CHAPTER THREE
WORD PERCEPTION

The question of how words are identified is clearly central to understanding reading. It has also been a major focus of research in cognitive psychology for the last 15 years. Much has been learned—not only facts but also a greater awareness of what the issues really are. However, there are some areas where our ability both to ask and answer questions is limited.

While laypeople undoubtedly differ in how they believe that word identification occurs and what its place is in the total process of reading, let us attempt a sketch of one commonsense view that emphasizes learning to read in order to raise some questions and indicate where the discussion will lead. Many people think that the question of how children learn to read their native language is central to understanding the reading process, including skilled reading. As we hope to show in this chapter, a naive version of this developmental perspective gives a misleading picture of skilled reading. However, we would like to present such a perspective to raise some important questions about reading. Some will be pretty much resolved in this chapter, but many will recur throughout the book.

When one starts out from a developmental perspective, it appears that recognizing the printed word is the central problem of reading. Presumably, the six-year-old child has a well-developed system for language understanding, and the major thing to be learned is how to plug the squiggles that are on the page into that system. If the child can learn to access the words of the spoken language from the written representation, then he or she should be able to understand the written representation. This suggests one central question about reading: *Is word recognition all that needs to be learned?*

If one looks at the beginning reader, this process of trying to identify written words is extremely effortful, and in fact some children fail to learn to do it fluently. In contrast, their processing of more complex speech than they are attempting to read appears to be relatively effortless and "biologically programmed." (Biologically programmed in the sense that all people without significant brain damage appear to be able to comprehend spoken language well.) While the adult (or even older child) can decode words with far less effort than the beginning reader, it might seem from this perspective that the processing of words is the "bottleneck" in the reading process: That is, in order to read, an "unnatural" and effortful step, visual word identification, has to be grafted onto a "natural" language understanding system designed by evolution. This leads to a second general question: *Is identifying words effortful and the rest of the reading process "automatic"?*

In addition to suggesting that word processing is effortful and the rest of the reading process is "automatic," this view also suggests that identifying words (especially with an alphabetic writing system) is largely a process of going from the letters to the appropriate sounds. While the previous chapter has made clear that the relationship in most alphabetic systems is more complex than going from individual letters to individual units of sounds, it would appear that if the beginning reader has the general idea of the alphabetic principle, he or she can go from the print to a sound, and for most words, the sound will be close enough to that in the spoken language to be able to access the correct meaning. This leads to a third question: *Are words identified by accessing the sound and then the meaning?*

If we think of the translation from letters to sound, we might think that reading is a letter-by-letter process (albeit fast). That is, the letters in a word might be processed *serially* (i.e., one at a time) from left to right in order to identify a word. This view is also concordant with the usual introspection that one's attention seems to sweep across the page smoothly from left to right. This leads to a fourth question: *Are letters in words processed serially or are words processed as wholes?*

Clearly, as readers become more proficient in reading, they become more fluent at identifying words. What exactly has been learned? A fifth question is: *Do skilled readers learn to apply the "rules of spelling" in a fluent way or do they learn specific associations between visual patterns and the sound and/or meaning of the word?*

Lastly, beginning readers seem to be aided quite a bit by having context for decoding the words. It is often more difficult for them to identify words in isolation than in a story. Especially if one believes that word identification is the bottleneck step in reading, this suggests that context may play a large role in reading. In fact, several writers (e.g., Goodman 1970) have suggested that reading is a sophisticated guessing game whereby the reader develops hypotheses about what is going to come up on the printed page and then tests these hypotheses by sampling the display. In such views, the process of reading words in isolation is only marginally related to that of reading words in text. This leads to a sixth question: *Does context radically affect the process of word identification?*

While the picture of reading presented above—associated with the naive layperson—is intuitively reasonable, most of it is incorrect as a theory of the skilled reader. In addition, while the reading process is far less well understood for beginning readers, much of it may be incorrect for them as well. In fact, much of what we know about word identification in skilled readers can be summarized by the following statements. (While not all people in the area would agree with them, most would.)

1. Word recognition is relatively automatic, and "higher order processes," such as constructing the correct syntactic structure, relating word meanings, and fitting the text into what the reader understands about the world, are what takes most of the reader's processing capacity.
2. Word recognition is not merely converting letters to sounds and then sounds to meaning. In fact, a defensible position is that converting to sound is largely irrelevant to the identification of words (although we believe otherwise). Most researchers do believe, however, that conversion to sound does play a part in the reading process after word identification—largely for its ability to aid short-term memory.
3. Words are not processed serially letter by letter. The letters in common short words appear to be processed *in parallel* (i.e., at the same time), although words are not learned as visual templates. Longer words may be processed differently, although not much is known about how they are processed.
4. Words are processed pretty much the same way in isolation as in text. While context somewhat affects the speed of processing words, its effects are surprisingly small.

(What is known about the other two questions—what people learn when they learn to read and whether they use rules of spelling in reading—is not easily summarized in a sentence or two. The latter question will be discussed in the present chapter and the former in chapters 9 and 10.)

The fourth point has important methodological implications, since it justifies the use of experiments in which words in isolation are identified to illuminate how words are processed in reading. Since context does have some effect, however, findings with isolated words can not be assumed to be perfect indicators of how word identification operates in reading text. While we will briefly allude in this chapter to how words are processed in context, we need to explain quite a bit of information about eye movements (in chapters 4 and 5) in order to discuss words in context. Chapters 4 through 8 will discuss how words in context are processed, although that issue is a major focus of Chapter 7. Since it is more difficult to study how words are processed in reading text (both because it requires more sophisticated equipment and because the experimenter has far less control over the situation), historically, most of the research on word perception has been on isolated words. Thus, isolated words will be our focus for this chapter.

How Long Does it Take to Identify a Word?

Before going on to discuss the questions raised earlier in detail, we will introduce the issues and familiarize you with the techniques used to get at those issues by asking the naive question of how long it takes to identify a word.

Response-time methods Let's start with something simple. We present a subject a word on some sort of visual display and measure how long it takes the subject to say the word aloud. If we can precisely control when the presentation of the word began and can precisely measure when the subject begins making the response, we would have a measure of something relevant. But is this measure the time to identify a word? What prevents us from concluding that it is?

First, we have measured the time it takes for something to emerge from the subject's mouth. What we are interested in, however, is the time that it took for the subject's brain to achieve a state that we call "identification." After the subject has identified the word, several other processes must take place for a response to occur: (a) the subject must decide what response is called for in the experiment—in this case, it is saying the name of the word; (b) the subject must retrieve the motor program for executing the response; (c) the command must be sent down nerve pathways to the mouth; (d) the muscles of the mouth and throat must execute the command. All of those processes take time. Thus the time we have measured is the time it takes to identify the word plus some excess baggage that we might want to simplify and call "decision time" and "response-execution time." This problem is clearly not unique to the naming task we have selected; it would be true of any response of the subject to a word.

In fact, it takes practiced subjects about 400 milliseconds (msec) to name common words. Thus we might feel that such an experiment would at least allow us to say that people identify a word in less than 400 msec. However, there is another basic problem facing us: What exactly do we mean by "identifying a word"? In reading, the important thing is getting to the meaning of the word. When the subject has named the word, does he or she necessarily know the meaning? The answer is clearly no. We can clearly name words that we do not know the meaning of and there are people with brain damage who can name a lot of words and nonwords and appear to have no idea of the meaning of what they are reading. Thus, naming (or more properly getting to the name of a word in memory) does not necessarily mean that its meaning has been accessed. Perhaps, however, for normal people reading words they know the meaning of, both of these events occur at roughly the same time. That seems reasonable, but it is by no means a foregone conclusion. In fact, one of the central questions of word processing is how those two events—accessing the name of a word and accessing the meaning of a word—are related.

Perhaps there is a better task for getting at whether the person has accessed the meaning of a word. Let's try a *categorization* task, such as asking the subject to judge whether the word is an animal or not. We would be certain then that the subject has accessed the meaning. However, to achieve that goal, we may have paid a big price. The naming task is relatively easy and effortless. While there is a decision stage, the decision of executing the vocal response /dawg/ when we see DOG seems natural, relatively quick, and relatively constant across words. However, the categorization task seems less so. In judging whether a word is an animal, subjects are relatively slow to respond "yes" to STARFISH and relatively slow to respond "no" to BACTERIA or even to ROSE, but quite quick to respond "no" to STONE. It is clear that the decision stage in the categorization task is much more intrusive: subjects need to do mental work after they have identified the meaning of the word in order to decide if it is in the appropriate category. In spite of all this, subjects can usually make these category decisions for relatively common instances within about 700 msec, so that we can really be sure that words are identified within about that time.

Is there no other simpler task that allows us to be sure that the subject has processed the meaning but doesn't involve an extensive decision stage afterwards? Unfortunately, no one has been clever enough to come up with one. One attempt to measure the time to identify a word that has been widely used is *lexical decision*. In the lexical-decision task, the subject is shown a letter string and asked to decide if it is a word or not. (Obviously, nonwords are used as well.) This task appears to be a bit simpler (and faster) than the categorization task. Moreover, we know that subjects must, in some sense, have identified the word when they know it is a word. However, we can't be at all sure that subjects know the meaning of a word at the moment that they know it is a word.

Brief presentation methods Perhaps methods that time the subject's response are not the best way. Instead one can think of timing the presentation. If we flashed a word briefly on the screen (let's say for 60 msec), and if the subject could still identify it (e.g., name it and give a synonym), would that mean that it only took the subject 60 msec to identify a word? (Note that here we do not necessarily time the subject's response; the time pressure is solely produced by the brief exposure.) There are several problems with making that conclusion. The first is that even though the stimulus is only physically present for 60 msec, the visual representation lasts longer than that. This is the phenomenon of *iconic memory* discussed in Chapter 1. The data on iconic memory suggest that the visual image would last for at least about 250 msec, although it would be fading over that interval.

Can we defeat iconic persistence? A procedure designed to do that is *masking*. After the presentation of the stimulus, a *pattern mask,* usually consisting of bits and pieces of letters or letterlike forms is presented in the same location. That is, the subject sees the word for 60 msec followed immediately by the mask. Subjectively the word looks like it disappears when the mask comes on. In spite of this, subjects with a little practice, which gets them used to this mildly bizarre situation, can identify words if they are exposed for about 60 msec (Adams 1979).

Does this demonstrate that it takes about 60 msec to identify a word? Well, it does seem to demonstrate that it takes at least that long, since you need a 60 msec dose of visual information to do it. However, masking is a complex phenomenon which is still far from understood. Even though the stimulus looks like it disappears after the mask comes on, the information may still be there for further processing. One possibility is that after 60 msec or so, the visual information from the masked stimulus is transferred to some kind of short-term memory where the mask cannot disrupt it so that the word identification processes can still operate on the information; however, information in this buffer need not lead to conscious perception. While this explanation may strike the reader as rather baroque, it is not at all far fetched. However, there is no compelling reason to accept such an explanation, so let's assume that the visual information (in some sense) disappears after 60 msec.

We must keep in mind that 60 msec is the time the stimulus is on, or, more importantly, the time before the mask appears. However, that doesn't mean that only 60 msec has elapsed between when the stimulus has come on and when the mask (in some fashion or other) tells the brain to let go of the first stimulus. We know it takes some time for the nerve impulses to travel from the eye, through the optic nerve, and into those regions of the brain that identify visual stimuli. (The exact regions are still not known, but there is some evidence that it is in areas known as "secondary visual cortex," [Petersen et al. 1988].) To pick a simple number, let's say that it takes about 50 msec. Thus, the word presented is moving up the visual pathways to the pattern-recognition system and the mask is following it 60 msec later. If the mask takes as long as the stimulus to get to those centers and interrupt processing of the stimulus, then the mask will arrive there 110 msec after the stimulus was presented. That is, the brain might only process the stimulus for 60 msec, but the total time elapsed between when the word appears and when the brain is forced by the mask to stop processing it would be 110 msec.

Assuming that the subject can identify the word, 60 msec plus this neural transmission time (assumed to be 50 msec in the above argument) would be a good estimate of how long it takes to process a word. How long, in fact, does neural transmission take? No one knows for sure, although 50 to 70 msec is probably not an unreasonable guess. (One other complication is that the masking stimulus' disrupting effect is likely to be transmitted to the brain a bit more quickly than the visual information about the word.) The bottom line, therefore, is that 60 msec is a lower-bound estimate for how long it takes to identify a word, although we should expect that it takes appreciably more time than that.

Estimates from reading text Perhaps we are making it all too complicated and artificial. Why don't we examine real reading? The typical college student reads at about 300 words per minute or 5 words per second. That means that words in text are, on average, processed in about a fifth of a second or 200 msec. There are several problems with this. The first is that there is time used up in reading that has little to do with word identification. If one is reading difficult text (and text can be difficult even without unusual or technical words), reading can be slowed down to one-half that speed or less. Thus, it is clear that reading is more than word identification. On the other hand, we can't be sure that the reader is really identifying every word. He or she may guess at individual words or even phrases, and it is very difficult to test whether all the words in the text have been identified. In spite of everything, we do seem to be converging on an estimate. The reaction time studies demonstrated that word identification probably takes less than 400 msec, the experiments with brief presentations demonstrated that it takes at least 60 msec, and the estimate from reading suggests a number something like 200 msec.

Physiological methods Perhaps we are wasting time with indirect methods. What about examining the brain itself to see when a word is recognized? The first problem is clearly that we have to study humans to study reading, and ethics prevent us from opening the skull to answer our question. Even if we could, we wouldn't know how to recognize "word identification" in the brain. We don't know for sure where it is and we also don't know whether we are looking for a pattern of increased electrical activity, decreased electrical activity, or some other, subtler change of state in the brain.

While there are several methods that can be used for studying human brain states without surgery, only one method at present, *evoked potentials,* is adequate for studying the time course of the processing of incoming stimuli. This method uses electrodes taped to the scalp and measures relatively gross electrical activity in the brain. The method is very imprecise for determining where the electrical activity is coming from. The signal is also very "noisy" so that the records from many trials have to be averaged in order to draw any conclusions. These records of electrical activity have certain relatively well-defined peaks that occur at certain approximate times after the stimulus is presented (although there is dispute about whether these peaks really are the same across tasks and subjects). If we forget about these disputes for the moment, there are several peaks of interest (e.g., Kutas and Hillyard 1980; Van Petten and Kutas 1987). One that occurs at about 300 to 400 msec appears to be the one that corresponds to making a decision about what response to select (see Figure 3.1). This is not what we are interested in; we are looking for something before that. There is a peak at 50 to 100 msec, but it appears to be associated more with the detection that "something has happened" rather than identification of a particular stimulus. The "negative peak" at about 150 to 200 msec appears to be a reasonable guess as something that might be associated with stimulus identification. For example, the size of this peak differs when an auditory stimulus is attended to than when it is not (Hillyard et al. 1973). However, from the data at hand, it is definitely not clear what this peak represents. At one extreme, it could be representing a relatively early stage of stimulus processing such as the recognition that the stimulus is a sequence of letters. At the other extreme, it could represent a relatively late stage of processing such as a reaction to the identification of a word rather than the identification of the word itself.

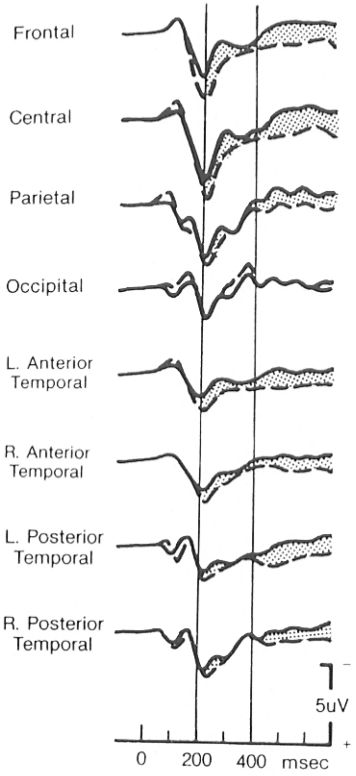


FIGURE 3-1 **Average evoked potentials to function and content words taken at several locations on the scalp. As can be seen in the graph, there is some indication that some of the functions differ at about 100–150 msec, indicating that the brain can discriminate these classes of words that quickly. (From Kutas and Hillyard 1983, with permission of the Psychonomic Society and the author.)**

However, it is suggestive that the estimates from these records also indicate a time something like 150 to 200 msec for word identification.

What then can we conclude from all the above methods? We are certain that 700 msec (categorization reaction time) is too slow for an estimate of the time to access the meaning of a printed word. We are pretty sure, but not certain, that 400 msec (naming time) is also too slow for an estimate. We are also pretty sure that 60 msec is too fast an estimate and that something like 150 msec is about right, although we shouldn't be too surprised if word identification took place in as little as 100 msec or as much as 200 msec after the word is first sensed by the eye.

Before going back to the questions raised at the beginning of the chapter, let us conclude this introduction by considering the question of whether the speed of identifying a word is influenced by the frequency of the word in the language. The frequency of the word is usually measured by taking some corpus of text that is assumed to be representative and actually counting the number of times that a particular word appears. The frequency count for American English that is the current standard is that of Francis and Kucera (1982). To give you some feel for frequency counts, words such as IRK, JADE, COVE, VANE, and PROD have counts of 2 to 5 per million, while words like CAT, COAT, GREET, and SQUARE all have frequencies greater than 30 per million.

The difference in lexical decision time between a high-frequency word such as COAT and a low-frequency word such as COVE is about 100 msec; however, the difference in naming times is considerably less, about 30 msec (Balota and Chumbley 1984). Clearly, both of these differences can not be estimates of how much longer it takes to identify a low-frequency word than to identify a high-frequency word. In reading, the difference in "fixation time" (i.e., the time the reader looks at a word) between high-frequency words and low-frequency words is also something like 30 msec (Inhoff and Rayner 1986). Thus, 30 msec seems like a better guess as the effect of frequency on the time to identify a word. (However, the effect may be appreciably bigger for *really* low-frequency words.) It should be pointed out that high-frequency words tend to be short and low-frequency words tend to be longer; however, the frequency effects reported above are obtained when the high- and low-frequency words are equated for the number of letters.

Let us now return to the questions raised at the beginning of the chapter. Our discussion of the time it took to access a word had several goals in mind. First, we wanted you to see how questions have to be sharpened in order to be answerable. For example, terms like *word identification* are not precise enough. Second, we wanted to introduce you to the experimental tools that we have available for answering questions about word processing. Third, we wanted to carry through an argument fully enough to see what sorts of "answers" we usually have. Each technique has its own problems, but if enough techniques appear to converge on a common answer, we can have a reasonable degree of confidence in our conclusion. (Sometimes we are luckier than others: techniques may converge or they may not.) Fourth, we wanted to give you one piece of the data that suggests that words in isolation are processed pretty much the same way as words in text: the times we estimated for word identification are not radically different in the two cases. (Rest assured that we will return to fully document this assertion later in the chapter and in Chapter 7.)

Is Word Processing Automatic?

The claim was made earlier that, perhaps contrary to intuition, the identification of the meaning of a printed word was a relatively automatic process and the part of reading that was using up most of the time and effort was likely to be the higher-order processes: those processes that put the meanings of words together so that sentence structure can be grasped, the meaning of sentences and paragraphs can be understood, and the intention and tone of the author can be comprehended.

Actually, the claim is a bit of an overstatement, since the case for the automaticity of word processing is not air-tight. However, the basic point will be seen to be valid: identifying the meanings of words is a rapid process for the skilled reader and is definitely not the "bottleneck" in reading. As we will see, this issue is important, since it shapes one's overall model of the reading process. For example, several influential models of reading view word identification as the hard and unnatural step and therefore assume that readers heavily depend on context to identify words (e.g., Goodman 1970; Smith 1971). On the other hand, context is likely not to be important in determining how quickly and accurately words are identified if identifying words is automatic.

You may be getting impatient by now and wondering what on earth we mean by "automatic". To the layperson, the word *automatic* connotes something that is rapid, involuntary, and effortless. We will adopt Posner and Snyder's (1975) three criteria for automaticity that somewhat formalize this intuition: (1) the person may be *unaware* of the process; (2) the execution of the process is not under the *conscious control* of the subject—that is, the subject's intention to perform the task may be irrelevant to whether it is done; (3) the process takes *no processing capacity*—that is, it uses no resources that other mental operations might also use. These criteria are offered as a tentative definition in order to structure discussion. However it may be too much to ask of word identification to satisfy all three criteria. The criteria may be too strict: Perhaps no process (with the possible exception of some reflexes) satisfies all three.

Is identification of words unconscious? How could we possibly identify the meaning of a word and be unaware of it? Consider the following experimental situation. A word is flashed briefly and is followed by a pattern mask, as described earlier, but the word is exposed for only about 20 msec before the pattern mask appears. (Remember, something like a 50 to 60 msec exposure is needed for the subject to be able to identify the word 100 percent of the time.) If the experimental situation is arranged carefully—the pattern is an effective mask, the word and mask are about the same brightness—subjects will say that they can't see the word, and they will be unable to perform above chance when asked to report whether or not a word was presented prior to the mask, but the meaning of the word will be identified.

How could we test to see that a word has been processed if the reader is unaware of its meaning? We need a subtler test than asking the subject to report the word. Let us assume that DOG is the word that is flashed briefly. One possibility is to present two words that can be clearly seen (let's say DOG and BOY) a short time after and ask the subject to choose which of the two was the briefly exposed word. Subjects are also at chance level on this test of recognition memory, so we still have no evidence that subjects have identified the word (Balota 1983).

As experimenters, we need to be even more devious, which necessitates explaining an experimental procedure known as *priming* (Meyer and Schvaneveldt 1971). For the moment, we will forget about brief exposures. In a priming experiment, the experimenter shows two words in sequence, the *prime* and then the *target*. The sequence might be the words DOG and CAT. The experimenter is primarily interested in how quickly the second word, the target, is processed. In particular, is CAT processed any more rapidly when a related word such as DOG precedes it than when an unrelated word such as FAN precedes it? If we measure the response time to judge that CAT is a word (the *lexical-decision* task), in fact subjects will usually be about 30 to 50 msec faster to respond "yes" when DOG is the prime for CAT than when FAN (an unrelated word) is. A similar, but somewhat smaller, effect can be obtained when the subject names the target. The precise interpretation of the priming effect is still hotly debated (see Chapter 7). However, for our present purposes, the important thing is that the priming effect demonstrates that the meaning of the prime has been processed, since the speed of processing the target is dependent on the meaning of the prime.

Now let's return to our situation where DOG is flashed for approximately 20 msec followed by a mask. Instead of asking subjects what they saw, we ask them to make a lexical decision on another letter string about 500 msec later. Amazingly enough, subjects will be faster to judge that CAT is a word when preceded by DOG than when preceded by FAN even though they have no awareness of seeing the priming word. Thus, the meaning of the priming word has been identified, since the time to judge the target is influenced by whether the prime is related to it in meaning, even though the subject is unaware of identifying the prime. While this phenomenon of "unconscious priming" is somewhat controversial, it has been replicated many times (e.g., Marcel 1983; Balota 1983; Fowler et al. 1981; Carr et al. 1982). Moreover, the size of the priming effect is usually unaffected by whether the subject can identify the target stimulus or not (e.g., Balota 1983; Carr et al. 1982).

We have been a bit vague about what "awareness" means, and this has been a subject of controversy. The standard criterion is that the subject is at chance level if asked to say the word. However, one experimenter (Marcel 1983) has reported priming even when the subject is no better than chance at distinguishing between whether a priming stimulus preceded the mask or nothing preceded the mask. While not all researchers are convinced that priming can be obtained even when the subject is totally unaware that a stimulus was present (Holender 1986; Cheesman and Merikle 1984), it is clear that the meaning of a word can be "looked up" by its visual representation without the conscious experience of perceiving the word. Thus, the identification of the meaning of visual words is automatic according to the first criterion outlined above.

Is intention to process a word important? In some sense, the experiments discussed above may have already made the point: the word's meaning is processed even though the subject is unaware of that fact. However, since the subject is trying to do well in the experiment and trying to see everything as well as possible, perhaps the subject is intending to process the stimulus (even if unaware of having processed it), and that is important for the meaning being extracted. Would subjects extract the meaning in the above priming experiments even if they were trying not to? We don't know for sure since the experiment has not been carefully done. However, there is clear evidence that the meaning of a word is extracted when the subject is trying hard *not* to process it. The standard experiment that demonstrates this is one in which subjects see a printed word written in colored ink and are supposed to name the color of the ink. One of the most interesting conditions is when the subject sees a color name such as RED printed in green ink. In such a case, the time to make the correct response, "green," is very slow compared to when the subject says "green" to a green color patch.

This phenomenon, which is commonly called the "Stroop effect" after its discoverer (Stroop 1935), is not a transient phenomenon. It is a large effect—subjects are usually about 200 msec slower to say "green" to RED written in green ink than to a green color patch—and it only decreases a bit with extended practice (Dyer 1973). Subjects know that the word is interfering, but they can't avoid processing it. A similar effect is obtained when subjects see the word CAT in the middle of a line drawing of a dog and attempt to name the line drawing (Rayner and Posnansky 1978). However, while this phenomenon tells us that something about the word is processed, we don't know for sure that it is its meaning. Perhaps it is only that the word form accesses the motor program to name the word.

One way to test whether the meaning of the word has been accessed is to compare the size of the interference effect when the word is a competing color word such as RED and when it is an unrelated word such as ANT. In fact, while both words interfere with saying "green," the interference effect is substantially greater for RED than for ANT (Klein 1964). Thus, it appears that the interference has two components: (1) since unrelated words interfere, there is competition between the name of the word and the name of the ink; (2) since color names interfere more, the meaning of the word competes with the meaning of the color. A second finding that reinforces this conclusion is that an associate of a competing color name such as BLOOD interferes with saying "green" more than an unrelated word such as ANT does (Dyer 1973).

Thus, the Stroop effect demonstrates that both the name and meaning of a word are processed by skilled readers even when they are trying hard *not* to process them. So we see that identifying words is automatic both in the sense that it may go on without awareness and in that it goes on even when the subject is trying not to do it. The evidence for the third aspect of automaticity—not requiring limited capacity—is not nearly as clear cut.

Does word identification take processing capacity? Before plunging ahead to answer this question, we need to discuss briefly what one means by "processing capacity," and how one would test for it. In our discussion of cognitive psychology in Chapter 1, we briefly touched upon the concept of "limited capacity." Most theories of cognition assume (either explicitly or tacitly) that many cognitive acts need some sort of attentive process and there is a finite amount of this attention which limits how much information can be processed at a time. Some processes, such as the normal control of breathing, are assumed not to need any attentional processing, while others, such as the processing of discourse, are assumed to need attentional processing, since it is very difficult to process two conversations at once.

The basic test for whether a process requires attentional capacity seems simple: A process can be assumed not to require attention if it can be done at the same time as another process and not interfere with it; accordingly, the process probably uses some attentional capacity if it interferes in some way with the performance of the other task. For example, if it were true that one could multiply two-digit numbers in one's head as rapidly while driving a car as while sitting in an easy chair, then one would want to conclude that either driving or mental arithmetic (or both) did not take any limited capacity. On the other hand, if the mental arithmetic slowed down during driving or more errors were made, then one would conclude that both tasks were using a pool of limited capacity or resources. Put simply, the test of limited capacity is usually whether two things can be done at the same time as well as one.

We will assume that a process demonstrates no need for capacity if it doesn't slow down or interfere with other processes. However, it is not necessarily the case that interference is due to limited resources being shared. To see this, let us return to the Stroop task. We found that it took longer to say "green" to RED printed in green ink than to say "green" to a green color patch. Does that mean that processing the form (RED) and the color (green) each used processing resources? This conclusion seems unlikely: the interference is different when the word was RED than when it was ANT, and it is not clear why it should take more resources to identify RED than to identify ANT. Moreover, if the subject sees GREEN in green ink, then the naming time is even faster than for the color patch (Hintzman et al. 1972).

Thus, interference is not necessarily the result of competition for attentional resources. It could result from a competition between two incompatible responses. A reasonable explanation of the Stroop effect is that both the color and form of the printed word are identified without needing any capacity, but after both are processed, they produce responses that compete with each other. (It is possible that settling this response competition does require limited capacity, but that is a side issue here.)

However, Keele (1972) has claimed that a variation of the Stroop task does argue that processing the meaning of the color and form do not require the same limited capacity. Instead of asking for a vocal response, Keele had subjects press keys to indicate the color of the ink (one key for red, one for green, one for yellow, and one for blue). He found an interference effect for competing color words but none for neutral words. Thus, the key-press task appears to have eliminated the interference caused by activating two names, but not the interference caused by activating two conflicting meanings.

Keele's argument is the following. The interference effect with color names demonstrates that the meaning of the word has been processed at least up to the level of knowing whether it is a color word or not; however, since there is no interference for neutral words such as ANT compared to the baseline condition, one can conclude that processing the word took no resources. The first part of the argument is valid, but the second is problematic. In the first place, the baseline in Keele's experiment was a repeated colored nonsense pattern, which he assumed took no capacity to process. But there is no reason to assume that processing resources were not needed to process the nonsense form. The problem, however, transcends the particular experiment. Even if a color patch were used, one could still argue that processing resources were needed to process its form. In addition, it is possible that the lack of interference observed could be because processing the color takes no resources.

While Keele's results are not conclusive, they are consistent with the hypothesis that processing a printed word does not require limited capacity. It would be more conclusive if we could demonstrate that two printed words could be processed as quickly as one. How could we test this? What we need is a task that presents the same response requirements when two words are presented as when one is, so that differences between the two (if observed) could be ascribed to the greater difficulty of identifying two words. (Naming clearly won't work.) One task that has been employed is *visual search*. The subject is presented with one or more words and asked to determine whether, for example, there is an animal name present (Karlin and Bower 1976; Pollatsek, Well, and Gott 1978). Thus, the response—a key press to indicate "yes" or "no"—is identical in all cases. The response time to decide whether the name of an animal is in the visual display takes longer the more words there are in the display (in fact, about 200 msec per word). Thus, it appears that processing words does take capacity.

However, the process that takes capacity may not be the identification of the meaning of the words; it may be the subsequent step needed to decide whether a word is in the experimentally specified category. Can we determine which? There is a similar experiment involving letters and digits. The subject is presented with from 1 to 6 characters and asked to respond whether a digit is present. In this task, the time to detect a digit when there are 5 letters present is virtually the same as when there are no letters present (Egeth, Jonides, and Wall 1972). Thus, it appears that subjects can process the meaning of 6 characters and, furthermore, categorize them as letters and digits as rapidly as it takes to process 1. The letter-digit experiment thus indicates that categorization per se does not necessarily take capacity. However, the categorization of characters as letters and digits could easily be more automatic than the categorization of words.

Unfortunately, there is no clear solution here at present, so let's try to remember what we know. First, if we look at the Stroop data, we know that the meaning of the word is processed along with the color, and as far as we can tell, accessing the meaning of the word does not take away resources from processing the color (or vice versa). If we look at the search data, we get the opposite picture: categorizing two words takes more time than categorizing one, so that some process associated with categorizing appears to take resources. The letter-digit data argue that the process that requires limited resources may not be categorization, although we don't know for sure.

Thus the available data don't allow us to conclude with any certainty that word processing is automatic in the sense that it takes no capacity. However, we have abundant data that, even if it takes resources to process a word, it doesn't take the processor with limited resources much time. First consider our discussion of recognizing words followed by masks. In those experiments, something like a 50 to 60 msec interval between word and mask was sufficient for recognizing words. Although one had to consider neural transmission time if one wanted to know how long it took to recognize a word after it was presented, the neural transmission time is irrelevant if one wants to know how long the central word-processing mechanism is actually involved in identifying a word. Thus, we would estimate that time to be about 60 msec. Similar estimates come from two other sources.

The first is an experiment that used a masking procedure, but while subjects were reading text (Rayner et al. 1981). When subjects read text, their eye movements were monitored (see chapters 4 and 5), and each time the eye came to rest, the text was exposed for a fixed amount of time until a mask came on, obliterating the text. Thus, on each fixation, the subject had only a limited amount of time to see the text. In spite of this, when the mask appeared 50 msec after the start of each fixation, subjects could not only read the text, but their reading rate was only slowed by about 15 percent compared to normal reading. This suggests that only a little over 50 msec is needed to identify a word.

A second line of evidence comes from the technique known as *rapid serial visual presentation,* which has the catchy abbreviation RSVP. In this technique, subjects see words appear one after the other in the same place on a video screen. The finding is that subjects can comprehend material even when it is presented at rates of 12 to 15 words per second (Forster 1970; Potter, Kroll, and Harris 1980). Inferences about time are a bit indirect with this technique, since subjects have some time to piece the material together after the sentence is finished and fill in details even if all words were not processed. Nevertheless, there is evidence that most words are processed (e.g., it makes a difference if words like *the* are presented or not) so that the subject does appear to be processing a word in about 70 to 80 msec or so (and doing at least some higher-order processing as well).

Thus, while we can't be completely sure that the identification of a word is completely automatic for a skilled reader, it appears to take at most about 60 to 70 msec of mental activity. Since the average skilled reader reads even the simplest text at about 300 words a minute, or about 200 msec per word, it thus appears that identifying the meaning of words takes at most something like one-third of the mental processing needed for reading. Even this estimate may overstate how much resources word identification takes: while word identification takes some time, it doesn't necessarily take resources away from other processes.

How Does the Processing of Words Relate to the Processing of Letters?

Physically, the description of a word (especially a printed word) seems obvious. If a two-year-old girl gave you a book and asked you to show her what a word was, your task would be quite simple. You would explain that the words were the physical entities between the spaces, and you could go on to explain that the letters were the little units inside words that were separated by the smaller spaces. (The task of explaining what letters are would be substantially more difficult with handwriting.) Other units, such as syllables, are not physically marked in any clear way, so that you would not be able to demonstrate what a syllable "looks like" to the child.

All of the above underscores what may seem obvious; words are units of meaning and so delineating them is important. Moreover, since English is an alphabetic language, delineating the letters is also important in the orthography. In an alphabetic language, it seems obvious that letters must be natural units in the *perception* of words. However, when one starts thinking hard about the question, it is not so obvious that letters are used as units in the perception of printed words, especially for skilled readers. Since the process of word recognition for skilled readers is so fast and automatic, it is possible that the process of letter identification is bypassed. In fact, Smith (1971) has claimed that skilled readers identify English words pretty much the same way they identify a picture: they recognize the word as a visual pattern through visual features and the fact that it is composed of letters is irrelevant to the perception of a word. That seems like a pretty extreme position, yet it is not easy to refute without experimental data.

The position at the other extreme also has its adherents (e.g., Gough 1972). Gough claimed not only that letters are used to recognize words but that words are read letter-by-letter serially from left to right, and the reader encodes the word as the sequence of letters. While this view has been criticized as being unfeasible because the processing would be too slow, experimental data suggest that letters can be scanned at about 10 msec per letter (Sperling 1963). Thus, the typical reading rate of 300 words per minute is not inconsistent with such a scanning process.

As you might suspect, we believe the truth to be in between these two extremes. In the remainder of this section we will describe two experiments that we believe rule out these extreme views, and then we will propose a relatively simple model of word processing that is consistent with these two experiments. The remainder of the chapter will then use the framework of this model to discuss word perception.

Ruling out the serial letter-by-letter model An obvious consequence of the assumption that letters are processed serially in order to perceive a word is that a single letter should be processed more quickly than a word. Sperling's (1963) experiment cited previously gives one estimate of the time to process each letter. He found (roughly) that if an array of unrelated letters was exposed for 10 msec (followed by a pattern mask) 1 letter could be reported, 2 letters could be reported given a 20 msec exposure, 3 letters reported given a 30 msec exposure, and 4 letters given a 40 msec exposure. (After this, short-term memory limitations came into play, and not many more than 4 letters could be reported even with longer exposure durations.) This experiment suggests that random letters are processed serially at the rate of 10 msec per letter.

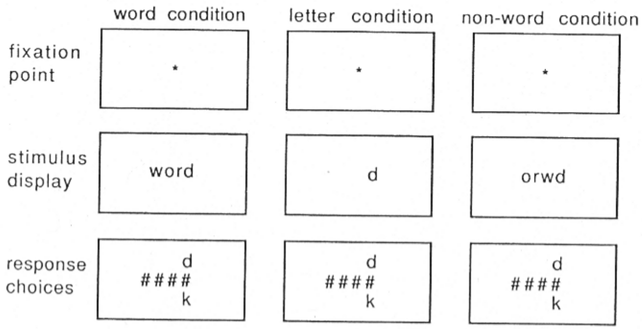
If letters in words were also processed serially, what would one predict for the time to identify a word? Since words are not composed of unrelated letters, then one might expect that processing a 4-letter word would require less time than processing all 4 letters of an unrelated letter string. For example, if the first letter of the word was a *T,* then the reader might expect the second letter to be an *H* or an E, and if it were, processing time for the second letter could be shortened by this expectation. But, and this is the important point, the serial letter-by-letter model of word processing predicts that it should take longer to process words than individual letters because processing the letters after the first letter will take some time, even if each of these letters is processed more rapidly than the first.

More than 100 years ago, Cattell (1886) tested this prediction by briefly exposing words and letters and asking subjects to report what they saw. In fact, subjects were better able to report the words than the letters! His experiment has several flaws, however. First, there was no mask presented after the words or letters, so that while the words and letters physically disappeared, the iconic representation (see Chapter 1) of the stimuli undoubtedly remained. Thus there may not have been much time pressure in encoding the visual information, and the errors observed may have been largely failures of short-term memory. Second, there was no control for guessing. Subjects may not have actually seen all the letters of a word but been able to guess fairly well from seeing a part of a word what the whole word was.

These factors would not explain why words were reported more accurately than letters, however. The first merely states that the icon may have lasted long enough to allow adequate time for both words and letters to be encoded. The second argues again why there may not have been much of a difference between words and letters. To explain why words were actually better than letters, one needs another factor. The one generally posited is that *words are more memorable than letters.* First of all, the words used in these experiments were usually concrete nouns, which should be more meaningful than letters. Second, in an experiment with many trials, the words are changed from trial to trial, whereas letters would have to be repeated. It could become very confusing to keep track of which letters were seen on which trials.

Cattell's experiment lay dormant until the cognitive psychology revolution of the 1960s, when Reicher (1969) replicated the experiment, attempting to remove possible artifacts (see also Johnston 1978; Johnston and McClelland 1974; Wheeler 1970). First, he used a pattern mask to control the effective stimulus presentation time. Second, he changed the task slightly to eliminate guessing as an explanation and to minimize the effect of memory. He presented a *target stimulus* (either a word such as *WORD,* a letter such as *D,* or a scrambled version of the word such as *ORWD*). The target stimulus was followed by the pattern mask and 2 probe letters, 1 above the critical letter of the target word and the other below it (see Figure 3.2). In this example, the probe letters would be *D* and *K,* and they would appear above and below where the *D* had been in either *WORD, ORWD,* or *D*. The

FIGURE 3-2 Sample displays from the experiment by Reicher (1969).



letters were chosen so that either would spell a word when combined with the other letters of the display (in this case, *word* or *work*). Thus, knowing or assuming that the target stimulus was a word would not allow the subject to perform above chance (50 percent correct) in the experiment.

Reicher found that the critical letter in the target word was reported more accurately than the same letter in isolation. He also found that the letter in isolation was reported with about the same accuracy as the letter in the nonword *ORWD*. Thus, Cattell's phenomenon appears to be real: letters in words are actually identified more accurately than letters in isolation. The phenomenon forces one of two conclusions. First, the serial model can't be correct if the errors in the experiment are due to limited encoding time forced by the mask, since it should take longer to encode 4 letters than 1. Alternatively, the serial model could be salvaged if one argued that the mask did not really impose perceptual difficulties and that the differences in the experiment were due to errors in short-term memory. While subsequent experiments have demonstrated that memory (in some sense) plays a part in the "word-superiority" effect (Mezrich 1973; Hawkins et al. 1976), there are several reasons to believe that Reicher's result is perceptual. First, the exposure duration is critical: Performance is about 100 percent even for the 4-letter nonwords (Adams 1979) if the target stimulus is exposed for 80 msec before the mask appears. Thus short-term memory, in some simple sense, can't be the limiting factor in these experiments. Second, the phenomenon that is probably most devastating for the serial model is that the *D* in the random string of letters *ORWD* is identified as accurately as the *D* in isolation. Since this condition should pose memory difficulties at least as great as the isolated letter condition, it would appear that there is no way that the serial model could explain why the target letters are reported equally accurately in the two conditions.

Words are not visual templates Reicher's experiment rules out the hypothesis that letters in words are processed serially, but what does it establish? One possibility is that the letters are processed in parallel and then are wired up to word detectors that automatically "fire" when the letters do so (see Figure 3.3). An attractive feature of this model is that it explains why letters in nonwords are processed as well as letters in isolation. However, it appears not to be able to explain why letters in words are actually processed better than letters in isolation. The other possibility is that skilled readers develop special templates for words and that they actually process words and letters by different systems that have nothing to do with each other (with the word system operating more rapidly). The latter alternative appears to be more attractive since it directly explains the word-superiority effect. However, we will argue (not just to be perverse) that the visual template model is almost certainly wrong and that versions of the other model can in fact explain the word-superiority effect (as well as much of the data from the word processing literature).

To get a sense of what is wrong with the visual-template model, let us consider what it would mean if the perception of a word did not go through the perception of its component letters. It would mean that there is some sort of visual template of the word DOG that would be compared against all visual patterns and if they were sufficiently similar to the visual template, then the visual pattern would be recognized as "dog." However, there is a problem here. We can recognize *dog* if it is written as *dog*, "DOG," and in innumerable different typefaces (including fairly strange ones used on computer screens). It is hard to see how one template or model could be used for all of these different varieties of typefaces. There are several ways one could attempt to salvage the model. First, one could postulate that there are separate templates for each variety of *dog* one encounters. Alternatively, one could posit that there are collections of "features" that define each word. Neither alternative seems very attractive in explaining how someone would be able to recognize words in a new typeface. In general, if the typeface is fairly novel the reader experiences initial difficulty, but then after a little while can read it about as well as a familiar typeface. One could say that a word printed in the novel typeface was close enough to the template of the word or had enough critical features so that it could be recognized. After sufficient experience, presumably the template would be altered or the critical features adjusted so that word would be recognized in the new typeface.

The problem with either of these explanations is the following. The reader has read several pages of the new typeface and has gotten used to it,

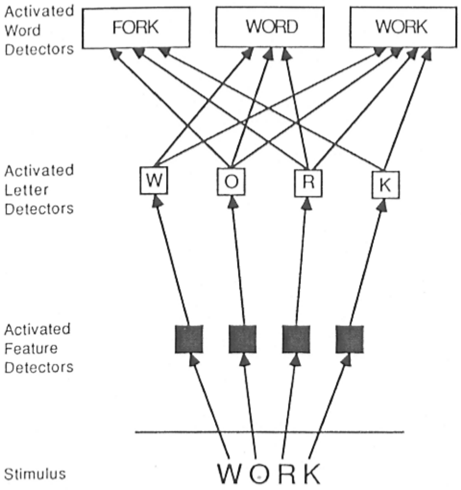


FIGURE 3-3 **Schematic model of a parallel model of word identification. In each position, letter features are identified, which in turn lead to the identification of letters and then the word. In this diagram, several words are activated, but the appropriate word, WORK, is activated most strongly, since it receives activation from all four component letter detectors. The model will subsequently be made more complex.**

but has not encountered the word lion *yet*. Both these models would predict (contrary to what would happen) that *lion* in the new typeface would present difficulties because its template (or critical features) have not yet become adjusted, since the perception of each word is separate from the perception of other words and the perception of its component letters. In contrast, the model that says that words are processed through their component letters would have no trouble with this kind of learning. Either a new template is learned for each letter in the new typeface or the critical features of the letters are adjusted when one encounters the new typeface, and once that happens, any combination of those letters should be perceived with little difficulty. (That is basically the alphabetic principle.)

Perhaps the clearest demonstration of this ability to overcome novel forms of words involves the use of text written in AlTeRnAtInG cAsE sUcH aS tHiS (Smith, Lott, and Cronnell 1969). In fact, given a little practice, readers encounter surprisingly little difficulty with such text and can read it as fast as normal text when the sizes of all the letters are equated and only a little slower than normal when the text looks like the above. The Smith et al. result could be criticized because measuring reading rate may not be a sufficiently sensitive measure of processing difficulties; if the text is easy enough, subjects may be able to guess words and letters in the alternating case condition well enough to get by. However, Coltheart and Freeman (1974) demonstrated that lexical decision times for words written in alternating case were only 12 msec slower than for words written in lower case.

How would these "whole word" theories attempt to explain this result? The template theory seems totally inadequate, since most subjects clearly have not encountered words in that form before. Smith, who espouses the feature theory, argued that the features for recognizing words are independent of the case of the letters. This seems pretty far-fetched, as it is hard to see what features certain upper and lower case letters such as *A,* and *a, R* and *r,* or *D* and *d* have in common. On the other hand, if the perception of words goes through the component letters, it is easy to see how the case manipulation poses little difficulty. (What we would have to add to the model is some explanation of why there was any difficulty with the alternating case text.) The fact that the case of the letters appears to be largely irrelevant to the perception of words has led to the widespread acceptance that word identification proceeds largely through case- and font-independent *abstract letter identities* (Besner, Coltheart, and Davelaar 1984; Coltheart 1981; Evett and Humphreys 1981; Rayner, McConkie and Zola 1980). It also leads to the conclusion that *word shape* is not an important cue for word identification (see Paap, Newsome, and Noel 1984, for corroborating evidence).

A model of word perception Let's briefly assess the argument so far. The experiments in tachistoscopic word recognition appear to rule out the hypothesis that letters in words (or even in nonwords) are processed serially. Second, it appears unlikely that the rapid perception of words is due to a visual template or set of features for each word (there will be additional evidence in a while against template theories). This appears to leave as the only reasonable contender the hypothesis that letters in words are processed in parallel and the encoding of a word goes through the component abstract letters. Since words are perceived better than either individual letters or strings of unrelated letters, however, one needs to say more than that letters are processed in parallel: In some way, the encoding of letters must be mutually facilitative.

Let us start out with a proof by blatant assertion: There are computer simulations of such a parallel encoding model that in fact can predict the word-superiority effect (McClelland and Rumelhart 1981; Paap et al. 1982; Rumelhart and McClelland 1982). While we make no claims that this type of model must be the way that the brain perceives words, it is at least plausible and is reasonably consistent with what we know about word perception. (However, it will need to be complicated as we will discuss later.) It is not easy to explain how such a model can explain the word-superiority effect, so if the following does not satisfy you, we recommend either going to the originals or taking our word for it. We will use the Paap et al. model, since it is somewhat easier to see how it works.

As can be seen in Figure 3.4, the model is quite simple. At the first stage of analysis there are visual features such as horizontal lines, edges, and corners. These feed into letter detectors at the second level of analysis and the letter detectors, in turn, feed into word detectors. (To simplify things, we leave out another level which would distinguish between case- and font-dependent letter detectors for *A, a,* a, and ***a*** which would all feed into a single abstract *A* detector.) The detectors work pretty much the way that individual neurons (nerve cells) work: if there is enough activity in the neurons feeding into a neuron, the neuron itself will become active. Thus, an *A* would be recognized in a given location if the features that constitute one of these representations of an *A* (such as horizontal lines, slanted lines, acute angle pointing up) are excited by the visual input. (Novel typefaces would presumably be learned either by creating new detectors for each of the letters or by modifying the features of one of the old detectors to be able to recognize the new ones as well.)

An important aspect of the model is that the activation is not all-or-none. If four features of a capital *A* are active, then the *A* detector is more active than if three features are, which in turn would produce more *A* activity than if two *A* features are active. The same applies at the word level. The *dog* detector would not need activity from all of the *d, o,* and *g* detectors to start its activity: there would be some activity in the *dog* detector given any activity in the component letters, and the more activity in the component letters, the more activity in *dog*. (Obviously, the word detector would have to know the spatial position of the component letters to be able to distinguish *dog* from *god*.)

Let us return to the word-superiority effect. Recall that the subject is briefly presented with a word or letter followed by a mask and errors are

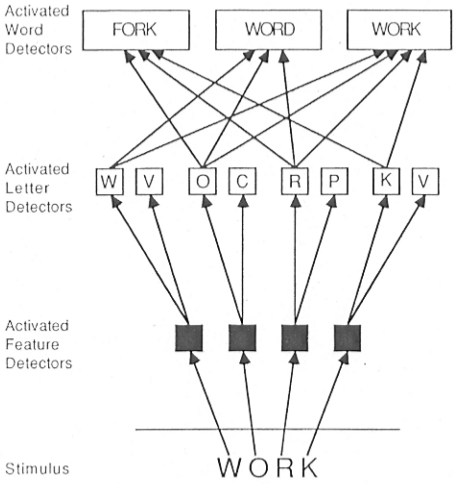


FIGURE 3-4 **Schematic model of a parallel model of word identification. As in the previous figure, there are three stages, feature detectors, letter detectors, and word detectors. However, in this version, more than one detector is activated at each level (e.g., *W* and *V* in the first letter position and several word detectors are activated). In a more realistic model, even more letter and word detectors would be activated. Moreover, what is not shown in the figure is that the activation for the detectors would not be equal: the detector for what is actually presented would get the highest activation, and the amount of activation for the other detectors would depend on their visual similarity to what was actually presented. In the model, only excitatory connections are shown. However, most such models would have inhibitory connections as well (e.g., evidence for *K* in the last position would inhibit firing of the detector for *WORD*).**

made. Thus, not all the visual features have been adequately processed. For the purposes of the argument, let us assume that a particular exposure duration sets the level of visual information so that each letter detector is only excited to 50 percent of its maximum level and the word detector is excited to a higher level (let's say 80 percent of its maximum level). How? Because of the redundancy of words: only a small fraction of the possible combinations of letters are words in the language. Thus, even a hint of *w*, *o*, *r*, and *k* may make the stimulus much more likely to be WORK than any other word (see Figure 3.4). In the Paap et al. model, letters are identified by reading off either the letter-detector activation directly or the word-detector activation. Since letters in isolation would have to be identified solely on the basis of letter-detector activity and letters in words can be identified either on the basis of word-detector or letter-detector activity, letters in words will be identified more accurately than letters in isolation even though the letter-detection level comes before the word-detection level!

We have sketched how the model explains the word-superiority effect. We have glossed over complexities, the major one being how the decision is actually made on what letter is actually in a given spatial position. The decision is complex since it must pay attention to activity at both the letter- and word-detector levels and it must also sort out competing activity (e.g., if a *K* is present, the "K" detector would be active, but so would the "H" and "N" detectors, since those letters share visual features with *K*). You may still find the explanation baffling, in that it is still hard to see how the letter in the word can be identified better than the letter in isolation *even when guessing has been controlled*. You may just have to take our word for it that such a model can in fact predict such a result. The model even makes good quantitative predictions about the size of the difference between identification of letters in words and letters in isolation.

There is a point that must be emphasized: The data we have been discussing is the percentage of correct identifications in a forced-choice between 2 letters. As the model makes clear (see Figure 3.4), letter identification is quite indirectly related to the activation of words and the activation of letters. Moreover, so far neither we nor Paap et al. have committed ourselves to what in the model would correspond to "perception of a word" or "perception of a letter." The simplest possibility is that the word (or letters) is perceived if the excitation in a particular detector exceeds a certain threshold, let's say 75 percent. That assumption leads to what may sound like an absurd prediction, namely that a word could be perceived before its component letters are perceived. However, the prediction may not be absurd. First, misspellings of words are sometimes (incorrectly) identified as the correctly spelled word and the reader is unaware of the misspelling (Ehrlich and Rayner 1981). Hence, the "word" is in some sense perceived before the component letters. Similarly, words can be misperceived (i.e., perceived as other words). Perhaps more strikingly, one can often be aware that a string such as "diffrence" is misspelled but take a while to discover what is wrong. All three examples suggest that perception of words and letters is somewhat independent processes (We will come back to these phenomena when we discuss the effects of context on word perception and proofreading in chapters 7 and 12.)

WORDS, SUBWORD UNITS, AND SOUND

We will assume for the rest of the book that the Paap et al. model gives an essentially correct picture of the relationship between word and letter identification: The letters in a word are processed in parallel and lead to identification of the word. But is the access to words that simple? In the simple model presented so far, each word is represented as a unit as in a mental *lexicon* or dictionary and access to each *lexical entry* or word detector in the lexicon is by a *direct visual route* in which the only psychologically real subword units are letters. In such a model, all other information about the word, such as its meaning, pronunciation, and etymology, is available only when the lexical entry has been accessed (just as the information about the pronunciation and meaning of WORK is found when the reader gets to the entry for "work" in a dictionary). Are rare words processed in the same way? What about words that the reader has never seen before? If the reader encounters something like MARD, does it have to be read letter by letter as if it was a random letter string or do its wordlike properties make it easier to perceive?

The issues are complex, and there is no definite answer to many of the questions. While the Paap et al. model in fact does a remarkably good job of handling a lot of data, we believe that there are phenomena that suggest that it is too simple and that there are other processes involved in the identification of words. In particular, we want to argue that a process in which the reader also uses the letters of a word to access the sound and then the meaning is also important in word recognition. That is, we believe that there are two routes to the lexicon—the *direct route* (exemplified by the Paap et al. model) and an *indirect route* (going through sound)—that are used by skilled readers to access the meaning of a word. While the central issue in word processing is how the meaning of a word is identified, it will be simpler to introduce the issues involved by asking how the pronunciation of a word is performed. We will even make a second detour initially to discuss how *pseudowords* (nonwords, like MAFER, that look like words) are pronounced. It will take a while, but we will return to the central issue—how we access a word's meaning.

The simplest possibility for how words are pronounced is that the "motor program" to pronounce the word is stored at the word detector or lexical entry. (An equivalent metaphor is that there is a direct link or pathway from the word detector to the motor program.) But how would such a model predict that a pseudoword such as MARD is pronounced? It would appear to require a totally different mechanism unless you posit that all conceivable words (even those never seen such as MARD) also have lexical entries. Since that alternative seems unlikely (and has not been seriously entertained), a more analytic mechanism is needed. Such a mechanism either would have to apply some rules of English pronunciation to novel letter strings or would have to be able to apply analogical rules from known words to novel strings.

Thus it appears that there is a sharp discontinuity between how words and nonwords are processed. What would a system such as that of Paap et al. do when it encountered a nonword? Would it first try to process the letter string as a word, and only when it decided that no word detector was sufficiently activated, would it then engage these other mechanisms for processing nonwords? If so, it would seem that processing a nonword should take a lot more time than processing a word. The facts of the matter are quite different, however. In the Reicher task, there is not only a word-superiority effect but a pseudoword-superiority effect! Letters in pseudowords like MARD are processed more accurately than letters in isolation (Baron and Thurston 1973; Hawkins et al. 1976). In most experiments, letters in words are processed a bit more accurately than letters in pseudowords, but there are several experiments in which there is no difference between the two (e.g., Baron and Thurston, but see Carr, Davidson, and Hawkins 1978). While it takes somewhat longer to pronounce pseudowords than words, it is still a very rapid process: The difference is 200 msec for unpracticed subjects but appreciably less for practiced subjects (Baron and Strawson 1976).

In many ways, this efficient processing of pseudowords should not be surprising. The organism should be equipped to deal with all plausible stimuli about as efficiently as ones actually encountered. This raises the possibility that we were wrong before and that words and nonwords are in fact processed by the same machinery. How can that be, since nonwords can't be processed by the same machinery that we have put forward to describe the encoding of words? We will argue that the way to resolve this paradox is that there are two sets of machinery used to process both words and nonwords, and thus our picture of word perception is incomplete. We will take a somewhat indirect route by first considering how nonwords would be processed.

Processing of pseudowords Since we have discussed the Reicher paradigm in some detail, we will consider the pseudoword-superiority effect before going on to the more complex question of how pseudowords (and words) are pronounced. One's immediate reaction to the pseudoword-superiority effect is that it would be impossible for a model such as that of Paap et al. to explain. However, as we shall see, the Paap et al. model was in fact constructed largely to explain the pseudoword-superiority effect! (On the other hand, the pseudoword-superiority effect certainly goes against the word-as-visual-template model.)

So far, the Paap et al. model could not explain how letters in pseudowords are identified almost as accurately as letters in words, but it can if one mechanism is added. As we will see, a similar mechanism will be added to explain pronunciation. The mechanism is as follows. When a letter string appears, it excites not only the lexical entry which is identical to it but its "neighbors" as well. Thus the stimulus WORK excites the lexical entry "work" as well as "word," "wore," "fork," and possibly other lexical entries that are visually similar (see Figure 3.4). The question of how *similarity* is defined is complex: The similarity of the stimulus and the word detector would undoubtedly depend on the number of letters that they had in common in the same positions and the visual similarity of the letters that differed, and might also depend on the position of the difference (e.g., a difference between the stimulus and the lexical entry in the first letter position might be more important than a difference in a later position). However similarity is defined (let us assume, for simplicity, that it merely depends on the number of letters in common), the more similar the letter string to the lexical entry, the greater the excitation. Thus, WORK would be identified correctly, since the excitation would be greatest for "work."

What about MARD? Even though it has no lexical entry, it has neighbors like "card," "ward," "mark," "mare," and "maid," and each of

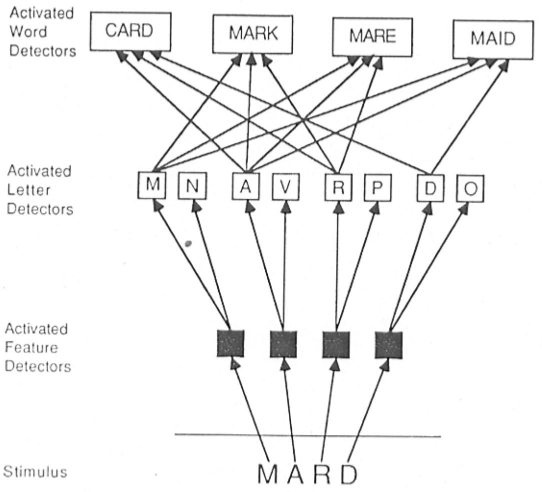


FIGURE 3-5 The same mechanism as in Figure 3-4 in the process of identifying a pseudoword. As in Figure 3-4, only some of the activated letter and word detectors are shown. What is also not shown in the model is an additional stage, which would "poll" the activated word detectors to find the most popular letters at each letter position (see text). There are more complex versions of this model, in which there are "letter cluster detectors" as well (e.g., a detector for WO in the initial position). As subsequent discussion will indicate, these additional detectors are needed in order to account for all of the phenomena involving pseudowords.

these entries would be excited to an appreciable degree (see Figure 3-5). To explain how a letter is detected, we need to make another assumption, namely that the subject can read letters off the lexical entries by polling them letter by letter. In the above example, most of the excited lexical entries would "vote" for *m* in the first position, *a* in the second position, *r* in the third position, and *d* in the fourth position, so that the letters of the pseudoword MARD could be read off of the word detectors. (The "read off" process would be in parallel, though). In the model, letters in real words are read off the word detectors in the same sort of way, so that there is some "noise" from the wrong votes of neighbors. However, it is still probably difficult to believe that this process would in fact lead to almost as accurate letter identification for pseudowords as for real words, so you will have to take our word that the model can in fact predict such a result. (The McClelland and Rumelhart model has a slightly different metaphor; it postulates feedback from the word detectors to the letter detectors, but this seems equivalent to the process of reading the letters off the word detectors described above.)

It thus seems that the direct lexical-entry model, if suitably modified, has a lot more flexibility and power than it appeared to at first. Note that while there are no subword units in the model other than letters, the process of "reading the letters off the word detectors" is analytic: the polling of the word detectors is letter by letter. This point will become even more important when we next consider how words (and pseudowords) are pronounced.

PRONOUNCING WORDS AND NONWORDS

When one thinks of pronouncing words and nonwords, it appears that there must be two different mechanisms involved. First, let us consider *irregular words,* those whose pronunciation could not be derived from applying any general knowledge of English pronunciation, e.g., ONE, TWO, CHOIR, WOMEN. In the case of ONE, for example, it appears that when the string is pronounced /won/, it must be because the lexical entry for the word has been accessed which in turn allows access to information about how to pronounce it (e.g., an auditory image, a motor program). A pseudoword such as MARD, however, can't be pronounced this way, unless one wants to postulate lexical entries for all conceivable words. It thus appears that a different system, perhaps a set of rules, allows a person to pronounce this or any new wordlike string of letters. In the case of MARD, the rules would be relatively simple: the three consonants can all be pronounced by general rules for the pronunciation of those consonants; the pronunciation of the vowel would be a bit more complex, however, since one would have to know the rule that makes *A* sound like /ah/ when followed by an *R*. Following this analysis, it would appear that *regular words* (i.e., those whose pronunciation is given by rules of English), such as TREE, could be pronounced by either of these mechanisms: the direct lexical access of pronunciation or the rule-generated system. It should be kept in mind that the rules that we are talking about are for the most part unconscious; the reader does not consciously apply them when pronouncing strings of letters in about half a second and may not even have conscious access to some of them upon longer reflection.

If this commonsense analysis were correct, what would we predict about the difficulty, or speed, or both, of pronouncing regular words, irregular words, and pseudowords? While the assumptions so far would not make any predictions about the relative ease of pronouncing irregular words and pseudowords, we might expect that irregular words are pronounced more quickly since the *direct lexical route* (i.e., the one in which the pronunciation is stored in the lexicon) would be faster than the rule system because the rules might be complex. This is in fact the case. What about regular vs. irregular words? The model we have outlined so far would predict that regular words are pronounced faster *if the rule system is used in the pronunciation of words*. There are two possible mechanisms for this: (1) The pronunciation of irregular words such as ONE could be slowed down by conflict—that is, the lexical system generates the pronunciation /won/, the rule system generates the pronunciation /own/, and there is conflict as in the Stroop paradigm discussed earlier; (2) The pronunciation of regular words could be speeded up by having two independent mechanisms generate the same pronunciation, and thus the response would be strengthened relative to responses that had only one mechanism feed into them. Needless to say, either or both of these hypotheses could be true.

In fact, a reliable finding (Baron and Strawson 1976; Seidenberg et al. 1984a) is that regular words are pronounced more rapidly than irregular words, so that it appears that both mechanisms are operative in the pronunciation of words. (If only the direct lexical mechanism were operative, then there would be no difference between regular and irregular words.) Moreover, there is little difference in naming time between regular and irregular words if they are frequent in the language, whereas there are clear differences between the two if they are less frequent in the language (Seidenberg et al. 1984a). That makes sense in terms of the dual-mechanism explanation. The indirect mechanism should be little influenced by the frequency in the language since it is based on "rules." However, the direct look-up should be faster for high-frequency words than for low-frequency words. Thus for high-frequency words, the direct route is fast enough to make the indirect route irrelevant. For lower-frequency words, the two routes are of more comparable speed; thus low-frequency regular words should be faster than low-frequency irregular words since there are two ways to access the sound.

This argument has failed to convince many researchers (e.g., Glushko 1979; Humphreys and Evett 1985) that there is in fact a rule-governed system that can access the pronunciations of words in time to affect fluent naming of words for reasons we will discuss presently. However, there are two issues that are often confused in this controversy that must be separated: (1) Are there two systems involved in word pronunciation? (2) If so, is the system that is not the direct lexical route a rule-governed system? We feel that most of the criticism is related to point (2) and not to (1).

One reason that the antagonists of a rule-governed system are unhappy with postulating such a system is that it is far from clear how to specify the rules of pronouncing English. For example, is the word DUMB regular or irregular? If it is regular, many of the rules would have to be fairly context-specific (e.g., "*B* is silent after *M*"). If one didn't allow such rules, then the pronunciation of *B* would be independent of context, and the silent *B* would be irregular. We probably need such contextual rules in our system to handle certain common things such as the ubiquitous "silent e," which is not itself pronounced but lengthens the sound of the previous vowel (e.g., FATE vs. FAT). However, how do the rules deal with COMB, COMBING, and COMBINE? We could call either COMBING or COMBINE irregular, which seems unsatisfactory, since neither seems irregular. If we stay within a rule framework, however, we now have to postulate more complex rules: the *B* in COMBING is silent because it is part of the syllable COMB, whereas the *B* in COMBINE is pronounced because it is in a different syllable than the *M*. Presumably, we would need a rule that knows that ING in COMBING is a suffix and for COMBINE knows that COM is a prefix. (The task of figuring out that COM is a prefix is made more difficult by the fact that BINE isn't a word.) For longer words, there are also difficult questions about how to construct general rules for assigning stress to syllables without creating tons of irregular words.

In sum, if we postulate a rule system for generating pronunciations, we have one of two options. Either we can postulate simple general rules (Coltheart 1978; Simon and Simon 1973), in which case a large percentage of English words will be irregular (50 percent by the most extreme count) or we can postulate very complex rules that seem more consonant with our intuitions about regularity (Venezky 1970) so that few words (5 percent) are irregular. However, no one has proposed a rule system that is particularly satisfactory, so that if we accept the existence of a rule system, we are largely accepting an article of faith. Moreover, as the above example indicates, such a system will likely have to have information about specific lexical information (e.g., ING is a suffix, COM is a prefix, BINE is a stem that can be combined with prefixes).

The critics of a rule system propose that a lexical system looking quite a bit like that of Paap et al. can handle the pronunciation of both words and nonwords. Let's see how it works. First, consider a nonword such as MARD. It excites a neighborhood of lexical entries as before, such as "ward," "card," "mart," "mark," "maid," "mare." Each of these lexical entries excites, in turn, the pronunciation of that word. The pronunciation of MARD is then generated by polling each unit of sound (phonemes) in turn. Thus, most of the neighbors vote for an /m/ sound in the first position, an /ah/ sound in the second, and so on. Thus, according to the model, the apparently rule-governed behavior of generating pronunciations to novel strings is not due to abstract rules but to computations of knowledge contained in the lexicon (Brooks 1977). This type of model has been termed an *analogical model,* since the pronunciation of the novel string is purportedly generated by analogy with known words. We are not completely happy with the term since it does not really capture the type of computation done on the lexicon to derive a pronunciation. However, for lack of a better term, we will refer to these models as analogical models.

This is in fact a very clever system for producing rulelike behavior, but (a) is there evidence for it? (b) does it really work? The strongest evidence for such a model is that there appear to be lexical influences on the pronunciation of nonwords (Glushko 1979). That is, a nonword such as BINT, which has word neighbors that are inconsistent in their pronunciation (e.g., PINT, HINT, MINT), takes longer to pronounce than one such as TADE whose neighbors are consistent. Similar effects are found with words (e.g., Glushko 1979). Words whose pronunciation is regular but that have irregular neighbors (e.g., GAVE which has HAVE as a neighbor) take longer to pronounce than those that have no irregular neighbors (e.g., COAT). Proponents of a rule system could argue that these effects are produced by differential strengths of rules: rules that are consistently applied are stronger than those that are not.

The problem with the simple analogical system described above is that it doesn't really work. One difficulty is that nonwords such as JOOV are easy to pronounce in spite of the fact that they have no near neighboring words (Coltheart 1981): There is no word beginning with JOO and none ending with OOV. To generate a pronunciation with such a model, one has to postulate that words such as GROOVE and JOIN are neighbors, but this stretches the idea of neighbor quite a bit. In addition, the analogical mechanism would have to be quite clever in knowing how to "line up" the appropriate elements so that the extra phoneme in GROOVE is taken care of. There would be similar problems in pronouncing most longer nonwords such as MARDTORK or anything from "Jabberwocky," which would have virtually no word neighbors. Accordingly, the analogical model actually proposed by Glushko was complicated by expanding the lexicon to include bits and pieces of words. In his model, virtually all subsets of words were in the lexicon (e.g., the "lexicon" would include WOR . . . , WO . . . , . . . ORK, W . . . . , . . . RK, in addition to WORK), and all of these units would have pronunciations attached. When a word or nonword appears, all of these units (both the word and subword units) are activated and a pronunciation is somehow computed from all of these firing units.

As you have probably surmised, such a model seems pretty close to a rule model. It is not too different to say that the *A* in MARDTORK is pronounced /ah/ because of a rule indicating that *A* followed by *R* is pronounced that way and to say that there are lots of ". . . AR . . ." entries excited which dictate the pronunciation /ahr/. Since neither model has been worked out in enough detail, it is hard to say whether they are just two different metaphors for thinking about "rules" or whether there is a principled difference between them. In either case, it is clear that the way that the human figures out the pronunciation of a new nonword is a very complex computation. Our present sympathies go with the analogical system, since it is a more satisfying explanation of the phenomenon than one in which complex rules are stored but are not open to consciousness (sometimes even after reflection). In addition, the notion of "rules" suggests some sort of serial combination process which is unlikely to be able to operate in the brief time span of word identification.

Some of the analogue theories have pushed the claim a bit too far, in our opinion, in arguing that there is really only one system that handles all pronunciation (Glushko 1981; Humphreys and Evett 1985), even the pronunciation of words. In the one-system view, when a word such as ONE is encountered, the lexical entries for ONE, all its parts, and all its neighbors' parts are all excited and the pronunciation is based on that total input. To account for the fact that ONE is pronounced correctly nearly all the time, the theory would have to postulate that the excitation of the item itself is enough stronger than the excitation of the neighbors, so that the conflicting pronunciations of the neighbors can only slow down but not overrule the pronunciation offered by the lexical entry "one." The important thing to keep in mind is that according to this view, the direct connection between a stimulus and its "own" lexical entry has no special status: it is only stronger than all the others.

This one-system view is problematic for several reasons. First of all, it would seem difficult to explain why instructions could change the output. If asked to pronounce HAVE, some would say /hahve/, but if asked to pronounce it according to the "rules of English," they could easily switch and say /hayve/. The most problematic result for such a view is that there are people with brain damage who seem to have selective damage for either the direct lexical route or the rule/analogy route. People termed *surface dyslexics* can come up with a pronunciation for virtually all words and nonwords (although understanding little); however, they mispronounce many irregular words by "regularizing" them, such as pronouncing ISLAND /izland/ (Marshall and Newcombe 1973). Thus, it appears that their problem is parsimoniously explained by postulating that the rule/analogy system is intact while the direct system is damaged. (It is not fully damaged, since not all irregular words are regularized.) On the other hand, *phonemic dyslexics* pronounce most words correctly, but are virtually unable to pronounce nonwords (Coltheart 1981). Their problem is parsimoniously explained by saying that their direct system is relatively intact but their rule-analogy system is almost completely damaged. The proponents of the single-system view argue that since the data from these patients are a bit more complicated than is presented here, it is not completely conclusive evidence for two systems. For most people including ourselves, however, the evidence is conclusive enough that the burden of proof is on the single-system theorists to show why the two-system theory is wrong. (We will discuss the dyslexic data again in Chapter 11.)

To summarize this long, but necessary, discussion of pronunciation, it appears that there are two systems active in pronouncing words: a direct lexical system, in which the pronunciation of words is looked up in the appropriate lexical entry, and a rule/analogy system, whereby the pronunciation of a word or nonword is generated either by a system of rules or by a complex computation on a set of lexical and sublexical neighbors, or perhaps a combination of both. Both the direct lexical system and the rule/analogy system appear to be operative at all times, since the pronunciation of regular words takes less time than the pronunciation of irregular words.

ARE WORDS ACCESSED THROUGH THEIR SOUND?

If we consider Figure 3.6, which summarizes our discussion of the previous section, we see that both the direct lexical and rule/analogy systems are active in determining how a string of letters is pronounced. We now consider the related question of which systems are involved when printed words access the lexicon. The central question of this section is raised by path 2-4

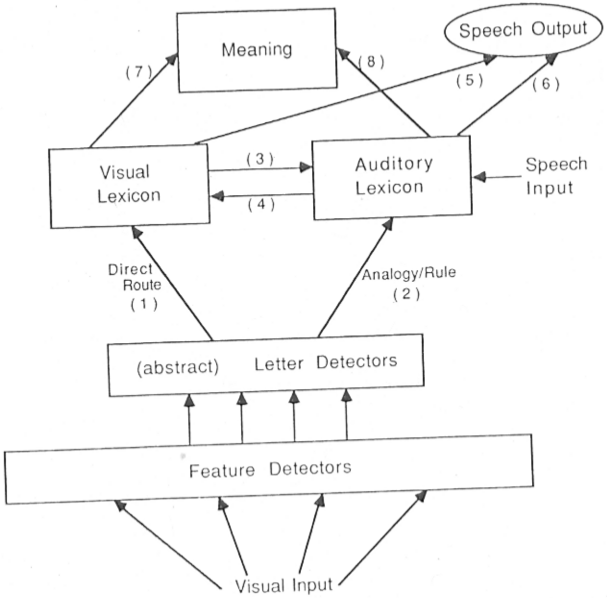


FIGURE 3-6 **Model of word recognition. Encoding of meaning would involve cooperative use of paths 1-7 and 2-8. Lexical decision would involve paths 1 and 2-4, although signals from the auditory lexicon would interfere with certain responses (as with pseudohomophones). Naming words would primarily involve routes 1-5 and 2-6, while naming pseudowords would primarily involve route 2-6. The analogy/rule route would include letter cluster detectors (or some equivalent) not shown in Figure 3-5.**

in the figure: Namely, does the rule/analogy system succeed in generating a pronunciation of a word in time for the sound generated to aid in lexical access? According to one extreme view (Gough 1972), the rule system is the only route by which lexical access occurs. (This can't be true, however, since readers would be unable to discriminate THERE from THEIR.) However, perhaps the most common view today is the opposite extreme: virtually all lexical access is by the direct visual route; the route going through rules (or analogies) to sound and then to the lexical entry plays a minor role, if any, in the access of word meanings. That position is reasonable, but not the only one possible. Let us consider the evidence.

The test of whether the rule/analogy system is involved in lexical access has focused on two empirical questions. The first is whether (as with pronunciation) the time to access the meaning of a regularly spelled word is less than that for an irregularly spelled word. If the rule system were involved, then one would expect a difference, since the rule system should be able to access the lexical entry for regular words but not for irregularly spelled words. The second is whether lexical entries are activated either by homophones or *pseudohomophones* (i.e., nonwords which when pronounced have the same sound as a real word such as PHOCKS). In fact, both questions have usually been examined with the lexical decision task: the former by testing whether the response time for regular words is shorter than for irregular words, and the latter by testing whether response time (to respond "nonword") is longer for pseudohomophones than for other pseudowords. For ease of future discussion, we will term these differences, if obtained, the *regularity effect* and the *pseudohomophone effect*, respectively. There are other tests for indirect access to the lexicon which we shall discuss subsequently.

Unfortunately, the data on the regularity effect has not been consistent: some studies have found a regularity effect (Bauer and Stanovich 1980; Parkin 1982) and others have not (Coltheart 1978). The pseudohomophone effect, in contrast, has been obtained pretty consistently (Rubenstein, Lewis, and Rubenstein 1971; Coltheart et al. 1977), although there are some failures to find it (Dennis, Besner, and Davelaar 1985). The relevance of the pseudohomophone effect has been questioned, however, on two grounds. The first is that, while the pseudohomophone effect establishes that the rule/analogy system is involved in deciding the lexicality of nonwords, it does not establish that the system is involved in the lexical access of real words (Coltheart et al. 1977). Second, there are claims that the pseudohomophone effect may be an artifact: pseudohomophones may look more like real words than other pseudowords. The second criticism, however, can be countered because it has been shown that *deep dyslexic* patients (similar to the phonemic dyslexics discussed previously), who appear to have a grossly impaired rule/analogy system, do not show the pseudohomophone effect (Patterson and Marcel 1977).

One clear demonstration of an effect of sound-based codes on lexical decision times for words comes from an experiment by Meyer, Schvaneveldt, and Ruddy (1974). They showed that the time to decide whether TOUCH was a word was slowed down when preceded by a word such as COUCH, which induces an incorrect expectation about the pronunciation, compared to when TOUCH was preceded by an unrelated word. (If a word such as BRIBE is preceded by a rhyming word such as TRIBE, there is a small facilitation effect compared to when BRIBE is preceded by an unrelated word.)

However, even the Meyer et al. effect has been dismissed by some researchers who level a more fundamental criticism at this research: Lexical decision time may be fundamentally flawed as a measure of lexical access time. There are two fairly different versions of this argument. The first is that lexical decision time may not reflect the time for one lexical entry to reach a sufficient level of activation, but instead reflects the total excitation in the lexicon (produced by a set of lexical entries in the "neighborhood"). The second is that the lexical decision judgment may cause people to perform certain "checking" strategies, such as going through the sound code after entering the lexicon, that are specific to the task and are irrelevant to accessing the meaning of the word (Balota and Chumbley 1984, 1985; Chumbley and Balota 1984; Seidenberg et al. 1984b).

Since the lexical-decision task probably does tap processes other than lexical access, it would be nice to have some convergent validity for the Meyer et al. effect. What would appear to be the simplest way to show that sound accesses meaning would be to demonstrate the Meyer at al. effect in a categorization task (i.e., the time to decide whether a visually displayed word is a member of a semantic category such as tools or furniture). The problem with this is that categorization time depends on factors other than the time to access the word, such as how good an exemplar of the category a word is (e.g., it is easier to judge a ROBIN to be a bird than to judge a TURKEY to be a bird). These variables overpower the relatively small effect found by Meyer et al., so that while there is a hint of the biasing effect in the categorization task, it is fairly weak (McMahon 1976). Another attempt to generalize the Meyer et al. effect was by Treiman, Freyd, and Baron (1983). They used sentences such as "He made a nasty hasty remark." Their finding was that people were slower to read sentences containing two "off rhymes" than matched sentences which contained synonyms of the words. While the Treiman et al. result indicates that some sort of recoding to sound occurs during reading, it is possible that the interference effect does not occur during lexical access, but during a later, comprehension stage when the phrase is understood (see Chapter 6).

While the lexical-decision task is flawed, it appears that most of the effects found using it can be found with other tasks as well. Accordingly, it is probably worth a bit of effort to go back and try to explain the previous lexical-decision results better. Perhaps the major finding to be confronted is that the regularity effect is quite small (if it exists at all). This indicates that the direct visual route is the dominant one for accessing the meaning of words. This certainly makes sense for English because there are so many irregular words: If we relied mainly on going from the letters to sound by rules and then to meaning, lexical access would be error-prone. This raises the question of the role of the rule system in the access of a word's meaning. How can the rule (or analogy) system be involved and yet not lead to all sorts of errors? One possibility is that the direct visual route is the only one normally used, and the rule-to-sound system is brought in only when lexical access fails by the visual route. Such a model could explain why there are occasionally regularity effects: Some words in the experiment are not within the sight vocabulary of the reader, and those words are the ones to give a regularity effect. However, the model does not explain either the pseudohomophone effect or sound-biasing effect. In the former case, it is not clear why the sound of the nonword should matter if accessing the rule system occurs only after lexical access fails. In the latter case, the sound of the previous word should be irrelevant if the direct visual route is used first.

A more convincing explanation of the data (and probably the model most widely believed) is the "horse-race" model (Coltheart 1978; Meyer and Gutschera 1975). In the horse-race model, each system (the direct visual route and the rule/analogy-to-sound route) works independently to come up with a candidate for a word. Thus, if ONE is presented, the direct visual route would access "*1*" while the rule/analogy route would access something else, perhaps "own." For regular words, both routes would access the same entry and there would be no problem. For irregular words, there is a problem for the system to decide which candidate is correct. The usual assumption is that if there is a conflict the system has to access again and "recheck," although such procedures are usually unspecified. The way proponents of the horse-race theory resolve the problem is to view competition as a race in which the lexical candidate, or horse, that is fastest wins, and postulate that the visual horse almost always wins.

In some sense, this version of the horse-race model is not too different from the direct-access model, since the direct visual route is functionally the route that produces lexical access in almost all cases. Where it differs is that the second route is actively working from the moment that the word appears rather than being activated only when lexical access fails. This difference allows the horse-race model to comfortably explain the general effects found in lexical decision. One would expect regularity effects to surface only when the direct route is slow. If we make the reasonable assumption that the direct visual system's speed in processing a word is more affected by the frequency of the word in the language than the speed of the rule system, then the horse-race model predicts that regularity effects should appear mainly for low-frequency words (which we have seen is true). The pseudohomophone effect is also easy to explain. Since both routes are working from the beginning, the pseudohomophone will activate a lexical entry through the rule system, whereas other pseudowords will not, thus leading to more errors and slower times in judging pseudohomophones as nonwords. The sound-biasing effect can be explained if we assume that the rule/analogy system can be biased by previous activation. Thus, if COUCH biases the rule system to interpret the OU in TOUCH to be pronounced /ow/, then the rule system will fail to find an entry for "touch" and thus slow lexical access relative to when such biasing is not present.

A phenomenon that argues strongly that sound-based coding is used in accessing meaning comes from the categorical-decision task. Meyer and Gutschera (1975) compared the ease of rejecting *pseudomembers* of a category to that of *nonmembers*. For example, if FRUIT is a category, PAIR is a pseudomember (i.e., a homophone of a member), and ROCK is a nonmember. They found that subjects made more errors with pseudomembers (i.e., falsely classified pseudomembers as members of the category) and were slower to respond when they were correct. Unfortunately, in the Meyer and Gutschera experiment, the pseudomembers also had more letters in common with members so that it is not clear that this result is because of visual similarity or because the pseudomembers are homophones of a category member. This has been remedied by Van Orden (1987) who controlled for visual similarity. The effect is surprisingly large: pseudomembers that differed from members by only one letter (e.g., MEET) were falsely classified as members of the category FOOD about 30 percent of the time, while nonmembers that also differed by one letter from members (e.g., MELT) were classified as members only about 5 percent of the time. Since the words were visible for 500 msec, subjects were misclassifying words a quarter of the time when they were clearly visible!

While this result clearly implicates phonological access in getting to a word's meaning, there are two possible access routes. One is the rule/analogy system to sound to lexicon route, but the other is a lexicon to sound to lexicon route (e.g., MEET activates its lexical entry by the direct visual route, which in turn activates the sound of the word that then activates "meat"). While both routes are possible, it is not at all clear why the latter route would cause so much interference, since it would seem that it should be much slower than the direct route: there are two extra steps for MEET to access the meaning of MEAT than for it to access the meaning of MEET. While the access of "meat" could be speeded up by priming from the category word FOOD, it is hard to see how access of "meat" could catch up to the access of "meet" if sound is accessed *only* after accessing the lexical entry for "meet." A follow-up experiment by Van Orden, Johnston, and Hale (1988) provides additional evidence for an indirect route to meaning: The same-sized effect is obtained when pseudohomophones are employed in a categorization task (e.g., SUTE is classified as an item of clothing as often as HAIR is classified as an animal).

The discussion of the pseudomember effect makes it clear that many of the details of the horse-race model have not been clearly worked out. One aspect, in particular, that has not been specified carefully is what happens after the first horse has won the race. Presumably, in cases where one horse wins by a large margin, the lexical-decision response is already programmed so that the second horse is irrelevant. But what about in reading for meaning? Is there interference when the second horse accesses a conflicting lexical entry, or does the winner inhibit all other lexical entries? A second aspect is that the nature of the conflict between competing horses has not been carefully worked out. How exactly does the rule/analogy horse slow down processing in the case of irregular words without producing errors?

Since most people assume that these details can be worked out, the horse-race model has been accepted as the most plausible model of lexical access. However, there is a facet of the model that should not go unnoticed. It argues that for words within our sight vocabulary, the fact that the alphabet was invented largely to capture the sound of words is virtually irrelevant. In fact, in a language like English in which there are so many irregular words, it seems to argue that the main function of the rule system is to interfere with identifying words. Since we find this aspect of the horse-race model a bit troubling, let us propose a slightly different dual-access model that seems a bit more satisfying, in that it will suggest that the involvement of the rule system aids lexical access.

In the *cooperative-access model* (Carr and Pollatsek 1985), both the direct visual and rule (or analogy) to sound routes are accessing the lexicon (as in the horse-race model). However, they are each activating a candidate set as in the Paap et al. model. Thus, when ONE is presented, the rule system activates not only "own" but also a set of words that are similar in sound such as "on," "wan," "won." (The direct route also activates a set of candidates exactly as in Paap et al.) The lexical entry that gets the most summed activation from the two systems is then identified as the word. Thus the two systems cooperate in exciting lexical entries in the visual and sound neighborhoods of the word rather than each sending forth a single lexical candidate, or horse. In the cooperative-access model, the rule system adds activation to the correct entry, although in the case of irregular words, it will add more activation to competing entries. Whether this extra activation can facilitate lexical access (on the average) depends on details of the decision process, details that have not been formulated precisely so far. Thus, we can't argue for sure that the indirect route would be functional in aiding lexical access in spite of the irregularity of English. Instead, we would like to present a process in language comprehension that indicates that such cooperative access does go on, and that a route that is more error-prone or "noisier" can in fact help a system that is less error-prone.

When we comprehend speech, there are usually two "routes": the sound of the speech and the visual information about the position of the mouth (primarily the lips). Both clearly convey information, since (a) we can comprehend speech when we don't see the speaker and (b) some deaf people can make reasonable sense out of "speech" just from reading the lips. Moreover, we have the clear impression that looking at the speaker helps in comprehension. Yet it is clear that reading the lips alone is inadequate for fully understanding speech (certain differences in the speech signal are not reflected in the external appearance of the face and mouth).

Thus it appears that the lip system must help the sound system in speech comprehension in much the way we are arguing that the analogy/rule sound system helps the direct visual system in reading. Furthermore, there is evidence that the lip information is integrated with the sound information in such a cooperative analysis. McGurk and McDonald (1976) have shown that if subjects see a videotape of a mouth saying /ga/ while they simultaneously hear /ba/ on the soundtrack, they will perceive the sound as /da/. This makes sense in terms of a cooperative computation model, since "da" is similar to both "ga" and "ba"; as a neighbor of both, "da" may get more total excitation from the two channels (sound and visual) than either "ga" or "ba."

Another point that needs to be made is that most irregular words in English are not that irregular. Most of the ones studied in the regularity-effect experiments merely have one irregular sound, usually a vowel (e.g., PINT). Even in the case of more irregular words such as ONE, while the rules would completely miss the initial /w/ sound, they would correctly predict that the final consonant sound was /n/ and would get a reasonable approximation to the vowel sound. The regularity effect might only get a clear test if the language included wildly irregular words (such as if DROON were a word and it was pronounced /step/). In fact, there appears to be a clear difference in naming times for wildly irregular words such as CHOIR and regular words (Baron and Strawson 1976; Seidenberg et al. 1984a). Thus some of the inconsistency between experiments on the regularity effect may be because some experiments used only the mildly irregular words and got negligible effects as a result. In fact, Seidenberg et al. (1984a) report large differences between what they dub "strange words" (such as CHOIR) and regular words, and smaller differences between "normal" irregular words such as PINT and regular words. They claim that the effect obtained with strange words is not due to their unusual pronunciation but to their unusual orthographic structure (many are "loan words" from other languages). Thus, they want to conclude from the same data that there is no regularity effect on high-frequency words and only a small regularity effect on low-frequency words: Strange words are slower not because they are irregular but because they look weird. Unfortunately, because there appear to be few, if any, words in English that look weird but are not irregular or vice versa, there is no easy way to resolve this problem. One possibility is to test subjects like the deep dyslexics or phonemic dyslexics who presumably do not use the sound system to access the lexicon. If Seidenberg et al. are correct, these subjects should find strange words hard to access because of their unusual orthography, whereas if access of strange words is slow for normal readers because of the irregularity of their pronunciation, the difference between strange and normal words should disappear for these subjects.

CROSS-CULTURAL STUDIES OF WORD PERCEPTION

Throughout the second half of this chapter, we have focused on two major issues. We have discussed how letters within words are processed and we have discussed the role that sound plays in word perception. We have argued that letters within words are processed in parallel, and presented a model of how letter recognition and word identification interact. We have also suggested that there are two routes to the lexicon, one that goes directly from the printed letters to the lexicon (a direct route) and one that involves initially transforming the printed letters into a sound representation and accessing the lexicon via the sound representation (an indirect route). To what extent are the conclusions that we have reached generalizable to languages other than English?

First, with respect to the issue of the importance of letter processing, we suspect that what we have said holds true for any alphabetic system. With logographic systems, the issue is something of a moot point because the printed characters represent word units. Hence, it is clear by definition that all of the characters have to be processed in such a system. With syllabaries, while we know of no direct evidence on the issue, we also suspect that the points we have made would hold true.

The second issue, the role of sound representations in accessing the lexicon, is more interesting and has been studied rather extensively (Henderson 1982, 1984; Hung and Tzeng 1981). Numerous studies have been conducted to compare (1) word perception in alphabetic systems with *shallow* orthographies (i.e., those with a close correspondence between letters and phonemes) like Serbo-Croatian (Feldman and Turvey 1983; Katz and Feldman 1983; Lukatela et al. 1980; Lukatela et al. 1978; Turvey, Feldman, and Lukatela 1984) with alphabetic systems with *deep* orthographies (i.e., those in which morphemic properties are more directly related by the writing system) like English, (2) syllabaries to English (Besner and Hildebrandt 1987; Morton and Sasanuma 1984), and (3) logographic systems to English (Tzeng, Hung, and Wang 1977). In addition, there has been some interest in comparing Hebrew (where critical information used in converting to the sound representation is not explicitly contained in the print) to English (Bentin, Bargai, and Katz 1984; Navon and Shimron 1981).

Generally, our impression is that the results of the studies mentioned above (and others) lead to the conclusion that the specific orthography may alter the extent to which a reader relies on one route or the other, but that the results are consistent with the hypothesis that cross-culturally there are two routes to the lexicon. The work done on Serbo-Croatian has led some investigators (Turvey et al. 1984) to argue that accessing the lexicon via the sound representation is not an optional strategy for readers of that language. However, there is reason to suspect that readers of that language can also access the lexicon via the direct route (Besner and Hildebrandt 1987; Seidenberg and Vidanovic 1985). The regularity of the letter-to-phoneme correspondence in a shallow orthography like Serbo-Croatian may simply lead readers to rely more heavily on the route through sound to the lexicon. Likewise, logographic systems (like the Japanese Kanji) might lead to a heavier reliance on the direct visual route than English (Morton and Sasanuma 1984). It has been suggested (Morton and Sasanuma 1984) that syllabic systems (like the Japanese Kana) have to be translated into a phonological code before lexical access is possible. However, Besner and Hildebrandt (1987) recently reported evidence consistent with the conclusion that the lexical access of words written in Kana can be achieved without reference to phonology. Finally, studies with Hebrew readers (Bentin, Bargai, and Katz 1984; Navon and Shimron 1981) show that although the direct route is very important in lexical access, the phonological route is used by these readers even though the print is more irregular than English in coding the cues used for making a translation into a phonological representation.

In summary, our argument is that while different writing systems may influence readers to rely more heavily on one route than the other, the present evidence suggests that both routes are used in all languages. Once readers have acquired the ability to decipher the written symbols, reading may be a culture-free cognitive activity (Gibson and Levin 1975; Hung and Tzeng 1981) in the sense that the writing system may have little effect on the process of reading. Thus, we believe that the points we will stress in the remainder of this book are generally true cross-culturally. Of course, differences in cultures and the structures of languages may have profound influences on how people comprehend both spoken and written discourse. Such concerns, however, are largely beyond the scope of the current book.

PROCESSING SIMPLE AND COMPLEX WORDS

Most of the literature on the identification of words that we have discussed has used short (3- to 6-letter) words. The word-superiority effect experiments (and simulations) have virtually all used 4-letter words. The regularity literature we have just discussed employs a somewhat wider range of words, but a majority are still 6 to 7 letters or fewer and have only one *morpheme* (unit of meaning). Moreover, almost all the words used were nouns with a sprinkling of verbs and adjectives. Thus, our picture of word identification is incomplete. In this section, we will explore two additional kinds of words. First, we will briefly discuss *function words* (prepositions, conjunctions, articles, and pronouns), for there is some evidence that they may be processed differently than the types of words we have discussed so far. We will then discuss what is known about the processing of complex words. Since the bulk of the research on word identification is not in these areas, our picture will remain sketchy.

Function Words

Psycholinguists often make a distinction between *function words,* which include prepositions, conjunctions, articles, and pronouns, and *content words,* which include nouns, verbs, and adjectives. While there is fairly general agreement that the two classes of words may be psychologically different, there is some uncertainty about the precise boundary between the two (e.g., most people don't know where to put adverbs). One way in which function and content words appear to be different is that function words are a *closed class;* that is, there is a relatively small number of them in the language (roughly a few hundred) and that is all. In contrast, content words are an *open class:* the number of nouns, verbs, and adjectives is not only large but not bounded, with new ones probably being invented each day. Another difference is that most content words "mean something" in a way that function words do not: A content word such as *tree* means something in isolation, but *and* means little in isolation. (However, locational prepositions, such as *above,* and pronouns seem to be about as meaningful as content words.) One possible test of "meaningfulness" would be whether an isolated word could be a meaningful utterance: One could envisage the noun *dog,* the verb *climb,* the adjective *red,* or even the abstract noun *democracy* being uttered in isolation to express something, but not the function word *of.* Most function words seem to have meaning only as joiners of content words—they are the glue that holds sentences together. Function words are also among the most frequent words in the language (*the* is the most frequent).

Much of the data suggesting that function words are special comes from the neuropsychological literature. Perhaps the most striking finding is that there are people with brain damage whose ability to read aloud and comprehend content words is virtually intact, but whose ability to read aloud and comprehend function words is markedly impaired (Coltheart, Patterson, and Marshall 1980). This pattern of deficits occurs for many of the phonemic dyslexics described earlier. It also occurs for many patients with *aphasia* (i.e., general language problems), especially for a class of patients with *Broca's aphasia.* (Patients with Broca's aphasia usually have difficulty uttering function words in spontaneous speech as well, so that their speech is "telegrammatic.") Phenomena such as these suggest that function words may be represented in a lexicon separate from content words.

One interesting experiment on the representation of function words involved patients with Broca's aphasia (Bradley, Garrett, and Zurif 1980). They found that normal subjects had no word-frequency effects with function words in a lexical decision task in contrast to the usually large word-frequency effect for content words. On the other hand, Broca's aphasics had a large word-frequency effect for both types of words. Such a result would indeed indicate a radically different storage or retrieval mechanism for function words. However, a subsequent experiment failed to replicate the Bradley, Garrett, and Zurif result. With normal subjects, Gordon and Caramazza (1982) obtained equal frequency effects on lexical decision time for open and closed class words. Thus, the exact nature of the distinction between open and closed class words is still unclear.

Complex Words

We mentioned above that there is no limit to the number of content words in a language. In fact, one of the striking aspects of human language is the generativeness of words. One way that new words are generated is to describe new places, concepts, or technological inventions. However, new words are also generated in profusion from old words. If by a "word" we mean something set off by spaces, there are languages in which new words are created every minute. For example, in German, one can either say, "the man who came over to dinner last Tuesday night" or, "the cameoverfordinnerlastTuesdaynight man." While English is not so extreme, new compound words such as *headroom* are probably being constructed each day. In fact, we are often not clear on what a word is: do you write *wire service* as one word, as two words, or do you hedge and hyphenate it?

The question we would like to raise is whether the lexical-access model proposed so far is really adequate to explain the full spectrum of words. One reason to believe that the parallel letter processing models we have considered so far may be inadequate for recognizing all words is that it may be unreasonable that all words whose meaning you know are actually stored in the lexicon. As mentioned earlier in the chapter, there are books that have tabulated the frequency of usage of English words (taken from a corpus of text such as magazine articles and books). There are many words that do not seem at all strange (e.g., ABUSIVE, CREASES, PONDER, THINNING) that have a frequency of usage of 1 part in 1 million. Even if one assumes that high school students have each read something like 4,000 pages of text a year for 10 years, and if a typical page has about 500 words on it, they have only read something like 4,000 × 10 × 500, or 20,000,000, words in their life. Thus, they have *seen* words in the 1 per million category only an *average* of 20 times in their lives. But because of statistical fluctuation, the chances are pretty good that there are many of these words in this category that they have never seen. Moreover, there are many forms of a lot of words. For the words listed above, are you confident that you have seen them in the past tense (if they are verbs) or the plural (if they are nouns)? Yet you would likely be able to recognize those words easily if you encountered them. Moreover, there are various forms of words (e.g., CHARACTER, CHARACTERISTIC, CHARACTERISTICS, CHARACTERISTICALLY). It seems not unreasonable that some of these forms are not actually stored, but instead are constructed from a "base form" and some sort of rule.

Even if all words are actually stored, however, there might be good reasons to have a more complex access procedure than a single-stage parallel look-up. First of all, there might be a limit to how many letters can be accessed in parallel by the visual system. Thus, there may be some sort of sequential access for longer words, whereby they are accessed a part at a time. Since short words appear to be processed in parallel, however, the most plausible size for the units of sequential access would appear to be larger than a single letter and perhaps on the order of 4 letters or so. A second reason for some sort of sequential access is that it might aid understanding the word. That is, almost any linguistic analysis would indicate that the meaning of the word ENDED is *end* + past tense. If lexical access were in two stages, the meaning might be understood as a part of the access process rather than requiring an additional step.

The most well-developed and interesting model of sequential access has been proposed by Taft and Forster (1975) and modified subsequently (e.g., Taft 1979, 1985, 1986). In the original version, which we are still most comfortable with, the first stage of lexical access is accessing the *root morpheme*. The way the root morpheme is defined is somewhat different for the two types of polymorphemic words. *Affixed words,* the first type, have a stem and prefixes and suffixes (e.g., *END*ING, IN*CLUDE, SELECT*IVE, UN*DO*ING). For these words, the root morpheme (in italics in the examples) is simply the stem to which the prefixes and suffixes are added. The second type is *compound words,* such as HEADSTAND and TOADSTOOL, which are made up of two essentially equal morphemes (both of which are usually words). Taft and Forster define the root morpheme of compound words to be the first morpheme. This definition goes against more

linguistically motivated definitions such as Selkirk's (1982) because for most compound words in English, the second morpheme is the "root": a *headstand* is a type of stand and a *footstool* is a type of stool.

According to the Taft and Forster (1975, 1976) model, both classes of complex words are accessed by the same basic process: (a) initial access is to the root morpheme; (b) subsequently the actual word is accessed. A "file drawer" metaphor might help to explain the idea (Forster 1976). Initial access of the root morpheme allows you access to a file drawer with all the words containing the root morpheme in it. For example, in the case of ENDING, the file drawer accessed would have all words with END as the root morpheme such as ENDED, ENDING, ENDPLAY, ENDGAME. When the file drawer is accessed, search for the lexical item is restricted to these items.

While Taft and Forster's model makes the decomposition of complex words the central focus of lexical access, they do not deal with the generativeness of forming words. In fact, they assume that all the forms are stored in the file drawer rather than being constructed through rules. Our prior discussion suggests that it is likely that some complex words are constructed rather than accessed; however, there is no data on this one way or the other. Thus, we will focus (as do the researchers in the area) on whether a two-stage model of access involving decomposition is a viable theory of lexical access.

Perhaps the experiment that gives the best feeling for Taft and Forster's model is one in which subjects made lexical decisions on prefixed words. Taft (1979) coded the words for both the "surface frequency" (the frequency of the compound word itself) and for the "root morpheme frequency" (the sum of the frequency of all words containing the root morpheme). If access of complex words were merely a look-up of each word in a separate lexical entry, then one would expect the surface frequency to predict lexical decision time. In fact, Taft found that lexical decision time was affected by the frequency of the root morpheme even when the surface frequency was equated. However, Taft also got an effect of surface frequency when root-morpheme frequency was equated. Bradley (1979) carried out a similar experiment using suffixed words. For two types (words ending in -MENT and -NESS), she got only effects of root-morpheme frequency, but for the third type (ending in -ION), she got effects of neither frequency.

The explanation for Taft's result seems pretty straightforward. If the first stage of lexical decision time is accessing the root morpheme (or file drawer), then the frequency of the root morpheme should be a major determiner of lexical access time. How does surface frequency come into play in Taft and Forster's model? They view search through the file drawer as a sequential process, so that all that matters is the relative frequency of the entries in the drawer (i.e., how far down the list is the entry you are searching for). Thus, if the target word is second in the file drawer, the time to find it in the file should be the same regardless of how much less frequent it is than the first word in the file. Thus, the lower the surface frequency, the lower the word should be in the file drawer. However, how far down it is depends not only on its own frequency but on other items in the file as well. Thus, it is hard to predict how far down a word should be just from its surface frequency, except to say that it *generally* should be further down. Thus, Taft's results seem more consonant with Taft and Forster's model than Bradley's, although Bradley's results are not necessarily inconsistent with the model. (No model would easily predict why there was no frequency effect for the -ION words.)

A second line of evidence that complex words are accessed in morphemic pieces comes from priming experiments (Stanners et al. 1979). As with the priming studies described earlier in the chapter, the lexical-decision task was employed; however, the priming differed in two significant ways. First, the "prime" is not merely associatively related to the target word (like DOCTOR-NURSE), but is either the word itself or a morphemically related word. Second, the interval between the prime and target is much longer than in the associative priming studies, where the interval is usually a second or less. In most of the morphemic priming experiments, the interval between the prime and target is at least 10 or so intervening items and thus about 10 to 20 seconds. (There would likely be no priming effect over that long an interval for associated words such as DOCTOR-NURSE [Gough, Alford, and Holley-Wilcox 1981].)

Stanners et al. (1979) observed that subjects were over 100 msec faster on repeated strings than on nonrepeated strings even over these long intervals between target and prime. The key finding was that the priming effect on a word such as START was as large when it was preceded by STARTED as when it was preceded by START. If we assume that the amount of priming indexes the similarity of the access of the target to the access of the prime, then this result suggests that lexical access of STARTED involves all the processes of the lexical access of START and then some additional ones. In contrast, consistent with the model, when the stem START is used as a prime and the whole word STARTED is used as a target, there was only partial priming (Stanners et al. 1979).

The story gets more complicated, as you might expect. If a verb is less transparently related to its root, then there is only a partial priming effect. For example, SPOKEN will prime SPEAK, but only about half as much as SPEAK primes SPEAK. In addition, *derivational* affixes (affixes that change the part of speech) appear to behave differently than *inflections* (affixes that change tense or number) even when the derivation is transparent (e.g., SELECTIVE primed SELECT only about half as much as SELECT did). Thus, Stanners et al. (1979) found that the priming effect depends both on the transparency of the relationship between the complex word and its root and on the type of morphemic relation.

These experiments clearly show that morphemic decomposition is implicated in the lexical-decision task. However, as mentioned earlier, some researchers have serious misgivings about the lexical-decision task as an index of lexical access (e.g., Balota and Chumbley 1984; Seidenberg et al. 1984b): They feel that the lexical-decision task is tapping postlexical-decision processes as well. Unfortunately, it may be difficult to tap lexical access for complex words employing other tasks. For example, naming latency (i.e., the time to *begin* the pronunciation of a word) may suffer from the opposite problem as lexical decision. That is, people may begin pronouncing complex words well before they have completely accessed them. Categorization tasks tend to be difficult, since many polymorphemic words are either difficult to categorize or the categorization may depend only on the root morpheme. Perhaps the best task to employ would be to measure the time to *finish* saying the word. However, it is technically much more difficult to measure the offset of a spoken word than the onset (since it is usually more gradual), and thus this technique has not been used.

Given the above problem, can we convince ourselves that the above experiments really do say something about decomposing complex words as a stage of lexical access? First, let's consider the priming experiments. If we find that a morphemically related word such as STARTED primes START, what possible explanations are there for the phenomenon that would not involve decomposition at the point of lexical access? The first possibility is that the whole priming phenomenon is postlexical. That is, that the faster time for END when it is repeated has nothing to do with the speed of accessing the word, but instead is speeding up the decision (i.e., responding "yes" to that thing earlier makes it easier to respond "yes" now). One argument against that argument is that one gets positive priming from nonword stems such as VOLVE (which get "no" responses) to word targets such as INVOLVE (Stanners, Neiser, and Painton 1979).

A second possibility is that priming is getting at lexical access, but that the decomposition effects occur after lexical access of the prime. That is, after lexically accessing STARTED, one then decomposes it as *start* + past tense in order to understand it; the *start* that is created in this postlexical decomposition is what primes START later. While such postlexical decomposition may be part of the effect, we feel that the data argue pretty strongly that it can't be the whole effect. First, if that were all that was going on, it is hard to see why SPOKEN wouldn't prime SPEAK as strongly as STARTED primes START. Perhaps there are two effects: this postlexical morphemic priming and priming just due to letter overlap. However, there is evidence that mere letter overlap has no effect: e.g., ARSON does not prime SON at all, while DISHONEST primes HONEST (Lima 1987b). Similarly, in a tachistoscopic recognition experiment, BORING is primed by BORE but not by BORN (Murrell and Morton 1974). Thus, it appears that at least some of the priming effect is due to decomposition of the word during lexical access.

Moreover, if decomposition is postlexical, it is hard to understand why STARTED should prime START as strongly as START itself does. The fact that START and STARTED prime START equally suggests that, as Taft and Forster predict, the first stage of lexical access for END and ENDED is virtually identical. On the other hand, Taft and Forster would not predict any priming from SPOKEN to SPEAK since the root for SPOKEN is either *spoken* or *spok-* while the root for SPEAK is *speak-*. Thus, it appears either that some of the priming effect is postlexical or that the sequential access is operating in a different fashion than postulated by Taft and Forster.

Suffixes in English can be divided into two classes: inflections and derivations. Inflections are suffixes that change the tense or number of verbs or the number of nouns but preserve the part of speech. On the other hand, derivations are suffixes that change the part of speech such as -LY, -NESS, -ITY, -IVE. As mentioned earlier, in the experiments of Stanners et al., inflections whose stems appear in the target word (e.g., *START*ED) produce full priming, but derivations whose stems appear in the target word (e.g., *SELECT*IVE) produce only partial priming. An unsolved problem is why this difference occurs. Taft and Forster have to explain such effects by postulating different orderings in the file drawer. Consider the inflected case. In the END- file drawer would be *end, ended, ending*. If we assume that *end* itself is first (because it is most frequent), then full priming from ENDED to END would occur; accessing *ended* would imply having accessed *end* since *end* is above *ended* in the file drawer. On the other hand, there may be many derivational forms such as SELECTIVE that may be at least as common as the uninflected versions (i.e., SELECTIVE may be at least as common as SELECT).

Another possible explanation for the difference is that there may be a qualitative difference in what happens when you get in the file drawer. In the case of derivations, there is a rule (admittedly with exceptions) for how you form past tense or plurals, for example, that applies to every word. In contrast, there are several ways to change verbs to adjectives (e.g., SELECT-SELECTIVE, DIFFER-DIFFERENT), which, in addition, have different meanings (e.g., the SELECTIVE noun is doing the selecting while the DIFFERENT noun is not doing the differing). Moreover, there are some verbs for which no appropriate derivation exists (e.g., PROCESS-?). This raises the possibility that inflections may be constructed by rule, whereas derivations are accessed by a specific entry in the file drawer. That is, access of all inflections entails accessing of the stem, whereas access of a derivation may not entail access of the root because they are different entries in the file drawer.

One problem for the morphemic decomposition model is how the system knows how to decompose the word. One possibility advanced by Taft (1979, 1985, 1986) is that there is an orthographic principle that defines what the first unit is rather than the first morpheme. His rule is (roughly) to take as many consonants as possible following the first vowel, and Taft dubbed this unit, which accesses the file drawer, the *BOSS* (Basic Orthographic Syllabic Structure). Much of the research on the BOSS is plagued with conflicting results, however. One piece of evidence (Taft 1979) for the BOSS is that splitting up even monomorphenic words such as LANTERN with the BOSS as the first chunk (e.g., LANT ERN) aids lexical decision times relative to splitting up the words according to phonetic syllables (i.e., LAN TERN) and to leaving them unsplit (e.g., LAN TERN). However, Lima and Pollatsek (1983) failed to replicate this effect. They found that any gap in the word was worse than if there was no gap, and furthermore found no difference between the BOSS division and the phonetic syllable division. A second paradigm devised by Lima and Pollatsek employed a priming technique, whereby the first part of a word was presented (e.g., LANT or LAN) and followed by the whole word 90 msec later. They similarly found no evidence for the BOSS (no difference between the priming effect for LANT and LAN) in monomorphemic words. Taft (1987) has recently found a difference between these conditions, however, although he used a larger interval between prime and target (250 msec). Another finding against the BOSS is that when compound words were used and the BOSS was different from the first morpheme (e.g., TEA is the first morpheme, but TEASP is the BOSS of TEASPOON), Lima and Pollatsek found that there was priming only when the morpheme was the prime. Thus, evidence for the BOSS is fairly shaky, and at present, it appears that the only safe statement we can make is that lexical access appears to involve (at least on some occasions) accessing the root morpheme.

You may have noticed that we have avoided talking about prefixes. Prefixed words present a problem, since the beginning of prefixed words is not the root. In addition, there are many "pseudoprefixed" words (e.g., REPERTOIRE) that look like prefixed words. Taft (1981) also developed a process to deal with prefixes. He hypothesized that if a word started with something that could be a prefix, the system assumed that it was in fact a prefix, stripped it off, and then located the root morpheme (e.g., RE would be stripped off REJUVENATE, and then the word looked up under the root which would be *juvenate* or some part). In the case of a pseudoprefixed word, Taft hypothesized that RE would be stripped from REPERTOIRE, the lexical entry "pertoire" searched for, and only when not located, would the lexical entry "repertoire" be searched for. He found that lexical decision times were faster for prefixed words than pseudoprefixed words; however, the difference was only about 30 msec—not plausibly the time to strip, search, and then search again. What seems more plausible is that access is going on in parallel, with Taft's decomposition route being faster than the direct route. The advantage of prefixed words over pseudoprefixed words also gives us hope for the lexical-decision task; Lima (1987a) got a similar advantage for prefixed words in a reading task, where her dependent variable was the amount of time readers fixated on the word.

A second result that argues for the dual-route theory is one obtained with suffixed and "pseudosuffixed" words. Here, there is no difference in the absolute time to make lexical decisions to the suffixed and pseudosuffixed words (e.g., SISTER vs. SENDER), but there is a route-priming effect whereby a pair of words of one class will be responded to more rapidly than a mixed pair (Manelis and Tharp 1976).

To summarize, it appears that the theory that all words are recognized through a direct look-up in the lexicon is probably too simple. At least some morphemically complex words appear to be accessed through a sequential look-up whereby the root morpheme is accessed followed by accessing the rest of the word. The data are less clear about whether there is some sort of orthographic principle that will account for how words are decomposed.

At present, this area of word processing is still quite undeveloped. One issue that has not been explored is the "grammar" of complex words (Selkirk 1982). For complex combinations such as *undeveloped,* which is both prefixed and inflected, there appear to be rules indicating how to compose the units. That is, *undeveloped* is formed by first going from *develop* to *developed* and then adding *un*. To see why this is true, consider *undevelop*. This is not a word, and if it were, it would have a different meaning than *undeveloped* (i.e., it would mean to actively undo something that was developed). This argument suggests a constructive aspect to processing complex words rather than (or in addition to) the relatively simple sequential look-up posited by Taft and Forster.

SUMMARY AND CONCLUSIONS

At the beginning of the chapter, we raised several questions about the processing of words. Some of them have turned out to have simpler answers than others. One question that was raised was whether word processing was an automatic process or whether identifying words was a major part of the mental effort that went into reading. We saw that words in isolation (at least relatively common and short words) could be identified without awareness and without intention, a seemingly automatic process. While it was far from clear that the activity involves no mental effort, the process of identifying words appears to be a relatively small part of the mental effort in reading *for the skilled reader*. We will discuss in detail the effects of context on word identification in subsequent chapters. The data from this chapter, however, allow us to make an educated guess as to what the answer will be. Since accessing the meaning of words is such an automatic and easy process when words are seen in isolation, we wouldn't expect context to speed up processing very much, if at all. We also would be surprised if identifying words in text was performed in a substantially different way than words in isolation, since it would seem wasteful to have two different machineries, each of which is so rapid and accurate.

The second conclusion is related to the first. Letters in words (at least short words) appear to be processed in parallel. To many in cognitive psychology, parallel processing is the hallmark of automatic processing. If something is done in series then it requires an attentive mechanism and therefore probably takes processing capacity.

The rest of our discussion did not yield any simple answers. In fact, the data make clear that the identification of words is a very complex process—much more than one might have supposed at the outset. A direct visual look-up, whereby the letters access a word in parallel, appears to be a necessary ingredient of fluent word identification. Otherwise, irregular words could not be recognized. However, using rules (or some equivalent analogue process) to access a word through a sound code does not seem to be necessary for accessing the meaning of a word, since irregularly spelled words can be identified easily. Moreover, certain patients who appear to be severely impaired in their ability to use rules to access sound codes are relatively normal in their ability to extract the meaning of content words. (Their fluency, however, may be impaired.)

Sound encoding appears to play some part in accessing the meaning of words in fluent reading; however, it is less clear exactly what the role is. In some views, the role is very minor: the sound system may matter only for processing a few low-frequency words (Seidenberg et al. 1984a). This conclusion is based on the small regularity effect which appears only for low-frequency words. There are data that lead to a different conclusion. First, even high-frequency words may be biased by sound codes (Meyer, Schvaneveldt, and Ruddy 1974) and words (or even pseudowords) can be misclassified as their homophones (Van Orden 1987). A major problem in deciding on the role of sound in word encoding is that irregularities in the language are usually quite minor, so that it is not clear whether one should expect a large regularity effect in the first place. There is no simple resolution to the problem. The position we have taken is that the data are consistent with a cooperative computation model, wherein entries in the lexicon are excited both by the direct visual route and by the indirect rule to sound route, with the recognized word being the entry that has accrued the most combined excitation. That is, we see the sound system as involved in most lexical access. At present, however, reasonable people can hold almost diametrically opposite views on the subject. The common ground for all positions is that direct visual access is important and that sound encoding plays some part.

There was also evidence that morphemically complex words are probably looked up in two stages, possibly with the help of morphemic rules (e.g., Taft and Forster 1975, 1976; Taft 1985). Thus word encoding appears to involve three systems—the direct visual route, a spelling-to-sound route, and a morphemic-decomposition route, i.e., a direct route and two more constructive processes. Since most of the evidence that word processing is automatic comes from the study of relatively short frequent words (i.e., those for which the direct route could predominate), it is possible that word processing is not so automatic for words whose access relies more heavily on the more constructive routes.

We should emphasize that this chapter has for the most part dealt with a relatively narrow window of word perception: We have discussed skilled readers of English reading print. However, our discussion of cross-cultural studies led us to conclude that the points we stressed were generally true for other writing systems. We have focused on English because it has been studied far more intensively than any other language. We have not discussed handwriting, since there is little data on recognition of handwriting. It is possible, however, that the perception of handwriting operates differently from print. First, since handwriting is often quite messy, sentential context may be more important in deciphering it than print. Second, since letters are not transparent visual units as in print, more constructive processes may be needed in addition to automatic letter detection.

Let us close with some comments on the relevance of the study of skilled readers to the process of learning to read. The better we understand the word-identification process in skilled readers, the better we understand what the *goal* of instruction should be. However, even a perfect understanding of the skilled reader may say little about the beginning reader. At one extreme, the adult reader may be exactly like the beginning reader, but may do everything much faster and in a much more ''automated'' way. At the other extreme, the beginning processes may be a crutch to get over some hurdle so that skilled reading may involve totally different processes than that of beginning reading. Thus, there may be little in the processes of skilled reading that indicates how the reader acquired those skills. We will discuss these issues in depth in chapters 9 and 10. The point we wish to leave you with is that much of the research and many of the issues in learning to read have been framed by the research on skilled reading of words that we have discussed in this chapter.

Library of Congress Cataloging-in-Publication Data
Rayner, Keith
The psychology of reading.
Bibliography
Includes index.
1. Reading, Psychology of. I. Pollatsek, Alexander. II. Title.
BF456.R2R33 1989 418 88-22419
ISBN 0-13-733007-3

Editorial/production supervision: Wordcrafters/Karen Winget
Cover design: Ben Santora
Manufacturing buyer: Ray Keating

A Division of Simon & Schuster
Englewood Cliffs, New Jersey 07632

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

ISBN 0-13-733007-3

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Simon & Schuster Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*