. One contributing factor appears to be a lack of *phonemic awareness*. Children's attention is normally on the meaning of what they hear and say, not on the individual words and sounds. In order to understand how the spellings of words map onto their spoken forms, children must begin to pay attention to smaller units of sound. For example, a child who is not aware that "bat" contains three units of sound, the first of which is the same as the first sound of "boy," will not understand why the printed form of "bat" contains three letters, the first of which is the same as the first letter of "boy." A number of tasks have been developed to tap phonemic awareness, ranging from counting phonemes (how many sounds do you hear in "bat?") to comparing phonemes (do "bat" and "boy" start with the same sound?) to deleting phonemes (what would you get if you took away the /b/ from "bat?"). Children's performance on such tests is an excellent predictor of their later reading success (see Adams 1990 for a review).

To teach phonemic awareness, one can take advantage of the fact that awareness of phonemes is the end-point of a long developmental process. The process begins with awareness of words and syllables and progresses to units that are smaller than syllables but larger than phonemes, including initial consonant clusters (e.g., the "bl" of "blast") and rimes (e.g., "ast"). Programs that use a gradual approach to teach phonemic awareness have been successful both in fostering phonemic awareness and improving later reading performance (see Adams 1990). Phonemic awareness instruction is particularly successful when it is closely integrated with reading instruction, allowing children to use their newly gained phonemic awareness skills in relating print to speech (e.g., Blachman 1987).

Another reason why children may have trouble benefiting from phonics instruction is that, when they first begin to learn to read, they believe that the links between printed words and spoken words are arbitrary and non-analytic (Byrne 1998, Frith 1985). For example, children may think that it is the color and shape of the McDonald's logo, not the letters it contains, that allow it to say "McDonald's." Young children are thought to be *logographic* readers,

treating printed words as holistic symbols. Children must break away from the logographic hypothesis in order to learn that the parts of printed words (the graphemes) represent the parts of spoken words (the phonemes) in a systematic fashion.

Yet another stumbling block to conventional phonics instruction involves the teachers rather than the students. Many teachers have little or no opportunity to learn about linguistics and the structure of written language. As a result, they may not provide optimal instruction (Moats 1994). Because teachers are themselves good readers, they tend to think about language in terms of how it is spelled rather than how it is pronounced. They may find it hard to put themselves in the place of a child who does not yet know the conventional writing system. For example, a teacher may think that there is an /ɪ/ (or "short i") sound in "girl" because the spelling of this word contains an "i." However, the spoken word does not actually contain the same /ɪ/ vowel as "bit" and it would be misleading to suggest to a child that it does. As Moats (1994: 99) states, "lower level language mastery is as essential for the literacy teacher as anatomy is for the physician. It is our obligation to enable teachers to acquire it."

To summarize, reading instruction that includes explicit attention to phonics generally works better than instruction that does not. However, there is room for improvement in traditional phonics programs. Research suggests that improvement can occur by better preparing children to benefit from phonics instruction and by better preparing teachers to teach it.

4 Learning to Spell

One aspect of whole language programs that is attractive to many teachers and parents is the focus on writing. In many whole language classrooms, children write each day in personal journals. Correct spelling is not stressed, with children instead being encouraged to invent spellings for words they do not know. It is assumed that invented spellings like "bo" for "blow," "grl" for "girl," and "wet" for "went" will give way to conventional spellings as children learn to read, and that direct instruction in spelling is not necessary. However, research shows that children are less likely to learn words' spellings from the reading of meaningful, connected text than from the study of isolated words. Research further shows that the correlation between reading ability and spelling ability is far from perfect – that there are a number of people who are good readers but poor spellers. For most children, learning to spell requires something above and beyond immersion in a sea of print. The benefits of spelling instruction are not confined to spelling itself. Such instruction can also foster reading and phonemic awareness. For example, as children practice spelling consonant clusters like "bl" they learn to analyze these clusters into their component

phonemes. Spelling instruction, like reading instruction, requires a teacher who is knowledgeable about children's errors and the linguistic reasons behind them. For example, a teacher who is aware that the middle part of "girl" is a syllabic /r/ sound rather than /ɪ/ followed by a separate /r/ will understand why young children frequently misspell this word as "grl." (See Read 1986 and Treiman 1993 for a discussion of children's common spelling errors and Treiman 1998 for a review of research on early spelling instruction.)

5 Dyslexia

Even with good instruction, some apparently normal children have great difficulty learning to read and spell. Such children are known as *dyslexics*. The popular view is that these children see letters and words backwards. As a result, they may misread "was" as "saw" or "day" as "bay." Similar errors occur in spelling, as when children write "bit" as "dit" or even "tid" (in what is known as *mirror writing*). However, research reveals that such mistakes do not constitute the majority of reading or spelling errors among dyslexics. Moreover, normal children make the same kinds of errors when they are first learning to read and write. Most researchers now believe that, in the great majority of cases, dyslexia is more a linguistic problem than a visual problem (Olson 1994, Vellutino 1987).

If dyslexia is a linguistic problem, what kind of linguistic problem is it? The most widely accepted hypothesis is that dyslexia reflects a weakness in the phonological component of language (Stanovich 1992, Olson 1994, Vellutino 1987). Dyslexics have difficulty becoming aware of the phonemic structure of spoken language and thus have trouble learning about the way in which spellings map onto sounds. Dyslexics' phonological problems also extend to remembering words and to producing them quickly and accurately. These problems are, in part, genetically based. For example, if one member of a pair of identical twins exhibits reading problems then the other member has an elevated chance of showing similar problems.

If dyslexia indeed stems from linguistic weaknesses, particularly weaknesses in the area of phonology, then teaching must attempt to remediate these problems. Instruction that centers on visual perception, such as exercises designed to improve eye tracking or binocular coordination, does not appear to be successful (Vellutino 1987). What is needed, instead, is an intensive reading program that includes a liberal dose of phonics. A successful program of this kind is provided at the Benchmark School (Gaskins 1998). Here, reading disabled children spend over four hours a day in literacy activities. These activities are designed to help the children become aware of the sounds in spoken words and how these sounds are represented with letters, as well as helping children use this knowledge in reading and writing connected text.

6 The Effects of Literacy

Does learning to read change people's basic cognitive or linguistic abilities? Some have suggested that literate individuals and societies differ greatly from non-literate ones, the former being more abstract, more rational, and more skeptical. Although research has not supported these grand claims, it has provided empirical evidence that literacy has certain cognitive consequences (see Stanovich 1993). For example, United States college students who read extensively have larger vocabularies and more knowledge about the world than their peers who do little reading in their free time. In our society, opportunities to learn new words arise more often while reading than while conversing or watching television.

Learning to read also appears to deepen and alter people's knowledge about language. Phonological awareness, and metalinguistic awareness in general, develop hand in hand with learning to read and write. Thus, preliterate children and illiterate adults tend to do poorly in tasks requiring access to the phonemic structure of language, although they do better on rhyming tasks and syllable-level tasks (Morais et al. 1986, Morais et al. 1979). Another effect of literacy is to color people's knowledge about the sounds of language. For example, seeing that words like "went" and "elephant" contain an "n" in their spellings, children may come to conceptualize /n/ after a vowel as a separate unit rather than as part of the vowel, as they did previously (Treiman et al. 1995). If people's ideas about spoken language are indeed influenced by their knowledge of written language, it may be difficult to study the structure or processing of spoken language without considering the written language (Derwing 1992).

7 Conclusions

Linguists have often assumed that speech is the primary form of language and that writing is secondary. This view implies that investigations of language and language processing should focus on spoken language and that there is little to be gained from studies of written language. This chapter has presented evidence, to the contrary, that the study of written language processing is interesting and informative in its own right. There are many questions to be answered about how people relate print to speech and about how children can best be taught to do so. This is an area in the study of language with important real-world applications. Moreover, it appears that written language takes on a life of its own once acquired, influencing the representation and processing of spoken language. The study of writing and written language processing can no longer be ignored within linguistics.