A TUTOR'S ANALYSIS OF THE ROLE OF PREDICTION IN READING

A growing number of analyses of reading are paying particular attention to the use that a reader must make of prior knowledge relevant to the material he/she is endeavoring to read (e.g., Goodman, 1968, 1970; Hochberg, 1970; Kolers, 1970; Smith, 1971, 1973). In formal psychological jargon this use of prior knowledge is frequently referred to as hypothesis testing; teachers know it more familiarly as guessing; and I shall refer to it as prediction. I believe that reading is impossible without prediction, and since it is only through reading that the opportunity to develop and employ skills of prediction must be a critical part of learning to read.

It is not necessary, however, that prediction be taught, for prediction is as much a part of spoken language comprehension as it is of reading. Children with sufficient verbal ability to understand written material that is read to them have both the competence and the experience to direct their ability in prediction to reading. My aim is to demonstrate that prediction is essential for reading, that everyone who can comprehend spoken language is capable of prediction, and that prediction is routinely practiced in reading by beginners as well as fluent readers.

Four reasons for predictions

- Individual words have too many meanings. Words in our language tend to be multiply ambiguous, and the most common words have the most meanings (Fries, 1940). Everyday words like come, go, have take, table, chair not only have a multiplicity of different meanings, they are often also ambiguous as to their grammatical function. How is the word house
pronounced? The word cannot even be articulated until the reader knows whether it is a noun or a verb. The most common words in any language, the prepositions, have so many different meanings they take up more space in dictionaries, that words in any other class. It should be noted, however, that speakers and writers are almost never aware of this potential ambiguity, and that listeners and readers are rarely aware of the multiplicity of possible meanings either.

• The spellings of words do not indicate how they should be pronounced. There are over 300 “spelling-to-sound correspondence rules” of English (Venezky, 1967), and there is no rule that will specify when any of these particular rules must apply, or when the spelling to be “sounded out” is an exception. The rules of phonics are highly complex. To take a very simple example, how should a word beginning with HO....be pronounced? The answer depends on whether the ho is followed by ...t, ...ot, ...ok, ...rizon, ...use, ...rse, ...pe, ...ney, ...ist, ...ur, ...nest, eleven different possibilities (all depending on what follows the initial letters, indicating that phonics must be applied from right to left).

• There is a limit to how much of the “visual information” of print the brain can process during reading. Flash a line of about thirty random letters on a screen for about a tenth of a second and the most an experienced reader will be able to recognize is four or five letters. This four-letter or five-letter limit in fact represents an entire second that it takes the brain to decide for anything else to be seen; a condition that can be characterized as “tunnel vision.” In other words, for as long as one is trying to identify letters one after the other, reading is an impossibly slow and restricted process (Smith and Holmes, 1971).

• The capacity of short-term memory (or “working memory”) is limited (Atkinson and Shiffrin, 1970; Simon, 1974). Not more than six or seven
unrelated items—say an unfamiliar telephone number—can be held in short-term memory at any one time. Try to overload an already filled short-term memory and other information will be lost. As a consequence, it is virtually impossible to read a word more than four or five letters long a letter at a time. By the time the end is reached, the beginning will be forgotten. It is similarly impossible to store the first words of a sentence while waiting to get to its end before making a decision about meaning. By the time the end of the sentence is reached, the beginning will have been forgotten.

Defining prediction

There is a common feature underlying the four “reasons for prediction” that have just been listed. In each case the brain is confronted by too many possibilities; it must decide among more alternatives that it can handle. Decision-making takes time, and there is a fundamental rule that applies to every aspect of decision-making, whether it involves the identification of a single letter or word in a line of type, or the comprehension of a sentence or an entire book. The fundamental rule is this: the greater the number of alternatives, the more time is required for a decision (Garner, 1962). Recognition is never instantaneous. We may be able to identify a letter or a word if comes from a small set of known alternatives -- when we know in advance that it is a vowel, or the name of a flower—but the same letter or word will be quite unrecognizable if it comes from a larger set of alternatives. The reason for this bottleneck is simple: the greater the number of possible alternatives, the more information the brain has to process in order to reach a decision. The art of fluent reading lies in the skilled reduction of the amount of visual information the brain has to process. If you know a letter will be either A or B, you need only a glimpse of that letter to decide which it is. But if the letter could be any one of the 26 letters of the alphabet, much more visual information will have to be taken into account.
My general definition of prediction is the prior elimination of unlikely alternatives. In the jargon of Information Theory, prediction is the reduction of uncertainty. The qualification “unlikely” in the preceding definition must be emphasized. “Prediction” in the sense in which I am using the word does not mean wild guessing, nor does it mean staking everything on a simple outcome. Rather prediction means the elimination from contention of those possibilities that are highly unlikely, and the examination first of those possibilities that are most likely. Such a procedure is highly efficient for making decisions involving language.

**Prediction in operation**

Imagine that I have written 26 letters of the alphabet on 26 index cards, one letter to each card and that I shuffle the pack of cards, select one at random, and ask you to guess what that card is. You could very rightly object that since every letter is equally probable, nothing you know could in any way increase your chances of making a correct guess. Whatever letter you might choose to guess, the probability that you will be correct is exactly the same, namely, one in 26. One the average you would expect to make 13 guesses before you are likely to be right.

However, letters do not occur randomly in the English language. Some have a much higher probability of occurrence than others, for example the most common letter E is forty times more likely than the least common letter Z. So when a letter is selected at random from English text, your prior knowledge of the language can obviously make a difference to your chances of making a correct guess.

It is easy to demonstrate that people can and do use their knowledge of the relative probabilities of English letters in this way, knowledge that often they are not aware they have. For example, one can ask an audience of several
hundred to write down their guess of what the first letter of a pre-selected six-letter word might be. In an example I demonstrated at the Reading ‘74 Conference at Your University the pre-selected word was STREAM. The majority of people will write E, T, A, I, O, N, S, H, R, D, L, or U, which happens to be the 12 most frequent English letters in order of frequency. Scarcely anyone will predict Z, or Y, or J. Usually S happens to be the most common guess for the initial letter of six-letter words, by about one person in eight (as opposed to the one in 26 that would be expected if guesses were made at random). Tell an audience that the first letter is indeed S, and fully half of them will correctly guess the second letter T first time, and fully half again will guess that the third letter is R. Most people will then correctly guess that the fourth letter is E, and go on to be incorrect with their guess that the following letter is another E, although they will be correct on their second attempt with A. These days, K is usually the guess for the final letter, with M the successful second guess. In other words, by using their prior knowledge of the relative frequency of letters and groups of letters in English, people rarely have to labor through a dozen or more unsuccessful guesses before they can decide what the next letter of an unknown word might be. The average number of guesses is about three. (The statistically computed average number of alternatives that successive letters of English words might be is seven or eight, [Shannon, 1951]). The effect of such prior knowledge is considerable. Most English words remain recognizable if every other letters is obliterated, demonstrating that we have scarcely have to look at most letters to identify them in words. A more graphic illustration of the saving that the prior elimination of unlikely alternatives can accomplish is that a single glance at a sequence of random words on a screen is usually sufficient to permit the recognition of two or three words, or twice as many letters than could be recognized if the letters flashed on the screen had been randomly selected.

But readers know far more about language than the relative likelihood of particular letters in isolated words. We can make excellent guesses about words in sentences. Take any book that happens to be handy, read the last couple of
lines of a right hand page, and then guess what the next word will be when you turn the page. You will not be right every time of course, but you will almost always guess a word that is possible. Remember, what is important is not to be absolutely correct, but to eliminate unlikely alternatives. Once again, statistical analyses of English texts have shown that although in theory an author might draw from a pool of fifty thousand words or more for the words to be used in a book, there are on the average no more than 250 alternatives available when he or she writes any particular word in that book (Shannon, 1951). The readers do not need to predict the exact word that will confront them. Nor need they predict more than a few words ahead. But if they can reduce the number of immediate alternatives from many thousands to a couple of hundred, they are taking a considerable burden from the limited information-processing capacity of the brain. Once again our illustrative experiment will demonstrate this saving: If the sequence of 30 letters flashed briefly on a screen comprises a single coherent sentence or meaningful phrase, then the viewer can usually see it all at one glance.

There have been hundred of experiments showing that sequences of letters and words are identified faster, more accurately, and with less visual information, the more they correspond to possible sequences in the English language (a classic example is Tulving and Gold, 1963). The experiments demonstrate not only that individuals, including children, have a considerable prior knowledge of language that enables them to eliminate many unlikely alternatives in advance, but that this knowledge is exercised automatically, without the individual’s awareness and without specific instructions to do so. But the prior rejection of unlikely alternatives is a characteristic of the way the human brain works. The reason we are rarely surprised by anything we see, even when we visit an unfamiliar setting, is that we always have a set of prior expectations about what will in fact see. We do not predict everything; we would be surprised to see a camel in the harbor or a submarine in the zoo, but not vice versa. Nor are your predictions over-specific; we rarely predict exactly what we shall see.
next. Instead, we quite automatically and subconsciously eliminate unlikely possibilities from consideration.

**The advantage of prediction**

Prediction in reading, I have argued, involves the prior reduction of uncertainty by the elimination of unlikely alternatives. We never make our decisions as if we had no prior expectation; recognition and comprehension in such circumstances would always be disruptively time-consuming and tedious. Instead we seek just enough information to decide among the alternatives that are most likely. As a result, the four limitations on reading that I have discussed as reasons for prediction are very easily overcome, and there are other advantages as well.

Most words have many meanings, but if we are predicting, then we are usually looking for only one meaning of any particular word. You may not be able to guess if the next word is going to be table or chair, sideboard, or coatrack, but if you know that it will refer to a piece of furniture, you will not even consider that table might be a set of numbers, or chair a verb. The reason neither speakers and writers are aware of the potential ambiguity of what they say is that they already know the meaning they are trying to express and do not consider alternative possibilities; they are embarrassed if a double-meaning is pointed out to them. Similarly listeners expect a certain meaning if they are following (or rather predicting) the sense of what they are trying to comprehend; hence puns are so excruciating when eventually we manage to see them. Words may have a multiplicity and grammatical function taken one at a time, but in meaningful sentences they are rarely ambiguous.

The pronunciation of words may not be predictable from their spellings, but if you know what a word is likely to be, it is not difficult to use “phonics” to confirm or reject a particular expectation. As all reading teachers know implicitly,
phonics is easy if you already have a good idea what the word is in the first place. If a child can predict that the next word is likely to be either cow, horse, or sheep, he or she will not need much knowledge of spelling-to-sound correspondences to decide which it is. It is in fact through such prediction that a mastery of useful phonic skills is acquired.

Obviously, prediction will speed up reading, and therefore help to overcome the limitation imposed by the brain’s rather sluggish rate of information processing. The fewer alternatives you consider, the faster you can read, and the more efficient will be the reading that you accomplish. Reading with prediction means that the brain does not have to waste time analyzing possibilities that could not possibly occur.

The limited capacity of short-term memory is overcome by filling it always with units as large and as meaningful as possible. Instead of being crammed uselessly with half-a-dozen unrelated letters, short-term memory can contain the same number of words, or better still, the meaning of one or more sentences. In fact, prediction works better at these broader levels; it is easier to predict meanings rather than specific words or letters, and very few letters or words need to be identified to test prediction about meanings.

The first of the bonus advantages of prediction in reading is that the reader is working already at the level of meaning. Reading is meaningful before the reader even begins. Instead of trying to slog through thickets of meaningless letters and words in the fond hope that eventually some nugget of comprehension will arise, the reader is looking for meaning all the time. If any possibility of meaning is to be found in a text, the predicting reader is the one who will find it.

The final advantage is of particular practical importance in many classrooms, namely that with prediction it does not matter if the reader’s
language does not exactly match that of the writer. Everyone can understand language that he or she could not possibly produce; that is why parents quickly learn to conduct their more intimate conversations out of the hearing their pre-school children. Yet the language ability of children and schools is all too often evaluated by the speech that they produce. Few readers, even adults, can succeed in threshing out the sound of a sentence, word for word, unless they have a good prior idea of what the sentence as a whole means. There is no way a child can be expected to identify words as a preliminary to getting the meaning if the words are in fact not among those he or she would choose to express such a meaning. But with prediction, a “one-to-one match” is not required. It will not matter if a child thinks the author has written “John ain’t got no candy”, rather than “John has no candy”, provided the meaning is understood and provided the teacher is not demanding literal word-for-word accuracy.

**Prediction in the classroom**

Two basic conditions must be met if a child is to be able to predict in the manner that is essential for learning to read. The first condition is that material from which children are expected to learn to read must be potentially meaningful to them, or otherwise there is no way they will be able to predict. The opposite of meaningfulness is nonsense, and anything that is nonsensical is unpredictable. Any material or activity that does not make sense to a child will make it more difficult to read.

But meaningfulness of materials and activities is not enough; children must also feel confident that they are at liberty to predict to make use of what they already know. With prediction there is a constant possibility of error, but then readers who read without ever making errors are not reading efficiently; they are processing far more information than is usually necessary. The child who will become a halting, inefficient reader is one who is afraid to make a mistake. The
worst strategy for any reader who is having difficulty understanding text is to slow down and make sure that every word is identified correctly.

The notion that prediction should be encouraged worries many teachers; it may sound as if a virtue is being made out of error. But one should distinguish prediction from reckless guessing. The guesser is usually the child trying to achieve what the teacher is demanding by getting every word right, no matter how little relation it bears to sense. A striking characteristic of older children with low reading ability is that they read as if they have no expectation or interest that the material might make sense, but are determined to get the words right at all costs.

Also, accuracy is overrated. There are only two possibilities for a mistake made during reading, either the mistake will make a difference to the meaning, or it will not. If the mistake will make no difference, if the child reads “house” instead of “apartment,” then it will make no difference. There is no need to worry. But if the mistake does make a difference, if the child reads “house” instead of “horse,” then the reader who is predicting will subsequently notice the anomaly, simply because he or she is following the meaning. The child who overlooks obvious errors of sense is not the child who rushes through to understand the gist of a passage but the one who tackles the passage one word at a time.

How then can prediction be taught? There are some obvious methods, such as encouraging a child to guess what a difficult word might be, and playing reading games where the teacher stops suddenly, or leaves an occasional word out, or makes an occasional deliberate mistake. But more important I think is that prediction should not be discouraged. Prediction is a natural aspect of language. The preferred strategies for a child who meets an unfamiliar word in an interesting story he or she is reading are the same as those for fluent readers: first skip, and second, guess. Sooner or later the child will have to predict if he or she is to become a fluent reader. Feedback is an essential part of all learning
activities, but it can come too soon, or too often. A child who pauses before identifying a word may not want the teacher to help “sound it out,” nor the rest of the class to tell what it is, he or she may in fact know what the word is and simply be wondering what it has to do with the rest of the sentence. A child who “makes a mistake” need not be “corrected” by having the teacher, or the rest of the class, say the right word immediately. If left alone he or she might self-correct in the following sentence, a far more valuable skill in reading than the blind ability to word-call. One of the beautiful advantages of reading sense is that it provides its own feedback; errors become self-evident.

One of the most formidable impediments to prediction, at all levels of reading, is anxiety. A child who is afraid to make a mistake is by definition anxious, and therefore unwilling to take the necessary risks of prediction. Individuals of any age labeled as reading problems will show anxiety, especially in situations where they feel they are being evaluated; their reluctance to predict will lead to laborious nonsensical reading, and their “difficulty” will become a self-fulfilling prophecy.

Prediction is not everything in reading. Other important considerations include the efficient use of short-term memory, the minimal use of visual cues, and the selection of an appropriate rate of speed for particular reading tasks, together with the acquisition of effective strategies for the identification of unfamiliar words from context. But these are all skills that come primarily through the practice of reading; they are fostered rather than taught (in fact, many teachers are not aware of the extent to which these skills are involved in reading). The advantage of prediction is that it facilitates precisely the kind of confident, successful, and meaningful reading experience through which all of the critical skills of reading are acquired.

Handout created by Frank Smith of the DVC Learning Center. Copyright 2002.