CHAPTER 3

Lexical semantic processing during speech comprehension

Helen E. Moss
Department of Experimental Psychology, University of Cambridge, UK

M. Gareth Gaskell
MRC Cognition and Brain Sciences Unit, Cambridge, UK

INTRODUCTION

In order to understand spoken language the listener must be able to identify the sequence of words in the speech stream and to access a range of different kinds of information about each of those words from the mental lexicon, allowing them to be interpreted and combined with other words to build up a representation of the utterance. The kinds of information that are made available by lexical access include, at least, morphological structure, syntactic structure, and meaning. In this chapter we focus on the latter; the processes by which the meanings of words, or their semantic representations, are activated as we hear them; processes that are at the heart of language comprehension.

Our study of the activation of meanings by spoken words must take into account the special, temporal characteristics of speech. Words must be recognised rapidly and efficiently—at a rate of several words a second in normal speech—because the signal fades quickly and the acoustic information for each word is immediately overwritten by the next word (Crowther & Morton, 1969). The speech signal is also inherently variable; the realisation of most sound segments varies as a function of the surrounding phonological context, and background noise is the norm rather than the exception (Repp, 1978; Warren & Marslen-Wilson, 1987). Speech is also continuous, and therefore, segmentation is a fundamental problem;
how do we tell when one word ends and the next begins (e.g. Cutler & Norris, 1988). In all these ways, spoken language differs from written text, where the signal is clear, less variable, and remains in view for as long as the reader wishes, and where word boundaries are conveniently marked by white spaces. Thus, although we assume that the target of lexical access is the same, whether the input is spoken or written (or, indeed, Braille or sign language), the mapping process from sound to meaning may differ in important ways from that from text to meaning, especially in terms of its time-course and the range of competitor words that are activated (see Holcomb & Anderson, 1993).

The study of the processes by which word meanings are activated provides a bridge between two bodies of psycholinguistic research. On the one hand, several models of spoken word recognition have been developed. However, these models are primarily concerned with the way in which the speech signal activates lexical form representations rather than with the access of word meanings. A second area of research has focused on the interpretation of words in discourse contexts (e.g. Garrod & Sanford, 1981). Such studies do not, however, address the issue of how the meanings initially become available for interpretation and combination. These two areas of research have largely been carried out in isolation from each other and one of the aims of this chapter is to bring together research that provides a basis for greater integration of the two sets of issues.

FRAMEWORK

Models of spoken word recognition

Before beginning the discussion of how the meanings of words are accessed during speech, we first need to consider briefly the nature of spoken word recognition in general. Most current models describe the process of spoken word recognition in terms of activation of processing units within a mental store or “lexicon”. Each processing unit corresponds to a specific word in the vocabulary. A processing unit increases in activation as a function of the degree of match between its phonological form and the incoming speech signal. At a certain point, one processing unit will reach a criterial level of activation, allowing that word to be recognised. On these models, there will be an initial phase during which many words in the lexicon are a potential match for the input, and so many processing units will be activated simultaneously. These words are all candidates for recognition, and the competition among them is resolved as further speech input is heard, which will continue to match some words, so increasing their activation, but will mismatch other words
so reducing their activation or allowing it to fade away (e.g. Luce, Pisoni & Goldinger, 1990; Marslen-Wilson, 1987; McClelland & Elman, 1986; Norris, 1994). The Cohort model (Marslen-Wilson, 1987; Marslen-Wilson & Welsh, 1978), which has perhaps been the most influential of these, claims that a set of word candidates is activated on the basis of approximately the first 100–150 ms of the speech input. This word-initial cohort consists of all the words with onsets consistent with the acoustic input. As the speech continues, words which continue to match the input are increasingly activated, while those words that mismatch the input at any point experience a sharp drop-off in their activation level. At a certain point in any word—its recognition point—the input will be consistent with only one candidate, and this word will be recognised. For example, as the speech input/tre/ is heard, this will activate a word initial cohort including, tress, trestle, tread, trend, trek, and trespass. When the following /s/ is heard, only tress, trestle, and trespass will continue to be activated, while the activation of the mismatching candidates (tread, trend and trek) will quickly start to decline. Then when the /p/ is heard, only the single word, trespass, continues to match and it can be recognised at this point. For many words the recognition point will be before the acoustic offset (as in the case of trespass), although this will not always be the case, especially for monosyllables, and following context will sometimes be necessary for identification (e.g. Bard, Shillcock, & Altmann, 1988). Other models such as TRACE (McClelland & Elman, 1986) and the Neighbourhood Activation Model (NAM; Luce et al., 1990) have similar processing assumptions to the Cohort model, but differ in a number of details. For example, NAM places less emphasis on the onset of the word, with the activation of candidates being a function of total degree of phonological overlap regardless of initial match (e.g. mountain would be a strong competitor of fountain, whereas on the Cohort model mountain would not enter the word-initial cohort and so would not be activated at all). TRACE also treats the beginnings and ends of words in similar ways, although competition still depends more strongly on overlap early on in the word. This is because TRACE implements competition in terms of a “rich get richer” mechanism. Words with a low activation value are subject to strong competition effects from their neighbours, whereas words with higher activations are more resilient to competitor effects. Since activations build up as speech comes in, this means that word-initial speech is more important in determining the final activation level of a word.

Although these models make detailed predictions about which words will be activated at any point in the duration of a word, these are primarily claims about the activation and identification of the phonological forms of words, rather than access to semantic information. For example,
in TRACE, the top-most level of representation consists of nodes corresponding to each word in the model's lexicon, but these simply identify the word as a specific phonological form, rather than its meaning(s). In general these models do not focus on when in the recognition process the meaning of a word is accessed, or whether the meanings of its competitors are also activated. Nevertheless, the Cohort model makes one crucial claim about the access of word meanings: that there is early activation of multiple semantic representations (Marslen-Wilson, 1987). What this means is that the meanings of all the words in the initial cohort will start to be activated in parallel, as soon as the first 100–150 ms of the word has been heard. So, for example, on hearing the initial sequence /tre/ the meanings of trespass, trestle, trend, trek, and so on will all begin to be activated, as well as their phonological forms. This is an essential premise for one of the other major claims of the model, which is that when a word is heard in context, sentential semantic constraints interact with the word recognition process to allow the appropriate word to be selected more rapidly than when the same word is heard in isolation. In order for this to be possible, the meanings of the words in the candidate set must be available for evaluation against the prior context. Any kind of interactive activation model must make a similar claim in order for higher level context to have an effect early in the word recognition process. This contrasts with accounts in which semantic information only becomes available for the correct candidate after it has been recognised, such as Forster's Search model, in which the meaning of a word is looked up in a master file only after it has been uniquely identified on the basis of its orthography (Forster, 1976) and Morton's logogen model, in which a processing unit has to reach a pre-set threshold before its meaning is accessed (Morton, 1969).

In this chapter we will consider the available evidence with respect to the issues of when semantic information becomes available during word recognition, and whether the meanings of competitor words are activated or not. We will also examine experimental data concerning more detailed questions about the structure and contents of the meanings accessed, whether there are differences in the time-course with which different kinds of semantic information become available, and whether all competitor meanings are fully activated, in order to start developing a more complete model of lexical semantic activation. We will then turn to the question of how the data can be accounted for in models of word representation and processing. In particular we will consider the claims of a new kind of model in which word forms and their meanings are not represented as single processing units, as in the models outlined earlier, but are distributed over patterns of activation across many processing nodes (Gaskell & Marslen-Wilson, 1997b). This kind of distributed, connectionist model
offers interesting new ways of thinking about how the meanings of words are activated during speech (see also Chapter 8 in this volume by Chater & Christiansen). Due to space limitations the main part of our discussion will focus on the recognition of single words heard in isolation. Although this simplifies the nature of the problem, a full consideration of lexical semantic access during speech must take into account the fact that words are not heard as single units but as part of larger discourse contexts, and therefore their meanings must be interpreted with respect to the meanings of both preceding and subsequent words. We return to this point at the end of the chapter, where we outline some of the ways in which the issues concerning single word recognition interact with the demands of semantics processing in context.

WHAT?

In this investigation of how word meanings are accessed we will start with the question of what it is that is actually being accessed: What is the nature of the semantic information that becomes available when we hear a word? One approach to this question is to try to map out the contents of the mental representations of word meanings that are the goal of the access process. The question of how the meanings of words are represented in our minds has taxed linguists, philosophers, and psychologists for centuries and has generated a huge body of debate and research (e.g. Locke, 1690/1981; Lyons, 1977). Almost all theories of semantic representation share the assumption that word meanings are componential; that is, they are made up of a collection of semantic attributes or features, rather than being unanalysable wholes. The majority of psychological evidence supports the componentiality claim, and we will assume that it is correct (see McNamara & Miller, 1989 for a review of this issue). The major debate over the years has been between “classical” theories of word meaning, in which the lexicon contains only the most rarefied of information, that is, the necessary and sufficient conditions for application of the meaning of a word (e.g. bachelor would be represented by the features [+human, +male, +adult, +unmarried] and nothing more; Katz, 1972), and theories in which a word’s meaning is made up of a richer set of information, including characteristic or typical as well as defining features. This class of account includes prototype and exemplar-based theories (see Smith & Medin, 1981 for a review). On the classical view there is a sharp dividing line between the meaning of a word that is stored in the mental lexicon, and other “encyclopaedic” information that lies outside the lexicon and is part of our general knowledge store. This distinction is not made so clearly on the other views, where certain semantic properties may be more or less central to the meaning of the
word, but they may all be part of the same kind of representation or "cut from the same cloth" (Jackendoff, 1983).¹

In spite of the enormous interest in word meanings, the true nature of our semantic representations remains elusive. Traditionally, psychologists have asked questions about people's knowledge of word meanings by asking them to define words, list their properties, or make judgements about their properties (e.g. McNamara & Sternberg, 1983). The problem with this approach is that it relies on subjects' metalinguistic judgements about meanings. Although these reflections tell us something about what people know, they do not necessarily correspond to the mental representations of word meaning that are involved in normal, automatic language comprehension. The idea that our mental lexicons contain static definitions of words, which can be the objects of introspection, is an appealing one, but is probably no more than an artifact of our literacy and familiarity with written dictionaries and encyclopaedias. Canonical definitions may have no part in language comprehension, but rather can be explicitly learned as a result of metalinguistic education. A study of definition generation described by Barsalou (1993) demonstrates the opacity of the relationship between our explicit generation of "meanings" and the underlying semantic representations. This study showed that there was a great deal of variability in the semantic properties listed for a word, both across subjects, and within subjects on different occasions. Although it is possible to interpret this as a result of variation and fuzziness in the representations themselves, a more plausible explanation is that, even if our semantic representations are available to conscious awareness, their format is unlikely to be a list of propositions that can simply be read out. Therefore there must be some kind of translation and construction process to produce a list of properties, and the variability may be located in these processes.

In the light of these problems, we need to try to find out about the nature of semantic representations in a way that does not involve explicitly asking people to reflect on the meanings of words. This can be done using a class of experimental procedures known as "on-line" tasks. These tasks require subjects to make a speeded reaction time response as close in time as possible to the process we are interested in investigating. The attention of the subjects is directed away from the process of interest because they are concentrating on the reaction time task. The speeded, implicit nature of these tasks enables them to tap into the fleeting,

¹ Some authors have distinguished between word meanings and conceptual representations. In line with most psychologists, we do not distinguish between the two, and use the terms interchangeably. For a discussion of how meaning and concepts could differ, see Murphy (1991).
automatic and largely unconscious processes that underlie language comprehension, rather than explicit, metalinguistic knowledge, so avoiding many of the problems associated with traditional methods of studying the mental representations of word meanings (Marslen-Wilson & Tyler, 1980; Tyler, 1992). The most valuable on-line paradigm for this purpose is semantic priming (e.g. Fischler, 1977; Meyer & Schvaneveldt, 1971; see Neely, 1991 for a review). In a priming task, subjects make a speeded response to a target word (e.g. lexical decision or naming), and reaction times are measured as a function of the preceding word. When the preceding prime word is related (e.g. cat–dog) subjects’ reaction times are significantly shorter than when it is unrelated (e.g. pen–dog). The majority of priming studies have focused on the issue of how contextual factors affect word recognition, by exploring the influence of the prime on recognition of the target word. However, by manipulating the nature of the relationship between prime and target, we can exploit the same task to chart the range of semantic information that becomes available when the prime word is heard. For example, if the prime canary facilitates responses to bird, we can infer that the semantic information that a canary belongs to the bird category was accessed when canary was heard. Likewise, if canary primes yellow and wings, this suggests that these perceptual attributes about what canaries look like were made available, and if canary primes sing, or cage, this shows that certain pieces of information about what canaries typically do, and where we see them, are also being activated (Hodgson, 1991; Moss, Ostrin, Tyler, & Marslen-Wilson, 1995a).

It is important to note here that the priming task can tell us about the information automatically accessed when a word is heard in isolation. If we assume that the full semantic representations of words are exhaustively accessed whenever a word is heard, then this is equivalent to a direct view onto the nature of the underlying semantic representation. This view of exhaustive, context-independent lexical semantic access has often been implicitly assumed in psycholinguistic research, and has been labelled the standard position by Williams (1988). However, it is possible that the semantic information made available when a word is heard is not the same on every occasion, and may not always correspond to the sum total of information that is represented in the lexicon (cf. Gerrig, 1986). We return to the issue of how variable this information may be in different contexts in the discussion at the end of this chapter. For the moment, the point is that we cannot necessarily treat the semantic information accessed for a word in isolation as an exhaustive or invariant reflection of the meaning stored in the lexicon. A more appropriate way to characterise the information activated for a word in isolation is as a set of default information that is generally made available when there is no supporting context. On some accounts, this corresponds to a core meaning, with
other more peripheral properties of the semantic representation being made available only when specifically relevant in a given context (Barsalou, 1982; Greenspan, 1986).\(^2\)

Most of the numerous priming studies that have been carried out over the last 20 years suggest that a wide range of semantic information is made available when a word is heard (or read).\(^3\) Reliable priming has been shown for words that belong to the same semantic category (category coordinate), such as cat–dog, chair–table (e.g. Hines, Czerwinski, Sawyer, & Dwyer, 1986; Moss et al., 1995a), words that stand in a superordinate–subordinate category relationship, such as bird–robin, table–furniture (e.g. Neely, Keefe, & Ross, 1989), antonym pairs such as hot–cold (e.g. Colombo & Williams, 1990), and pairs of words that share functional properties and relations such as bee–hive, broom–floor (e.g. Chiarello, Burgess, Richard, & Pollack, 1990; Moss et al., 1995a) or perceptual properties, like cherry–ball (Schreuder, Flores d’Arcais, & Glazenburg, 1984). Given this pattern of results we might want to conclude that as we hear or read a word, a rich set of default information rapidly and automatically becomes available to us, including considerably more than the restricted set of defining features identified by classical theories of word meaning. The meaning activated within a few milliseconds of hearing a word includes, at least, information about the semantic category the word belongs to, and some kind of link or connection to members of the same category, as well as information about its perceptual and functional properties.

There are, however, two important methodological points that need to be borne in mind in interpreting the results of these priming studies. The first important point is that we need to be certain that the priming task is measuring the effect of the semantic relationship between the prime and target word, rather than some other kind of relationship between them. A possible problem here is that many priming studies have confounded semantic relations with an associative relationship. By associative relationship, we mean the probability that subjects will give the target as a response to the prime in a free association test (e.g. Moss & Older, 1996). For example, cat and dog are strongly associated, whereas pig and horse

\(^2\) However, the term core meaning is not always used to refer to the properties that are automatically activated in language comprehension. McNamara and Miller (1989) call such properties the immediate level of meaning, and use the term core meaning to refer to additional information that people may be able to retrieve in order to make fine judgements about the real nature of referents.

\(^3\) Most priming studies have looked at written rather than spoken words. However, because we have assumed that the semantic representation accessed by the written and spoken word is the same, we can consider evidence from both modalities at this point. It is when we move onto the time-course of activation and the role of the competitor environment, in the following sections, that the two modalities diverge.
are not, even though both pairs share the same kind of semantic relations of category co-membership. If we assume that the mental lexicon contains a representation of the phonological form of each word (i.e. what it sounds like) as well as of its meaning, it is possible that associated words are linked not only semantically, but also at the level of their phonological forms. These form-level links may build up because of the likelihood of hearing the words together in the language (Chiarello et al., 1990; Shelton & Martin, 1992; Tanenhaus & Lucas, 1987). For example, the word *dog* often follows shortly after *cat*, as has been demonstrated in several studies of large language corpora (e.g. Spence & Owens, 1990). This high probability of co-occurrence may lead to facilitation of recognition of a word when preceded by an associated word (e.g. *dog* preceded by *cat*). Priming between a pair of words that are strongly associated does not necessarily reflect the activation of the meanings of the prime and target, as indicated by the fact that words that often occur together but which have little semantic relation prime each other significantly (e.g. *pillar*—*society*, *stork*—*baby*; Moss, Hare, Day, & Tyler, 1994). However, several studies have now shown that semantic priming does hold across the range of relations described previously, even when any associative connection is deliberately avoided (Chiarello et al., 1990; Fischler, 1977; Lupker, 1984; Moss, Ostrin, Tyler, & Marslen-Wilson, 1995a; Shelton & Martin, 1992; Zwitserlood & Schreipers, 1995). Therefore, we conclude that associative priming between pairs of words that frequently occur together may be a separate mechanism that can operate in addition to priming at the semantic level (Moss et al., 1994), but that semantic priming also exists in its own right, and provides a window onto the semantic information that is activated when a word is heard.

The second, perhaps more problematic point is that priming may involve non-automatic processes in addition to the implicit effect of activation of the meaning of the prime on the recognition of the target, and that certain tasks are more susceptible to these strategic effects than others (e.g. Neely, 1991). For example, priming of lexical decision involves a component of semantic matching (also termed coherence checking) of the target’s meaning against that of the prime, after both words have been recognised (Neely et al., 1989). However, this does not undermine the use of the task to probe the meaning activated by the prime, as long as the target can only be matched against the semantic information initially made available when the prime was heard. Indeed, the semantic matching process may be an integral part of normal language comprehension, in which the meaning of each word is evaluated against the meanings of the previous words, rather than a conscious strategy that is peculiar to the priming paradigm (Hodgson, 1991). A more difficult possibility to address is that the facilitation of the target could be due to backward priming (Koriat, 1981). This
may operate if the prime has not been fully processed when the target is presented, and the target can act as a context for the prime as well as vice versa. If this is the case, then priming may not reflect the information that is automatically accessed by default for the prime heard in isolation, but may be a product of activation of its meaning in the context of the following target word, which may or may not be the same thing. For example, the information “used to sweep floors” may not always be accessed when we hear the word broom, but may be activated by backward priming when we heard broom followed by the target context floor. However, backward priming has only been demonstrated for associative relations and it is not clear that it applies to purely semantically related primes and targets.

One of the problems in evaluating the role of purely automatic semantic priming is that there is debate as to which paradigms are susceptible to strategic priming and which are not. Although it is generally thought that a low proportion of related pairs and a short interval between prime and target is effective in minimising strategic effects, Shelton and Martin (1992) have recently argued that any paradigm in which prime and target are presented to explicit pairs encourages post-lexical matching and backward priming. They advocate a version of the priming task in which the primes and targets are presented in a continuous list without gaps between each prime–target pair. However, this paradigm is not without its own problems (Moss et al., 1995a). The issue of the extent to which different tasks pick up a purely automatic priming effect remains an open one. A related issue is whether different tasks reflect different processing pathways or levels within the recognition system (see Balota, Paul, & Spieler, Chapter 2). For example, a lexical decision task may focus attention on the semantic level to a greater extent than a naming task. However, it is not yet clear whether this account applies to spoken word recognition in the same way as visual word recognition. Reading is a task that is learned relatively late in life, and there is an extensive literature detailing the possible different “routes” that may be available to the reader (e.g. sounding out from orthography to phonology, recognition of whole words, or direct semantic access). There is no evidence that speech recognition has the same potential for dissociable pathways. Therefore, priming paradigms provide our most transparent window onto the activation of lexical semantic information, although it will probably never be possible to eliminate strategic effects entirely.

Lexical semantic ambiguity

The discussion so far has concerned the semantic information that is retrieved in the case of words that have a single meaning. An additional set of questions arises when we extend our investigation to those words
which have two or more separate meanings, such as bank, plane, and rose. Are all the possible meanings of these ambiguous words automatically accessed when we hear them, or is only one of the meanings made available on any given occasion? This is an important question, since a surprisingly large proportion of words (at least in the English language) are ambiguous in this way, and it is one that has received a great deal of attention in psycholinguistic research.

Several experiments have suggested that both meanings of an ambiguous word are activated when it is heard in isolation. For example, Holley-Wilcox and Blank (1980) presented subjects with an ambiguous word as a prime, followed by a target related to one of its two meanings. Lexical decisions to both targets were facilitated as much as when the prime was a related unambiguous word. However, the activation level appears to be modulated by the relative frequency of use of the alternative meanings. When the meanings are polarised, such that one is much more frequently used than the other, priming was observed for the higher frequency (or dominant) meaning but not for the lower frequency (subordinate) meaning (Simpson, 1984). More recently, Simpson and Burgess (1985) have argued that the subordinate meaning is activated, but to a lesser extent and with a slower time-course than the dominant meaning. The results of these experiments indicate that when we encounter an ambiguous word, in the absence of any other information, all possible meanings are made available, with the degree of activation determined by relative frequency. However, the limitations of these studies are first, that they have been carried out with visually presented materials, so that we cannot determine the time-course of activation of the meanings during a spoken word, and, second, that associative priming has not been ruled out, so that we cannot be sure that the results reflect simultaneous access to multiple conceptual semantic representations rather than the co-occurrence of ambiguous primes with target words related to both meanings.

Most of the interest in ambiguous words has centred around the activation of their meanings when heard in context, rather than in isolation. The key question is whether the exhaustive activation of all possible meanings is automatic and context-independent, or whether the biases and constraints of the prior semantic context can guide access such that only the appropriate meaning is activated. This is just one aspect of the broader debate about whether the language comprehension system operates in a purely bottom-up, data-driven fashion, or whether contextual information can interact with lower level processes (e.g. Tanenhaus & Lucas, 1987). A number of early studies using a sentential cross-modal priming paradigm suggested that access of the meanings of an ambiguous word is not guided by context. For example, Swinney (1979) presented an ambiguous word in a spoken context which biased towards one of the
meanings. Visual targets for lexical decision were presented shortly after the offset of the ambiguous prime. Targets related to both the context-appropriate and context-inappropriate meanings were facilitated, indicating automatic exhaustive access. However, when a longer delay (e.g. 100 ms) is introduced between prime and target, only the relevant meaning supports priming, indicating rapid decay or suppression of the meaning that does not fit with the semantic context (see Simpson, 1984 for a review). More recently, however, several researchers have pointed out some of the weaknesses of the earlier studies, and shown that the picture is more complex. Most importantly, the nature of contextual bias was not well controlled in the earlier studies (Tabossi, Colombo, & Job, 1987). Tabossi et al. found that when the sentential semantic contexts biased towards specific semantic features of one of the meanings of an ambiguous prime word rather than making one meaning more plausible in a general way, only targets related to the appropriate meaning primed significantly, at least for the subordinate meanings. Simpson and Krueger (1991) also found context-dependent activation of meaning when contexts were strongly constraining. Thus, the nature of the semantic context, as well as meaning dominance, appears to determine the degree to which alternative meanings will be activated when an ambiguous word is heard.

Summary

Results of priming experiments do not support the view that only necessary and sufficient features of a word's meaning are activated when it is heard in isolation. Rather, it seems that a wider range of information is made available although confirmation of this awaits further experiments in which possible associative relations and strategic backward priming are completely eliminated. One interesting finding that is emerging from our recent research is that functional properties of meaning support particularly robust priming effects, suggesting that they may be especially salient in the mental representations of word meanings. By "functional properties", we mean information about what a thing typically does, how it behaves, what it is used for, and where it is found, and so on (for example, bee—honey, desk—work, broom—floor). These relations support priming across a wide range of paradigms, including the continuous list version of the lexical decision task, claimed by Shelton and Martin (1992) to eliminate strategic priming effects (Moss et al., 1995a). Further evidence for the central role of functional information comes from neuropsychological studies of patients with semantic impairments. We studied a patient, PP, who had a profound semantic deficit known as semantic dementia (Moss, Tyler, Hodges, & Patterson, 1995b). PP showed robust normal priming for functionally related word pairs (e.g. hammer—nail, broom—
floor), but no priming for category coordinates (e.g. cat–dog, spade–rake). At the time of testing, this patient had little or no conscious awareness of the meanings of words, and so would have been unable to carry out any kind of strategic processing (for example, when asked what was her favourite kind of food, she replied, "food, food, I wish I knew what that was..."). The implicit facilitation of functionally related words in this case suggests strongly that this information is automatically activated when words are heard, and may be more resistant to brain damage than other kinds of information. We have found that several patients with semantic processing deficits show the same pattern of effects, with a tendency for greater priming for functional relations of various kinds than for other more formal kinds of semantic relation such as category coordinates or super-ordinates (Moss & Tyler, 1995). A related finding is that young children also weight the importance of functional properties of objects more heavily than that of form-based properties (Nelson, 1974; although see Flores d’Arcais, Schreuder, & Glazenberg, 1985 for an alternative account).

The robustness of functional property priming suggests that the information activated is that which is most relevant to the use of a word’s referent. This is plausible, given that when we are listening to spoken language we are generally trying to interpret it in terms of its relevance to ourselves. For example, in talking about chairs and tables, we are generally considering their use for sitting on or eating from, rather than the fact that they belong to the category furniture, or artificial object, or that they are in the same category as beds and sideboards. This claim is consistent with McNamara and Miller’s (1989) view that one of the major determinants of the salience of semantic properties is their “goal-relevance”. It is also unsurprising that function should have a special status when we consider how word meanings fit into larger scale mental representations of world knowledge. Many theories of knowledge representation stress the importance of scripts or schemata, that encode information about the nature and structure of the events and situations we encounter in the world (e.g. Garrod & Sanford, 1981; Schank & Abelson, 1977). Within such scripts, the functional role of each individual element is represented. Script elements will often correspond to the meanings of single words. If the aim of language comprehension is to interpret the incoming message with reference to the background knowledge we have stored as scripts, then the

---

4 This emphasis on functional properties runs counter to a widely held view in the neuropsychological literature that when semantic memory is impaired, superordinate information is less vulnerable to damage than detailed semantic properties (including functional properties, e.g. Warrington & Shallice, 1984). However, these conclusions are based on data from off-line tasks, and so may be probing knowledge in a different way from implicit semantic priming studies (Moss & Tyler, 1995).
crucial semantic information that we need about the meaning of each word is that which allows us to determine its functional role (or more correctly, the functional role of its referent in the world).

However, at this stage we have not yet determined whether the centrality of functional information generalises to all types of word. So far, all of the prime words that we have tested in our functional relation condition have been artifacts such as furniture, vehicles, and so on, rather than natural kinds such as animals and birds. It has been argued that functional information is at the core of our concepts for artifacts, in a way that is not true of natural kinds such as living things (e.g. Atran, 1989; Keil, 1986). For natural kinds, category membership may be more central. There is also a body of neuropsychological and developmental literature that suggests that perceptual properties may be especially important for living things (Keil, 1986; Warrington & Shallice, 1984). Therefore, we might expect that when prime words refer to living things, we will not see automatic activation of functional information. On the other hand, although living things rarely have well-defined functions or purposes in the same sense as artificial objects, we have recently started to investigate the “biological” functional properties of living things (such as breathing, moving, eating, for animals). These have been argued to be central to our representations (e.g. Keil, 1986) and appear to be well-preserved for patients with semantic impairments (Tyler & Moss, 1997; Moss, Tyler, Durrant-Peatfield & Bunn, 1998). Another limitation of all of this priming research is that the vast majority of primes are concrete nouns; it is possible that different semantic properties will be activated for abstract words (e.g. truth, luck) and for words from other syntactic classes, such as verbs and adjectives. However, these have only rarely been studied, especially in the current context of investigating the nature of semantic properties that are activated as the words are heard.

WHEN?

The previous section gives us some idea of the kinds of semantic information that are activated for a word at, or shortly after, its offset. However, it has been demonstrated in a variety of different psycholinguistic paradigms that the lexical access process starts considerably earlier than the end of a spoken word, at around 100–150 ms from word onset (e.g. Marslen-Wilson & Tyler, 1980; Tyler & Wessels, 1983). We now need to examine whether the meanings of words also start to be activated at this early stage (as claimed on the Cohort model), or whether there is a delay between activation of a word’s phonological form and access to its meaning, such that meanings only become available after the word has been uniquely identified. It is important to examine the time-course of the
activation of word meanings because the nature of speech recognition is fundamentally constrained by the fact that speech is distributed in time and is fast-fading. The question of how early lexical semantic information becomes available also has important consequences for the nature of language comprehension. If the meaning is available before a word is recognised, this would allow the semantic properties of a word to be evaluated against the prior context, and may speed recognition of a word that is highly congruent, and/or hinder the recognition of a word that is unexpected or anomalous. Early semantic activation would in turn lead to rapid integration of the word into the ongoing message-level representation, so providing contextual constraints that can help recognition of the next word along. Thus, the earliness of semantic activation in spoken word processing is closely related to the possible nature of contextual influences on word recognition.

One source of evidence that the meanings of words are made available at an early stage in the recognition process comes from studies demonstrating the influence of semantic variables on the latency of reaction times in word recognition tasks such as naming and lexical decision (Balota, Ferraro, & Connor, 1991). For example, several studies have shown that words with concrete meanings (i.e. whose referents can be experienced through the senses, e.g. table, dog, helicopter) are recognised more quickly than those with more abstract meanings (e.g. truth, luck, attitude; de Groot, 1989; James, 1975). Although most of the studies in this area have concerned the recognition of written words, we have demonstrated the same effect for spoken words using both naming and lexical decision tasks (Tyler, Voice, & Moss, 1996). This suggests that meanings are available before a word has been uniquely identified, and can influence the time course of recognition. The fact that concrete words are recognised more quickly is explained in different ways according to different models of word meanings. On some accounts, concrete words have dual semantic representations (both visual and verbal), whereas abstract words have only the single verbal representation (Paivio, 1986); on other accounts the difference between abstract and concrete words is quantitative rather than qualitative, with concrete words having more semantic properties than abstract words (Jones, 1985; Plaut & Shallice, 1993). The key point for all of these accounts is that there is a richer semantic representation for concrete words, which starts to be accessed early in the recognition process and facilitates the identification of the word. A similar claim has been made for the facilitatory effect of

5 Gernsbacher (1984) argued that the concreteness effect in lexical decision was confounded with rated word familiarity. However, a number of studies have now shown a concreteness effect when familiarity is controlled (as well as written word frequency), at least for low-familiarity words (e.g. Strain, Patterson, & Seidenberg, 1995).
multiple meanings. Some studies have demonstrated faster reaction times for words with many meanings (e.g. Kellas, Ferraro, & Simpson, 1988). Although this effect is less well documented, and appears to be less reliable (Rueckl, 1995) than the concreteness effect, the proposed basis for the effect is similar. The faster recognition times for semantically ambiguous words are thought to be a result of a richer set of semantic information becoming available early in the course of recognition, and so speeding identification of the word. Although the exact mechanism behind the facilitatory effects of a richer semantic representation is not clear, the general claim is that there is feedback of activation from the semantic level to the lexical form level, and more semantic information means more activation (the more is better principle, Balota et al., 1991). We return to this question later in the chapter.

A second source of support for the view that semantic information becomes available early in the duration of spoken words comes from studies that measure event-related brain potentials (ERPs). This technique uses electrodes placed on the scalp to record electrical activity in different areas of the brain while subjects are engaged in a cognitive task. The advantage of ERP studies is that they measure a response over which subjects have no voluntary control, and so cannot be affected by people’s conscious strategies or metalinguistic knowledge. Several studies have identified a consistent “N400 effect”, which appears to reflect the ease with which a word can be integrated into the prior context. The effect is that words that are consistent or predictable in their context (which may be a sentence or a one-word prime as in the semantic priming experiments described earlier) produce a smaller negative-going ERP, which peaks at approximately 400 ms after the onset of the stimulus, than do context-inconsistent or unprimed words (e.g. Brown & Hagoort, 1993).6

Holcomb and Neville (1990) examined the time course of the N400 effect for words presented in either the visual or auditory modality, as a function of whether the target word was preceded by a related or unrelated prime. The N400 reduction was found for both modalities, but it could be detected sooner after the onset of the target for spoken words. The effect could be detected as early as 200 ms from stimulus onset (as

6 These authors have argued that the N400 priming effect reflects post-lexical integration of the meaning of the target with the meaning of the prior context, rather than automatic activation of lexical items. This is comparable to Neely’s semantic matching account of semantic priming discussed earlier. This appears to be consistent with the view that post-lexical semantic integration should be seen as one of the automatic processes involved in normal language comprehension rather than as a task-specific conscious strategy (Hodgson, 1991), although this is not the conclusion reached by Brown and Hagoort (1993) who state that integration is a controlled process that can be guided by the subjects’ awareness of the information content of the discourse.
opposed to about 300 ms from onset for written words). Given that the duration of all of the spoken word targets was more than 200 ms, and that semantic information about the target word must be available for comparison with the prior context to produce the N400 effect, this indicates that semantic information is available before the acoustic offset of the target words.

Perhaps the most striking evidence for the early activation of spoken word meanings comes from a series of cross-modal priming studies described by Marslen-Wilson (1987). Subjects listened to primes that were either complete words (e.g. captain) or fragments of words that have been cut off before they become unique (e.g. /kapt.../). The fragment prime facilitated subjects' lexical decisions to target words related to the complete word (e.g. ship), which suggests that the meaning of captain has been activated when only part of the word has been heard. The same fragment /kapt.../ is also consistent with the cohort competitor, captive, and Marslen-Wilson (1987) found that targets related to this competitor (e.g. guard) were also significantly primed. But when the prime was the complete word, captain, ship was strongly facilitated, but guard was not. This indicates that early in the duration of the word, the semantic representations of multiple cohort competitors are activated, but as soon as further acoustic input is heard, which mismatches with any of those candidates (the final /n/ segment in the captain/captive example) the activation of the competitor meaning rapidly decays. In a further cross-modal priming study, Zwitserlood (1989) showed that multiple activation can be detected as early as 130 ms from the onset of the prime word. This suggests that semantic information starts to be activated as soon as lexical access begins.

One potential problem with the Marslen-Wilson and Zwitserlood studies is that the target words in each case were associates of the primes (e.g. captive-guard, captain-ship). It is possible that the early automatic priming observed in these studies is form-based associative priming, and does not reflect activation of semantic representations by the word-initial fragment. However, in a more recent study, Zwitserlood and Schrieffers (1995) found priming as early as 3.6 phonemes from the onset of the prime for targets which were non-associated category coordinates such as pig-horse, ruling out a purely form-based explanation of the results.\(^7\) We

---

\(^7\) Zwitserlood and Schrieffers (1995) also reported priming for non-associated targets for shorter prime fragments (mean 2.6 speech segments), but only when a delay of approximately 100 ms was introduced between prime and target. Although the authors claim that this is due to the time taken for activation to build up within the system, it is also possible that the inter-stimulus interval allows second-pass processes to operate (cf. Marslen-Wilson, Moss, & van Halen, 1996).
have also demonstrated significant priming for non-associated targets at the Isolation Point of words.\textsuperscript{8} The mean duration of the fragment from onset to Isolation Point was 331 ms, approximately half the duration of the complete words (mean 622 ms). In one experiment we carried out a direct contrast between associated and non-associated category coordinates, and we found that non-associated pairs such as silver-bronze prime as well as strongly associated pairs (silver-gold) (Moss, McCormick, & Tyler, 1997). Therefore, we can conclude that the results from the cross-modal priming experiments reflect early activation of semantic information rather than associative priming.

Taken together, the studies described provide compelling evidence that at least some semantic information is made available very early in the duration of a spoken word. We turn now to an additional question concerning the time course of activation of semantic information. Does the full semantic representation of the word start to be accessed at the same time, or is there earlier activation of some kinds of information than others? Given our earlier argument that functional properties are especially salient aspects of word meanings that play a central role in interpretation, we might also predict that this kind of information will be activated more rapidly than other semantic properties, at least for artificial objects.

However, in one of the few sets of studies to address this issue, Schreuder and colleagues have claimed that perceptual properties are the first to become available, at least for concrete nouns (Flores d’Arcais et al., 1985; Schreuder et al., 1984). They carried out visual priming experiments, in which the target shared either a “perceptual property” with the prime (e.g. cherry-ball, which share the property of being round), or a “conceptual property” (e.g. cherry-banana, sharing the property of being fruit), or both (cherry-apple, which are both round and fruit). Schreuder et al. (1984) found that when subjects were asked to name the targets, there was significant facilitation in the perceptual condition, but not in the conceptual condition, but when the reaction time task was lexical decision, targets sharing conceptual relations were significantly facilitated, but those sharing perceptual relations showed a reduced effect. The assumption is that word naming reflects activation at an earlier point in time than the lexical decision task, and therefore that perceptual properties undergo a

\textsuperscript{8} The Isolation Point is the point in the duration of a spoken word where people start to recognise what the word is going to be, but are not yet confident because there are still one or two other possible candidates available. It can be identified empirically in a gating task, where subjects are played out increasingly long fragments of a word from its onset, and asked to indicate what they think the word is and how sure they are that they are right (e.g. Tyler & Wessells, 1983).
process of rapid activation and decay when the prime is processed, whereas conceptual properties are accessed more slowly. This finding was extended by Flores d'Arcais et al. (1985). The authors claim that these results support their model of semantic representation in which concepts include two distinct kinds of properties—perceptual elements and knowledge-based or functional elements. Perceptual elements are based on our direct experience of objects in the world, and therefore their activation is also more direct and more rapid than that of functional elements, which are not apparent in the world but have to be learned or inferred.

Although this study suggests an intriguing insight into the time course of activation of semantic information it is not clear what kind of semantic information underlies the "conceptual property relation". The word pairs in this condition are all members of the same semantic category and many are probably strongly associated. Flores d'Arcais et al. (1984) describe the properties as "functional" when implementing the results in their model, but priming for these word pairs could be supported by information other than functional properties, such as category links via the superordinate (e.g. cherry-fruit-apple) or form-based associative co-occurrence. Another problem is that the targets were presented following a visual prime at an SOA of 400 ms. Thus, although naming may tap into an earlier stage of processing than lexical decision, it is still the case that subjects have seen the whole of the prime word by the time they start to respond to the target, and so the results cannot tell us about the activation of different semantic properties that may take place throughout the duration of a spoken word.

We undertook a cross-modal priming study to investigate the time-course of activation of different kinds of semantic information, that avoids the problems of the Schreuder et al. (1984) study (Moss, McCormick, & Tyler, 1995). We presented subjects with spoken prime words, which were played out either to their offset or up to the Isolation Point. Visual targets were displayed for lexical decision at one of these two points. The targets corresponded to either a perceptual property (e.g. spade-handle, aeroplane-wing) or a functional property (e.g. rifle-shoot, helicopter-fly) of the prime word. In both cases, the target describes a property without introducing the additional possible semantic connections that are involved in category relations. Moreover, the target words were not strong associates of the primes. All of the prime words were artificial.

---

9 In common with the majority of other psychological studies the materials in our perceptual condition refer only to the visual sense modality (e.g. colour, shapes, visible parts) and do not include auditory or olfactory properties. The latter are clearly important attributes of many objects, but most are unsuitable for a priming study because they cannot be described with a single word.
objects. If Schreuder et al.’s claim is correct, and perceptual properties are always activated more rapidly than functional properties, then we should see priming for the perceptual properties at an earlier point in the duration of the prime. However, if functional properties are more salient than perceptual properties for artificial objects, as suggested by our earlier priming results, and by developmental (e.g. Keil, 1986) and neuropsychological evidence (e.g. Warrington & Shallice, 1984), this would be more consistent with functional properties becoming available more quickly and so showing priming at the earlier point.

The results supported the latter hypothesis. Functional properties showed significant priming both when presented at prime offset and at the Isolation Point of the prime. This indicates that functional properties are activated early in the duration of a word referring to an artificial object—before the word has been uniquely identified—and continue to increase in activation as the rest of the word is heard. Perceptual properties, in contrast, showed no facilitation at the Isolation Point, although a priming effect emerged by the acoustic offset of the prime word. This shows a slower activation function for semantic information about what an object looks like than about what it is used for (see Fig. 3.1), and is not consistent with Schreuder et al.’s model on which perceptual elements of the

![Graph showing priming effects for perceptual and functional properties of words referring to artificial objects](image-url)

**FIG. 3.1** Priming effects for perceptual and functional properties of words referring to artificial objects (Moss, McCormick, & Tyler, 1997). Reprinted with permission.
semantic representation are activated before other kinds of information. The results are, however, consistent with our conclusions concerning the role of functional semantic information in language comprehension, as outlined in the previous section. The early availability of this kind of information allows the rapid assignment of functional roles to the elements in the discourse and provides an additional source of contextual constraint for the interpretation of upcoming words. However, as in our previous discussion, we have to bear in mind that the results to date all concern the time-course of semantic activation for artificial objects and will not necessarily extend to other classes of word.

WHICH?

The data presented in the previous section indicate that the meaning of a word becomes available very early in the duration of the speech input. In the introduction we highlighted the point that on most current models of word recognition, a set of potential word candidates are activated, not just the single word that is eventually recognised. It is therefore possible that semantic activation begins to be activated for a whole range of competitor words, as the speech input begins to be heard. However, the different models of word recognition differ with respect to which words will be activated as competitors. While the Cohort model claims that only words that overlap at the onset will be activated, other models such as the Neighbourhood Activation Model (Luce et al., 1990) and the TRACE model (McClelland & Elman, 1986) claim that the activation of competitor words will be a function of overall phonological overlap between the speech input and the phonological form of the word, regardless of whether there is a complete match at word onset. Therefore, a third set of important questions about lexical semantic processing is whether the meanings of all competitor words are activated and, if so, which words count as competitors.

Turning first to the claims of the Cohort model, there is good evidence that the meanings of at least some word-initial competitors are activated early in the duration of a word. It has already been mentioned that, in Marslen-Wilson’s (1987) study, word-initial fragments supported priming not only for targets related to the word that was actually being spoken, but also for targets related to competitors that overlapped at the onset. For example, /kapt../ primed ship (related to captain) as well as guard (related to captive). This effect was also demonstrated in a later study by Zwitserlood (1989) in which the fragments were embedded in a range of sentential semantic contexts. In this study, Zwitserlood found that even when fragments such as /kapt../ were heard in sentences that made one of the candidates more plausible than the other, both targets were still
facilitated until after the Isolation Point of the word, indicating activation of the meanings of competitor words, even when they were inappropriate in the current context. However, context effects did start to operate before the offset of the prime, by facilitating the selection of the correct word from the set of candidates activated on the basis of the sensory input. These findings are consistent with the claim of the Cohort model, that the semantic representations of multiple lexical items are activated as the first 150 ms or so of the input is heard, and that the meanings of these competitor candidates can then be evaluated against the context. This will allow the correct word to be recognised more quickly if it fits in well with the contextual constraints. Notice, however, that the prior context does not have any effect on whether the meanings of competitors are initially activated, but only on the process of selection among those meanings.

Although Marslen-Wilson and Zwitserlood's experiments show that the meanings of at least two word-initial competitor words are activated, further research is needed in this area before we can determine whether meanings are activated for the full set of cohort candidates. To go back to the example of captain, is it the case that the meanings of cat, catapult, cap, cab, character, and so on, are all activated in addition to captain and captive when the input /ka.../ has been heard, or is a greater amount of sensory information required before semantic information starts to be activated for the full range of competitors? Two factors which might influence the activation of competitor meanings are: (1) the relative frequency of competitors, and (2) the total number of competitors in the cohort. There is some evidence that relative frequency modulates the activation of multiple-word candidates. For example, Marslen-Wilson (1990) used cross-modal priming to examine the activations of word pairs such as road and robe, which diverge on their final segment. When these words were presented auditorily with the word-final consonants cut off, the amount of time by which the visual target word was facilitated depended on its frequency; so high-frequency words such as road were primed more strongly than their low-frequency competitors. This frequency effect diminished when the prime words were presented in unambiguous form (i.e. with their final segments intact). Even though Marslen-Wilson's experiments measured the activation levels of the forms rather than the meanings of competitor pairs, it is plausible that the meanings of high-frequency competitor words will show a similar head-start in the time-course of activation. In fact, Zwitserlood (1989) found a similar transient effect of relative frequency in the cross-modal priming study discussed previously; that is, at the earlier presentation points, prime words that were of higher frequency than their cohort competitor facilitated the associatively related targets more than did those primes
that were lower in frequency than their competitor. By the later points this difference had disappeared. However, the number of items in each frequency band was very small and the effect was not statistically significant. No effect of the relative frequency of the prime word to its competitors was found in the study by Moss, McCormick, and Tyler (1997), although the earliest presentation point was the Isolation Point of the prime, by which time such frequency effects may have passed. These results suggest that any effects of relative frequency of words in their competitor sets on the activation of word meanings are small and short-lived.

The number of competitors in a cohort may also affect the activation of the meanings of the members. Whereas it is plausible that all the meanings of a small cohort may be activated in parallel (e.g. squid and squint), the same may not be true when the cohort contains hundreds of words (e.g. cat, cab, can, captain, catapult, catalogue, catch, character, cancel, etc.). There is little empirical evidence relating to this issue, since most studies so far have examined quite small cohort neighbourhoods. However, Zwitserlood and Schriefers' (1995) experiments showed semantic priming for short fragments of words, which were consistent with a cohort of 12.7 words on average, although only when a delay of approximately 100 ms was introduced between prime and target.

We now turn to the question of whether there is activation of competitor words that overlap with the speech input but which mismatch at the onset, as predicted by the Neighbourhood Activation and TRACE models of spoken word recognition. The usual definition of a competitor (or neighbour) on these models is any word that diverges by only one phoneme in any position. So, for example, the competitors of cat would include bat, rat, pat, cot and cut as well as cab and cap. Several studies have examined the activation of competitors that do not overlap at onset by looking at the effectiveness of rhyme-priming (i.e. words that share all their phonological information with the exception of the initial phoneme; e.g. Connine, Blasko & Titone, 1993; Marslen-Wilson et al., 1996; Milberg, Blumstein & Dworetzky, 1987). For example, when we hear the word cattle, is the meaning of battle activated? This has been tested in priming experiments where the target is a word related to a rhyme of the prime. If the meaning of battle is activated by the input cattle, then hearing the prime cattle should facilitate the target war (related to battle but not to cattle). Although there have been some apparently conflicting results, the conclusion appears to be that there is no immediate activation of the meanings of rhyming competitors, even when the rhymes are different only by one phonetic feature of the initial segment (e.g. blank does not prime wood, even though it differs from plank by only the voicing feature of the first phoneme). In fact, the meaning of a word does not
appear to be activated even if the input differs by less than a full phoneme at the onset, but is ambiguous between two possible phonemes (e.g. an input p/blank that has a voicing value half way between a /p/ and a /b/ does not prime wood (related to plank) or page (related to blank; Marslen-Wilson et al., 1996). The lack of activation for rhyme competitors has been shown in experiments where the visual target for lexical decision is presented immediately at the offset of the auditory prime, so reflecting the activation state of competitors immediately after the word has been heard. In other studies, where the target has been presented at a delay from the offset of the prime, or where subjects’ reaction times have been particularly slow, some facilitation by rhyme primes has been found (Connine et al., 1993; Milberg et al., 1987). A direct comparison between immediate and delayed priming was carried out by Andruski, Blumstein, and Burton (1994). They found an immediate inhibitory effect of a subphonetic manipulation of the initial consonants of words on recognition of a related target. However, this effect disappeared when the target was presented at a delay of 250 ms. This suggests some kind of recovery process, which operates when the initial comparison between speech input and lexical forms fails. This may be similar to a backward priming effect; subjects try to reinterpret the prime input, partly on the basis of the target which hints towards a meaning other than that which may have been initially activated on the first pass (Marslen-Wilson et al., 1996). Therefore, while there is good evidence for activation of the meanings of at least some word-initial cohort competitors, there is no evidence for immediate activation of the meanings of competitors that mismatch at word onset.10

A third type of potential competitor comprises words that are embedded in longer words. In the case of words that are embedded at the onset of a longer word (e.g. cat in catapult), we can assume that the meanings of both would be activated in the same way as any other cohort competitors, since both are consistent with the first part of the input. However, when a word is embedded in the middle or at the end of a longer word, the situation may be more complicated. There is good evidence of this kind of competition from word-spotting tasks, where subjects have to press a button as soon as they recognise a word. For example, McQueen, Norris, and Cutler (1994) showed that subjects are

10 Because these studies were primarily concerned with the access of lexical form representations by competitor inputs, the targets for lexical decision were strong associates of the prime as well as being semantically related. As argued earlier, this may confound co-occurrence based associative priming with activation of semantic information in the mental lexicon. However, in this case, targets which are both semantically and associatively related to the competitor word are failing to prime, and therefore it is unlikely that targets which are semantically related only would prime either.
slower to monitor for a short word such as mess when it is embedded in another word (e.g. domestic) than when it is embedded in a non-word. But does this imply that the meaning of the embedded word is activated as the speech is heard? Evidence from a cross-modal priming study (Shillcock, 1990) suggests that it is: When a target related to the embedded word bone (e.g. rib) was presented for lexical decision at the offset of trombone there was a small but significant priming effect. This suggests that listeners continue to access additional possible word meanings, even though when the /b/ is heard, the input is still consistent with a word from the onset (trombone). Although this effect is predicted by models such as TRACE (McClelland & Elman, 1986) and Shortlist (Norris, 1994), in which lexical access procedures are initiated by each incoming phoneme, it is not consistent with the Cohort model, which claims that lexical access is initiated only at word onset and that one of the major cues to the onset of a word is the identification of the end of the previous word (so the /b/ in trombone should not be treated as a word onset, because trom is not a complete word). However, this finding should be treated with some caution, since a replication of Shillcock’s result using cross-modal repetition priming found no facilitation of the embedded word target (Marslen-Wilson, Tyler, Waksler, & Older, 1994).

In summary, the set of competitors for which semantic information is made available seems to depend on an exact (if transient) match between the available sensory evidence and lexical form representations. The meanings of word-initial cohort competitors are activated simultaneously for as long as they match the speech input. There is also some evidence that competing meanings are activated for embedded words that do not share word onsets. However, the meanings of words that rhyme with, or possess a global resemblance to, the input do not appear to be activated during first-pass processing, although they can be recovered quite rapidly with a second pass. We cannot yet be sure whether the meanings of all members of the initial cohort are accessed as soon as the beginning of the word is heard, or whether only a subset of the most frequent words start to be activated at the earliest point. The results of our cross-modal priming study (Moss, McCormick, & Tyler, 1997) suggest that the answer to this question may be modulated by the kind of semantic information concerned. Functional properties of artifacts seem to be activated by the Isolation Point of a word, regardless of the relative frequency of the word in its cohort, but perceptual properties are not activated until the offset of a word. Although there is some evidence that the relative frequency within its competitor set may affect the rise time of activation, this seems to be a small and fleeting effect.
HOW?

Having considered the evidence concerning the nature of semantic information accessed, when it is accessed and the activation of competitor meanings, we turn to the task of developing a model of lexical representation and processing that can accommodate these findings. As a first step the data argue strongly against any model of word recognition in which semantic information is activated for only a single word following the point of recognition. It is clear that at least some aspects of the meanings of words are available before recognition point, perhaps as early as 130 ms from the acoustic onset, and that meanings of at least some competitor words are also activated. This rules out Forster’s search model and Morton’s logogen model as adequate accounts of semantic activation processes during spoken word recognition, because both of these claim that the meaning of a single word is accessed only after the point at which it has been uniquely recognised on the basis of its match to the speech signal (see Introduction). Rather, we need a model that allows early activation of multiple semantic representations. The evidence also suggests that the range of meanings activated is largely determined by word-initial match to the acoustic input, and not just overall phonological overlap. This is more consistent with the Cohort model in which the competitor set is determined by the onset of the word, than with models like TRACE and NAM, where a word-initial match is not crucial.

As outlined earlier, neither the Cohort model, nor the other activation/competition based models of word recognition, have considered the semantic representation of words in any detail. There are two different approaches to semantic representation (and to representation within the cognitive system in general) that could be adopted. The first is a localist approach, as exemplified by Collins and Loftus’s (1975) semantic network model, in which the meaning of each word in the lexicon is represented by a single processing node. These nodes are connected by activation links to nodes corresponding to word forms, which in turn are activated by the acoustic input. The nodes within the semantic level are connected to other nodes for words which are related. This kind of localist connectionist approach is very similar to that adopted for the phonetic feature, phoneme and word-form levels in the TRACE model, and would be a natural extension of such a model. More recently, however, several researchers have rejected localist representation in favour of distributed models. In these models, a unit within the language system, such as the

---

11 These models were originally proposed to account for written rather than spoken word recognition, and it is a separate question whether they can account for activation of meanings during reading.
meaning of a word, is not represented by a single node within a network, but by a pattern of activation across a large set of nodes. These nodes may correspond to smaller units within the system (e.g., semantic features) or there may be no one-to-one correspondence between any node and any identifiable element of the system. In the following sections we discuss the interpretation of the data concerning lexical semantic activation according to a localist and distributed model, and the differences between them.

**Localist models.** The structure of a localist semantic network and its relation to the levels of representation generally believed to be involved in word recognition are shown in Fig. 3.2. When lexical form nodes start to be activated by the acoustic input, this activation passes on to the semantic level to activate the corresponding meaning node. To account for the earliness of semantic activation this activation must propagate in a cascaded fashion on the basis of partial activation at the form level, rather than requiring a threshold to be reached first. To account for the activation of meanings of multiple competitors, we can also assume that activation passes in parallel from the form to the meaning nodes of all words that are consistent with the speech input so far. On this model even the largest cohort of words could be activated in parallel because each node is separately represented and has no adverse effects of the

![Diagram](https://via.placeholder.com/150)

**FIG. 3.2** Representation of semantic and associative relations in a localist network.
activation of other meaning nodes. Finally, to account for the apparent influence of meaningfulness on word recognition, there must also be feedback of activation from the meaning to the form levels, in the same way as feedback is assumed to operate at other levels of the system.

In the semantic network model, activation also spreads out from the semantic nodes to other words with related meanings. For example, if the meaning of *cat* is activated, activation will spread to the models for *dog, kitten, tiger, animal, purr* and so on. This automatic spread of activation has been one of the most influential accounts of semantic priming effects, such as those described in earlier sections. If spread of activation is the basis of semantic priming, then the links that exist between nodes must include many different kinds of semantic relation. In particular, the links between words that are related by functional properties (at least for artificial objects) should be particularly strong. In terms of the network, this would mean that pairs of words like *broom* and *floor* or *party* and *music* would be more strongly linked than pairs with other kinds of relation such as perceptual relations. In this model, associative priming is captured by activation links between nodes at the lexical form level rather than within the semantic level (see Fig. 3.2). The basis for the formation of these links would be the frequent co-occurrence of the words in the language.

Although the basis for the formation and strengthening of links between associated words is sound and can be implemented easily using simple learning mechanisms (e.g. Hebbian learning; Hebb, 1949), it is not clear what principles encapsulate the linkage between purely semantically related words. The pattern of priming effects we find for perceptual and functional properties of artificial objects suggests that contextual relevance is an important factor: Functional properties are most relevant for the integration of the word’s meaning with its utterance context. It follows that the formation or strengthening of links in a semantic network should be based in part on this relevance, so that the immediately useful properties of a word are swiftly retrieved when a word is activated.

The effects of semantic properties on the time-course of word recognition are interpretable as top-down facilitation from the semantic network to the word-form level in a model such as TRACE. For example, the effect of semantic ambiguity on recognition time (Kellas et al., 1988) can be accommodated as an effect of strength of feedback from the semantic level. A word with multiple meanings is connected to more than one node in the semantic network. If, as activation propagates through the network, these nodes feed back activation to the word-form level, then their facilitatory effect should be proportional to the number of meanings, producing a small advantage for semantically ambiguous words (Balota et al., 1991; Kellas et al., 1988). The effect of concreteness on word
recognition can also be explained in terms of feedback, based on the assumption that concrete words are represented twice—in a verbal system and an image-based system—whereas abstract words are only represented verbally (Bleasedale, 1987; Paivio, 1986).

In summary, a localist implementation of semantic representations in a word recognition model could account for many of the empirical findings concerning the activation of word meanings. The separable effects of associative and semantic priming are explained neatly in terms of a two-level spreading activation system. Similarly, an advantage for words with multiple meanings in lexical decision is easily accommodated in the localist framework. However, the localist approach, although tenable, seems less productive when we focus on effects of semantic structure, such as concreteness effects or functional/perceptual distinctions. In these cases, it seems somewhat clumsy to represent fine details of the structure of word meanings purely in terms of their links to other words. It may be more profitable instead to turn to distributed models of word recognition, which stress the componential nature of word meaning and allow partial activations of selected properties to develop.

*Distributed models.* Several distributed connectionist models of language comprehension mechanisms have been developed. These models share the common representational assumption that elements of the language system are captured as patterns of activation over large numbers of processing units, rather than in a single node. This system of representation makes processes such as generalisation to novel elements and pattern completion simple to perform (Hinton, McClelland, & Rúmelhart, 1986). A further advantage of distributed systems is that when damage is simulated (by removing units or connections between units) performance degrades gradually, as is generally found in brain-damaged patients (e.g. Plaut & Shallice, 1993).

The distributed representation approach seems particularly appropriate in the case of semantic representations. As we outlined earlier, most theories of word meanings claim that meanings are not stored as unanalyzable wholes, but rather are composed of a range of different kinds of information and semantic features. In a distributed account, it is possible to model a word's meaning over a set of processing nodes which each correspond to an individual semantic property. However, by assigning each node of the network to a labelled "microfeature", there is still an element of localist representation in the system; it has simply been moved from one level of representation (word meanings) to a more fine-grained level (semantic features). It is conceivable that even this level of localism is unrealistic and that there is in fact no level of representation at which single units have a fixed meaning (Clark, 1993).
A number of connectionist models have demonstrated aspects of lexical processing within the distributed framework (e.g. Joordens & Besner, 1994; Kawamoto, Farrar, & Kello, 1994; Masson, 1995; Plaut & Shallice, 1993). Microfeatural accounts of distributed representations ensure that words with similar meanings will have similar representations. For example, the words teapot and coffeepot might share such features as [container], [handle], [spout], and differ on relatively few features. Distributed models of lexical processing exploit these similarities in modelling semantic priming data. Masson (1995) modelled access to semantic information during word recognition in terms of a mapping from a set of perceptual (orthographic) nodes onto a set of semantic nodes, which encoded the distributed semantic representations of the words. The network employed a simple learning algorithm to memorise the semantic patterns, allowing the network to develop “attractors” for each word. The time taken to access the meaning of a word could then be modelled in terms of the number of processing cycles needed to settle into the correct semantic state. Priming occurs when a word with a similar semantic representation is presented before the target. The network settles into the semantic state corresponding to the prime word and when the target is presented the network is quick to settle into the semantic state for the target, due to the close proximity of the prime’s semantic state.

Although representational similarity may provide a good basis for explaining purely semantic priming, it cannot easily be extended to cover associative priming. This can occur even when there is little or no similarity between the meanings of the two words, as in pillar and society or elbow and grease. In addition, priming based on representational similarity is necessarily symmetrical, whereas associative priming is often highly asymmetrical. Associative priming has, however, been modelled within the distributed framework using recurrent networks which exploit co-occurrence statistics in order to reduce output error (Moss et al., 1994; Plaut, 1995). In these models, associative priming is based on the assumption that associatively related words often co-occur in sentence contexts, as in the phrases elbow grease and pillar of society (Fischler, 1977). These co-occurrences can be learned by a network and then employed in the prediction of incoming words. The associative priming effect is then realised as an advantage in the recognition of a contextually predictable word over a contextually unpredictable word. Moreover, there is some doubt as to whether representational similarity alone can be the full explanation of purely semantic priming. If priming results from the representations of prime and target involving similar patterns of activation over semantic units, then we should see the most robust effects for those word pairs that have the most similar meanings, that is, synonyms and close category coordinates. However, our research suggests that word pairs connected by
goal-relevant functional properties may in fact prime more strongly. The degree of semantic overlap between functionally related word pairs such as *broom-floor* or *shampoo-hair* intuitively seems to be less than that between close category coordinates such as *broom-mop* or *shampoo-conditioner*. The explanation of such effects within a distributed model will depend on more sophisticated accounts of the structure of the semantic representations of words such as *broom*, and how the functional properties are captured within them.

Distributed representations are convenient for modelling partial activations and so are able to explain why some properties of a word become active more quickly than others (Borowsky & Masson, 1996; Kawamoto, 1993). Such behaviour is particularly valuable in the explanation of the differential time-course effects found in the activation of perceptual and functional properties of artificial objects (Moss, McCormick, & Tyler, 1997). This can be explained as an example of pattern completion from a partial representation in an attractor-type network (e.g. Masson, 1995). As a word is heard, a partial semantic activation develops, consisting, in the case of an artificial object, of its functional properties. The partial nature of this representation is transient, since the network quickly fills in the remaining features to complete the pattern, activating both perceptual and functional properties. However, as in the localist description of this kind of effect, we need to ask what drives the initial activation for functional properties. Somehow, contextual relevance must be built into the learning phase of such a model to allow functional properties of artificial object words to be activated more quickly.

The concrete–abstract distinction has been captured in distributed systems as a qualitative representational difference. Plaut and Shallice (1993) used a microfetural semantic representation for which concrete words had roughly four times as many features as abstract words. The psychological motivation for this difference was based on the finding that the number of predicates subjects can generate for a word is highly correlated with its imageability (Jones, 1985).\(^{12}\) In addition, the meanings of concrete words are believed to be more stable across contexts than abstract words (Saffran & Schwartz, 1994). These findings were used to argue that the context-independent semantic representation generated when an isolated word is read or heard is richer for concrete than for abstract words. Plaut and Shallice's objective was to simulate the neuropsychological data on deep dyslexia by examining network performance under damage. However, this representational system may also be able to explain the effect of concreteness found on lexical decision time for both spoken and written words. This is because there is greater redundancy in

---

\(^{12}\) Imageability is, in turn, highly correlated with concreteness.
the richer semantic representation used for concrete words. This redundancy makes a full semantic representation easier to construct given the partial information that develops during the time-course of perception of a word.

The fact that words often have multiple meanings is more troublesome for distributed models. Each meaning is represented as a pattern over all the processing units, which implies that only one meaning can be properly represented at any one time. A distributed model of semantic ambiguity could easily choose to activate only one meaning for an ambiguous word, based on frequency or contextual factors. However, the experimental evidence suggests that in many cases two or more meanings are activated simultaneously (Simpson, 1984). The only way to accommodate this in a distributed model is to construct a blend of the meanings, which is broadly similar to both meanings, by settling into a pattern that is somewhere between the two, and so shares some properties of both meanings (Joordens & Besner, 1994). It is, as yet, unclear whether modelling ambiguity in this way allows the lexical decision advantage for ambiguous words to be simulated. Recent attempts to simulate this effect have had mixed results (e.g. Borowsky & Masson, 1996; Joordens & Besner, 1994; Kawamoto et al., 1994), with attention focusing on the representational basis required in order to make a lexical decision. A deeper understanding of the lexical decision task is necessary in order to resolve this debate.

A distributed model of speech perception. The models just described have all employed a static form representation as input in order to model the lexical processing involved in word recognition. However, in the case of speech perception we have argued that a more realistic representation is needed, which captures the transient nature of the incoming speech. Gaskell and Marslen-Wilson (1997b) have recently developed a model which incorporates many of the key processing claims of the Cohort model into a distributed connectionist architecture, and which includes a level of semantic representation (see Fig. 3.3). The model employs a simple recurrent network architecture (Elman, 1991) to map from a stream of phonetic features, representing incoming speech, to distributed lexical representations. The recurrent connections allow the network to recognise patterns entering the network over time, meaning that internal representations of words can be learned by the network.

In comparison to the speech perception models we have discussed earlier on in this chapter, such as TRACE and Cohort, the scope of the model is extended: It captures the time-course of activation of meaning and other lexical knowledge rather than simply the identification of word forms. Indeed, the model no longer incorporates a representation of form that is separable from the lexical entry. Both form and meaning are
products of the perceptual process and are represented at essentially the same level in the system.

Nonetheless, many of the properties of earlier models of speech perception are retained, but implemented in a distributed framework. A crucial feature of the model is the early activation of multiple semantic representations, as in the Cohort model. As we have discussed, Cohort predicts that once a small amount of word-initial speech has been processed, the meanings of all words matching so far are activated, allowing contextual information to influence the process of selection from the remaining candidates. Because Cohort is a localist model, employing independent word-form representations, these meanings can also be accessed independently. For example, given the onset /fr/, the meanings of the words freeze and frost would be activated, making elements such as cold available for integration with the sentential context. However, the meanings of fry and frazzle would also be activated, whose semantic representations would include the feature hot. According to Cohort, both these (and many other) concepts would be simultaneously active during the presentation of the speech.

However, in the distributed model all these meanings must be encoded in the same distributed representational space. As we have seen, this implies that when many meanings are activated simultaneously they will interfere, producing a semantic blend of the different features. Obviously,
the details of this interference depend on the theory of semantic representation chosen, but, as a general rule, its strength depends on the number of competing words and their relative frequency. This gives a rather different picture of the time-course of activation of word meaning. Instead of many meanings being activated early on, with a gradual filtering process as subsequent information eliminates candidates, the activation of semantics increases as the number of matching candidates drops. The distributed model predicts that at the beginning of a word, the semantic representations of the thousands of word candidates interfere strongly, cancelling out any semantic activation. As speech eliminates mismatching candidates, the interference between the remaining candidates decreases, increasing the potential for priming of related words. The recognition point of a word is marked by the isolation of a single coherent semantic representation.

The distributed model of Gaskell and Marslen-Wilson does not predict that multiple meanings cannot be accessed in parallel, but it does suggest that their informativeness depends strongly on the number of meanings being represented. Thus, we may well get semantic priming for speech tokens before recognition point (e.g. Zwitserlood & Schriefers, 1995), but the strength of the priming should depend strongly on the number of word candidates remaining and their relative frequencies. Evidence for this pattern of semantic activation was found in a study by Gaskell and Marslen-Wilson (1997a) using a cross-modal semantic priming.

The integration of semantic processing and lexical access also offers the potential for explaining a wider range of data. The differences we find in the time course of activation of different types of semantic knowledge (Moss, McCormick, & Tyler, 1997) can be explained in terms of the influence of semantic structure on the interference between semantic patterns. Broad semantic distinctions such as between perceptual and functional properties or between concrete and abstract words should be reflected in the distributed semantic structure of the words and are therefore able to affect the blending and interference between semantic representations during lexical access.

DISCUSSION

In this chapter we have examined evidence concerning the nature of semantic information activated when a word is heard, the time-course of its activation and the activation of meanings of competitors for recognition. The major conclusions are that: (1) a wide range of information is automatically activated when a word is heard in isolation—certainly more than defining features, and with functional attributes being particularly salient, at least for concrete artifact nouns; (2) word meanings start
to be activated early in the word's duration, probably on the basis of the first CV or CCV (this means that there is no delay between the activation of lexical forms and the activation of semantic contents; moreover, it may be possible to chart different activation curves for different properties of the word's meaning); and (3) the meanings of a number of candidates for word recognition are all activated as a word starts to be heard. The available evidence suggests that the competitor space is determined mainly by the onset of the word, with less evidence for the activation of meanings of competitors that do not overlap at word onset.

Next we examined how these characteristics of lexical semantic access could be accounted for in a model of word recognition and semantic representation. We argued that the data are consistent with the main processing claims of the Cohort model but that these claims can be implemented in two different kinds of architecture—either localist or distributed representation. These approaches both capture the claims about multiple semantic access, activation, and competition, but the different representational assumptions have important consequences for how these processes operate. This highlights the important point that issues concerning lexical semantic processing are intricately bound up with issues concerning lexical semantic representation, and the two cannot be considered in isolation. We have discussed how such an integrated model could accommodate these findings, considering both localist and distributed representations. However, we are not yet in a position to reject one of these approaches to semantic representation in favour of the other. Indeed it is slightly disturbing that two such dissimilar theories are proving so difficult to distinguish experimentally (although see Masson, 1995 for a step in this direction).

In concluding, it is important to emphasise that we have only been able to discuss the access of the meanings of words heard in isolation. However, the goal of language comprehension is, of course, not to understand individual unconnected words, but to interpret the meanings of words as they are heard in sentential and discourse contexts. We have, however, referred to the interaction of lexical semantic processing and the higher level context at a number of points. For example, claims about the earliness of activation of semantic information clearly have implications for the ability of contextual constraints to influence the word recognition process. Also, the apparently central role of functional properties of word meaning are readily accounted for by the important part they play in the integration of word meanings into larger knowledge structures such as scripts.

Contextual constraints on the processing of word meanings probably have their most noticeable role in the disambiguation of words that
have more than one meaning. But the role of context is ubiquitous, and applies to unambiguous, just as much as to ambiguous words. Indeed, we should probably think of ambiguity as a matter of degree, rather than an all-or-none state, ranging from subtle variations on a single meaning, through words with different senses and metaphorical extensions, to the classical homophones that have two unrelated meanings attached to a single sound sequence. Even in the former case—words with a single meaning—the constellation of semantic properties activated for the word in any given utterance may differ according to the nature of the context. The relevance of different properties in different contexts has been clearly demonstrated in a number of experimental studies. For example, Barclay, Bransford, Franks, McCarrell, and Nitsch (1974) showed that different properties of a word like piano were more or less effective in cueing recall depending on whether piano had been read in the context of playing a piano (piano as a musical instrument) or lifting a piano (piano as a heavy wooden object). The important question is how such “semantic flexibility” comes about; does the prior context guide the access process such that only relevant properties are activated, or are all default properties automatically activated, followed rapidly by a process of selection? Both of these positions have been argued in the literature with some experimental support for both views; for example, Whitney, McKay, Kellas, and Emerson (1985) argued in favour of the automatic context-independent view on the basis of a priming experiment in which target words were facilitated whether they were related to a contextually relevant or irrelevant semantic property of the prime in different sentences. In contrast, Tabossi (1988) found priming only for relevant semantic features in a similar paradigm, but with better controlled and more strongly constraining sentence contexts. An alternative account is that in certain kinds of semantic contexts, the level of access is “deeper” with more properties being activated than in other contexts, although not necessarily limited to relevant properties only (Moss & Marslen-Wilson, 1993). Thus, the debate about the processes by which context modulates the meanings of words continues. However, the consideration of the representation and processing of words in isolation provides an essential framework against which to study these contextual effects.

ACKNOWLEDGEMENTS

This research was supported by a British Academy Postdoctoral Research Fellowship to Helen Moss and a programme grant from the Medical Research Council of Great Britain to William Marslen-Wilson and Lorraine K. Tyler.
REFERENCES


