CHAPTER FOUR

Speech Processes in Skilled Reading

The role of speech processes in reading might have been treated under lexical access (Chapter 2) or under comprehension (Chapter 3). Or, from the point of view of some reading researchers, it might be omitted altogether. Instead it is here the subject of a separate chapter intended to suggest a theoretical description of speech processes in reading that links speech processes to lexical access *and* comprehension. The proposal is that speech processes occur automatically as part of lexical access and that they support comprehension. But first we will consider some alternative possibilities.

E. B. Huey (1908/1968) was the first psychologist to write a book on reading. He had the idea that private reading is accompanied by a "silent inner voice." Many readers have the same impression that some implicit speech activity accompanies skilled reading. Nevertheless, it has proved difficult even to convincingly demonstrate, let alone to explicate, a role for speech processes in reading. Some have argued that skilled reading is solely a visual and semantic process and that speech plays no serious functional role. These visual theories suggest at most an optional role for speech (Allport, 1977; Barron & Baron, 1977; Davelaar, Coltheart, Besner, & Jonasson, 1978; Frederiksen & Kroll, 1976). The reason for this view is that most researchers have focused on speech recoding as the major speech issue for reading. (See McCusker, Bias, & Hillinger [1981] for a review.)

SPEECH RECODING

It has been common to ask this question: On encountering a printed word, does the reader transform the visual input into a spoken form in order to access its meaning? This is the question of speech recoding. Examples of recoding are actually commonplace. When a child just learning to read encounters the word *cat*, it is not difficult to imagine the word identification process having an intermediate stage: *cat* → /kæt/ → oh! furry Morris! Even adults show recoding, especially in reading a foreign language they have not completely mastered. For example, suppose a reader's German is about good enough to be able to translate *Die Frau ohne Schatten*. In this case, to be "about good enough" means being able to transform the written German into a speech form and then recognize at least part of it. *Frau* → /fraw/ → of course that's *woman!* And *Schatten* /šatən/. If it sounds a bit like *shadow*, that helps quite a bit. Words like *die* and *ohne* are perhaps different—one merely apprehends their syntactic value (if the language is somewhat familiar) just as one apprehends their English syntactic equivalents, *the* and *without*. (For some readers there may also be useful top-down knowledge, namely, that Richard Strauss wrote an opera of this name. This knowledge may make it easier to bypass speech recoding and to translate the phrase immediately into something meaningful.)

The example is important to remind us that there are commonplace occasions in which speech recoding occurs. When reading is new, when a language is incompletely learned, and perhaps when words are unfamiliar, speech recoding does seem to be a process used by the reader. The question for skilled reading, however, is whether such a process is *ordinarily* operative. There are three logical possibilities: Speech recoding does occur, it may occur optionally, or it does not occur. The last possibility is ruled out by common sense. The weight of the evidence appears to be that speech recoding may occur as an optional strategy in skilled reading, but that it will most often not occur. This is the view represented in Figure 4-1.

According to Figure 4-1, a visually presented word is encoded as a string of letters, which are used directly to access the word's meaning in memory. Of course both direct access and speech recoding are much simplified in the figure. For example, the visual representation sufficient for direct access may be some momentary synthesis of the letters generated by the reader (Massaro, 1975). The important fact is that information-processing models with very few exceptions (notably Gough [1972]) assume that speech recoding is optional and ordinarily plays little role in lexical access.

The experimental evidence supporting such a model has come from different tasks. Lexical decision experiments have provided some of the most important results. One conclusion is that the time it takes to decide that a nonword (pseudoword) is in fact not a word is affected by the phonetic value of the nonword. *Brane* is not a word in English, but it can be pronounced exactly like the real word *brain*. That is, it has the same

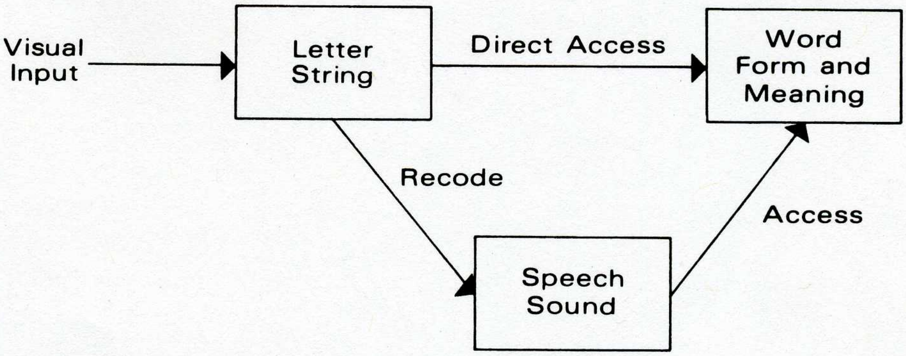

Figure 4-1. The conventional view of direct access versus mediated access. A word's memory location is accessed either from the visually encoded letter string (direct access) or following a recoding of the letter string into sound.

phonetic value. Compared with a pseudoword with equal visual similarity to *brain*, *brane* takes longer to decide. That is, *brane* takes longer than *brone* (Coltheart, 1978). But notice that this is not a matter of lexical access since neither *brane* nor *brone* is a word. (Lexical access, by definition, involves words.) Such a result merely demonstrates that the pseudoword is recoded into speech and that if its sound matches that of a real word, then the subject has to recheck to make sure it is not really a word.

The critical comparison, at least according to Colheart (1978), is between words that are homophones and words that are not. (Homophones are words that have the same phonetic value, such as *groan* and *grown.)* When careful controls for word frequency are made, homophones are not faster in lexical decision than nonhomophones (Coltheart, Davelaar, Jonasson, & Besner, 1977). By one interpretation (Coltheart et al., 1977; also Rubenstein, Lewis, & Rubenstein, 1971), a homophone should be quicker to decide if a word's speech sound is encoded. This is because there would be two chances instead of just one to have a lexical access. So a lexical decision should be faster with a homophone. The fact that it is not is taken to support the idea that recoding does not occur (Coltheart, 1978).

According to one interpretation of the model of Figure 4-1, the reader may attempt lexical access along both routes (Meyer & Ruddy, 1973). This is a "horse race" model; the route that results in lexical access is the "winner." The test for this hypothesis has been the comparison of "regular" words with "exception" words. Regular words are those that have more predictable correspondences between their letters and their phonemes. Exception words are those that are irregular, at least in the sense of having less common correspondences between letters and phonemes. For example *rough* is an exception word relative to *raft* because the phoneme /ʌ/ is usually spelled *u* (as in *luck*) and the phoneme /f/ is usually spelled *f*. The horse race model predicts that regular words will be accessed more quickly than exception words. This is because for regular words, the speech recoding route sometimes may win as letters are recoded into phonemes. For exception words, the visual route should be the only alternative. Unfortunately, the results of experiments which have compared regular with exception words are mixed. Some show an advantage for regular words (Barron, 1981; Glushko, 1981; Stanovich & Bauer, 1978), but others do not (Bauer & Stanovich, 1980; Coltheart, Besner, Jonasson, & Davelaar, 1979).

Actually, a model proposed by Glushko (1981) helps explain conflicting results. In contrast with the model of Figure 4-1, Glushko suggests that lexical access is visual and that activation of phonological information occurs *after* access. According to Glushko's model, the regularity effect appears just when the "neighbors" of the regular word have the same pronunciation it has. For example, access to *rate* would activate such neighbors as *mate* and *late* that have the same pronunciation. However, if a regular word, such as *save*, is accessed, some of its neighbors, at least *have*, would have a different pronunciation. Thus whether a regular word results in a faster lexical decision than an exception word depends on the extent to which words spelled similarly are pronounced similarly. In fact, this result was obtained by Bauer & Stanovich (1980).

One other line of evidence must be mentioned—the effects of "priming" on lexical decision. The critical case is deciding on, say, *couch* immediately following a decision on *touch*. Although these two words are spelled identically except for the initial letter, they have different pronunciations. Research by Meyer, Schvaneveldt, and Ruddy (1974) showed that lexical decisions in such cases (e.g., the decision *touch*) were slower than decisions on words preceded by other words that shared both their pronunciation and their spelling. However, as with the regularity experiments, the research on this "negative primacy" effect is not consistent (cf. Hillinger, 1980). Furthermore, a clear prediction from this line of reasoning is that there should be a positive primacy effect. For example, *grown* should be accessed more quickly when preceded by *groan* than when preceded by *green*. Evidence again is mixed (Davelaar et al., 1978; Hillinger, 1980). Thus evidence concerning lexical access is inconclusive. Sometimes there seems to be a prior stage of speech recoding and sometimes not.

Some intriguing research on this issue has been reported by Lukatela, Popadic, Ognjenovic, and Turvey (1980), who took advantage of an unusual bi-alphabetic situation: Serbo-Croatian (spoken in Yugoslavia) is written in two different alphabets, Roman and Cyrillic. Each alphabet has some letters which are unique to it. However there are some letters which are the same shape in the two alphabets but have different phonetic values. For example, *p* occurs in both alphabets. It corresponds to /p/ in Roman and /r/ in Cyrillic. Lukatela et al. had subjects make lexical decisions under various conditions. The most important case involved a letter string that had two different pronunciations in the two alphabets but was a word in only *one* of the alphabets. On such a string, subjects' decision times were slowed down compared with a case where the letter string received the same pronunciation in each alphabet. Lukatela et al. attribute this result to a level of phonological processing. The letter string produces one sound in one alphabet and a different sound in the other. The resulting conflict slows decision time. This result seems to implicate some phonological process, but it does not mean that individual letters were recoded prior to access. It suggests an automatic phonetic activation process that cannot be suppressed.

In summary, there is evidence that some speech processes may accompany lexical access. But the consensus is that these processes are not necessary as a stage of processing that precedes lexical access.

COMPREHENSION AND MEMORY

Even if lexical access can occur without a stage of speech recoding, it is possible that subsequent reading processes make use of speech. The most likely candidate is the immediate memory for what has been encoded. In the discussion on comprehension in the last chapter, the immediate or short-term memory for a text was assumed to contain propositions. That is, the focus was on the reader's representation of encoded text meanings. However, there is more to the contents of short-term memory than meaning abstractions. There are also words and perhaps the speech sounds of the words.

In the verbal efficiency theory of reading (see Chapters 6 and 7), there is a special status of immediate memory. In particular, it holds a fairly exact record—within its limitations—of what is read. This means that the words themselves, not merely their meaning abstractions, are available to the reader for a brief period of time. There is an important value to the reader in this verbatim record: It helps secure reference.

Reference Securing

When a reader encounters a written word, lexical access may provide all the meaning information that is needed. Indeed, it is fair to assume that the reader's initial encounter with a word will result in an attempt to immediately encode the word's meaning—not merely its meaning as represented in semantic memory, but also its semantic role in the sentence. For example in sentence (1) and sentence (2) the identical word *theory* has two very different semantic roles.

(1) *The theory explained how reading works.*
(2) *The theory of reading was strongly criticized.*

In sentence (1) theory is an *instrument* of explanation, whereas in sentence (2) it is a *recipient* of criticism. These semantic roles are sometimes referred to as *cases* following the example of case grammars (Fillmore, 1968). These roles are in fact part of the propositions that are encoded.

Thus the *immediacy* assumption (Just & Carpenter, 1980) is that words are fully encoded with respect to basic meaning and with respect to their semantic roles as soon as they are encountered. By this assumption, lexical access results in the complete encoding of a word fully sensitive to its use in the sentence. There will be times when a more tentative encoding occurs or the reader has to reinterpret the encoding in response to information later in the sentence. However, whenever possible the reader immediately tries for a full encoding, according to the immediacy principle.

However, the reference-securing problem is that the initial access to the word may not be sufficient to specify the encoding ultimately needed. For example, consider the following short text:

(3) *John takes Pepe, his dachshund, to football games with him.*
(4) *Believe it or not, they make him buy an extra ticket.*
(5) *They make quite a pair, especially when Pepe jumps on John's shoulder to get a better view of the action.*

In reading *dachshund* in sentence (3), how does the reader encode the word? Semantically, encoding it as [dog] may be sufficient. In fact, later the reader may very well remember only that it was a dog, not that it was a *dachshund*. However, notice that there are interpretations of sentences (4) and (5) that are enhanced if the code is more than [dog]. The anomaly of requiring an extra ticket is that a dachshund is not such a big dog. It would not seem so strange to demand that a husky or St. Bernard pay for a seat. Similarly sentence (5) will be understood differently depending on whether the reader's code is [large dog] or [small dog].

Of course, the semantic encoding can be enriched exactly along these lines, i.e., the code can include semantic category information and essential attribute information. The problem, of course, is how is the reader to anticipate what attributes are essential to encode? In the dachshund example, subsequent text may not need the size of the dog but rather its characteristic coloring, ear size, or personality characteristics.

The *reference-securing principle* is that these potential problems of encoding are solved by a reference-secured code that is part of the reader's activated memory. Such a code allows access to the information stored with a word in memory even after the word has been accessed. This code is the *name* of the concept. Since needed semantic information for a word will be connected in memory by links to the word's name, the name itself will serve nicely as a reference-secured code, regardless of the exact nature of the needed semantic information.

The value of a name-based reference code is even more clearly seen for words that really do need context for their interpretation—for example, most verbs, the more abstract nouns and adjectives, and words whose function is strictly syntactic. Consider again sentence (3). When the verb *takes* occurs, there has not been enough context to secure its reference. *Take*, like most verbs, is very context dependent. Compare, for example, *take a bath* with *take a piece of cake*. In sentence (3) *take* has roughly the meaning of *[x* causes *y* to accompany him to *z]*, but it depends critically on encountering the word *with*. This does not happen until six words later. If the word *take* (it's "name") is available in memory over this six-word span, there is no problem and reference is secured.

One interesting feature of the reference-securing hypothesis is that the name code may be incomplete and still be helpful. This is because it is part of a code that has partly redundant components. In the case of the dachshund, the code would be something like (dog, small, /daks . . ./). This would be sufficient to secure the dachshund reference and discriminate it from both poodle and dalmatian. In some cases, merely the initial phoneme of a word is enough to secure the reference. In other cases, more phonetic information is required (see Perfetti & McCutchen, 1982).

Finally, this reference-securing process may apply more typically to words with referential (semantic) content than to words that have only syntactic functions, such as auxiliary verbs and articles, and prepositions (so called "function" words). Such words serve mainly to coordinate words that need to be reference-secured. In fact, the evidence from eye movement research is that syntactic words are fixated only about half as often as content words (Rayner, 1977; Just & Carpenter, 1980). Furthermore there are differences between syntactic words and content words in lexical decisions. Decisions for content words are a function of their printed frequency, whereas decisions for syntactic words are not (Bradley, 1978). It is possible that syntactic function words are not processed by the same mechanism that processes content words. Their reference may be secured only indirectly as part of a referring phrase. However, the main claim of the reference-securing hypothesis is that reference securing is achieved by the activation of phonetic information that allows the securing of a later reference.

A MODEL OF AUTOMATIC SPEECH ACTIVATION

If name codes are useful for reference securing and if phonetic information is part of a name code, then reading is more efficient if phonetic information is part of lexical access. Of course, if a text remains available to the reader, regressive eye fixations can reaccess the word as necessary. But such regressive eye movements do not seem to be very frequent. An efficient system would automatically activate phonetic information during access and use regressive reaccessing only when necessary.

The key process is activation as part of lexical access. Figure 4-2 shows the sketch of the model described in Perfetti and McCutchen (1982) adopted from Rumelhart and McClelland (1981). The question of speech recoding becomes irrelevant in this model. Phonetic activation is not a first step to lexical access; rather, it is part of the access, sometimes reaching a high level prior to the completion of access and sometimes not. The former would look like "recoding" and the latter would not.

The model in Figure 4-2 represents a display of information available over a brief period. It is, in part, an extension of the interactive model described in Chapter 2. There, letters, features, and words were represented at different levels. Figure 4-2 ignores the feature level and shows instead a phoneme level. The horizontal time line represents a brief period, less than a second.

A single word is activated by the lexical access processes described in Chapter 2. Visual information activates elements in memory. The first enclosed box indicates mutual activation of words, letters, and phonemes. As this is a fully interactive model (see Chapter 2), activation flows from the letter level to the word level and vice versa. Equally important, activation flows from the letter level to the phoneme level (and vice versa) *and* from the word level to the phoneme level. This means that as a word is identified, its constituent phonemes are activated. If the identification process is slow, the phoneme level gets a lot of activation from the letter level before a word decision is made. In this case, word identification will be affected by phoneme activation. If the word is identified fairly rapidly, then a phoneme activation will lag behind by a few milliseconds. The eye can move forward to its next fixation during this period.

What is critical is the assumption that some phonetic information linked to the word has been activated by the time the next word is accessed.

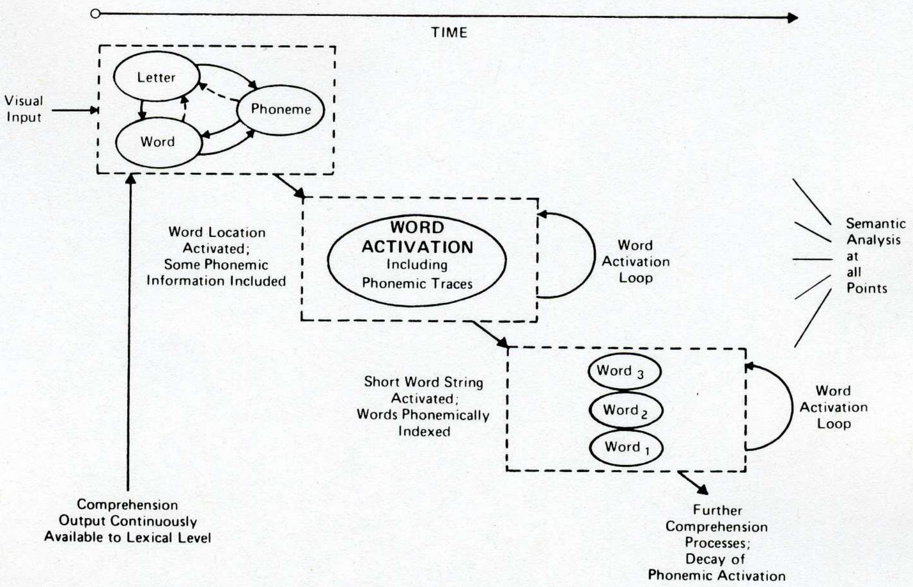

Figure 4-2. Schematic view of an automatic speech-activation model. Visual perception of letters activates both words and phonemes consistent with the current status of letter identification. Following Rumelhart and McClelland (1981), there is feedback (e.g., word-to-letter and word-to-phoneme) as well as feedforward in this process. Automatic phonemic activation occurs during this process and as a result of a word identification decision being made (center box). Thus, a word can be identified before phonemic activation has achieved a high level, but not without some activation occurring. Continuous processing of the word (center box) and reprocessing it as part of a word string (bottom right box) maintain phonemic activation. The processes are continuous (the boxes are not stages) and semantic analysis does not await speech information.

This is represented by the middle frame of Figure 4-2. The word activation loop represents the potential for continuous reprocessing of a word based on its speech sounds. The last frame shows what is activated following two more visual lexical accesses. All words are activated simultaneously and are secured by phonetic information. There are decay functions for this information, and the activation loop helps retard this decay.

Comprehension, in this model, occurs at all points during the process. A word's meaning can be immediately encoded as far as possible. However, there is a memory for the word itself in the form of a phonetic name that can be reaccessed. This name is also part of the propositional encoding in memory. Thus, automatic phonetic activation serves the reader's comprehension. It makes a word available in memory for at least a short time.

Comprehension and Subvocalization

There is evidence for the assumption that comprehension and memory processes are supported by speech sounds. In fact, a general property of short-term memory is that it relies heavily on speech sounds (e.g., Atkinson & Shiffrin, 1968). It may be quite possible for short-term memory to sustain a visual code rather than a speech code (Kroll, Parks, Parkinson, Bieber & Johnson, 1970), and for the congenitally deaf, one must assume a nonspeech coding system. However, except in such special circumstances, short-term memory for both visual and auditory information includes a speech-based code. Perhaps the original experiment of Conrad (1964) demonstrates this fact most simply. When subjects were presented printed letters and then asked to write them down, their errors were based on phonetic similarity more than visual similarity; for example, *F* was confused more with S than with *P*, and *V* was confused more with *B* than with *N*.

More relevant for ordinary reading, however, are studies in which the subject does some reading while engaged in tasks that reduce the chances for subvocalization. The logic of such studies is that speech processes, in the form of subvocalization, are helpful to memory, i.e., the reader pronounces the words silently during reading. But subvocalization is not quite what is meant by the assumption that speech processes support memory and comprehension in general. Subvocalization, in fact, may not be typical of skilled reading unless the text is of some difficulty (Hardyck & Petrinovich, 1970).

The research of Hardyck and Petrinovich (1970) and Locke and Fehr (1970), among others, has used measures of electromyographic (EMG) activity (see also McGuigan, 1970; Sokolov, 1972). Electrodes are attached to muscles involved in speech production, and measurements are taken of the amount of speech muscular activity during silent reading. Hardyck and Petrinovich found that the amount of EMG activity from the larynx could be controlled by readers who were given feedback by a bell that rang when their EMG activity increased. When readers kept their laryngeal EMG at a low level, the comprehension suffered for difficult passages but not for easy ones. The speech mechanism appears to be very specific to the sounds that would be produced by subvocalization. Thus Locke and Fehr found labial EMG activity to be increased during processing of words in a memory task with labial consonants, such as *p*, *b*, and *m*.

Such research demonstrates that speech muscle activation may accompany silent reading on some occasions. However, it does not suggest that such activity is characteristic of normal reading, only that it is characteristic of difficult reading or for memory rehearsal. Furthermore, it cannot directly address whether subvocalization activity is functional in the reader's attempt to remember and comprehend or whether it is merely a by-product of this attempt. Research that has addressed this issue has required subjects to produce overt speech while reading. This concurrent vocalization should interfere with subvocalization and, if subvocalization is functional for memory and comprehension, some adverse effects should be seen. The results of this research are quite mixed. Some experiments find concurrent vocalization to reduce comprehension (Kleiman, 1975; Levy, 1975; Slowiaczek & Clifton, 1980) while others do not (Levy, 1978). Still other studies find comprehension accuracy, but not comprehension times, to be affected (Baddeley & Lewis, 1981). The conflicting results are partly a matter of how demanding a given comprehension measure is (Slowiaczek & Clifton, 1980). They are also dependent on the extent to which the concurrent vocalization task makes demands on central processing (Baddeley & Lewis, 1981; Waters, 1981). For example Waters had an assessment of how much a concurrent vocalization task interfered with monitoring a tone (i.e., with a nonreading task). It interfered with reading no more than with the nonreading task. Overall, the important conclusion seems to be this: Silent reading is affected by a reader's concurrent vocalization. It affects both comprehension and memory. But it does so largely because it uses the same limited-capacity working memory that is needed to assemble and integrate propositions. For a more thorough discussion of this see McCutchen and Perfetti (1982) and Perfetti and McCutchen (1982).

This last point, that concurrent vocalization interferes with reading insofar as it uses processing resources, implies one of two things. Either speech processes themselves are not functional in reading or subvocalization is not the relevant speech process in reading. The second conclusion may be the correct one. We have observed previously that subvocalization is not a characteristic part of most reading, although it may accompany difficult reading. Baddeley and Lewis (1981) suggest that subvocalization serves an *articulatory loop*, a rehearsal mechanism of working memory (Baddeley & Hitch, 1974). When concurrent vocalization has an effect on reading it is because it has interrupted the use of this rehearsal loop by the reader. However, according to Baddeley and Lewis (1981), this will not be a problem for ordinary comprehension but only for a process which has to use the rehearsal loop. These include specific memory tasks and other experimental tasks that demand some reprocessing of words.

What then remains of the speech processes that are automatically activated by lexical access? These cannot be the processes of a rehearsal loop because the latter is an optional part of processing. It may be "turned off" for most reading. If we look again at Figure 4-2, we can see the general solution to this problem. The middle part of the Figure 4-2 represents automatic word activation. It happens routinely and makes available to the reader a phonetic code. The lower right part of the figure represents automatic activation of additional words. (Phenomenologically, this activation may produce something like what Baddeley and Lewis (1981) called an "acoustic image.") Additionally, the word activation loop is available for activated words, as shown in Figure 4-2. These loops are optional. Concurrent vocalization can interfere with these loops but apparently not with the initial speech activation. Indeed, if this initial activation is automatic it should be resistant to interference (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977).

The relationship between automatic speech activation and other implicit speech processes may be considered as an activation continuum. Under ordinary circumstances, speech mechanisms in reading produce only very pale copies of actual speech. In fact they are not always good copies, because they may include only some of the phonemes, especially the initial phonemes of a word. Under more demanding circumstances, the speech mechanism in reading produces better copies, more similar to actual speech, although of course still unspoken. This general state of affairs can be represented by an activation continuum, as shown in Figure 4-3.

The speech activation continuum represents the degree of activation of a speech mechanism. Normal silent reading activates the mechanism above its resting level. However the representation available at this level is abstract (phoneme based) and impoverished. Increasing activation brings speech motor commands to near threshold level. This is the level at which EMG recordings reveal evidence of speech muscle activation. At a higher level, some of the motor commands get executed and subvocal speech is produced. Actual vocalization may occur at the highest level as voicing commands are produced. By this view, the speech process most used in reading is a lower level of activation that results from lexical access automatically. Vocalization is irrelevant for this activity but not for memory-maintaining activity.

Evidence for speech activation

Since evidence for low levels of speech activation cannot come from concurrent vocalization tasks, such evidence must be sought elsewhere. In fact, evidence for the assumption that speech processes are *automatically* activated is in short supply. However there are some demonstrations that speech processes are activated in silent reading.

One demonstration comes from the visual tongue-twister effect (McCutchen & Perfetti, 1982). This research assumed that the speech codes activated by words have two features consistent with the speech activa-

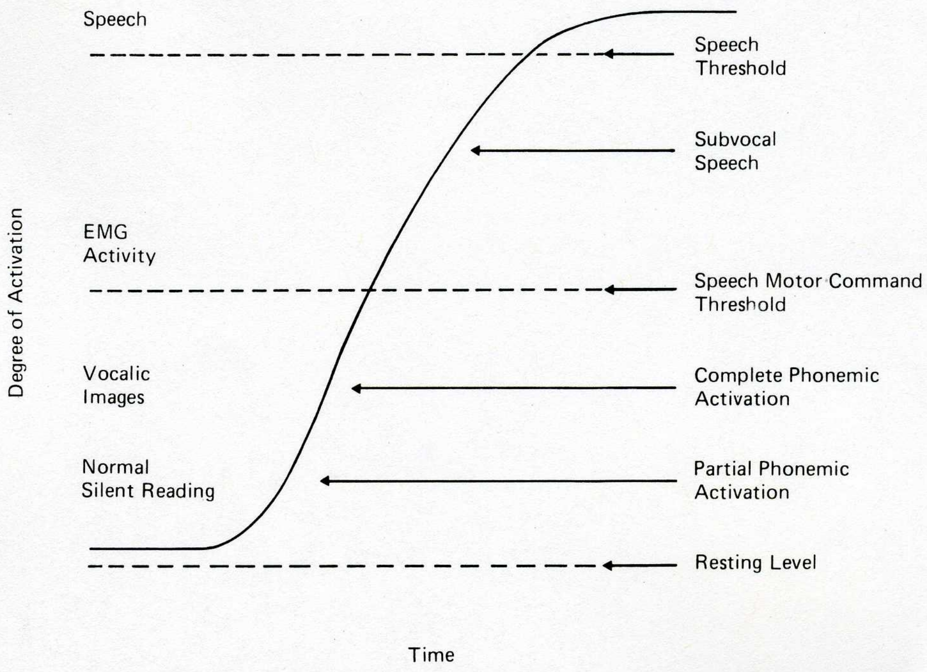

Figure 4-3. Speech activation continuum with thresholds for motor activity. The degree of speech activation increases from a resting level through low levels of phoneme activation through subvocal and true speech. From Perfetti & McCutchen (1982).

tion model discussed in this chapter: (1) They include consonants as well as vowels, and (2) they may be incomplete, in which case they include word-initial phonemes. These assumptions in part reflect the fact that consonants are high in information and that lexical access includes a priority for word beginnings (see Perfetti & McCutchen, 1982). McCutchen and Perfetti further assumed, following the reference-securing assumption, that the distinctiveness of a phonetic code would be threatened by repetition of that code within a sentence. It would take more processing to secure the reference of words that started with the same phoneme. Thus tongue twisters, sentences that repeat initial consonants, e.g., *Peter Piper picked a peck of pickled peppers*, should take longer to read, even in silent reading.

This is what the McCutchen and Perfetti experiments found. The semantic acceptability of sentences such as (1) took longer to decide than sentences such as (2).

(1) *The detective discovered the danger and decided to dig for details.*
(2) *The investigator knew the hazard and chose to search for answers.*

These sentences have nearly identical syntactic structures and comparable semantic interpretations. Thus, the tongue-twister effect is not due to semantics.

The possibility that the effect is visual, i.e., that the reader loses his place during visual processes, is lessened by other manipulations. In one, the tongue-twister effect is found for sentences containing *different* word-initial letters which correspond to phonemes sharing a phonetic feature. For example, both *d* and *t* occurred in some sentences (/d/ and /t/ are both alveolars), and *b* and *p* both occurred in other sentences (/p/ and /b/ are bilabials). Also in a later experiment with children, the visual factor was reduced by having some words in uppercase and some in lowercase. Thus the tongue-twister effect seems to be due to phonetic activation during silent reading. However, it does not really demonstrate that the activation is automatic.

Another demonstration, perhaps more convincing for the assumption that activation is automatic, comes from Petrick and Potter (1979). They presented words one at a time very rapidly to subjects, at a rate of 12 words per second. Subjects were presented with a probe word immediately following the sentence, with the task of deciding whether it had occurred in the sentence. The key result is for probes that had not actually occurred. When a probe was phonetically similar to a word that had actually occurred, subjects took longer to reject it (and they made more errors). Importantly, this did not happen for probes that were only visually similar to words from the sentence. Thus, the interpretation is that subjects rapidly access automatically activated phonetic codes as well as semantic ones (semantically related probes took even longer to reject). Since the effect of a phonetically related probe occurred within 80 milliseconds of the final word of the sentence, it is plausible that the phonetic activation had taken place when the word in the sentence was accessed. The alternative possibility is that after the sentence is read, phonetic activation occurs as the probe initiates a memory search.

Perhaps the most direct demonstrations that phonetic activation may occur automatically during lexical access come from experiments presenting single words. The Lukatela et al. (1980) research with Cyrillic and Roman alphabets (described earlier) suggests that automatic phonetic activation occurs and cannot be suppressed. In such situations there is no opportunity for a later memory process. Navon and Shimron (1981) took advantage of some interesting characteristics of the Hebrew alphabet to provide such a demonstration. Hebrew marks vowel information by diacritic symbols beneath letters, which are largely consonants. However, except for the first school instruction in reading and for certain traditional works, most Hebrew texts omit the vowel symbols. Consonants carry the needed information, and vowel symbols are essentially redundant.

Thus, it was possible for Navon and Shimron (1981) to compose words in Hebrew script that either had redundant vowel symbols or no vowel symbols. Furthermore, some of the vowel symbols were incorrect, i.e., they could accompany some consonants but not the particular ones they were paired with in the experiment. Within this class of incorrectly paired symbols, some maintained the correct vowel sound despite being wrongly paired, while others did not. The interesting result of Navon and Shimron is that the time for skilled readers to read the word aloud was not affected by its having an incorrect vowel symbol, unless the incorrect symbol did not maintain the right vowel sound. By ruling out a visual explanation of this result, Navon and Shimron (1981) suggest that it reflects an automatic activation of letter sounds. This happens with or without the Hebrew vowel symbol and is interfered with only when the symbol triggers an incorrect sound. Of course, it is possible that *naming* the word, not the lexical access, is what requires the phonetic activation.

The final demonstration involves the backward visual masking of words. Naish (1980) and Perfetti and Bell (unpublished experiments) have demonstrated that phonetic effects occur in backward visual masking. The experimental situation used by Perfetti and Bell is illustrated in Table 4-1. A word is briefly exposed for 33 milliseconds. It is followed immediately by a second word (actually a pseudoword), which is on for 20 milliseconds. This pseudoword is followed by a row of *X*'s. Thus, the pseudoword masks the first word, and the *X*'s (a pattern mask) mask the second word. The subject is to identify the first word. In this context, masking is essentially a process in which one visual display interrupts the processing of an earlier display by erasing its stimulus features. (For theoretical discussions of masking processes see Breitmeyer & Ganz, 1976; and Turvey, 1973).

The critical feature in these experiments is the relation of the pseudoword mask to the word target (e.g., *main*). Because the mask interrupts processing of the target, identification is difficult and subjects make er-

TABLE 4-1 Word-masking Experiment

Trial event	*Duration (in msec)*
main	30
MAYN or MARN or CRUB	25
XXXXXXX	until subject responds

rors. But when the mask has many of the same letters as the target (e.g., MARN), identification is much improved compared with the case in which the mask does not share letters with the target (e.g., CRUB). In the Perfetti & Bell experiments, this enhancement of target identification occurs even when the target is in lowercase and the mask is in uppercase. Most important is the fact that identification is also improved if the mask has many of the same phonemes as the target, even if the letters are not the same. For example, when the target word is *quote* and the pseudoword is *KWOAT* there is high phonetic overlap not due entirely to letter overlap. (In fact *quote* and *KWOAT* are pseudohomophones. They could have identical pronunciations.) Identification of the target word is enhanced in these cases, compared with a word sharing only letters. Thus, there is an enhancement of recognition when the mask preserves either the letters or the phonemes of the word. The word's initial access has been accomplished by something like the model shown in Figure 4-2 (see also Figure 2-2). Its phonemes are partly activated when the pseudoword mask interrupts processing. As the masking word is processed, its features are activated. If these features, including phonetic ones, are similar to those already activated by the target word, then these previously activated features continue to receive some activation. Thus, the experiments of Naish (1980) and Perfetti and Bell demonstrate that a word presented too briefly to be easily recognized has its phonemes as well as its graphemes activated. This seems to be automatic phonetic activation during lexical access.

SUMMARY

The question of whether skilled silent reading includes implicit speech processes has been difficult to answer, despite its appearance as one of the earliest theoretical issues of reading. Much research has examined whether a stage of speech recoding occurs prior to lexical access and has seemed to demonstrate that it does not in skilled reading.

On the other hand, there are reasons to assume that speech processes occur as support for memory and comprehension. For example, the reader's need to secure reference is facilitated by having a phonetically indexed name code. This code can be automatically activated as part of lexical access, consistent with an interactive model of lexical access as presented in Chapter 2. The outcome of this activation process is a speech code that is not quite the same as a covert vocalization. Subvocalization may be the result of especially high levels of speech activation in response to memory demands.

There is evidence for speech activation during silent reading from sentence-reading tasks. Evidence from word-reading tasks may more directly demonstrate that activation is automatic.