Journal of Memory and Language **46,** 245–266 (2002)
doi:10.1006/jmla.2001.2810, available online at http://www.academicpress.com on IDEAL®

Making Sense of Semantic Ambiguity: Semantic Competition in Lexical Access

Jennifer Rodd

MRC Cognition and Brain Sciences Unit, Cambridge, United Kingdom

Gareth Gaskell

Department of Psychology, University of York, York, United Kingdom

William Marslen-Wilson

MRC Cognition and Brain Sciences Unit, Cambridge, United Kingdom

There have been several reports in the literature of faster visual lexical decisions to words that are semantically ambiguous. All current models of this ambiguity advantage assume that it is the presence of multiple unrelated meanings that produce this benefit. A set of three lexical decision experiments reported here challenge this assumption. We contrast the ambiguity seen in words like *bark*, which have multiple unrelated meanings, with words that have multiple related word senses (e.g., *twist*). In all three experiments we find that while multiple word senses do produce faster responses, ambiguity between multiple meanings delays recognition. These results suggest that, while competition between the multiple meanings of ambiguous words delays their recognition, the rich semantic representations associated with words with many senses facilitate their recognition.

Key Words: lexical ambiguity; polysemy; distributed semantic representations.

Many words are semantically ambiguous, and can refer to more than one concept. For example, *bark* can refer either to a part of a tree or to the sound made by a dog. To understand such words, we must select one of these different interpretations, normally on the basis of the context in which the word occurs.

Words can be ambiguous in different ways; a word like *bark* has two semantically unrelated meanings, which seem to share the same written and spoken form purely by chance. More common than this type of accidental ambiguity is the systematic ambiguity between related word senses. For example, the word *twist* has a range of dictionary definitions including *to make into a coil or spiral*, *to operate by turning*, *to alter the shape of*, *to misconstrue the meaning of*, *to wrench or sprain*, and *to squirm or writhe*. The meaning of this word varies systematically according to the context in which the word is used; for example, there are important differences between what it means *to twist an ankle* compared with *to twist the truth*. However, although the meaning of the word is ambiguous between these different interpretations, the interpretations are closely related to each other both etymologically and semantically; this is quite unlike the ambiguity for a word like *bark*.

The linguistic literature makes a distinction between these two types of ambiguity and refers to them as homonymy and polysemy (Cruse, 1986; Lyons, 1977, 1981). Homonyms, such as the two meanings of *bark*, are considered to be different words that, by chance, share the same orthographic and phonological form. Specifically, homographs are different words that share the same orthographic form, while homophones share the same phonological form. On the other hand, a

This research was funded by a MRC Studentship awarded to J. Rodd. Some of the research in this paper was reported at the Twenty-first and Twenty-second Annual Conferences of the Cognitive Science Society (August 1999 and 2000).

Address correspondence and reprint requests to Jennifer Rodd, Centre for Speech and Language, Department of Experimental Psychology, University of Cambridge, Cambridge, CB3 3EB, UK. E-mail: jrodd@csl.psychol.cam.ac.uk.

0749-596X/02 $35.00

polysemous word like *twist* is considered to be a single word that has more than one sense. Despite this linguistic distinction between homonymy and polysemy, psychologists have often used the two terms interchangeably (see Klein and Murphy (2001) for a discussion of this issue).

All standard dictionaries respect this distinction; lexicographers decide whether different usages of the same orthography should correspond to different lexical entries, or different senses within a single entry. Many criteria (e.g., etymological, semantic, and syntactic) have been suggested to operationalize this distinction between senses and meanings. However, it is generally agreed that while the polysemy/homonymy distinction appears easy to formulate, it can be difficult to apply with consistency and reliability; people often disagree about whether two usages of a word are sufficiently related that they should be considered to be senses of a polysemous word rather than homonyms (Lyons, 1977, 1981; Kilgarriff, 1992). While there may not always be a clear distinction between these two types of ambiguity, it is important to remember that words that are described as semantically ambiguous can vary between these two extremes and that our mental representations of these two types of words are likely to be very different.

Semantic ambiguity is very common in language, and our ability to understand ambiguous words is an important property of our language processing system. Evidence about how ambiguity affects human language performance can provide important constraints on models of word recognition. In particular, models of word recognition have been required to accommodate evidence that visual lexical decisions are faster for ambiguous words. In this paper, we evaluate the evidence of an ambiguity advantage in the light of the distinction between word meanings and word senses. In particular, we argue that while earlier studies in the literature show that semantic ambiguity can produce a processing advantage, it is not clear whether this is caused by ambiguity between multiple word meanings or between multiple word senses. Despite this, all current accounts of the ambiguity advantage assume that the advantage is produced by multiple, unrelated meanings; the experiments reported here investigate whether this assumption is correct.

THE AMBIGUITY ADVANTAGE

The *ambiguity advantage* is the finding that visual lexical decisions are faster for words that are semantically ambiguous. Early reports of an ambiguity advantage came from Rubenstein, Garfield, Millikan (1970) and Jastrzembski (1981), who found faster visual lexical decisions for ambiguous words than for unambiguous words matched for overall frequency. However, Gernsbacher (1984) discussed a possible confound between ambiguity and familiarity in these experiments; words with more than one meaning are typically more familiar. She found no effect of ambiguity over and above familiarity. Since then, however, several papers have reported an ambiguity advantage in visual lexical decision experiments using stimuli that were controlled for familiarity (e.g., Azuma & Van Orden, 1997; Borowsky & Masson, 1996; Hino & Lupker, 1996; Kellas, Ferraro, & Simpson, 1988; Millis & Button, 1989; Pexman & Lupker, 1999). Although these studies vary in the robustness of the effects they report, their cumulative weight has had the effect of establishing the ambiguity advantage as an important constraint on theories of lexical representation and lexical access.

Interestingly, a robust ambiguity advantage has only been observed using lexical decision. For word naming the ambiguity disadvantage has been very unreliable (see Borowsky and Masson (1996) for a discussion of this issue). Further, on a range of other tasks in which it is necessary to disambiguate the meaning of the ambiguous word, there is a clear ambiguity disadvantage. For example, when eye-movement measures are used for reading words in context, and if the context is neutral and the ambiguous word has two meanings of approximately equal frequency, then there is a disadvantage for ambiguous words compared with unambiguous words (see Rayner (1998) for a review). Additionally, priming studies have shown that, even in an inappropriate context, both meanings of an ambiguous word seem to be accessed (Swinney, 1979; Onifer & Swinney, 1981). Therefore, for an ambiguous word presented in a sentence context it appears that both meanings of the word are initially activated and that this produces longer reading times for these words. This suggests that the ambiguity advantage may emerge only in situations where it is not necessary to integrate the meaning of the word into a coherent semantic representation of a sentence. For this reason, many of the early accounts of the ambiguity advantage assumed that it is a presemantic effect, and that participants are performing the task without disambiguating the ambiguous word.

One interpretation of the ambiguity advantage has been that ambiguous words benefit from having multiple entries within the lexicon. For example, Kellas et al. (1988) assume that words are represented by individual nodes within an inhibitory lexical network. They suggest that while the multiple nodes of an ambiguous word do not inhibit each other, they both act independently to inhibit all other competing entries, and this increased inhibition of competitors produces the faster recognition times. A related account (e.g., Jastrzembski, 1981) assumes that the benefit arises from the presence of noise or probabilistic activation; because ambiguous words are assumed to have multiple entries, they benefit from having more than one competitor in the race for recognition. On average, by a particular point in time, one of these competitors is more likely to have reached the threshold for recognition than a word that has only one entry in the race.

These accounts of the ambiguity advantage predict that the effect will be seen for any ambiguous words whose meanings are sufficiently unrelated that they correspond to separate entries in the mental lexicon; they make no specific predictions about what should happen for words with multiple senses, as it is not clear whether related word senses would correspond to separate entries within the network.

An alternative view of word recognition is that words compete to activate a representation of their meaning. Several recent models of both spoken and visual word recognition have taken this approach (Gaskell & Marslen-Wilson, 1997; Hinton & Shallice, 1991; Joordens & Besner, 1994; Plaut, 1997; Plaut & Shallice, 1993). Rather than including localist lexical representations, these models use distributed lexical representations; each word is represented as a unique pattern of activation across a set of orthographic/phonological and semantic units.

Within models of this type, the orthographic pattern of an ambiguous word must be associated with multiple semantic patterns corresponding to its different meanings. When the orthographic pattern is presented to the network, the network will try to simultaneously instantiate the word’s two meanings across the same set of semantic units. These competing semantic representations will interfere with each other, and this interference is likely to increase the time it takes for a stable pattern of activation to be produced. At first glance, therefore, the ambiguity advantage is inconsistent with the predictions of these models.

In response to this inconsistency, there have been several attempts to show, with varying degrees of success, that this class of model can show an advantage for ambiguous words. Joordens and Besner (1994) and Borowsky and Masson (1996) both suggest that because ambiguous words have more than one meaning, on average the randomly determined initial state will be closer to a valid finishing state for ambiguous words, and this could reduce the time it takes for the network to settle. Kawamoto, Farrar, and Kello (1994) suggested that if an error-correcting learning algorithm was used to learn the mapping from orthography to semantics and then to compensate for the increased error produced by the ambiguous words in the semantic units, stronger connections are formed between the orthographic units. If lexical decisions are made on the basis of orthographic representations, then this could improve performance for ambiguous words.

These accounts of how the ambiguity advantage might arise from a model incorporating distributed semantic representations all predict that the effect should be strongest when the meanings of the ambiguous words are unrelated. In the proximity advantage account of Joordens and Besner (1994) and Borowsky and Masson

(1996), the benefit from having two meanings will be maximal when the states of semantic activation corresponding to the two meanings are furthest apart, i.e., when the meanings are semantically unrelated. Similarly, according to the Kawamoto et al. (1994), the ambiguity advantage is driven by the error produced during the learning of the meanings of ambiguous words. Therefore, the effect of ambiguity should be greatest when this error is maximal, i.e., when the meanings of the ambiguous words are highly unrelated.

In summary, all current accounts of the ambiguity advantage assume that it is ambiguity between unrelated meanings that produces the ambiguity advantage. None of these models explicitly predict what the effect of multiple word senses should be. For those models in which the benefit arises because of the presence of multiple localist lexical entries for ambiguous words, the presence of a benefit for words with multiple senses will depend on whether multiple senses are represented as separate entries within the network. Those models that involve distributed semantic representations predict that words with multiple senses may show a processing advantage, but that this should be reduced compared with words with multiple meanings.

In the following section we analyze in detail the stimuli used in previous studies that show a robust ambiguity advantage. This may help us to determine whether the assumption that the ambiguity advantage reflects a benefit for words that have unrelated meanings is correct, and to determine whether multiple word senses may also play a role. In particular, we look in detail at the stimuli used by Millis and Button (1989), Azuma and Van Orden (1997), and Borowsky and Masson (1996). These are three representative studies which show robust effects of ambiguity.

WORD SENSES AND WORD MEANINGS

As mentioned earlier, lexicographers routinely distinguish between word meanings and word senses when they structure dictionary entries. These dictionary entries provide a simple, yet reliable way to classify words as being ambiguous between multiple meanings or between multiple senses. We have used the entries in The Online Wordsmyth English Dictionary–Thesaurus (Parks, Ray, & Bland, 1998).[1] As we report later, the classifications made in this dictionary correspond closely to participants' judgements about the relatedness of the meanings of ambiguous words.

Looking first at the stimuli used by Millis and Button (1989) and Azuma and Van Orden (1997), neither study makes the direct contrast between ambiguous and unambiguous words; words with many meanings are compared with words with few meanings. Words were assigned to these groups by counting the number of meanings that participants could provide for each word. Crucially, both studies count highly related word senses as separate meanings. This can be demonstrated by example.

Millis and Button (1989) use *tell* as an example of a word that has many meanings. Participants produced up to four meanings for this word. These were *to inform*, *to explain*, *to understand*, and *to relate in detail*. Although there are clearly important differences between these four definitions, these differences are relatively subtle; all four definitions relate to a single core meaning of the word, to do with providing information. All these definitions are included as senses within a single entry in the Wordsmyth dictionary. This is just one of several examples of high-ambiguity items that are ambiguous between multiple word senses rather than between unrelated word meanings.

We compared the groups of high- and low-ambiguity words in the two experiments reported by Millis and Button (1989) that found a significant ambiguity advantage, and found that they do not differ in their number of Wordsmyth entries ($t(46) = .5$, $p > .6$) (see Table 1). In contrast, the two groups of words did differ significantly in the total number of senses they are given in the Wordsmyth dictionary ($t(46) = 4.4$,

---

[1] This particular dictionary was chosen because it reliably separates semantically unrelated meanings into distinct lexical entries, but unlike some other dictionaries it does not require that senses within an entry have the same syntactic class. This reflects the intuitions of participants that meanings from different syntactic classes can be highly related (Azuma, 1996).

TABLE 1

Mean Number of Dictionary Entries and Senses for Stimuli

	Stimulus group	Dictionary entries	Total senses	Dominant meaning senses
Millis & Button (1989)	Few meanings	1.2	6.9	6.4
	Many meanings	1.3	12.8	11.5
Azuma & Van Orden (1997)	Few meanings	1.9	9.6	7.2
	Many meanings	1.7	18.6	13.6
Borowsky & Masson (1996)	Unambiguous	1.0	6.8	6.6
	Ambiguous	1.8	12.1	8.8

$p < .001$), and in the number of dictionary senses of the dominant meaning of each word ($t(46) = 4.0$, $p < .001$).[2] Therefore, the high-ambiguity words used by Millis and Button (1989) have more senses than the low-ambiguity words, but crucially, they do not have more unrelated meanings. This suggests that the ambiguity advantage seen in this study should be interpreted as a benefit for words that have many related word senses, and not a benefit for unrelated meanings.

Azuma and Van Orden (1997) also compared words with few (2–4) and many (6–10) meanings. Again, items were assigned to these groups on the basis of the total number of meanings provided by participants, and highly related dictionary senses were counted as separate meanings. A different group of participants rated the relatedness of these meanings. For those words classed as having unrelated meanings, there was a benefit for those with many meanings over those with few meanings. However, it cannot be assumed that these words only have unrelated meanings. The relatedness measure used by Azuma and Van Orden (1997) was derived from the relatedness of the words' dominant meanings with each of its subordinate meanings. Therefore (as noted by the authors), this measure contains no information about the relatedness of the different subordinate meanings. This means that we cannot be certain that the high-ambiguity words have more meanings that are semantically unrelated than the low-ambiguity words.

Analysis of the dictionary entries for these stimuli shows a similar pattern to that seen for the Millis and Button (1989) stimuli. First, the two groups of words did not differ significantly in their number of dictionary entries. In fact the high-ambiguity words have slightly fewer dictionary entries than the low-ambiguity words. Second, as with the Millis and Button (1989) stimuli, the high-ambiguity words did have a significantly higher total number of senses within these entries ($t(33) = 4.6$, $p < .001$) and a higher number of senses for the dominant meaning of each word ($t(33) = 3.2$, $p < .005$) (see Table 1).

Finally, let us look at the stimuli used by Borowsky and Masson (1996). Their ambiguous and unambiguous words were taken from Fera, Joordens, Balota, Ferraro, Besner (1992), who asked participants to rate whether a word had no meaning, one meaning, or more than one meaning. This is the same procedure that was used by Kellas et al. (1988) and Hino and Lupker (1996). We chose to look in detail at the Borowsky and Masson (1996) stimuli because their result was the one of these where the effect of ambiguity was statistically significant and because they used the largest set of words.

The stimuli used by Borowsky and Masson (1996) appear to provide a clear comparison between words that people consider to be ambiguous and unambiguous. An analysis of the number of senses and meanings given in the Wordsmyth Dictionary for the 128 words used

[2] For those words with only one entry in the dictionary, the dominant meaning was simply this single meaning. For those words with multiple entries in the dictionary, the dominant meaning was determined by asking a group of 38 participants to provide associates for each word and then selecting the meaning for which the higher proportion of associates were related. This procedure was used by Twilley et al. (1994) to produce dominance measures for ambiguous words.

in their experiment revealed that the two groups of words do differ significantly in their number of dictionary entries ($t(126) = 7.3$, $p < .001$); the ambiguous words had, on average, more meanings than the unambiguous words. However the two groups of words also differed significantly in their total number of senses within these entries ($t(126) = 5.6$, $p < .001$) and in the number of senses that the dominant meaning of each word was given in the dictionary ($t(126) = 2.2$, $p < .05$); again the high-ambiguity words had more word senses (see Table 1). It is possible that, as with Millis and Button (1989) and Azuma and Van Orden (1997), the ambiguity advantage shown by Borowsky and Masson (1996) may reflect an advantage for words with large clusters of related word senses.

In all three of these studies, the high-ambiguity stimuli have more related word senses than the low-ambiguity words. In contrast, only one of these studies showed a difference in the number of dictionary entries. This is surprising. The Millis and Button (1989) and Azuma and Van Orden (1997) studies defined high-ambiguity words as those for which participants could generate many definitions. Therefore we might have expected these words to be ambiguous both in terms of number of senses and number of meanings, and yet they seem to differ only in their number of word sense. Why is there a bias in their stimuli toward ambiguity between multiple meanings rather than multiple senses?

We believe that this bias reflects the fact that multiple senses are simply more frequent in the language than multiple meanings. This is supported by an analysis of the 4930 entries in the Wordsmyth dictionary that have word-form frequencies of greater than 10 per million in the CELEX lexical database (Baayen, Piepenbrock, & Van Rijn, 1993). While only 7.4% of these word-forms correspond to more than one entry in the dictionary, 84% of the entries have more than one sense. Further, 37% of the entries have five or more senses. These figures show how common the systematic ambiguity between word senses is, compared with the accidental ambiguity between unrelated meanings. Therefore, when words are selected for an experiment as being ambiguous, without a distinction between word senses and word meanings, it is likely that a high proportion of these words will show the more common ambiguity between different word senses. Importantly, this explanation suggests that any experiment looking at ambiguity without explicitly making the distinction between word meanings and word senses is likely to be influenced by this bias.

Overall, these analyses suggest that multiple senses, and not multiple meanings, were crucial in producing the ambiguity advantage. In contrast, as described above, all current explanations of the ambiguity advantage assume that it is unrelated meanings that produce the processing benefit. We explore the potentially different effects of different types of ambiguity in the three experiments reported below.

EXPERIMENT 1

In this first experiment, lexical decisions for a large set of ambiguous and unambiguous words are analyzed using multiple regression analyses to determine the effects of multiple meanings, multiple senses, and meaning relatedness. To the extent that different effects emerge for these factors, this would provide the basis for further investigations.

Method

Participants. The participants were 25 members of the MRC Cognition and Brain Sciences Unit subject panel. All had English as their first language and had normal or corrected-to-normal vision.

Stimuli and design. One hundred twenty-four ambiguous words were selected to be included in the experiment. One hundred thirteen were taken from the Twilley, Dixon, Taylor, and Clark (1994) homograph norms. While most of the selected words had only two meanings, a few words with three meanings were included where the third meaning had a meaning probability in the Twilley et al. (1994) norms of less than 0.1. The remaining 11 ambiguous words did not appear in the norms, but were considered to have similar properties.

For half these ambiguous words, the two meanings correspond to separate entries in the Wordsmyth dictionary, and are therefore ambiguous between two meanings according to the criteria used by lexicographers. The remaining ambiguous words were judged to have more than one meaning by Twilley et al. (1994), but these meanings corresponded to a single entry in the Wordsmyth dictionary. For these words it is not clear whether their different interpretations should be classed as meanings or senses; their inclusion will allow us to look for effects of the relatedness of the meanings of ambiguous words.

Sixty unambiguous words were included in the experiment.[3] Only three of the unambiguous words had more than one entry in the Wordsmyth dictionary (*frog, bus, prayer*), the second entry for these three was considered to be sufficiently obscure that the words could be considered unambiguous.

All the stimuli were pretested for concreteness and familiarity, variables that are known to influence visual lexical decisions. Relatedness ratings and dominance ratings were also obtained for the ambiguous words. These ratings were made by four separate groups of participants who were either members of the MRC Cognition and Brain Sciences Unit subject panel or students at Cambridge University. Each of these ratings was made by a minimum of 24 participants. The three variables were rated on a 7-point scale as used by Gilhooly and Logie (1980).

For the concreteness ratings, participants made separate ratings for each of the two meanings of the ambiguous words. The words appeared in a rating booklet together with a word associate. This associate made it clear which meaning of the word was to be rated. For each word, an associate was selected for each of its two meanings (for example bark–dog and bark–tree). Word associates were also given for the unambiguous words to make the procedure consistent for the two word types. All the associates were taken from association norms (Twilley et al., 1994; Moss & Older, 1996).

For the relatedness ratings, raters were given each ambiguous word, together with short definitions of its two meanings, and asked to rate how related the two meanings were on a 7-point scale. Eleven additional ambiguous words that according to Azuma (1996) have highly related meanings were also included in the booklet. These were included to help participants use the whole range of ratings, as most of the ambiguous words had highly unrelated meanings.

The mean relatedness rating across all participants and items was 2.64. This low value reflects the fact that participants saw many of the pairs as completely unrelated; a rating of 1 was used more than any other rating. This is expected as these words were all selected to have meanings that are sufficiently unrelated that they should be considered separate meanings, and not senses within a single meaning. In all analyses of the response time data, rather than using the mean relatedness ratings, the inverse of these values was used; this made the measure more sensitive to small changes at the lower end of the scale.

The dominance scores were derived using the procedure used by Twilley et al. (1994); a group of 38 participants provided word associates for each word, and then the associates were classified in terms of which meaning they were related. The dominance score for each word was the proportion of valid responses that correspond to its dominant meaning. For example, a score of 1 would be given to a word with a highly dominant meaning, while a balanced homograph would have a score of 0.5.

The summary statistics of the stimuli are given in Table 2; the words themselves are listed in Appendix A. The nonword stimuli were pseudohomophones and had a similar distribution of lengths to the word stimuli. We decided to use pseudohomophones following the finding of Azuma and Van Orden (1997), who found a

---

[3] The imbalance between the number of unambiguous words and ambiguous words reflects the fact that this experiment was also designed to investigate the effects of the relative concreteness of the two meanings using a design similar to that used by Rubenstein et al. (1970). However the analyses of the differences between the groups showed no effects of relative concreteness, but simply revealed a main effect of concreteness and so are not reported here. The high proportion of ambiguous words also increases the set of words used in the analysis of meaning relatedness.

TABLE 2

Experiment 1: Descriptive Statistics for Stimuli

	Unambiguous	Ambiguous
N	60	124
Length	5.07	5.00
Log frequency	5.26	5.49
Familiarity	3.91	3.98
Word senses	2.83	7.43
Concreteness, meaning 1	5.13	5.45
Concreteness, meaning 2	—	3.91
Mean concreteness	5.13	4.68
Neighborhood	3.55	4.84

significant effect of ambiguity using pseudohomophone nonword foils but not when the foils were word-like nonwords.[4]

Procedure. All the stimulus items were pseudorandomly divided into four lists, such that each list contained approximately the same number of words of each stimulus type. Some items were then swapped between lists, to avoid having any ambiguous word occurring within the same list as an item that might bias participants toward one of its meanings. Participants were presented with the four lists in a pseudorandom order such that each possible order was seen by at least one participant. Within the lists, the order in which stimulus items were presented was randomized for each participant. All the participants saw all of the stimulus materials.

For each of the word and nonword stimuli, the participants were presented with a fixation point in the center of a computer screen for 500 ms, followed by the stimulus item. Their task was to decide whether each item was a word or a nonword; recognition was signaled with the dominant hand, nonrecognition with the other hand. As soon as the participant responded, the word was replaced with a new fixation point.

A practice run, consisting of 30 items not used in the analysis, was given to familiarize participants with the task. Each of the four lists was presented in a separate block of trials. Participants were given a short break after each block. Each block began with five stimuli not included in the analysis.

Results

The data from two participants were removed from the analysis because they had mean response latencies greater than 1000 ms. Individual responses longer than 1200 ms were also not included in the analysis; for the word data this meant that 1.2% of the data points were removed from the analysis. As recommended by Ratcliff (1993), all analyses were also performed on the inverse response times; for these analyses, all correct responses were included. These analyses are only reported where they differed in significance levels from the untransformed data.

Multiple Regression Analyses

The response time data for all 184 words were entered in a simultaneous multiple regression analysis. Ambiguity, number of word senses, word frequency, familiarity, length, lexical neighborhood, and mean concreteness were all included as predictors. A summary of the regression analysis can be seen in Table 3.

There are two crucial results in this analysis. First, there was a significant effect of ambiguity; ambiguous words were responded to more slowly than unambiguous words. Second, this ambiguity disadvantage was accompanied by a significant benefit for words that have many senses. All the other predictor variables except word length also accounted for unique variabil-

TABLE 3

Experiment 1: Summary of Regression Analysis for All Words

Predictor variable	Standardized coefficient	*t*
Ambiguity	.18	2.7**
Word senses	−.17	−2.4*
Frequency	−.29	−3.2**
Familiarity	−.26	−2.9**
Length	.13	1.6
Neighborhood	.34	4.3***
Concreteness	−.30	−5.0***

Note. $df = 177$, (*)$p < .1$, *$p < .05$, **$p < .01$, ***$p < .001$.

---

[4] We have repeated Experiment 1 using word-like nonwords. Consistent with the findings of Azuma and Van Orden (1997) and Pexman and Lupker (1999) we found a significant, but reduced, effect of ambiguity that was consistent with these results. The effect of relatedness was not significant.

TABLE 4

Experiment 1: Summary of Regression Analysis for Ambiguous Words

Predictor variable	Standardized coefficient	t
Frequency	−.45	−3.9***
Familiarity	−.10	−.8
Length	.18	1.9(*)
Neighborhood	.30	3.0**
Concreteness	−.21	−2.8**
Word senses	−.14	−1.7(*)
Relatedness	.17	2.3*
Dominance	−.10	−1.4

Note. $df = 117$, (*)$p < .1$, *$p < .05$, **$p < .01$, ***$p < .001$.

ity in response times (the effect of length was marginal in the analysis of the inverse response times ($p < .1$).

A second simultaneous multiple regression analysis was then carried out on the response times for the 124 ambiguous words in order to look for an effect of the relatedness of their meanings. Dominance scores were also entered in this analysis. A summary of the regression analysis can be seen in Table 4. As in the earlier multiple regression analysis of this data, frequency, lexical neighborhood and concreteness accounted for unique variability in response times. The effects of word length and number of word senses were marginal, and familiarity did not account for any unique variance in this smaller set of words. The effect of dominance was also not significant. Importantly, relatedness did account for unique variability in response times; ambiguous words were responded to faster when their meanings were judged to be semantically related.

Analyses of Variance

To provide further evidence for the effects of ambiguity and relatedness, ANOVA/ANCOVAs were also performed. From the set of 124 homographs, two sets of 27 homographs were selected, containing related and unrelated homographs respectively. They were selected by using only those homographs with a relatedness score of either less than 1.9 or greater than 3.4. A few homographs were then removed so that the two groups were matched for frequency, mean concreteness, length and familiarity. It is worth noting that even the homographs classified as having related meanings were not given particularly high relatedness ratings (see Appendix B); the mean relatedness score was 4.4 on the 7-point scale. Twenty-three of the 27 unrelated homographs had two meanings corresponding to separate Wordsmyth entries, for the related homographs only three words had two entries. This shows a high level of agreement between the relatedness judgements made by participants and the decisions of lexicographers about whether different usages of a word should be classed as separate dictionary entries. In a separate analysis, not reported here, we grouped the words according to whether they have one or two entries in Wordsmyth; this showed a very similar pattern of results to when the relatedness rating were used to classify the words.

A group of 43 nonhomographs was selected to be matched to the two homograph groups on frequency, concreteness, length, and familiarity. The properties of these words are given in Table 5; the words themselves are given in Appendix B. Although the groups were not matched in advance for neighborhood size (N; Colheart, Davelaar, Jonasson, & Besner, 1977) the words in the three groups did not significantly differ on this variable; $F_2(2,94) = 1.38$, $p = .2$.

The response times for these three groups of words were submitted to separate ANOVA/ANCOVA analyses, with items and participants as

TABLE 5

Experiment 1: Descriptive Statistics for Groups of Related and Unrelated Homographs

Group	N	Relatedness	Log frequency	Concreteness	Length	Familiarity	Neighborhood
Unrelated	27	1.37	5.46	4.85	5.04	4.03	5.04
Related	27	4.39	5.44	4.80	5.00	3.94	4.52
Nonhomographs	43		5.43	4.84	5.02	3.99	3.16

the random variables. The mean response times are given in Table 6.

In the participants analysis, the effect of group was significant; $F_1(2,44) = 4.79$, $p < .05$. In the items analysis, using the log-transformed frequency, familiarity, mean concreteness, and length as covariates, the effect of group was marginal; $F_2(2,90) = 2.88$, $p < .07$. Multiple comparisons were made between the individual groups, using the Newman–Keuls procedure. Responses to the group of nonhomographs were faster than to the group of homographs with unrelated meanings; this difference was significant in the participants analysis and marginal in the items analysis; $q_1(3,44) = 4.15$, $p < .05$, $q_2(3,90) = 3.22$, $p < .07$. The related homographs were significantly faster than the unrelated homographs in the participants analysis; again, the difference was marginal in the items analysis; $q_1(2,44) = 3.28$, $p < .05$, $q_2(2,90) = 2.53$, $p < .08$. The difference between the nonhomographs and the homographs with related meanings was nonsignificant in both analyses ($p > .5$). The error data showed no significant effect of group in either analysis; $F_1(2,44) < 1$, $F_2(2,92) < 1$.

These results confirm the findings of the regression analysis; homographs with related meanings are responded to more quickly than homographs with highly unrelated meanings. Further, they show that homographs are responded to more slowly than matched nonhomographs only when their meanings are judged to be unrelated.

Discussion

Three interesting results have emerged from this experiment. First, the analysis of the response times for this set of 184 words showed a significant ambiguity disadvantage; words with one meaning were responded to significantly faster than words with two meanings. This is in contrast with previous reports of an ambiguity advantage, and suggests that recognition of ambiguous words is delayed by competition between their different meanings. Second, this disadvantage for multiple meanings was accompanied by an advantage for words with multiple senses. This confirms our suggestion that previous reports of an ambiguity advantage should be interpreted as an advantage for multiple senses rather than multiple meanings. Finally, the significant effect of relatedness shows that the disadvantage for semantic ambiguity is modulated by meaning relatedness, such that it is maximal when the different meanings of the word are semantically unrelated; this replicates the effect of relatedness seen by Azuma and Van Orden (1997). The implications of this result will be discussed in the General Discussion.

TABLE 6

Experiment 1: Mean Lexical Decision Times, Analysis Using Relatedness Ratings

		RT (ms)		
Ambiguity	Relatedness	Mean	SD	Error (%)
Unambiguous		556	133	4.25
Ambiguous	Unrelated	577	136	4.83
Ambiguous	Related	561	134	3.54
Nonwords		636	155	7.14

EXPERIMENT 2

Experiment 1 suggests that the two types of semantic ambiguity have very different effects on lexical decision performance. While multiple meanings delay recognition, multiple senses produce a processing benefit. This result is clearly controversial; all existing models of the ambiguity advantage have assumed that multiple meanings produce faster visual lexical decisions. Experiment 2 attempts to replicate the contrasting effects of ambiguity seen in the multiple regression analysis of Experiment 1, using a factorial design to directly compare the effects of lexical ambiguity and multiple word senses.

Method

Participants. The participants were 25 members of the MRC Cognition and Brain Sciences Unit subject panel. All had English as their first language, and had normal or corrected-to-normal vision.

Stimuli and design. The word stimuli were selected to conform to a 2 × 2 factorial design where the two factors were ambiguity and number of senses. Words were classed as being unambiguous if they had only one entry in the Wordsmyth Dictionary (Parks et al., 1998) and as ambiguous if they had two or more entries. Two measures of the number of senses were used. These were the total number of word senses listed in the Wordsmyth dictionary for all the entries for that word, and the total number of senses given in the WordNet lexical database (Fellbaum, 1998).

Thirty-two words were selected to fill each cell of the factorial design such that the number of word meanings was matched across each level of number of word senses, and the total number of word senses was matched across each level of the number of word meanings. Therefore, unlike Experiment 1, the numbers of ambiguous and unambiguous words used in this experiment were equal. Of the words used in this experiment, 16% were also used in Experiment 1.

The four groups of words were matched for CELEX frequency (log- transformed), length (number of letters), number of syllables, concreteness and familiarity. Concreteness scores were obtained from a rating pretest in which the words were rated on a 7-point scale by 25 participants who were members of the MRC Cognition and Brain Sciences Unit subject panel and who did not participate in the lexical decision experiment. The familiarity ratings were made on a similar 7-point scale by 23 participants from the same population. The groups were not explicitly matched for neighborhood size; however, the number of words in CELEX that differed from each word by only one letter (*N*; Coltheart et al., 1977) did not differ significantly between the groups ($p > .3$).

The properties of the words are summarized in Table 7; the words themselves are listed in Appendix B. The nonword stimuli were pseudohomophones. They were chosen to be as wordlike as possible and had a similar distribution of lengths to the word stimuli.

Procedure. The stimulus items were pseudorandomly divided into four lists; each list contained approximately the same number of words from each stimulus group. Some items were then swapped between lists, to avoid having any ambiguous word occurring within the same list as an item that might bias participants toward one of its meanings. Participants were presented with the four lists in a random order. Within the lists, the order in which stimulus items were presented was also randomized for each participant. All participants saw all of the stimulus materials.

For each of the word and nonword stimuli, the participants were presented with a fixation point in the centre of a computer screen for 500 ms, followed by the stimulus item. As soon as the participant responded, the word was replaced with a new fixation point. Participants were told to decide whether each string of letters was a real English word, and to respond as quickly as possible without making mistakes. Real words were signaled with the dominant hand, nonwords with the other hand.

TABLE 7

Experiment 2: Descriptive Statistics for Stimuli

	Ambiguous few senses	Ambiguous many senses	Unambiguous few senses	Unambiguous many senses
Example	Pupil	Slip	Cage	Mask
N	32	32	32	32
Wordsmyth meanings	2.03	2.09	1.00	1.00
Wordsmyth senses	5.19	14.22	5.25	14.41
WordNet senses	4.88	11.84	5.00	11.19
Frequency	5.40	5.43	5.43	5.50
Concreteness	5.19	5.07	5.06	5.05
Familiarity	4.11	4.24	4.17	4.24
Length	4.47	4.41	4.47	4.53
Syllables	1.19	1.09	1.16	1.09
Neighbors	6.03	7.78	5.91	6.25

A practice session, consisting of 64 items not used in the analysis, was given to familiarize participants with the task. Each of the four lists was presented in a separate block of trials. Participants were given a short break after each block. Each block began with 10 stimuli not included in the analysis.

Results

The data from two participants were removed from the analysis because of error rates of greater than 10%. Incorrect responses were not included in the analysis. The overall error rate for responses was 3.6% (ranging from 0.8 to 7.7% for each participant). Responses longer than 1200 ms were also not included in the analysis; for the word data this meant that 1.1% of the data points were removed from the analysis. As with Experiment 1, all analyses were also performed on the inverse response times; for these analyses all correct responses were included. These analyses are reported only when they differ in significance from the analysis of the untransformed data.

Mean values were calculated separately across participants and items. The participant means were subjected to ANOVA, and the item means were subjected to ANCOVA. The mean response times are given in Table 8.

The ANCOVA revealed significant effects of familiarity ($F(1,121) = 11.4$, $p < .001$) and marginal effects of length ($F(1,121) = 3.65$, $p < .06$) and frequency ($F(1,121) = 2.72$, $p = .1$). The effects of concreteness and neighborhood were nonsignificant ($p > .5$), and so these variables were removed from the ANCOVA.

TABLE 8

Experiment 2: Mean Response Times (RT) and Percentage Error Rates

		RT (ms)		
Ambiguity	Senses	Mean	SD	Error (%)
Ambiguous	Few	587	143	4.08
Ambiguous	Many	578	135	1.77
Unambiguous	Few	586	141	2.99
Unambiguous	Many	567	129	1.63
Nonwords		659	143	3.92

The main effect of the number of senses was significant ($F_1(1,22) = 14.6$, $p < .001$; $F_2(1,121) = 4.42$, $p < .05$); words with many senses were responded to 14 ms faster than words with few senses. Ambiguous words were responded to 6 ms slower than unambiguous words, although this effect of ambiguity was not significant in either analysis ($F_1(1,22) = 2.9$, $p > .1$, $F_2(1,121) = 1.3$, $p > .2$). In the analysis of inverse response times, the effect of ambiguity was marginal in the participants analysis ($F_1(1,22) = 3.8$, $p < .07$), but was again not significant in the items analysis ($F_2(1,121) = 1.5$, $p > .2$). There was no significant interaction between these two variables in either analysis ($p > .2$).

The error data also showed a significant effect of the number of senses; fewer errors were made to words with many senses ($F_1(1,22) = 12.2$, $p < .005$; $F_2(1,121) = 5.19$, $p < .05$). In the error data neither the effect of ambiguity nor the interaction between the two variables reached significance (all $p > .4$).

Discussion

This experiment shows that words with many senses are responded to faster and with fewer errors than words with few senses. This replicates the significant word senses benefit seen in Experiment 1. This advantage for multiple senses was shown alongside a small disadvantage for words with multiple meanings. Although this difference was not significant, there was no indication of the kind of advantage for words with multiple meanings that has previously been reported.

EXPERIMENT 3

Experiments 1 and 2 have shown that word senses and word meanings have very different effects on visual lexical decisions. Ambiguity between multiple meanings produces a disadvantage, while multiple senses produce faster responses. This experiment investigates whether these contrasting effects of ambiguity are also present in the auditory domain, using the same factorial design as Experiment 2.

We argue that the ambiguity effects seen in Experiments 1 and 2 reflect the influence of amodal semantic representations on visual word recognition. If this is the case, then it is of interest to see whether the same pattern emerges in the auditory domain. It is possible that semantic information plays a similar role across the two domains, but it is also possible that the temporal characteristics of speech may reduce the role of semantic information on spoken word recognition compared with visual word recognition. However, we do know that the semantic information relating to spoken words is accessed early on, prior to the word becoming unique (Marslen-Wilson, 1987; Zwitserlood, 1989), and this makes it at least possible that auditory lexical decision will show an influence of semantic ambiguity.

A study by Holley-Wilcox (1977) (cited in & McCusker, Hillinger, Bias, 1981) supports this idea that it is possible to detect effects of semantic ambiguity using auditory lexical decision. They found that auditory lexical decisions were significantly slower for homophones like *plane* and *plain*, which although sharing the same phonology are spelled differently, than for nonhomophones. This result can be explained by assuming that competition between the different meanings of the homophones is slowing the recognition. This is consistent with the ambiguity disadvantage seen the visual domain in Experiment 1 and suggests that competition between the different meanings of ambiguous words does play a role in the auditory domain.

However, a possible problem with using homophones that are not homographs to look for semantic ambiguity effects is that there may be interference between their different orthographic representations. Although it may seem unlikely that interference between orthographic representations should affect an auditory task, this idea is supported by Ziegler and Ferrand (1998). They found slower auditory lexical decisions for words whose rimes could be spelled in more than one way (e.g., *sleep*). This raises the possibility that orthographic interference may have contributed to the finding of Holley-Wilcox (1977) and makes it preferable to avoid such items in any experiment designed to show the effects of semantic ambiguity. Therefore, Experiment 3 uses the same ambiguous words as the visual experiment, words like *bark* that share both orthography and phonology, and differ only in their meanings.

As well as allowing us to investigate the pattern of ambiguity in the auditory domain, this experiment also allows us to check whether the pseudohomophone nonwords used in Experiment 2 were crucial to obtaining the observed pattern of results. It is not yet clear how these nonwords affect lexical processing, and so for us to argue that these ambiguity effects have important implications for models of word recognition, they should be shown in the absence of pseudohomophones.

However, the primary aim of this experiment is to try and replicate the ambiguity disadvantage, which was significant in Experiment 1, but not in Experiment 2.

Method

Participants. The participants were 26 students at Cambridge University. All had English as their first language.

Stimuli and design. The word stimuli were selected to conform to the same 2 × 2 factorial design as in Experiment 2. Seventy-seven percent of the words selected were also used in Experiment 2. Twenty-three words were selected to fill each cell of the factorial design; the number of words in each cell is smaller than that used in Experiment 2 because of the additional phonological constraints used to match the groups.

The four groups of words were matched for CELEX frequency (log-transformed), number of phonemes, the phoneme at which the word becomes unique, actual length of the words in ms, concreteness and familiarity. Concreteness and familiarity scores were taken from the pretest described in Experiment 2. All words had only one syllable.

The properties of the words are summarized in Table 9; the words themselves are listed in Appendix C. The nonword stimuli were created to be as word-like as possible, and had a similar distribution of lengths to the word stimuli.

Procedure. The organisation of the stimuli into four blocks of trials followed the same procedure as Experiment 2. The onset of each item was 1000 ms after the participants' response to the preceding item. If the participant did not re-

TABLE 9

Experiment 3: Descriptive Statistics for Stimuli

	Ambiguous few senses	Ambiguous many senses	Unambiguous few senses	Unambiguous many senses
N	23	23	23	23
Wordsmyth meanings	2.04	2.13	1.00	1.00
Wordsymth senses	5.43	13.61	3.59	14.00
WordNet senses	5.00	11.43	4.43	10.17
Frequency	5.30	5.34	5.42	5.43
Concreteness	5.11	5.01	5.08	4.99
Familiarity	4.17	4.30	4.31	4.33
Length	610	602	601	605
Phonemes	3.43	3.43	3.52	3.56
Uniqueness	3.70	3.87	3.78	3.74

spond within 3000 ms of the onset of a word, the next item was presented. Participants were told to decide whether each sound was a real English word and to respond as quickly as possible without making mistakes. Real words were signaled with the dominant hand, nonwords with the other hand.

Results

The data from four participants were removed from the analysis because of error rates of greater than 10%. Incorrect responses were not included in the analysis. The overall error rate for responses was 5.8%. Responses longer than 1500 ms were also not included in the analysis; for the word data this meant that 2% of the data points were removed from the analysis. As with Experiments 1 and 2, all analyses were also performed on the inverse response times; for these analyses all correct responses were included. These analyses did not differ in significance levels from the untransformed data and so are not reported.

Mean values were calculated separately across participants and items. The participant means were subjected to an ANOVA, and the item means were subjected to an ANCOVA. The mean response times are given in Table 10.

The ANCOVA revealed significant effects of familiarity ($F(1,86) = 4.6$, $p < 0.05$) and length ($F(1,86) = 236$, $p < 0.001$). Concreteness, frequency, number of phonemes, and uniqueness point were not significant predictors of response times ($p > .2$); these variables were not included in the ANCOVA.

There was a significant effect of the number of senses ($F_1(1,21) = 20.7$, $p < .001$; $F_2(1,86) = 6.6$, $p < .05$). Words with many senses were responded to 33 msec faster than words with few senses. The effect of ambiguity was also significant ($F_1(1,21) = 22.4$, $p < .001$; $F_2(1,86) = 4.7$, $p < .005$). Ambiguous words were responded to 29 msec slower than unambiguous words. The interaction between these two variables was significant in the subjects analysis but not in the items analysis $F_1(1,21) = 16.5$, $p < .001$; $F_2(1,86) = 2.3$, $p > .1$.

The error data showed a similar pattern of results to the response time data. Fewer errors were made to words with many senses, although this difference was only significant in the subjects analysis and marginal in the items analysis; ($F_1(1,21) = 10.5$, $p < .005$; $F_2(1,86) = 2.7$, $p < .1$). Fewer errors were also made to unambiguous words compared with ambiguous words, although this difference was only marginal in the subjects analysis and did not approach signifi-

TABLE 10

Experiment 3: Mean Response Times (RT) and Percentage Error Rates

		RT (ms)		
Ambiguity	Senses	Mean	SD	Error (%)
Ambiguous	Few	986	176	8.3
Ambiguous	Many	935	174	4.3
Unambiguous	Few	939	167	5.7
Unambiguous	Many	924	186	3.6
Nonwords		1031	173	6.0

cance in the items analysis; ($F_1(1,21) = 4.2$, $p < .06$; $F_2(1,86) = 0.7$, $p > .4$). The interaction between the two variables was not significant in either analysis ($p > .4$).

Discussion

This experiment has shown that the pattern of ambiguity effects in the auditory domain is essentially the same as in the visual domain; an advantage for words with many senses coexists with a disadvantage for words with multiple meanings. A second important feature of this experiment is that it shows the effects of multiple senses and multiple meanings without the use of pseudohomophones. This demonstrates that these nonwords are not necessary to see the pattern of results seen in Experiment 2 and suggests that the ambiguity effects we have demonstrated are pervasive in word recognition.

GENERAL DISCUSSION

The results of these three experiments represent an important challenge to accepted views of how semantic ambiguity affects recognition of isolated words. Previous reports of an ambiguity advantage have been interpreted as showing that there is a processing advantage for words that have multiple meanings. A range of models has been put forward to explain how this advantage might arise.

Our analyses of the stimuli used in three of the clearest demonstrations of this effect suggested to us that the accepted interpretation might be incorrect and that related word senses and not unrelated meanings might be responsible for this processing advantage. The results of the three experiments reported here support this view. In all three experiments we found a significant benefit for words that have many senses, compared with words with few senses. In contrast, ambiguity between unrelated meanings consistently produced a processing disadvantage; this ambiguity disadvantage was significant in Experiments 1 and 3. We now consider the implications for models of word recognition.

The Ambiguity Disadvantage

We have already discussed how models of word recognition have tried to explain the apparent advantage for words with multiple meanings, but our data suggest that they must accommodate exactly the reverse effect. This challenge is less problematic than might be expected.

The ambiguity disadvantage is a natural prediction of models in which words compete for the activation of semantic representations (Gaskell & Marslen-Wilson, 1997; Hinton & Shallice, 1991; Joordens & Besner, 1994; Plaut, 1997; Plaut & Shallice, 1993). As discussed earlier, in these models interference between the different meanings of ambiguous words would delay their recognition relative to an unambiguous word. As noted by Joordens & Besner (1994), an ambiguity advantage can be produced by these models only if an additional mechanism is present to overcome this semantic competition. Our results suggest that no such mechanism is required.

The ambiguity disadvantage reported here removes a major hurdle for models in which words compete to activate distributed semantic representations. The ambiguity disadvantage naturally emerges from the semantic competition present in such models and has been shown in a model of this type where a simple least mean square algorithm was used to learn the mapping between distributed orthographic and semantic representations (Rodd, 2000).

This new pattern of results can also be accommodated by those models in which words compete to activate abstract word nodes within a lexical network. Earlier, we discussed how these models could produce an ambiguity advantage by assuming either that ambiguous words are more efficient at inhibiting competitors, or that they benefit from having multiple competitors in the race for recognition. Interestingly, these models can just as easily accommodate a disadvantage for words with multiple meanings. As in all experiments of this type, the ambiguous words and unambiguous words were matched on total frequency. This means that the frequency of each meaning of the ambiguous words is on average half that of the unambiguous word. This difference in the frequency of the word meanings could produce faster lexical decisions for the unambiguous words. Further, if lateral inhibition were present between all word nodes within the lexical network, including the nodes corresponding to the different meanings of an ambiguous word, this would act to slow the recognition of ambiguous words.

Therefore, it appears that both classes of models considered here can accommodate the ambiguity disadvantage. The question that remains is whether the ambiguity disadvantage should simply be explained in terms of an effect of frequency of word meanings, or whether we can claim that it provides evidence of competition between the different meanings of words. We suggested earlier that looking for an effect of meaning relatedness might help us to determine the mechanisms by which any observed ambiguity effects might arise.

In Experiment 1 we found that, at least in the visual domain, the ambiguity disadvantage is modulated by the relatedness of the two meanings of the ambiguous words; within the words that we classified as ambiguous between multiple meanings, there was a benefit for those words whose meanings were moderately related. This suggests that the ambiguity disadvantage cannot be explained solely as the results of a frequency bias; this account cannot allow semantic factors to modulate the size of the ambiguity disadvantage. Similarly, the relatedness effect suggests that the effect cannot be explained entirely as resulting from lateral inhibition between abstract word nodes; if the effect was entirely presemantic, there would be no mechanism by which the semantic relationship between the two meanings of a word could play a role.

The only way to explain the relatedness effect in a nonsemantic way is to assume that a sufficient number of the words that we classed as ambiguous between different meanings, were in fact ambiguous between multiple senses; we think that this is unlikely. Therefore, we believe that the modulation of the ambiguity disadvantage by meaning relatedness is evidence of the active involvement of semantic representations in the process of lexical competition. This is consistent with models of word recognition in which words compete to activate semantic representations (Gaskell & Marslen-Wilson, 1997; Hinton & Shallice, 1991; Joordens & Besner, 1994; Plaut, 1997; Plaut & Shallice, 1993). In these models, the ambiguity disadvantage arises because of the difficulty in mapping a single orthographic or phonological pattern of activation to multiple patterns of semantic activation. The different possible semantic patterns interfere with each other, and the additional time that it takes for this competition to be resolved produces the ambiguity disadvantage. If the ambiguity disadvantage is indeed caused by this interference between competing semantic patterns, then we would expect to see an effect of meaning relatedness. The level of interference is related to the degree of overlap between the two patterns, such that any semantic features shared by the two patterns will reduce the interference. As with the ambiguity disadvantage, this relatedness effect has been simulated in a model where a simple least mean square algorithm was used to learn the mapping between distributed orthographic and semantic representations (Rodd, 2000).

In summary, the ambiguity disadvantage reported here, together with the relatedness effect, can most easily be interpreted as evidence that competition to activate a distributed semantic representation is a fundamental part of the word recognition process.

The Sense Advantage

All three experiments reported here found that lexical decision times are faster for words with many dictionary senses than for words with only a few senses. This result is somewhat counterintuitive. Given that ambiguity between multiple meanings produces a processing disadvantage, why should ambiguity between multiple senses produce the reverse effect? If we accept that the ambiguity disadvantage reflects interference between the different meanings of ambiguous words, then although the interference between different senses would be reduced relative to words with multiple meanings, this interference would surely slow recognition relative to unambiguous words. The result is equally problematic for models in which words compete to activate abstract word nodes. If we assume that different word senses correspond to different word nodes, then we would expect multiple senses to delay recognition in the same way as multiple meanings. Alternatively, if we assume that multiple senses correspond to a single lexical node, we would expect them to be recognized as quickly but not faster than unambiguous words.

This idea that multiple senses might be expected to show a similar, but possibly reduced, ambiguity disadvantage is reinforced by recent evidence from Klein and Murphy (2001). They embedded polysemous words in two different phrases which biased the reader interpretation to either the same or different senses of the word; they found that sense consistency aided both memory and comprehension. From this they concluded that there are separate representations for the multiple senses of polysemous words. If the representations of the meanings of different word senses are sufficiently independent to produce this pattern of results, then we would expect the interaction between multiple senses to delay recognition in a similar, although reduced, way to that seen for words with multiple meanings. This suggests that an additional mechanism is necessary to explain the word sense advantage, and that this mechanism would need to be sufficiently strong to overcome any effects of semantic competition between different word senses.

One possible explanation of the word sense benefit is that words with many senses and words with few senses differ in the amount of semantic information contained in their semantic representation. In other words, a word with many senses can be considered to be semantically rich. This is essentially the same argument that Plaut and Shallice (1993) put forward to account for the processing benefit of concrete words over abstract words. In their computational account of the concreteness effect, the difference between abstract and concrete words is reflected in the number of semantic features in a distributed semantic representation; abstract words were given fewer semantic features than concrete words. These additional features produce more stable representations, which in turn lead to faster settling times for words with more semantic features. Such an account of the words sense effect would need to assume that this benefit for semantic richness is sufficient to overcome any disadvantage caused by the ambiguity of these words.

This is related to the Schwanenflugel, Harnishfeger, and Stowe (1988) *context availability* account of the concreteness effect. They claim that contextual information about words is necessary for the integration that occurs in comprehension, and that concrete words are processed more easily because of the ease with which contextual information can be accessed. Schwanenflugel et al. (1988) report evidence that concreteness has an effect on lexical decision time only when it is confounded with contextual availability; when contextual availability, frequency, and familiarity were partialled out, concreteness did not significantly predict response times, whereas contextual availability accounted for a significant proportion of the variance independent of frequency and familiarity. It is possible that contextual information may be more readily available for words that have many senses and which are used in a wider range of contexts.

A third possible explanation of the sense effect is that it is a direct result of using a task in which words are presented in isolation without a sentential or semantic context. As already stated, words with many senses can be used in a wide variety of contexts, and it is possible that this experience results in the development of a relatively context independent representation of the word. On the other hand, words with few senses are used in a far more restricted range of contexts and so may develop representations whose access is more dependent on the word appearing in the appropriate context. This difference in the extent to which the lexical representation of the words are context independent may be important when participants are asked to recognize the words without any context. Presumably this task is more difficult for the words with fewer senses whose representations are more context dependent.

A final possible explanation is that the word sense benefit reflects differences in the attractor basins that develop within a distributed semantic network. As noted by Kawamoto (1993), the different senses of a word correspond to a set of highly correlated patterns of semantic activation, and these senses will collectively create a broad, shallow basin of attraction, containing more than one stable state, corresponding to each different sense. It is plausible that under certain conditions, settling into the correct attractor may be quicker for such a broad attractors, compared with the steep, narrow attractor basins that would develop for words with only few senses. This could potentially explain the opposite effects of the two types of ambiguity; while multiple meanings correspond to separate attractor basins, multiple senses correspond to multiple stable states within a single, broad attractor basin. This hypothesis needs to be assessed by performing the appropriate network simulations to determine the conditions under which such a pattern of effects might arise.

Further, this hypothesis can be extended to suggest that the word sense benefit might emerge only in lexical decision, and not in a range of other tasks. In lexical decision, participants may be able to respond correctly on the basis of the general familiarity produced by accessing a representation of the word's core meaning, and it is not necessary for them to disambiguate between a word's different senses. In terms of attractor structure, it is possible that the response is made as soon as the pattern of semantic activation has entered the broad attractor basin corresponding to the word's meaning, but before the activation has settled into a stable state corresponding to a particular sense of a word. This explanation of the word sense benefit predicts that, if we look for an effect of word senses on a task which requires the retrieval of a particular sense of the word, then the sense benefit should be eliminated and possibly reversed due to the need to disambiguate between individual word senses (cf. Klein & Murphy, 2001).

CONCLUSION

The data reported here require us to reconsider how semantic ambiguity affects the recognition of isolated words. While we do not consider these data to be inconsistent with the existing ambiguity advantage data, they do contradict the accepted interpretation of these data. We have found that multiple meanings delay recognition, while multiple senses produce a processing advantage. We therefore claim that the ambiguity advantage reported in previous studies should be interpreted as showing a benefit for words with many senses; this is consistent both with the data reported here and our analysis of the stimuli from previous studies.

Our claim that ambiguity between multiple meanings can slow lexical decisions is entirely new. Yet it can apparently be incorporated into most current models of word recognition by assuming that there is competition between the different meanings of ambiguous words or that there is an advantage for the more frequent meanings of unambiguous words compared with ambiguous words matched on overall frequency. However, the finding that the ambiguity disadvantage is modulated by meaning relatedness suggests that the effect must, at least in part, be due to competition to activate a distributed semantic representation. It is less clear how the word sense benefit should be interpreted, and further work is required to determine the cause of this intriguing effect. Nonetheless, the overriding message from this series of experiments is that the word recognition process is intimately tied in with the competitive process by which the stored meanings of words are retrieved.

APPENDIX A

Experiment 1 Stimuli

Ambiguous words				Unambiguous words	
admit	advance	affair	arms	bus	fee
article	badger	bark	batter	baby	fun
blind	bonnet	bowl	boxer	clay	sane
bridge	broke	bulb	cabinet	coal	grow
calf	can	cane	case	frog	seek
chance	charm	chest	china	goat	item
clog	company	craft	cricket	lung	task
deed	degree	dense	digit	hill	vote
dry	express	feet	fence	tent	warn
firm	fling	free	glare	lake	poet
glass	grain	hamper	horn	tiger	alone
interest	jumper	kid	kind	apple	fraud
last	late	lean	left	bible	grief
letter	lie	like	limp	brain	dozen
lobby	marble	march	maroon	cider	unite
might	nail	net	novel	cigar	urban
odd	organ	palm	panel	glove	thief
park	patient	peer	picket	hotel	throw
pine	pitcher	poach	poker	lorry	amuse
pole	pride	pupil	ram	metal	brutal
rare	rate	reflect	refrain	ocean	misery
ruler	sack	safe	sage	river	prayer
scrap	screen	seal	season	cattle	terror
second	sense	sentence	shed	forest	winter
sign	spade	speaker	spell	weapon	dollar
stable	staff	stag	stalk	rabbit	travel
stamp	staple	static	stern	throat	destroy
store	strand	straw	swallow	custard	kingdom
swear	temple	tend	tense	diamond	citizen
term	toast	trial	trust		
uniform	vent	watch	yard		

APPENDIX B

Experiment 2 Stimulus Groups

Ambiguous		Unambiguous	
Few senses	Many senses	Few senses	Many senses
ash	angle	ant	belt
calf	bark	bandage	bend
chap	blow	bet	bite
cricket	boil	bone	burn
cuff	bowl	bulk	dip
fleet	bust	cage	drain
fudge	clip	cake	feather
hide	clutch	carton	flash
lime	compound	crew	grip
loaf	duck	crude	hammer
loom	flush	deaf	hang
mint	fold	farm	hook
mole	gag	feast	load
novel	gum	foam	loop

Ambiguous		Unambiguous	
Few senses	Many senses	Few senses	Many senses
page	hail	harsh	mask
pen	jam	heap	nest
pine	jar	hinge	pinch
poach	lap	hurdle	roll
port	lean	join	saddle
prune	lock	lump	scan
pupil	pitch	path	shade
rare	scale	profit	slice
rash	seal	request	slide
rifle	slip	rust	smash
stable	spell	silk	sour
stern	stall	slim	spin
stunt	stem	slot	steam
tend	strain	snake	sway
tense	strand	soap	thread
toast	stud	spy	tread
utter	swallow	stain	whip
yard	tap	trot	wire

APPENDIX C

Experiment 3 Stimulus Groups

Ambiguous		Unambiguous	
Many senses	Few senses	Many senses	Few senses
bark	calf	belt	ant
boil	chap	bite	bone
clutch	fleet	bounce	crude
duck	fudge	chill	farm
fit	hide	dip	feast
flush	loaf	drain	fog
fly	mint	hook	grin
fold	mole	kick	growl
fret	page	loop	guess
gag	pen	mask	harsh
gum	pine	nest	hinge
hail	poach	shade	loud
jam	port	slide	rust
jar	prune	smash	shirt
lean	rare	snap	silk
lock	rash	soak	sip
seal	sage	spin	slot
slip	stern	steam	snow
spell	stunt	sway	soap
stall	tend	thread	spy
stem	tense	tread	stain
stick	toast	wheel	task
stud	yard	wire	winch

REFERENCES

Azuma, T. (1996). Familiarity and relatedness of word meanings: Ratings for 110 homographs. *Behavior Research Methods, Instruments and Computers*, **28,** 109–124.

Azuma, T., & Van Orden, G. C. (1997). Why safe is better than fast: The relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language*, **36,** 484–504.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database* [CD-ROM] Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Borowsky, R., & Masson, M. E. J. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning Memory and Cognition*, **22,** 63–85.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1997). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

Cruse, A. D. (1986). *Lexical semantics*. Cambridge, England: Cambridge University Press.

Fellbaum, C. (Ed.) (1998). *Wordnet: An electronic lexical database*. Cambridge, MA: MIT Press.

Fera, P., Joordens, S., Balota, D. A., Ferraro, F. R., & Besner, D. (1992). *Ambiguity in meaning and phonology: Effects on naming*. Paper presented at the 33rd annual meeting of the Psychonomic Society.

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, **12,** 613–656.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, **113,** 254–281.

Gilhooly, K., & Logie, R. (1980). Methods and designs: Age of acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, **12,** 395–427.

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming—An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, **22,** 1331–1356.

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, **98,** 74–95.

Holley-Wilcox, P. (1977). *The effect of homophony with auditory presentation of stimuli*. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, May 1977.

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology*, **13,** 278–305.

Joordens, S., & Besner, D. (1994). When banking on meaning is not (yet) money in the bank—Explorations in connectionist modeling. *Journal of Experimental Psychology: Learning Memory and Cognition*, **20,** 1051–1062.

Kawamoto, A. H. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language*, **32,** 474–516.

Kawamoto, A. H., Farrar, W. T., & Kello, C. T. (1994). When two meanings are better than one: Modeling the ambiguity advantage using a recurrent distributed network. *Journal of Experimental Psychology: Human Perception and Performance*, **20,** 1233–1247.

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the timecourse of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **14,** 601–609.

Kilgarriff, A. (1992). *Polysemy*. Unpublished doctoral dissertation, University of Sussex.

Klein, D. E., & Murphy, G. L. (2001). The representation of polysemous words. *Journal of Memory and Language*, **45,** 259–282.

Lyons, J. (1977). *Semantics*. Cambridge, England: Cambridge University Press.

Lyons, J. (1981). *Language and linguistics*. Cambridge, England: Cambridge University Press.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition*, **25,** 71–102.

McCusker, L. X., Hillinger, M. L., & Bias, R. G. (1981). Phonological recoding and reading. *Psychological Bulletin*, **89,** 217–245.

Millis, M. L., & Button, S. B. (1989). The effect of polysemy on lexical decision time: now you see it, now you don't. *Memory & Cognition*, **17,** 141–147.

Moss, H., & Older, L. (1996). *Birkbeck word association norms*. Hove: Psychology Press.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension—Effects of frequency of meaning and contextual bias. *Memory & Cognition*, **9,** 1–39.

Parks, R., Ray, J., & Bland, S. (1998). *Wordsmyth English dictionary—Thesaurus*. [ONLINE]. Available: http://www.wordsmyth.net/ [1999, February 1], University of Chicago.

Pexman, P. M., & Lupker, S. J. (1999). Ambiguity and visual word recognition: Can feedback explain both homophone and polysemy effects? *Canadian Journal of Experimental Psychology*, **53,** 323–334.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, **12,** 765–805.

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, **10,** 377–500.

Ratcliff, R. (1993). Methods for dealing with reaction-time outliers. *Psychological Bulletin*, **114,** 510–532.

Rayner, K. (1998). Eye movements in reading and information processing: 20 Years of research. *Psychological Bulletin*, **124,** 372–422.

Rodd, J. M. (2000). *Semantic representation and lexical competition: Evidence from ambiguity*. Unpublished doctoral dissertation, University of Cambridge.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, **9,** 487–494.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decision for abstract and concrete words. *Journal of Memory and Language*, **27,** 499–520.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, **18,** 645–659.

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory & Cognition*, **22,** 111–126.

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, **5,** 683–689.

Zwitserlood, P. (1989). The locus of effects of sentential–semantic context in spoken-word processing. *Cognition*, **32,** 25–64.

(Received October 27, 1999)
(Revision received March 7, 2001)