The prominent view of visual word recognition theorists is that readers must first match the visual stimulus to some internal representation before meaning of the stimulus word becomes available. Very simply, a word must be recognized before its meaning can be accessed. For example, in Morton's classic logogen model (Morton, 1969), word recognition devices (logogens) must receive sufficient activation via featural detectors before the word is identified. The meaning of the word does not become available before this theoretical threshold has been reached (see Becker, 1980; Forster, 1979; Norris, 1986, for similar views).

The notion that recognition must precede meaning access would appear to be a reasonable assumption, i.e., how could a reader access the referent of a word without first determining the identity of the word that was presented? In the present chapter we challenge this assumption. We first present a brief review of the tasks available that provide the major data bases for the available models of word recognition. Based on the studies reviewed in this section, we suggest that theorists do not have available an adequate measure of recognition processing that is unequivocally devoid of meaning access. We then turn to a review of the literature to determine whether there are any studies available that provide compelling evidence that meaning can contribute to the word identification process. Finally, we present a brief discussion of a theoretical framework that predicts a meaning-level influence in word identification.

Before turning to a more detailed discussion of these issues, there are two points that need to be noted. First, the emphasis in the present discussion is on studies that involve isolated word recognition or word recognition in neutral or unrelated contexts. When words are presented in associatively/semantically related contexts, performance in word recognition tasks is facilitated compared to
when words are presented in neutral or unrelated contexts. Such priming effects would appear to indicate that meaning access, via a related context, has taken place before word recognition is completed. Although there are alternative accounts of such priming effects (see Neely, 1990, for a review), such effects and the accompanying theories are only minimally related to the present thesis. Clearly, the fact that meaning can be accessed via a related context does not presuppose that meaning of a word in an isolated context can influence the word recognition process. It is this latter issue that is the focus of present discussion.

The second point that needs to be discussed is an operationalization of the terms word identification and word access. Here we equate word identification with the processes that lead up to and include the decision to execute a response (i.e., the Magic Moment, see Balota, 1990). Word access involves accessing an internal representation that corresponds to a word. For example, in Morton's logogen model, lexical access would be reflected by activation at a lexical representation that may or may not be sufficient to exceed the threshold needed to execute a response, whereas, lexical identification would be reflected by the activation at a lexical representation that is sufficient to execute a response.

Here, as in all interpretations that we are aware of, lexical identification is primarily operationalized via three major tasks: lexical decision, pronunciation, and threshold identification. Hence, we operationalize lexical identification as (a) the time it takes, after stimulus presentation, for the subject to produce a sound that triggers a voice-operated relay (pronunciation), (b) the time it takes, after stimulus presentation, to produce a word/nonword response that triggers a microswitch (lexical decision), or (c) the stimulus conditions (e.g., duration, intensity, etc.) necessary to identify a stimulus (threshold identification). Because these three tasks provide the major data base for current models of word identification, it is necessary to briefly specify any limits that are inherent in these tasks concerning their use as windows into the magic moment of word identification.

LEXICAL DECISION, PRONUNCIATION, AND THRESHOLD IDENTIFICATION: THE TOOLS OF THE TRADE

On the surface, these three tasks would seem to be faithful reflections of word identification. For example, making a lexical decision would appear to involve the point in time when an internal representation has been sufficiently matched by stimulus-driven information. Thus, the button press in the lexical decision task (LDT) would appear to be a reasonable reflection of word identification. Likewise, it would appear that naming a word simply involves a match between stimulus-driven processing of the stimulus word and some lexical representation. Once this match is completed, the appropriate sequence of motor codes is en-
gaged for output. Thus, the onset of pronunciation would also appear to be a reasonable reflection of word identification. Finally, with respect to the threshold identification task, the stimulus duration needed to identify a briefly presented stimulus word would appear to reflect the amount of stimulus information needed to surpass the lexical threshold. Hence, accuracy at a given stimulus duration may be used as a metric for the ease of word identification. Although at first glance these three tasks are tantalizingly simple, there are aspects of the tasks that question whether they can be used as sources of data to build premeaning access models of word identification. We shall now turn to a brief discussion of these aspects.

Lexical Decision Performance

First, let us consider the LDT. It is now clear that the LDT cannot be viewed as a faithful reflection of only processes leading up to and including word identification (e.g., Balota & Chumley, 1984; Balota & Lorch, 1986; Besner & McCann, 1987; Chumley & Balota, 1984; de Groot, 1983; Keefe & Neely, 1990; Lorch, Balota, & Stamm, 1986; Neely, 1990; Neely & Keefe, 1989; Neely, Keefe, & Ross, 1989; Seidenberg, Waters, Sanders, & Langer, 1984; West & Stanovich, 1982). The major problem with this task is very simple. The LDT is not only a word identification task but is also a discrimination task, in which subjects are forced to discriminate words from nonwords. Often one finds that variables that presumably only influence word identification processes are confounded with the discrimination component of this task. For example, consider the finding that high-frequency words produce considerably faster lexical decisions than low-frequency words. Here, manipulations of stimulus frequency/familiarity are confounded with a type of information, familiarity, that would be available to the reader, and very helpful in making the word/nonword discrimination, i.e., words are more familiar than nonwords. Specifically, compared to high-frequency words (e.g., CAKE), low-frequency words (e.g., SHAM) are more similar to, and therefore more difficult to discriminate from, nonwords (e.g., ROLM) on the relevant familiarity dimension. A similar discrimination confounding occurs in semantic priming studies. Briefly, if there is a relationship between the prime and the target then the target must be a word, however, if there is no relationship between the prime and the target then the target can either be a word, primed by an unrelated word, or a nonword. Hence, the discrimination is more difficult on unrelated trials, thereby slowing response latency (see, e.g., Balota & Lorch, 1986).

The major point to note concerning the LDT is that there is often a confounding between the manipulation of interest and the information that subjects can use to make the word/nonword discrimination. Subjects are very sensitive to such confoundings (see Chumley & Balota, 1984), and these confoundings can produce exaggerated influences of variables, thereby misdirecting theories of word
recognition. Most importantly, for the present discussion, these confoundings cause concern regarding the utility of the LDT as an unequivocal reflection of lexical identification without meaning access.

Pronunciation Performance

The pronunciation task does not involve the problem of subjects making a discrimination between words and nonwords, and therefore, there should be little contribution of contaminating decision processes on overall response latency. However, even in the pronunciation task a variable could potentially play a role at many different loci in this task. Specifically, in pronouncing a word, the reader not only has to recognize the stimulus but also has to output the recognized word. Thus, a variable could have an impact on identification processes and/or on processes after identification that are tied to the output of the response. Consider the impact of word frequency. Balota and Chumbley (1985) used a delayed-pronunciation task to tease apart the impact of frequency on word identification from its impact on output processes. In the delayed-pronunciation task, subjects are given sufficient time to recognize the word and then are presented a cue to pronounce the word aloud. If frequency only influences the identification stage in this task, then one should not find a frequency effect after subjects have had sufficient time to recognize the stimulus. However, Balota and Chumbley found that subjects still produced a significant frequency effect in this task, even though they were given up to 1400 ms to recognize the stimulus (also see Connine, Mullennix, Shermoff, & Yelen, 1990).

More recently, Balota, Boland, and Shields (1989) have demonstrated that prime-target associations can influence both onset latencies and production durations in a delayed-pronunciation task. For example, in one of their experiments, they reported that subjects were both faster to begin their production of two related words, compared to two unrelated words, even though they had 1400 ms to recognize the stimuli (also see Dallas & Merkle, 1976; Midgley-West, 1979). Interestingly, Balota et al. also found that the production durations were shorter for related words than unrelated words.

The important point for the present discussion is quite simple. Because variables appear to influence processes after subjects have sufficient time to recognize the stimulus, one cannot unequivocally use data from the pronunciation task as a metric of processes that lead up to word identification, and hence, take place before meaning access.

Threshold Identification Performance

Possibly, the threshold identification task is a pure measure of lexical identification. This task presumably taps the amount of stimulus information necessary to surpass some lexical threshold to make the correct stimulus identification. Unfor-
fortunately, there is again an inherent problem that often produces interpretive difficulties when this task is employed.

Consider the perceptual problem a subject encounters while attempting to identify a highly degraded stimulus. The subject may be able to determine that the word is short and most likely begins with the letters th. Because the subject is typically not penalized for guessing, the subject proceeds to make a best guess at the target identification. Now, what types of information might the subject use to make such a guess? One obvious piece of information may be frequency information. That is, when presented with such a guessing situation, the subject may be more likely to guess the high-frequency word thread over the low-frequency word thump (see Catlin, 1973). If frequency influences threshold identification in this situation, one cannot conclude that frequency only influences lexical identification, but one must also consider its influence on the guessing biases that subjects may bring into the experimental situation. This concern with threshold identification is a general methodological concern in perceptual studies and falls under the general heading of sophisticated guessing. The point is quite simple, one cannot infer that a variable is influencing lexical identification in the threshold identification task, without taking into consideration all the potential guessing biases that subjects may bring into the experimental setting. The use of preexisting memory information to maximize accuracy in a perceptually degraded situation will most likely not be highly related to the processes that are involved in fluent word identification.

In sum, the tools of the trade in word recognition research that theorists use to build models that assume mandatory word identification before meaning access have characteristics that do not allow unequivocal interpretation of the data obtained from these tasks. The important point here is that if one cannot disentangle word identification from meaning access, then it may be more fruitful to consider a framework that might allow meaning to contribute to word identification.

On the other hand, it is possible that there is no evidence available even from lexical decision, pronunciation, and threshold identification tasks that indicate that meaning can contribute to word identification processing. If this were the case then one might at least feel cautiously comfortable with the extant models. We now turn to an evaluation of the relevant literature.

THE CONTRIBUTION OF MEANING TO WORD IDENTIFICATION: A REVIEW OF THE LITERATURE

As noted earlier, it is important to operationalize word identification. Here, as in the vast majority of word recognition studies, we shall operationalize word identification as the ability to (a) name a word aloud, (b) discriminate words from nonwords, or (c) identify a degraded stimulus. Hence, we will be using the
above-criticized tasks in developing our review, and acknowledging caveats in interpretations along the way. The important point is that models of word recognition have also relied on these tasks and have come to the conclusion that meaning does not influence word identification. We are simply evaluating the relevant literature regarding this theoretical predisposition.

In the present review we address the four following areas: (a) the influence of pre-existing meaning representations (e.g., concreteness, contextual availability, and polysemy) in word identification; (b) associative reaction time as a task and as a predictor of lexical decision/pronunciation performance; (c) the influence of attaching meaning to previously nonlexical strings; (d) meaning access without word identification. A summary of the findings concerning each of these four areas along with a description of the tasks and potential confounds is displayed in Table 8.1.

The Influence of Pre-existing Meaning Representations on Word Identification

The studies reviewed in this section address whether there are meaning-level characteristics of words that influence word identification. Consider, for example, concreteness effects in word identification. If one could provide evidence that concrete words are recognized faster, or slower, than abstract words, then this would suggest that the fact that a word can be the object of a sense verb (i.e., touch, see, hear, etc.) is sufficient to modulate word identification. This would be an example of a meaning-level variable (semantic class) participating in word identification processes.

Before we turn to this literature, it should be noted that a major difficulty in this area, along with other areas reviewed below, is whether there has been a confounding between the target variable (e.g., concreteness) and some other variable that may influence task performance. Gernsbacher (1984) has pointed out that many of the studies that have addressed meaning variables in isolated word recognition have not adequately controlled for the familiarity of the stimulus. Hence, this will be a general theme cutting across the studies reviewed in this and the following sections.

(1) Concreteness Effects

The question of whether there are differences in word recognition for semantically abstract versus concrete nouns has been prominent in the word recognition literature. In fact, the concrete/abstract distinction has been the most widely investigated dimension regarding semantic effects in word identification. Despite the widespread interest in this area, at least until recently, the results in this area have been equivocal. We first describe studies involving threshold identification and then turn to studies involving lexical decision and pronunciation performance.

It should be noted that in most studies discussed in this section concreteness has been confounded with imageability (see Boles, 1983, 1989). Because these
### TABLE 8.1
The Contribution of Meaning to Word Identification: A Review of the Literature

<table>
<thead>
<tr>
<th>Study</th>
<th>Task</th>
<th>Effect</th>
<th>Potential Confounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winnick &amp; Kressel (1965)</td>
<td>TL^a</td>
<td>$A_{rt} + C_{rt}$</td>
<td>RB, d, CA, j, f</td>
</tr>
<tr>
<td>Paivio &amp; O’Neill (1970)</td>
<td>TI</td>
<td>$A_{rt} &lt; C_{rt}$</td>
<td>repetition of items, CA, RB</td>
</tr>
<tr>
<td>Richards (1976)</td>
<td>TI</td>
<td>$C_{rt} &lt; A_{rt}$</td>
<td>FNM, g, I, RB</td>
</tr>
<tr>
<td>Boles (1983)</td>
<td>word/digit report; word identification in left/right visual fields</td>
<td>familiarity of word had more impact than I or C</td>
<td>RB, CA</td>
</tr>
<tr>
<td>Boles (1989)</td>
<td>TL/incidental learning</td>
<td>In Expt. 1, main effects for I and interaction with visual field primarily for intentional learning instructions</td>
<td>FNM, RB, CA</td>
</tr>
<tr>
<td>Rubenstein, Garfield, &amp; Millikan (1970)</td>
<td>LDT</td>
<td>$C_{rt} &lt; A_{rt}$</td>
<td>stimulus problems (see Clark, 1973), CA, I</td>
</tr>
<tr>
<td>James (1975)</td>
<td>Expt. 1 &amp; 2 (LDT)</td>
<td>$C_{rt} &lt; A_{rt}$ for LF words; $C_{rt} = A_{rt}$ for HF words</td>
<td>FNM, I</td>
</tr>
<tr>
<td>James (1975)</td>
<td>Expt. 3 &amp; 4 (Naming)</td>
<td>$C_{rt} &lt; A_{rt}$ for both LF &amp; HF words</td>
<td>FNM (E1 - E3), (although familiarity manipulated episodically in E4)</td>
</tr>
<tr>
<td>Richardson (1976)</td>
<td>LDT/Naming</td>
<td>$A_{rt} = C_{rt}$</td>
<td>FNM, CA</td>
</tr>
<tr>
<td>deGroot (1989)</td>
<td>LDT</td>
<td>replicated James (1975), E1 &amp; E2</td>
<td>FNM, CA</td>
</tr>
<tr>
<td>Day (1977)</td>
<td>LDT</td>
<td>$C_{rt} &lt; A_{rt}$ only in left visual field; $C_{rt} = A_{rt}$ only in right visual field</td>
<td>FNM, CA, repetition of items, I</td>
</tr>
<tr>
<td>Kroll &amp; Mervis (1986)</td>
<td>LDT</td>
<td>$E1, E2$: conditions blocked and mixed; $C_{rt} &lt; A_{rt}$; $E1$: conditions blocked A following C ($C_{rt} &lt; A_{rt}$); conditions blocked C following A ($C_{rt} = A_{rt}$)</td>
<td>FNM, CA, I</td>
</tr>
<tr>
<td>Bleasdale (1987)</td>
<td>Expt. 1 (Naming)</td>
<td>$C_{rt} &lt; A_{rt}$</td>
<td>CA</td>
</tr>
<tr>
<td></td>
<td>Expt. 2-4 (LDT)</td>
<td>$C_{rt} &lt; A_{rt}$</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Study</th>
<th>Task</th>
<th>Effect</th>
<th>Potential Confounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coltheart, Patterson, &amp; Marshall (1980)</td>
<td>RA (U)</td>
<td>C %E = A %E</td>
<td>FNM, CA, RB</td>
</tr>
<tr>
<td>Richardson (1975a)</td>
<td>TI</td>
<td>C %E = A %E</td>
<td>FNM, CA</td>
</tr>
<tr>
<td>Richardson (1975b)</td>
<td>TI</td>
<td>C %E = A %E</td>
<td>FNM, CA</td>
</tr>
<tr>
<td>Shallice &amp; Warrington (1975)</td>
<td>RA (U)</td>
<td>C %E = A %E</td>
<td>FNM, CA, RB</td>
</tr>
<tr>
<td>Shallice &amp; Warrington (1980)</td>
<td>RA (U)</td>
<td>C %E = A %E</td>
<td>FNM, CA, RB</td>
</tr>
<tr>
<td>Seymour (1990)</td>
<td>selective interference (Naming)</td>
<td>C rt = A rt</td>
<td>FNM, CA, RB</td>
</tr>
<tr>
<td>Coltheart (1980)</td>
<td>RA (U)</td>
<td>similar to Day (1977)</td>
<td>FNM, CA, RB</td>
</tr>
<tr>
<td>Saffran, Boggy, Schwartz, &amp; Marin (1980)</td>
<td>RA (U)</td>
<td>C rt &lt; A rt</td>
<td>FNM, CA, RB</td>
</tr>
<tr>
<td>Gernsbacher (1984)</td>
<td>LDT</td>
<td>main effect for familiarity (F), no main effect for C, no F x C interaction</td>
<td>CA</td>
</tr>
<tr>
<td>Schwanenflugel, Hamishfeger, &amp; Stowe (1988)</td>
<td>LDT</td>
<td>C rt &lt; A rt</td>
<td>CA</td>
</tr>
<tr>
<td>Schwanenflugel, Hamishfeger, &amp; Stowe (1988)</td>
<td>LDT</td>
<td>C rt = A rt (contextual availability unconfounded)</td>
<td></td>
</tr>
</tbody>
</table>

2. Polysomic Effects

<table>
<thead>
<tr>
<th>Study</th>
<th>Task</th>
<th>Effect</th>
<th>Potential Confounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rubenstein, Garfield, &amp; Millikan (1970)</td>
<td>LDT</td>
<td>homograph rt &lt; non-homograph rt</td>
<td>item sampling (see Clark, 1973), CA, FNM</td>
</tr>
<tr>
<td>Rubenstein, Lewis, &amp; Rubenstein (1971b)</td>
<td>LDT</td>
<td>homograph rt &lt; non-homograph rt</td>
<td>item sampling (see Clark, 1973), CA, FNM</td>
</tr>
<tr>
<td>Jastrzembski (1981)</td>
<td>LDT</td>
<td>words with more entries faster than words with few entries in dictionary</td>
<td>meaning metric not adequate, CA, FNM</td>
</tr>
</tbody>
</table>

(continued)
### TABLE 8.1 (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Task</th>
<th>Effect</th>
<th>Potential Confounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kellas, Ferraro, &amp; Simpson (1988)</td>
<td>LDT</td>
<td>ambiguous word rt faster than unambiguous word rt; ambiguous words demand less attentional resources than unambiguous words</td>
<td>CA</td>
</tr>
<tr>
<td>Kellas, Simpson, &amp; Ferraro (1988)</td>
<td>LDT</td>
<td>(same as above)</td>
<td>CA</td>
</tr>
<tr>
<td>Millis &amp; Button (1989)</td>
<td>LDT</td>
<td>words with many meanings responded to faster than words with few meanings</td>
<td>CA</td>
</tr>
</tbody>
</table>

**Associative Reaction Time as a Task and as a Predictor of Lexical Decision and Pronunciation Performance**

| deGroot (1989)      | LDT/Naming           | HF C words responded to faster than LF C words; C better predictor than frequency | FNM, CA             |
| Chumbley & Balota (1984) | LDT                  | speed to produce associate of target word good predictor of LDT            | FNM                 |

**The Impact of Adding Meaning to Previously Nonmeaningful Strings**

| Whittlesea & Cantwell (1987) | letter detection    | letters in nonwords given meaning detected better than letters in nonwords not given meaning |                     |
| Forster (1983)               | LDT, masked repetition priming | obsolete words assigned meaning produce more repetition priming than obsolete words not given meaning |                     |

**Failure to “Recognize” Words that have Activated Meaning Representations**

| Balota (1983)              | TPj (LDT)            | priming even though primes were presented below subjects’ threshold |                     |
| Carr & Dagenbach (1990)    | TP (LDT)             | (see Balota, 1983)                                                  |                     |

(continued)
TABLE 8.1 (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Task</th>
<th>Effect</th>
<th>Potential Confounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dagenbach, Carr, &amp; Wilhelmsen (1989)</td>
<td>supra.TP (LDT); masked priming (LDT)</td>
<td>(see Balota, 1983)</td>
<td></td>
</tr>
<tr>
<td>Fowler, Wolford, Slade, &amp; Tassinary (1981)</td>
<td>TP (LDT)</td>
<td>(see Balota, 1983)</td>
<td></td>
</tr>
<tr>
<td>Marcel (1983)</td>
<td>TP (LDT)</td>
<td>(see Balota, 1983)</td>
<td></td>
</tr>
</tbody>
</table>

Further Impacts of Meaning on Early Perceptual Processing

<table>
<thead>
<tr>
<th>Study</th>
<th>Task</th>
<th>Effect</th>
<th>Potential Confounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schendel &amp; Shaw (1976)</td>
<td>line segment detection</td>
<td>line segment from a letter detected better than line segment presented alone</td>
<td></td>
</tr>
<tr>
<td>Weinstein &amp; Harris (1974)</td>
<td>line segment detection</td>
<td>line segment in three-dimensional forms better detected than line segment in two-dimensional forms</td>
<td></td>
</tr>
<tr>
<td>Virzi &amp; Eghe (1984)</td>
<td>name target word or target word color</td>
<td>report colors/words that were not present in display (illusory conjunctions at meaning level)</td>
<td></td>
</tr>
</tbody>
</table>

TT = Threshold Identification
A = Abstract Words
C = Concrete Words
RB = Response Bias
CA = Contextual Availability not measured
I = Imageability
FNM = Familiarity Not Measured
RA(U) = Reading Aloud (Untimed)
WFNM = Word Frequency Not Measured
TP = Threshold Priming

Two factors are highly correlated, unless otherwise specified, one can assume that the concreteness effects also involve contaminated imageability effects. Of course, for the present discussion, an imageability effect would also support the notion that underlying referential information was modulating word recognition. Therefore, at this level, the contamination of concreteness and imageability does not strongly constrain our arguments regarding meaning influences in word recognition.

Concreteness Effects in Threshold Identification Studies. Winnick and Kressel (1965) investigated the influence of word frequency and concreteness on
threshold identification performance. The independent variable in this study was the number of 10 ms exposures that subjects required before they correctly identified the target word. The results indicated that there was no difference between the number of exposures needed to identify abstract and concrete words; both required 9.65 exposures.

Paivio and O’Neill (1970) investigated imageability/concreteness, word frequency, and the meaningfulness of target words in a threshold identification task. In their study, the stimuli were originally presented for 10 ms and then incremented by 5 ms exposure durations until the subject could correctly identify the target word. The results revealed main effects of frequency, imageability/concreteness, and meaningfulness. Thresholds for abstract words were lower than for concrete words. However, further analyses indicated that familiarity, as reflected by untried familiarity ratings by a different group of subjects, was correlated with concreteness, and that concreteness of the stimulus did not account for any variance above-and-beyond the influence of familiarity. However, even with the impact of familiarity partialled out, the correlation between meaningfulness and mean threshold identification, although reduced, was still significant.

A third study of recognition thresholds was conducted by Richards (1976). Richards used the same procedure that Paivio and O’Neill used and found main effects of both word frequency and concreteness across two experiments, with no interactions. Unlike Paivio and O’Neill, Richards found that concrete words had lower identification thresholds than abstract words. Unfortunately, Richards did not control for the familiarity of his stimuli.

A fourth series of experiments has been reported by Boles (1983, 1989). Boles was interested in the influence of imageability, concreteness, and familiarity on word identification in left and right visual fields. The visual field variable is important because it is possible that abstract words only have a single verbal-based representation, whereas, concrete words (and/or high imageable words) have both a verbal-based and an image-based representation, consistent with arguments by Paivio, among others, concerning dual codes. (We return to this issue below.) In the Boles (1983) study, words were presented for 100 ms for an identification response. The results of a series of five experiments yielded an impact of only word familiarity. Thus, Boles concluded that imageability and concreteness did not influence word identification nor did these variables interact with visual field of presentation, and therefore, involve different hemispheric representations.

More recently, however, Boles (1989) has modified this position. In this work, Boles has found that memory instructions can mediate these semantic effects. In his first experiment, subjects participated in a threshold identification study under intentional memory instructions. The results yielded main effects of familiarity and imageability, along with evidence that imageability modulated the right visual field advantage reported in this literature. That is, the right visual
field advantage was smaller for high-imageable stimuli compared to low-imageable stimuli. In the second experiment, subjects were given incidental learning instructions. The results of this study yielded main effects of concreteness and imageability that were in large part eliminated when familiarity of the stimuli was partialled out. Hence, Boles argues that imageability effects and their interaction with visual field can be obtained under intentional learning instructions in a threshold identification task, but not under incidental learning instructions. It should also be pointed out, that concreteness did not have an influence on threshold identification above and beyond familiarity and imageability in any of the Boles studies.

In sum, it is difficult to reach a firm conclusion regarding the threshold identification studies. Beyond the inconsistencies in the data, a further problem is the potential for response biases to play a role in performance. For example, it is possible that when presented a degraded stimulus, subjects are more likely to guess that a concrete word was presented rather than an abstract word. Although interesting, such a response bias would not support an influence of concreteness on word identification. On the other hand, one could argue that simply because concreteness does not account for any variance above-and-beyond the impact of familiarity in threshold identification studies, as Paivio and O’Neill and Boles found, this does not necessarily discount the impact of concreteness on word identification. It is possible that subjects use concreteness in both ratings of familiarity and the processes involved in threshold identification. We now turn to studies that involve response latency measures.

**Concreteness Effects in Lexical Decision and Naming.** An early study that explored the effect of concreteness along with a number of other variables (e.g., word frequency and polysemy) in lexical decision performance was conducted by Rubenstein, Garfield, and Millikan (1970). These researchers proposed that word frequency exerted its influence on lower-order processes while concreteness exerted its influence on higher-order processes. They found main effects of word frequency and polysemy as well as an interaction between concreteness and polysemy. Interestingly, they found that latencies for homographs having one concrete meaning and one abstract meaning were actually shorter than for homographs where both meanings were considered concrete. Because frequency, word length, number of meanings, and relative frequency of meanings were equated across the two sets of homographs, these results would appear to provide evidence for a true meaning effect. Therefore, at this level, it would appear that two semantic variables (concreteness and polysemy) work in consort to modulate word identification performance.

In an often-cited paper, James (1975) also manipulated word frequency and concreteness in a LDT. The results of his first experiment yielded main effects of frequency and concreteness along with an interaction between the two variables. The interaction indicated that low- and medium-frequency abstract words produced slower response latencies than concrete words at the same frequency
levels, whereas there was no impact of concreteness for high-frequency words. This pattern was replicated in his second experiment when the lexical decision list contained nonwords that were homophones of words (e.g., brane). James eliminated any influence of concreteness in his third experiment when only nonpronounceable nonwords (e.g., cbmre) were presented within the list. Of course, this would be predicted because subjects could simply base their discrimination in the LDT on the acceptability of the orthographic code instead of relying on word identification processing.

James also eliminated the concreteness effect in his fourth experiment. This experiment involved two parts. In the first part, subjects were presented the target words in a sentence generation task; in the second part they received the same targets in a LDT. In this experiment, the concreteness effects were eliminated in lexical decision performance. Overall, the results from the James' study indicate that concreteness can modulate word identification for medium- to low-frequency words when subjects cannot rely on orthographic regularity or a high degree of familiarity (due to the earlier sentence completion task) to make lexical decisions. Clearly, the constraints of the LDT are extremely important in interpreting James' results, as he readily acknowledged.

Richardson (1976) addressed concreteness, imageability and lexical complexity effects in both pronunciation and lexical decision tasks. In Richardson's study, lexical complexity simply involved a distinction between suffixed and nonsuffixed nouns. Across three different tasks (pronunciation, lexical decision with pronounceable nonwords, and lexical decision with nonpronounceable nonwords) there were no effects of concreteness, imageability, or lexical complexity, even though there were large effects of word frequency and length in letters. Interestingly, however, it should be noted that non-suffixed concrete words were pronounced 29 ms faster than abstract words when collapsing across imageability levels. In addition, a difference of 28 ms was found between concrete and abstract words in the LDT that contained pronounceable nonwords. However, consistent with the James' study, only a 7 ms effect of concreteness was attained for the LDT when nonpronounceable nonwords were used. Unfortunately, it is difficult to interpret these effects because they were not directly tested by Richardson and the power to detect such effects was relatively low—each experimental condition had only ten subjects, with 10 observations per subject per cell.

Recently, deGroot (1989, Experiment 4) replicated the interaction between concreteness/imageability and word frequency in lexical decision performance with pronounceable nonwords that was reported by James. In addition, in her Experiment 5, she reported a pronunciation study that also produced a significant effect of concreteness/imageability that was additive with the effect of frequency. Thus, deGroot argues that the interaction between concreteness and frequency may be limited to the LDT. More importantly, the deGroot study indicates that concreteness/imageability can influence both lexical decision and pronunciation performance.

Day (1977) addressed whether concreteness effects in the LDT would be
modulated by their visual field (and hence hemisphere) of encoding. As noted above, this is an intriguing extension of dual coding theory. Day reported three lexical decision experiments, all with pronounceable nonwords. The results of these experiments were quite clear. In each experiment, Day reported an interaction between concreteness and visual field, such that when the words were presented to the right visual field (left hemisphere) there was no impact of concreteness, whereas, when the words were presented to the left visual field (right hemisphere) concrete words produced faster lexical decisions than abstract words. Hence, the Day results suggest that not only does concreteness influence lexical decision performance, but it has a specialized right hemisphere impact.

More recently, Kroll and Merves (1986) also addressed the possibility that concrete and abstract words differ in their representational format. In order to test the dual-representation assumption, in their first experiment, Kroll and Merves presented pure lists of concrete words or pure lists of abstract words. The notion was that subjects may be able to attend to the different representational formats for the abstract and concrete words. The results of this experiment indicated that concrete words produced significantly faster (by items but not by subjects) lexical decisions than abstract words. Like James (1975) and Day (1977), these authors found that concrete words produced faster response latencies than abstract words. One might argue based on the results of this experiment that subjects were attending to the different representational formats of the concrete and abstract words. However, in a second experiment, where concreteness involved a within-subjects manipulation, a similar effect of concreteness was found. Hence, Kroll and Merves could not attribute the results of the first experiment to simply attending to different representational formats because a similar result was found when mixed lists were used. Interestingly, in both experiments Kroll and Merves reported an interaction between frequency and concreteness of the same nature reported by James (1975) and deGroot (1989), i.e., the concreteness effect was largest for low-frequency words. In their final experiment Kroll and Merves again blocked abstract and concrete words in a LDT, but now subjects received both blocks in a counterbalanced order across subjects. The results indicated that when subjects were given a pure list of concrete words first and then a pure list of abstract words, there was a large impact of concreteness. However, when abstract words preceded concrete words, there was no impact of concreteness. Instead of attributing these sequencing effects to selectively attending to different representational sources due to exposure to either only abstract or only concrete words, Kroll and Merves suggested that these effects may simply reflect strategic post-lexical effects in the LDT. Even if these results are due to strategic processes, it is clearly the case that subjects can be induced to utilize the concreteness dimension, in a speeded task, thereby suggesting that this is a readily available source of information, and that the recognition of this dimension can, at least, modulate a major task that provides data for models of word identification, i.e., lexical decision performance.
Like Kroll and Mervis, Bleasdale (1987) investigated the effects of concreteness in order to address differences in representational formats, i.e., single vs. dual representational systems. Bleasdale employed a priming paradigm in both the pronunciation task and the LDT. As noted earlier, because we are primarily interested in isolated word recognition processes, it is beyond the scope of the present discussion to detail the influence of related versus unrelated contexts. However, Bleasdale’s study is pertinent because in each of the four experiments, there was a neutral context condition (i.e., the neutral prime was BLANK) and an unrelated context condition. There is no a priori reason to expect such nonrelated contexts to differentially influence word identification processes for concrete and abstract words. Hence, one can use the concreteness effects in the neutral and unrelated prime conditions to determine whether there is any concreteness effect for isolated word recognition. In this light, the Bleasdale results are highly consistent and informative. In the first experiment, which involved a pronunciation task, there was a 10 ms advantage for concrete words over abstract words in the neutral prime condition and a 19 ms advantage in the unrelated prime condition. Although this difference was not directly tested, there was a main effect of target concreteness in this experiment and did not interact with prime context. A similar pattern was found in the three subsequent experiments that employed the LDT. That is, concrete words produced faster response latencies than abstract words by 27 ms for the neutral and 12 ms for the unrelated conditions in Experiment 2, 18 ms for the neutral and 26 ms for the unrelated conditions in Experiment 3, and 23 ms for the neutral and 26 ms for the unrelated conditions in Experiment 4. There are two further aspects of the Bleasdale study that are important to note here. First, the study appeared to produce an effect of concreteness in the pronunciation task; a task that does not involve the binary decision component. Second, the stimuli appeared to be well-controlled on dimensions that have been confounded in the past studies (see Gernsbacher, 1984; Schwanenflugel, Harnishfeger, & Stowe, 1988). Before turning to a discussion of these confoundings, we briefly describe some intriguing results regarding concreteness effects in acquired and developmental dyslexics.

Concreteness Effects in Acquired and Developmental Dyslexics. Individuals who have specific reading disabilities due to brain trauma are referred to as acquired dyslexics, whereas, individuals who have specific deficits in reading performance that cannot be attributed to specific brain trauma are referred to as developmental dyslexics. Interestingly, there is evidence from both groups of subjects that the concreteness dimension may modulate isolated word recognition performance.

First, consider the performance of the group of acquired dyslexics that has been referred to as deep dyslexics. One hallmark symptom of this disorder is the inability to pronounce nonwords. Interestingly, these individuals produce a number of semantic errors in isolated reading such that they often report an
associate to the presented word instead of the target word itself. More important-
ly, however, for the present discussion, concrete nouns are the most spared of all
word types, with abstract words, adjectives, adverbs, and function words pro-
ducing the most errors (Marshall & Newcombe, 1980). The concreteness effect
has now been reported in a number of studies (Coltheart, Patterson, & Marshall,
1980; Richardson, 1975a, 1975b; Shallice and Warrington, 1975, 1980). Thus,
these results suggest that a brain trauma can selectively interfere with an indi-
vidual’s ability to recognize, i.e., read aloud, abstract words.

Interestingly, Seymour (1990) has recently reported a selective interference
with the production of abstract words in a developmental dyslexic. In a case
study reported by Seymour, subject JB showed a dramatic difference in vocal
reaction time for abstract and concrete words in a reading aloud task. Abstract
words took nearly twice as long to say aloud as concrete words. Moreover, error
rates for reading abstract words were more than twice as high as for concrete
words.

Although there are a number of different accounts of the concreteness effect in
dyslexics, one of the more interesting accounts is that there is a selective break-
down in right hemisphere performance in these subjects (see Coltheart, 1980;
Saffran, Bogyo, Schwartz, & Marin, 1980). This, of course, is quite consistent
with Day’s (1977) finding, reviewed earlier, that abstract words produced consid-
erably slower lexical decisions when presented to the right hemisphere, but not to
the left (see however, Boles, 1983, 1989).

Unfortunately, as in studies outside the area of the dyslexia literature, stimulus
selection factors make it difficult to draw strong conclusions from the data. That
is, in many of the studies familiarity and contextual availability, see discussion
below, have not been adequately controlled. Moreover, because most of the
studies involved untimed reading, response biases may play a role. Thus, al-
though suggestive and, in large part, consistent with studies with normal subjects
reviewed above, one must be cautious in extending results obtained with these
patients to general models of word recognition performance.

Concreteness Effects and Contaminating Variables. Gernsbacher (1984) re-
viewed the inconsistent patterns of data in the word recognition literature regard-
ing the impact of concreteness and word frequency. She argued that one variable
that appeared to be consistently confounded in these studies was item familiarity.
In order to demonstrate the importance of familiarity in these past studies,
Gernsbacher reported two lexical decision experiments (one with pronounceable
nonwords and one with nonpronounceable nonwords) in which she factorially
crossed familiarity with concreteness. Familiarity produced a significant effect,
but concreteness did not, and there was no evidence of an interaction. Hence, it
appears that familiarity may have been modulating performance instead of con-
creteness in, at least, some of the above studies. This argument is also consistent
with the arguments by Paivio and O’Neill and Boles, reviewed earlier, who suggested that there was no impact of concreteness above and beyond familiarity in threshold identification performance.

However, there are two points to note here regarding the problem of familiarity. First, it is not clear precisely what subjects use to rate items on a familiarity dimension. As noted before, it is possible that subjects may use a type of meaning information, e.g., concreteness, that is also used in lexical identification. By partialling out familiarity, one may also be inadvertently partialling out correlated meaning variables. Second, and more importantly, there is one study (Bleasdale, 1987) that indicates that one consistently finds a concreteness effect when familiarity is well-controlled. Thus, even if one argues that familiarity ratings do not involve access to meaning representations, which we believe is a rather tenuous argument, there is at least some evidence for concreteness effects that cannot simply be dismissed as familiarity effects.

More recently, Schwanenflugel, Harmishfeger, and Stowe (1988) have also challenged the observation that concreteness per se modulates word recognition performance. Schwanenflugel et al. suggest that, in addition to familiarity, concreteness is correlated with a dimension they refer to as contextual availability. These authors argue that it is important to access contextual information for the necessary integration that occurs across words in comprehension. Moreover, they argue that abstract words will, in general, have less contextual availability than concrete words. In their first experiment, they nicely demonstrated that when concreteness is confounded with contextual availability, concrete words produce faster response latencies in an LDT than abstract words, however, when concreteness is decoupled from contextual availability, there is a significant effect of contextual availability, but no effect of concreteness. In the second experiment, they found that contextual availability accounted for a significant proportion of the variance, independent of the impact of rated familiarity and word frequency. Moreover, the results of this analysis indicated that concreteness did not predict a significant proportion of the variance in lexical decision performance after contextual availability, familiarity, and frequency were partialled out. A similar pattern was found in their third experiment which involved a sentence context manipulation. Here, we again focus on the results of the neutral context condition. The results of this condition nicely replicate the results of their second experiment, i.e., concreteness does not have an impact above and beyond contextual availability.

As with familiarity ratings, one must ask what types of information subjects use in making contextual availability ratings. It may in fact be the case that subjects access the concreteness dimension when making contextual availability ratings. Although this would appear to be a priori less likely than in the case of familiarity ratings, it is still a possibility. However, it should also be emphasized here that even if contextual availability ratings include a concreteness dimension,
the Schwanenflugel et al. study nicely demonstrates that contextual availability ratings predict word identification performance above-and-beyond any influence of concreteness.

**Overall Conclusions Regarding Concreteness Effects.** Several generalizations can be made concerning the reviewed literature. First, stimulus selection factors qualify the conclusions regarding concreteness effects discussed in all studies, with the exception of Boles (1983, 1989), Bleasdale (1987), Germsbacher (1984), and Schwanenflugel et al. (1988). Even the later studies must be interpreted with caution. For example, Bleasdale argues that he controls his stimuli on the familiarity dimension, and handles the context availability issue via his contextual manipulations. Bleasdale argues that contextual availability cannot adequately explain the priming results obtained in his study. In all four experiments conducted by Bleasdale, the Target Concreteness by Prime Context interaction failed to reach significance. If contextual availability is the crucial factor then one would expect greater facilitation for abstract words and this was not the case. Clearly, however, it would have been useful to obtain contextual availability ratings for the Bleasdale stimuli to address whether there are influences of concreteness above-and-beyond contextual availability.

Although there are concerns in this area whether concreteness if the only factor that is modulating performance, the results from this area suggests that some type of semantic variable does influence isolated word recognition. Even if, as Schwanenflugel et al. have argued, contextual availability is the crucial factor in studies that have produced concreteness effects, then there is still evidence that a semantic variable influences isolated word recognition. Moreover, as already noted, the fact that untimed familiarity and contextual availability ratings predict speeded word recognition performance does not preclude a concreteness effect impact in normal word identification, unless one has evidence concerning the specific dimensions that subjects actually use to rate words on familiarity and contextual availability dimensions.

**(2) Polysemy Effects**

Just as the impact of concreteness has played an important role in word recognition research, there is also a considerable literature base that has addressed the impact of multiple meanings on word recognition. Most of this literature deals with the issue of multiple versus direct access of ambiguous word meanings (e.g., Balota & Duchek, 1991; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979; Tabossi, 1988), and involves the presentation of contextual constraints to direct the language processing system to a specific meaning of an ambiguous word. However, there are also a series of studies that simply address the impact of the sheer number of meanings available for a given word on word identification. If there is evidence that words with multiple meanings (e.g., yard)
are recognized differently than words with only a single meaning (e.g., *paper*), this could be viewed as suggesting that meaning is participating in the word recognition process.

One of the first investigations concerning how polysemy influences word recognition was reported by Rubenstein, Garfield, and Millikan (1970; see also Rubenstein, Lewis, & Rubenstein, 1971). These authors found that homographs presented in isolation produced faster lexical decisions than nonhomographs. Rubenstein et al. argued that ambiguous words involve lexical representations for each meaning, and hence, assuming a parallel access model, can be recognized more quickly. Unfortunately, Clark (1973) pointed out that the Rubenstein et al. results can easily be attributed to peculiarities of a few of the list items that Rubenstein et al. utilized. That is, Clark (1973) conducted statistical tests to determine whether the results would generalize beyond the specific language sample employed. The results of Clark’s analyses were quite clear. Effects that had been significant (i.e., the homography effect) in the original analyses across subjects did not produce significant effects across items (see Forster & Bednall, 1976, for a similar observation and discussion of this stimulus sampling problem).

Jastrzembski (1981; Jastrzembski & Stanners, 1975) found that the number of available meanings associated with a word can be a strong predictor of lexical decision performance beyond simple word frequency. Using the meaning metric of number of dictionary entries, Jastrzembski found that words possessing many (greater than 10) meanings produced faster lexical decisions compared to words possessing few (less than 4) meanings. Although these results are intriguing, there are two reasons for concern. First, the metric used in the Jastrzembski studies was the number of dictionary entries. Clearly the word *FUDGE* is polysemous. However, based on the number of dictionary entries, *FUDGE* possesses approximately 13 entries. It is difficult to imagine that subjects have knowledge of much more than 2 or 3 such meanings. A second concern about the Jastrzembski studies is that again familiarity of the stimuli appear to be confounded. This was nicely demonstrated by Gernsbacher (1984) who found that number of dictionary entries did not predict performance above-and-beyond familiarity. Again, however, one must question what types of information subjects rely on to make untimed familiarity ratings. It is at least possible that meaningfulness may play a role in such familiarity ratings.

There are three recent studies that provide more compelling evidence that the number of meanings available for a given word can influence word identification performance. First, Kellas, Ferraro, and Simpson (1988) found that ambiguous words presented in isolation in a LDT produced faster response latencies than unambiguous words. The stimuli used in this study were well-controlled for familiarity, frequency of occurrence, bigram frequency, number of syllables, and number of letters. Moreover, the results of this study indicated, via the use of a dual-task methodology (i.e., the response latency to detect a tone), that ambiguous words also demanded less attentional resources for processing as early as 90
ms after stimulus presentation. This general pattern was replicated in a second experiment in which attentional load was manipulated via a short-term memory load of 7 digits. These attentional effects have also been extended to an elderly sample (Kellas, Simpson, & Ferraro, 1988). Thus, it appears that the attentional resources demanded in word recognition are relatively less for ambiguous words compared to unambiguous words. Moreover, these meaning effects occur very early (i.e., at least as early as 90 ms) in word recognition.

Millis and Button (1989) have also reported evidence that polysemy can influence isolated LDT performance. In this study, the impact of polysemy was found when it was operationally defined as the number of meanings subjects can access from their memory. This metric was assessed in the following manner. One group of subjects generated meanings for the stimulus set, and a second group of subjects participated in a LDT with a selected set of target stimuli. The results indicated that, compared to words that only produced few meanings, words that produced many meanings in the meaning generation task, produced significantly faster response latencies in lexical decision performance. Because Millis and Button both (1) directly obtained measures of meaning availability from their subjects and (2) equated familiarity across their high- and low-meaning stimuli they avoided both concerns that were raised before with the Jastrzembski (1981) study.

Finally, we should note that recent work in our own lab has extended this pattern to the pronunciation task. That is, one might argue that number of meanings may influence lexical decision performance because meaningfulness is a relevant dimension in this task. However, we have found that pronunciation is also significantly influenced by the number of meanings available, with the same set of highly controlled stimuli used by Kellas et al. described earlier.

Conclusions Regarding the Impact of Polysemy in Word Recognition. Just as in the case of the impact of concreteness, we find that the early work on polysemy needs to be cautiously interpreted in light of stimulus confounds. However, the more recent work which controls for such confounds appears to suggest that polysemy can produce a reliable impact in both lexical decision and pronunciation performance. These results provide further support for the notion that meaning, now reflected by the sheer number of meanings available, can participate in isolated word recognition performance.

On the other hand, one might argue that these results may only reflect the fact that each meaning of a word has a unique lexical representation. Hence, based on a race horse model, one would predict that polysemous words that have multiple lexical representations would produce faster response latencies than words with only single lexical representations. Such a model does not need to assume any direct influence of meaning representation on word recognition. There are, at least two responses to this argument. The first response is simply based on parsimony. That is, we are unaware of any data that demands that words with
multiple meanings also have multiple lexical representations. Although such redundancy is possible, it might be computationally inefficient. Second, and more importantly, if one assumes that lexical representation is based on meaning representation then one loses the strict distinction between the lexicon and the system that represents meaning. As noted, this is precisely the distinction that we are concerned with.

Finally, it should be noted that one of the important lessons to be learned from this area is that the metric used to assess meaning availability is crucial in determining an impact of polysemy. That is, simply the number of dictionary entries appears to be an inappropriate measure, while the number of meanings generated by a sample of subjects from the target population seems to be a more powerful predictor of performance (see Kellas, Ferraro, & Simpson, 1988; Millis & Button, 1989, for a more detailed discussion of this issue).

**Associative Reaction Time as a Task and as a Predictor of Lexical Decision and Pronunciation Performance**

The time a subject takes to produce an associate to a given stimulus word should reflect processes involved both in word recognition and in meaning access. Moreover, because accessing meaning representations for a given stimulus word is also highly relevant to processes involved in language processing, performance in this task may provide interesting insights into more "normal" word processing demands.

First, consider the impact of concreteness/imageability in an associative reaction time task. DeGroot (1989) recently reported that concreteness/imageability significantly predicted associative response latency, above-and-beyond the impact of word frequency. Specifically, subjects were faster to produce an associate to target words that had relatively high concreteness/imageability ratings compared to targets with relatively low concreteness/imageability ratings. In fact, in DeGroot's experiments, the effect of concreteness was much stronger than the impact of frequency, which produced only a minimal effect. Thus, in a task that may more faithfully reflect processes involved in processing words in a similar fashion as in language processing; frequency had relatively little effect, compared to the rather large effect of a meaning variable, i.e., concreteness/imageability.

Chumley and Balota (1984) argued that a second use of associative word response latency is to use performance in this task as a predictor of performance in other word recognition tasks. If indeed associative response latency predicted performance in presumably pure word identification tasks, then one could argue that the time taken to access meaning is a predictor of word identification. The results of the Chumley and Balota study indicated that lexical decision performance could be nicely predicted from the speed to produce an associate to a given target word from a different group of subjects after the influence of lexical
variables (e.g., frequency, length, and naming performance) had been partialled out. Balota and Chumbley (1984) found a similar pattern. In this latter study, the time taken to produce the category name for a given exemplar nicely predicted lexical decision performance, after variables related to lexical access had been partialled out. Thus, it appears that the time to access a meaning of a word predicts performance in the LDT.

In sum, the results of the data reviewed in this section indicate that when subjects are asked to access meaning via an associative response latency task that performance in this task (1) is highly influenced by a meaning variable, i.e., concreteness/imageability, and (2) nicely predicts lexical decision performance, after variables that are related to current notions of lexical identification have been partialled out.

The Impact of Adding Meaning to Previously Nonmeaningful Strings

There are two studies in the literature that are quite intriguing because they appear to indicate that providing meaning for a previously nonmeaningful string can influence word recognition processes, above-and-beyond the impact of simple exposure to the stimulus. First, consider the study by Whittlesea and Cantwell (1987). These researchers were interested in, among other issues, the impact of meaning on the word-superiority effect. In their third experiment, they presented subjects with 24 pronounceable nonwords; half of which were assigned meaning and half were presented in a simple letter checking task to equate visual processing. The meanings assigned to the nonwords represented lexical gaps, i.e., meanings in which there were no obvious lexical representations, e.g., WALLEN—the sound a dam makes before breaking. After the subjects received the meaning assignment and visual processing tasks, they participated in a Reicher- (1969) type letter detection task, in which the nonwords were presented for 30 ms and immediately followed by a pattern mask. The subject's task was to identify as many letters as possible from the masked nonword. The results indicated that subjects could report significantly more letters from nonwords that were assigned meaning compared to nonwords that only received visual processing.

It is important to note here that mere exposure to the stimuli does not appear to be the critical factor modulating performance in the letter identification task. For example, one might argue that the meaning assignment conditions may have an impact via an increase in familiarity compared to the visual control condition. Hence, one may not need to appeal to an impact of meaning on letter identification. However, a comparison across Experiments 2 and 3 diminishes the plausibility of this account. In the first part of Experiment 2 subjects only received the meaning assignment task for half of the stimuli. The remaining half of the stimuli did not receive any visual processing until the letter identification task. In Experiment 2, the meaning superiority effect was 15%, i.e., the difference in the
probability of identifying a letter from a nonword that was assigned a meaning versus the probability of identifying a letter from a nonword that was not assigned a meaning. In Experiment 3, subjects received the same meaning assignment task used in Experiment 2, but, as noted above, also received a visual processing task for the remaining nonwords. The meaning superiority effect was not substantially reduced in this third experiment, i.e., it was still 12%. Hence, it appears that mere exposure to the stimuli is not the potent factor in producing the meaning superiority effect, but rather it appears to be the impact of meaning assignment.

Forster (1985) has also reported an impact of meaning assignment on nonword performance. In this case, Forster presented obsolete words such as *holimonth* (defined as, *a month of holidays*) in a masked repetition priming task. These stimuli were used because they were actually words, and hence, have the orthographic constraints of lexical items. However, these words were so obsolete that it was unlikely that any of the subjects knew their corresponding meanings.

In order to fully understand the Forster study, it is necessary to first briefly describe the masked repetition paradigm. On each trial subjects were presented the following sequence. First, an unrelated word was presented for 500 ms, second either the repetition prime or a visually unrelated prime was presented for 60 ms, and finally the target was presented for 500 ms. Because the prime stimulus is both forward masked by the first unrelated word and backward masked by the target, Forster has argued that this paradigm provides a pure measure of lexical access processing.

Forster first demonstrated that the obsolete words did not produce any repetition priming effect. This was consistent with other findings by Forster that have indicated that there is little if any masked repetition priming for nonword stimuli in the LDT. Of course, this would be expected because nonwords presumably do not have lexical representations. After this was demonstrated, the obsolete words were then assigned meaning. Now, the results yielded large repetition priming effects. Presumably, the assignment of meaning produced lexical representations. Of course, one could also argue that simply an increase in familiarity was producing this effect. However, there were aspects of this study that decrease the plausibility of this argument. Specifically, the stimuli were repeated across several blocks of trials. If familiarity was modulating the effect then one would expect the effect to diminish as subjects became more and more familiar with the stimuli. However, this was not the case. The masked repetition priming effect remained quite constant across blocks after meaning was assigned to the stimuli. Thus, the Forster results appear to reflect an all-or-none impact of meaning availability.

In sum, both the Whittlesea and Cantwell study and the Forster study indicate that providing meaning for previously meaningless letter strings can influence early word processing, i.e., letter identification and masked repetition priming in lexical decision performance. We would suggest that it is at least possible that
these results support the notion that meaning was participating in the word identification process. Of course, there is an alternative account. That is, it is possible that the effects are not simply due to meaning but rather are due to lexicalization. That is, by providing a meaning for a stimulus one is not only providing a meaning but also providing a lexical representation. Hence, the effects may only indirectly be due to meaning per se. However, there are aspects of both studies that suggest that the effects are not simply due to lexicalization. That is, as discussed earlier, it does not appear that mere visual exposure to the stimuli was modulating the effects. A lexically-based account would predict that such visual exposure would modulate performance in both studies. Hence, it appears that there must be a factor above-and-beyond visual exposure and lexicalization. We suggest that this factor is the association between the letter string and meaning. Of course, one might argue that lexicalization is the same as associating meaning to a nonword string. However, if one makes this argument then this would appear to weaken the distinction between lexical identification and meaning access, and, as noted before, this distinction is one of our major concerns with current models of word identification.

**Failure to “Recognize” Words that have Activated Meaning Representations**

If we assume that lexical decision and naming performance are accurate reflections of word recognition without meaning access, then a discussion of studies of threshold semantic priming is relevant. In threshold priming studies (see for example, Balota, 1983; Carr & Dagenbach, 1990; Dagenbach, Carr, & Wilhelmsen, 1989; Fowler, Wolford, Slade, & Tassinary, 1981; Marcel, 1983), subjects are first presented a threshold-setting task, in which stimuli are presented for brief durations (e.g., 10 or 15 ms) and followed by pattern masks. In the first session, the researcher attempts to determine the threshold at which subjects are at chance levels of performance at making discriminations between the presentation of a stimulus word versus the presentation of a blank field. After subjects participate in this session, they participate in a semantic priming LDT with the primes presented under the threshold conditions that were defined in Session 1. The interesting finding in these studies is that one still finds evidence for semantic priming even though the primes are presented so briefly that subjects are at or near chance presence/absence detection thresholds.

What is the importance of this finding? These results suggest that meaning has been accessed, even though subjects cannot make a discrimination between the presence of a word and the presence of a blank stimulus field. Thus, if we assume that lexical decision and naming performance can be used as metrics for lexical access without meaning access then these data are problematic. Clearly, if subjects cannot discriminate between the presence and absence of a stimulus, it is unlikely that they could discriminate between words and nonwords or be able to name the stimulus aloud. Thus, these results suggest that meaning has been
accessed (as indicated by the threshold semantic priming effects), even though word identification processing has not been sufficient to perform one of the tasks that have been used to build models of premeaning word identification.

It should be noted that there is some debate in the literature concerning whether researchers who have purported to have obtained threshold priming effects have provided adequate estimates of presence/absence detection thresholds (see Holender, 1986, and the accompanying commentaries for a detailed discussion). Although important, this debate does not strongly compromise the present arguments. That is, as noted, even if subjects were not at a detection level threshold in the earlier-cited studies, it is unlikely that subjects were at a level of visual analysis that would yield accurate lexical decision or pronunciation performance. Specifically, if subjects have difficulty indicating whether something or nothing was presented on a given trial, then it is unlikely that they could make a lexical decision or correctly pronounce a stimulus word under such highly masked conditions. This obviously runs counter to the suggestion that these tasks are a reflection of a premeaning access component of word recognition. It appears that meaning can be accessed very early and without the full visual record of the stimulus available for conscious report.

Further Impacts of Meaning on Early Perceptual Processing

We turn now to three studies that have demonstrated the impact of meaning variables on simple line detection, and on illusory conjunctions. Although these studies do not exclusively deal with lexical processing, they are of interest because they demonstrate influences of meaning variables on measures of early perceptual processing, and hence, further bolster the plausibility of such effects.

Schendel and Shaw (1976) reported an interesting study in which they found that subjects could more accurately recognize a line segment when it was part of a meaningful letter of the alphabet compared to when the same line segment was presented in isolation. These authors wanted to provide evidence that would extend and generalize the word-superiority effect reported by Reichel (1969) and Wheeler (1970). Schendel and Shaw operationally defined “context” as a specified letter of the alphabet (H, N, R, Z, etc.) and the “target” as either a horizontal, vertical, or slanted fragment taken from the letter (e.g., I in H, in F, etc.). Subjects first fixated the letter context (mean exposure duration = 23 ms) and then were shown a two-alternative forced-choice fragment display. The results indicated that subjects were more likely to choose the fragment that was part of the target letter string, compared to a visually matched fragment. Hence, not only does word level information influence the detection of letters, as indicated by the word superiority effect, but also letter level information influences the detection of the lines that make up the letters.

Weinstein and Harris (1974) found a similar finding with two- and three-
dimensional structures. They found that subjects detected a line segment faster when it was embedded in a meaningful three-dimensional structure as compared to a less-meaningful two-dimensional structure. In this study “meaningfulness” was operationally defined along Gestalt dimensions. Basically the results indicated that the more unitary a particular display appeared to the subject, the more accurately they could identify a target line segment embedded in that display. This object superiority effect is quite intriguing in the present context, because it extends the meaning influence on perceptual processing to pictorial stimuli. Hence, as discussed below, the impact of meaning in perceptual processing may not simply be tied to lexical strings but may be a general characteristic of the pattern recognition system.

Finally, Virzi and Egeth (1984) reported two experiments that address the impact of meaning-level information on the occurrences of illusory conjunctions. Illusory conjunctions involve the incorrect joining of two or more features from separate items in an array. Treisman and her colleagues (see Treisman & Gormican, 1988, for a review of this literature) have argued that such conjunctions involve early perceptual codes that have undergone relatively little processing, and that neither meaning assignment nor object recognition has occurred. Virzi and Egeth provided evidence that meaning can influence the probability of the occurrence of such conjunctions. They found evidence for meaning-level illusory conjunctions in an experimental paradigm used by Treisman and Schmidt (1982) to demonstrate purely perceptual conjunctions. The meaning variable in the Virzi and Egeth study was either the color of a word or the color that the word represented. For example, if the word RED was presented in blue ink and the word ALONE was presented in green ink, subjects would sometimes report the color word BLUE or the color RED. Neither of these stimuli were directly presented, but both could be produced by meaning-based illusory conjunctions. These results are quite intriguing because they suggest that phenomena that presumably reflect early integration of featural information, i.e., illusory conjunctions, can also be influenced by meaning level information.

In sum, the studies reviewed in this section suggest that there may be meaning-level effects in nonlexical aspects of perception. Hence, the meaning-level effects in the identification of words may not be an isolated characteristic of lexical processing, but may be a more general characteristic of the perceptual system.

**GENERAL DISCUSSION**

We have now reviewed the literature on the impact of meaning variables on early perceptual, primarily lexical, processing. We believe that there is sufficient evidence to reconsider models of lexical processing that place primary emphasis on processes that lead up to a measurable point in time, referred to as word identifi-
cation, that is devoid of meaning access. We have argued for such a reconsideration on two grounds. First, the tasks that are used to measure a premeaning word identification stage are not devoid of what would appear to be post-lexical processing. Second, even if one assumes that the available tasks are windows into only premeaning lexical identification processes, the general thrust of the literature involving these tasks (displayed in Table 8.1) is inconsistent with a premeaning word identification stage of processing.

Of course, it would be useful to provide an alternative modeling framework that would allow early meaning-level influences. We now turn to a brief discussion of such a framework. This is followed by a discussion of how this framework might handle the meaning-level influences reviewed earlier.

**How Could Access of Meaning Occur Before Lexical Identification?**

At one level, there appears to be a logical inconsistency in our arguments. Very simply, how could a reader know the meaning of the stimulus before the reader has identified the stimulus? The answer to this question is rather simple, and has been used in the past to answer a similar question produced by the word-superiority effect. In the latter case, the question is how can one know what the word is before one recognizes the letters that make up the word? The obvious answer in both cases is to rely on notions of cascadic processes and partial activation. That is, all one needs to assume is that there is partial activation of word-level or, in the present case, meaning-level units which in turn partially activate the representations that produced those partial activations. This of course is simply a slight modification of the basic interactive activation framework originally proposed by McClelland and Rumelhart (1981; Rumelhart & McClelland, 1982) to account for the word-superiority effect. As shown in Fig. 8.1,

![Diagram](attachment:image.png)

**FIG. 8.1.** Potential interactive activation framework for word recognition including meaning-level influences.
we are simply adding an additional level referring to meaning-level analyses to this framework. We see nothing inherently different between adding meaning-level information to this framework and adding word-level information to this framework (similar extensions can be made to other models of lexical processing, see Balota, 1990).

There are two points to note about this framework. First, this framework primarily involves a serial structure. That is, there is primarily a bottom-up flow of information from features, to letters, to words, to meanings. Meaning- and word-level information can have a reinforcing effect at lower levels only after earlier aspects of the processing system have been sufficiently activated to pass on activation to these higher levels. Second, the model relies quite heavily on the notion of continuous flow of activation. That is, activation does not need to accumulate at a given level before it spreads to a higher or lower level. This, of course, simply reflects the cascadic nature of the system.

The addition of meaning-level influences within such a framework provides the basis for some intriguing empirical questions. For example, once one has clearly established a meaning-level effect (e.g., contextual availability in lexical decision), then one might wish to limit the time for meaning level information to play a role. In this way one should be able to track the time-course of such information, as McClelland and Rumelhart did with respect to the time-course of word-level information. Second, it would be interesting to determine whether there is an analog of neighborhood effects at the meaning level. Specifically, McClelland and Rumelhart provided evidence that orthographically similar neighbors at the word level worked in consort to influence letter recognition. It would be interesting to determine whether words that share meanings with other words produce similar influences. We return to this possibility below.

Accounting for Meaning-Level Effects

Of course, a more important question that needs to be addressed is why one finds the particular pattern of meaning-level effects obtained in past studies. Up to this point, we have simply described studies that have demonstrated such an impact, without specifying why one might expect a particular pattern. We now turn to a brief discussion of such an account.

First, if there is an impact of concreteness above-and-beyond familiarity and contextual availability, as Blesdale has argued, one may appeal to the notion that concrete words have an “extra” (or possibly stronger) imaginal representation compared to abstract words. This extra meaning representation should serve as an extra top-down source of activation. Such a possibility is consistent with the data and arguments presented by Day (1977). This account simply is based on the assumption that more-means-better when considering the impact of meaning on word recognition. In addition, the impact of contextual availability and polysemy may be accounted for in a similar manner. That is, in both the case of
contextual availability and polysemy the evidence suggests that words with more meaning representations are recognized more quickly. Again, the premise of more-means-better is instantiated.

Second, the impact of meaning without conscious word identification (threshold semantic priming effects) simply reflects the continuous activation aspect of the framework. That is, meanings are partially activated even though sufficient information has not been activated for the subject to make a presence/absence decision. Part of this effect is probably due to the attentional disruption produced by the pattern mask at the visual level codes that are necessary in making the presence/absence decision. Threshold semantic priming effects simply suggest that there is partial information transfer before the subject can fully access such codes.

Third, the predictive power of associative response latency for lexical decision performance may simply reflect the fact that the ease of accessing a meaning level representation is reflected by the associative response latency measure. If this were the case, then one would expect more of a top-down influence for words that have meaning information readily accessible. Thus, one should expect associative response latency to be a strong predictor of lexical decision performance, as reported by Chumbley and Balota.

Finally, the impact of adding meaning to a nonword string, could be accounted for by the more-means-better principle. The notion here is that the addition of meaning would provide a level of representation that would benefit both word processing (Forster, 1985) and letter processing (Whittlesea & Cantwell, 1987).

Of course, it is not surprising that the more-means-better assumption would be useful in providing a surface-level account for the reviewed literature. However, one might argue that in some cases more could produce meaning-level inhibition. For example, it would appear counterproductive for both meanings of an ambiguous word to be simultaneously activated. Interestingly, this appears to be dependent upon the time-course of activation at the meaning level. For example, in his review of the literature on lexical ambiguity resolution, Simpson (1984) has argued for automatic ordered (by frequency of usage) access of multiple interpretations of ambiguous words, but then selective access of the contextually determined interpretation. Moreover, Simpson and Burgess (1985) have provided evidence that at longer intervals a selected interpretation can inhibit a nonselected interpretation of an ambiguous word. As indicated, both of these patterns of data appear to be highly dependent upon the time-course of activation. Thus, the more-means-better principle, at least with respect to polysemous items, may primarily reflect processes early in word recognition. The importance of this observation is very simple. In order to adequately model the semantic impact in word identification, it appears crucial to include assumptions regarding the temporal course of activation across levels within the system.
What Might the Meaning Representation for Words Involve?

Of course, one must address what more means in the present context. The topic of meaning representation is quite controversial and has been at the center of considerable theoretical debate. For example, one consistent problem that has occurred in the literature on word meaning is whether it should be represented as a single core meaning or a list of semantic features. Under both approaches it would appear that the meaning of a word involves a core set of semantic information that is, in some sense, different from the individual episodic experiences with a given word.

We would support an alternative approach to meaning representation that follows from suggestions made by Hintzman (1986, also see Medin & Schaffer, 1978; Schwanenflugel & Shoben, 1983). Although Hintzman's model was primarily developed to account for prototype learning, it can also be extended to meaning representations. Hintzman presents an alternative approach to account for the classic Posner and Keele (1968) findings regarding prototype abstraction. He argues that each of the episodic experiences with the instances of a prototype produces a unique episodic memory trace that consists of a list of primitive features. The higher delayed recognition memory that Posner and Keele found for nonpresented prototypes, compared to presented exemplars, simply reflects the retrieval process when the test item is presented. Hintzman argues that when the test item is presented it partially activates “all” of the episodic traces. The result of this activation process is an echo that represents all of the episodic traces in memory. If there is sufficient overlap between the echo and the test item, the item is recognized. In some cases, it is possible that the similarity of the echo to the test item may be stronger when a prototype is presented that partially overlaps with many episodic traces (i.e., the instances), than when an instance is presented that strongly overlaps with a single episodic trace. Thus, the culmination of many different, but related, episodic memory traces can produce the relatively high recognition confidence for a stimulus (i.e., the prototype) that was never directly presented.

It is important to emphasize here that this model suggests that meaning is produced at retrieval via the interaction of all previously stored memory traces. Meaning of a word is not simply a static representation that is represented by (a) a set of semantic features, (b) a prototype representation, or (c) the definition that the subjects look up in their mental dictionary, but rather, involves the accumulation of individual episodic traces that are produced by individual experiences with the word. Each of these experiences (i.e., the word and its context) contributes to the evolving meaning that we attribute to a given word.

There are a number of important implications of this approach. First, one does not have to appeal to different memory systems to account for different memory types. That is, all experiences produce episodic traces, and the stability
of some information, such as general definitions of categories (e.g., Fruit, Animal, Mineral, etc.) are simply due to the convergence of these episodic traces, as opposed to the abstraction of core meanings. In addition, this framework nicely handles the contextual dependency that we find in language (e.g., see Roth & Shoben, 1983). Static notions concerning meaning representation have difficulty handling such contextual dependency. Of course, one may be able to argue that different static meanings are accessed in different contexts but then it seems that one has to specify a different meaning whenever one provides a different context. The account here for contextual dependency is simply that when a target word and its context is presented it produces a unique echo. The echo from these episodic traces is the meaning. This meaning will primarily be a function of the context that the word has most typically been embedded within.

Of course, there are many questions that need to be addressed if the Hintzman-type approach would be extended to word meaning. For example, one must define what is encoded in an episodic trace. Are there primitive dimensions of such traces and, if so, how is this different from semantic features? Although there are clearly many questions that need to be resolved, we believe that the notion that word meaning is a constantly evolving characteristic of episodic traces is an intriguing and potentially fruitful approach to word meaning.

Implications for the Modularity of the Language Processing System

Our contention that meaning influences word recognition is problematic for models that assume a modular (strictly encapsulated) word processing system that limits meaning analyses to associative connections between words within a lexical network. However, we should point out that the processing system displayed in Fig. 8.1 does have an overall serial structure. Thus, although we allow for top-down influences from meaning-level analyses, we do not wish to argue against the importance of a highly structured language processing architecture. In this sense, the more important issue is the degree of interaction among the various levels. Moreover, in order to understand the degree of interaction, as noted before, one must consider the time-course of such an interaction. In this light, hopefully, the present chapter helps to frame useful empirical and theoretical questions regarding the temporal course of meaning-level variables in word recognition.

CONCLUSION

One of the major motivations for the present chapter has been functional in nature. That is, researchers in the area of word recognition have developed models that primarily emphasize orthographic and phonological aspects of lex-
ical processing, that are relatively devoid of meaning-level impacts. We believe that the available evidence suggests that meaning may contribute to early processes in word identification. Because of this current state of affairs, we feel that it is prudent to begin considering models that allow meaning-level influences in word identification. That is, we believe that the more important issue in word recognition research is to provide an accurate model of how words convey meaning than how orthography and phonology lead to the magical word identification point.

ACKNOWLEDGMENTS

We thank Christine Chiarello and Paula Schwanenflugel for their helpful comments on an earlier version of this chapter. This work was, in part, supported by NIA grant RO1 A607406.

REFERENCES


