Journal of Memory and Language **45,** 259–282 (2001)
doi:10.1006/jmla.2001.2779, available online at http://www.academicpress.com on IDEAL®

The Representation of Polysemous Words

Devorah E. Klein and Gregory L. Murphy

Department of Psychology and Beckman Institute, University of Illinois at Urbana-Champaign

Words that have a number of related senses are polysemous. For example, *paper* refers to both a substance and a publication printed on that substance. Five experiments investigated whether different senses are represented distinctly in the lexicon or if there is a common, core meaning. In all experiments, a polysemous word was used twice, in phrases that selected the same or different senses. Experiment 1 showed that sense consistency aided memory for the polysemous word. Experiment 2 extended this result to a timed sensicality judgment task. Experiment 3 demonstrated that the effects for polysemous words were very similar to those for homonyms. Experiment 4 ruled out the possibility of modifier–modifier priming. Experiment 5 showed that sense consistency facilitates comprehension relative to a neutral baseline, while sense inconsistency inhibits comprehension. These experiments provide evidence that polysemous words have separate representations for each sense and that any core meaning is minimal.

Key Words: polysemy; lexical semantics; psycholinguistics; ambiguity.

Much research in lexical representation has compared homonyms with unambiguous words. Homonyms usually arise through a historical accident in which two different word meanings converge on the same phonological representation, or in which a single word diverges into very different meanings. A typical example is *bank*, with the unrelated meanings of a financial institution and the side of a river. Most words do not have such unrelated meanings, but linguists and psycholinguists studying lexical meaning have nonetheless identified a wide range of meanings within individual unambiguous words. For example, the word *paper* may refer to a substance made out of wood pulp, a blank sheet of that substance, a daily publication, or an article that is printed on that substance. The meaning of *paper* has been extended so far that it is now possible for students to turn in a paper by handing in a disk or sending a file electronically, so that wood pulp is in no way involved. Unlike the meanings of a homonym, however, these uses of the word *paper* are related to one another and clearly arose through a process of extension of similar meanings rather than through an arbitrary historical coincidence (Clark & Clark, 1979; Sweetser, 1990). This phenomenon of words having multiple related senses is called *polysemy*.[1] Rather than being an exception, polysemy can be found in most content words to at least some degree.

Whereas the alternative meanings of a homonym like *bank* have no obvious relation, the senses of polysemous words are clearly related. Certain semantic relations between a word's senses appear over and over in polysemy (see Lehrer, 1990; Nunberg, 1979), for example, object/substance, object/representation of that object, type/token, and text/object containing that text. To illustrate one, the object/substance relation is found when the same word is used to refer to an object and the substance that makes it up, often becoming a mass noun in the second case. For example, an oak is a tree that is the source of some oak (wood) used to make a table; a chicken is an animal that may end up as

This research was supported by NIMH grant MH41704 and an NSF graduate fellowship. We are grateful to Gary Dell and Kristine Onishi for insightful comments and to Karin Humphreys for linguistic assistance as well as helpful comments.

Address correspondence to Gregory L. Murphy, Beckman Institute, University of Illinois, 405 No. Mathews Ave., Urbana, IL 61801. E-mail: gmurphy@s.psych.uiuc.edu.

[1] We will follow linguistic usage in calling the different meanings of polysemous words *senses*, generally referring to homonyms as having different *meanings*. However, *meaning* will also be used as a term covering both of these cases. As we will discuss, the distinction between a polysemous sense and a homonymous meaning can sometimes be difficult to draw.

0749-596X/01 $35.00

some baked chicken (the food). If we were to hear of a new species of plant called a *delgar*, we could say both "There is a delgar growing in my yard" (individual plant) and "This pen is made out of delgar" (the substance derived from that plant). Thus, these forms of polysemy are highly productive, and they are used quite easily when new words enter the lexicon (Murphy, 1997). For example, the word *book* can be used to refer both to a physical object containing a text and to the content of that text. The same form of polysemy is present in recently invented words for new information-storage devices such as *videotape*, *CD*, and *DVD*; for example, *That CD is cracked* (object) and *That CD is brilliant* (content of the CD).

The outlines of a theory of homonymic representation are fairly clear. The different meanings of *bank* or *calf* are considered to be different words, so it is generally believed that they are represented by different *lemmas* (lexical units—see Levelt, 1989). In lexicology, there also seems to be a belief that homonyms are different words, as indicated by separate dictionary entries (Zgusta, 1971, p. 74; also shown below). This is not to say that it is understood exactly how listeners identify which meaning of a homonym is intended but that there is fairly good agreement that these meanings are lexically separated. There is no such agreement for polysemy. Should the sense of *paper* meaning "substance made from wood pulp" be in the same lexical entry as "sheet of writing material [made from that substance]"? Should the two senses of *book* meaning "object with print" and "a particular text" be in the same lemma? Although such senses seem closely related, they are sometimes ontologically different things. In the sentence "Your book is green," a physical object is being talked about, but in "Your book was difficult to understand," the textual content is being referred to. If the senses are switched across sentences, the sentences may no longer be true: The textual content was not green; the object was not difficult to understand. Linguists have varied in their approach to this problem, ranging from suggesting that there is a single represented sense that accounts for all these uses of a word (Ruhl, 1989) to arguing that each distinguishable sense is separately represented (e.g., approaches based on lexicology, such as Zgusta, 1971).

Thus, the questions of how many senses are represented, how they are linked in memory, and how they are coordinated in processing are the critical issues surrounding polysemy. The present experiments aim to provide data that will constrain accounts of the representation and processing of polysemy.

Given the differences just described between polysemy and homonymy, it is obviously critical to keep these two phenomena distinct. Polysemy is the normal, expected presence of related senses in a word, such as an object and the substance making up that object, and homonymy is the unpredictable coincidence of two different words having the same name. Unfortunately, psychologists have not been very good at keeping these two terms separate. In particular, the term *polysemy*, which is used in linguistics to refer to a word having related senses (e.g., Cruse, 1986; Geeraerts, 1993), is often used as a synonym for *ambiguity* (including homonymy) in the psychological literature. For example, Hino and Lupker (1996) titled their article "Effects of Polysemy . . ." and then refer to their stimuli as "ambiguous" and "unambiguous." This is fairly typical of the terminology in the psychological literature. Furthermore, studies of ambiguity have sometimes combined these two phenomena in their experimental designs by treating homonyms and polysemous words as both "ambiguous." In the present article, we will use the term *polysemy* to refer to the phenomenon of related senses in otherwise unambiguous words (i.e., following linguistic practice), and we will use *ambiguity* to refer to words that are homonyms or homographs.

The Problem of Polysemous Representation

Whereas homonyms are different words that happen to share the same name, it is not entirely clear how polysemous words, whose senses are more closely related, are represented. In an influential paper, Nunberg (1979) argued against the idea that all distinct senses should be represented in the lexicon. Instead, he proposed that pragmatic principles could be used to derive

word senses from others. For example, if *dog* refers to a class of animals, one could easily derive the use of this word to refer to an individual animal in this class. We should emphasize that Nunberg's argument concerned how a linguistic theory of the lexicon should represent different senses. It was not presented as a psychological theory of representation and processing. Nonetheless, one can readily construct a psychological theory from this proposal. According to this theory, all that is represented is a core meaning of a word. The different polysemous extensions are generated on the fly, using pragmatics and plausible reasoning. Thus, on this view, different senses are not prestored but are rather computed from contextual features.

A similar account was proposed by Caramazza and Grober (1976), who identified 26 separate but related senses for the word *line*. They suggested that these senses are all related to a core meaning and that "it is precisely the core meaning that is stored in the psychological representation for the meaning of *line*" (p. 188). They argued against the notion that each sense is explicitly stored in the mental lexicon. Ruhl (1989) went so far as to argue that there is a single, defining sense for words (even most homonyms), with distinct senses neither created or stored.

Lehrer (1990) agreed with Nunberg that much polysemy can be predicted through general principles of meaning extension, but she also noted that these principles sometimes fail. She argued that the lexicon is simply unpredictable to some degree and that language users must learn which words can be extended in which ways, rather than relying entirely on pragmatic principles. That is, at least some senses must be explicitly represented (see also Rice, 1992). In lexicology, Zgusta (1971, p. 66) argued that it is usually impossible to find a single basic sense of a word from which the other senses can be derived. Thus, within the linguistic literature there is a variety of views on how explicit the lexicon must be—whether each word sense must be represented or instead is derived from a more basic or core meaning.

Few psychological studies have addressed this issue. Williams (1992) found that contextually irrelevant senses of polysemous words are active even over long delays in a lexical decision task. He compared this to results from the homonym literature, which show that priming for the contextually irrelevant meaning of homonymous words is short-lived. Therefore Williams argued that the senses of polysemous words cannot be represented independently, as homonym meanings are (see also Durkin & Manning, 1989). One possibility, then, is that the polysemous senses are connected through a common core.

Additional support for the core concept view of polysemy can be found in the work of Anderson and Ortony (1975), who argued that understanding is more than finding the correct lexical entry in a semantic associative network. According to their work, a polysemous word derives rich representations from its sentence, and both context and world knowledge must be involved in deriving those representations. Thus, they argued that semantic memory is not rich enough to explain how polysemous words are interpreted. Like Caramazza and Grober (1976), they seem to be suggesting that the lexical network contains core information, and other information necessary to understand the exact sense of the word is supplied by context.

Core-meaning theories suggest a view of polysemous senses as being somewhat ephemeral. Lexical meanings can be augmented or extended in a given context, but those extended senses are not permanently stored in the lexicon. Even having created an extension once does not make a subsequent creation easier on a simple core-meaning account, since it is the core meaning that is retained. (We will discuss more elaborate views in the General Discussion.)

Another view of polysemy representation is one that is much closer to homonym representation. According to this view, common senses would have separate entries connected to the same lemma. For example, the writing material sense of *paper* might be its core meaning, but other senses such as a daily periodical or an article would be represented distinctly. Presumably, these senses would all be linked to the same lemma, unlike homonyms. Such views can vary considerably, depending on how many senses

they claim are represented and whether one of these senses is picked out as being the core meaning. An important question surrounding this approach is how to decide which senses are distinct (e.g., *dog* used to refer to the class of dogs vs an individual dog? *dog* as an animal vs a kind of meat?) and when a sense rises above a mere nonce usage to deserve full representation. Many linguists appear to take the view that some reasonable number of senses are represented, rather than only a single (core) meaning or every possible sense (e.g., Cruse, 1986; Deane, 1988; Langacker, 1987; Rice, 1992; Tuggy, 1993).

As already mentioned, there is very little experimental evidence to support either the core or a multiple-sense theory or to provide constraints on either view, and what evidence exists is often muddied by the use of homonyms in the polysemous stimuli. One study that focused solely on polysemy was that of Murphy (1997), which showed that novel extensions of a word that were closely related to previously known senses were more acceptable than were more distantly related extensions. This suggests that polysemy could develop by the construction of a chain of extensions, each building on its predecessors (as proposed by Heine, 1992; Lakoff, 1987; Malt, Sloman, Gennari, Shi, & Wang, 1999). Furthermore, this is evidence that people can create novel extensions of words in context and that not all senses of a word need to be prestored to be understood—a conclusion reinforced by research on nonce uses of existing words (e.g., Clark & Gerrig, 1983; Gerrig, 1989)—which is consistent with a core view. However, although novel uses of a word may be comprehensible, this does not mean that commonly encountered senses are not stored. Murphy used novel words and novel extensions, but it is possible that many senses of actual words are represented in memory. It is our goal to investigate this question.

The main question being investigated in the present experiments is the degree to which different senses of polysemous words use the same or different representations. If polysemous words have only a core meaning, specific senses being derived online, then different uses of a word would have highly shared representations. Conversely, if each sense of a polysemous word is encoded and represented separately, then the representation of one use of the word might not overlap that of a different use of the same word. There are also a number of intermediate possibilities, in which a core part of the meaning is shared by most senses, varying in how much information is in the core and how much is in the senses. For example, one might believe that there is a very abstract core that is present in all the uses of a word, but each sense provides considerable detail to flesh out its particular meaning. Alternatively, one might propose that the core is a rich representation of the prototypical use of the word, and the senses provide only the minimal information that distinguishes them. Along with each representational view there are a number of possible processing accounts as well. For example, one might argue that the core sense is activated first, since it is common to all or many uses; one could argue that the core and other senses are activated in parallel; one might claim that all senses are activated until the correct one is selected, or that only the most likely sense is activated.

The problem in beginning an investigation of polysemy is that there are few explicit models of the representation and comprehension processes of polysemous words. Linguistic approaches virtually never discuss processing issues. The psychological literature has focused on the single-sense notion but has not explored most of the other possibilities. It is clearly impossible to sort through all these possibilities in a single study. Our approach has been to collect data that will act as constraints on all of these theories. We will argue that the results indeed rule out some of them, even if they do not yet determine which one is correct. Our strategy was to investigate the amount of overlap in different senses, using a priming technique in which one use of a word was followed by a subsequent use that involved the same or a different sense. Differences between these conditions indicate the amount of overlap of the senses' semantic representations. The question is whether there is enough semantic commonality across different uses of a word to indicate

a possible shared core meaning. As we have just pointed out, even if one believes in a core, it might range from being the entire stored representation of a word to a minimal, abstract component. If our results show that different senses have considerable semantic overlap, then single-sense views or any view with a substantial core will be supported. In contrast, if the results show minimal overlap, then separate-sense views or a view with a minimal core will be supported. The initial experiments begin by looking at whether there are distinctions between different senses of a single word. Later experiments attempt to gauge the size of such distinctions by comparing polysemous words to homonyms.

Although our results will not be able to narrow the field down to a single model, it is nonetheless important to begin to perform empirical work that will elucidate the representation of polysemy, because of its implications for our understanding of lexical processing and representation. Gerrig (1986) pointed out that considerable psycholinguistic research addresses how meaning is used in lexical access and discourse comprehension, yet there is little agreement on exactly what semantic information is included in lexical representations. For example, psychologists argue about whether meaning is selectively or exhaustively accessed (especially for homonyms) during comprehension, yet neither they nor linguists agree on just what that meaning is—what is and is not included in a word's semantic representation. Without an understanding of the content of lexical representations, it is impossible to decide whether access is selective, or in what ways it is and is not selective. If Caramazza and Grober (1976) are correct in saying that there are (at least) 26 distinct senses of the word *line* and 40 senses for *run*, for example, one might be reluctant to propose that all these senses are accessed each time the word is encountered. In contrast, if the representation of such polysemous words is a core sense or only a few senses, then exhaustive access seems a likely possibility. Thus, specifying the representation of polysemous words is a necessary part of explaining how meaning is represented and involved in production and comprehension.

EXPERIMENT 1

The first experiment used memory performance as a measure of the representation of polysemous senses. In particular, it investigated whether people are better able to recognize a word used in the same sense or a different sense than its original presentation. The experiment was based on a paradigm developed by Light and Carter-Sobell (1970). In their study, subjects saw phrases like *traffic jam* and then saw a phrase like *strawberry jam* and were asked if they had previously seen the word *jam*. Performance was worse in such cases than when they saw *raspberry jam* followed by *strawberry jam*, which uses *jam* in the same way. In Light and Carter-Sobell's (1970) experiment, most of the test words would be counted as homonyms rather than as polysemous words. It is perhaps not surprising that using a word in one way makes it difficult to retrieve a memory of the word used in an unrelated way. However, it is not so clear that the same effect would be found with the highly related senses of polysemous words. For example, would one find the same decline in recognition if one were to use the word *paper* to mean newspaper and wrapping paper? If the two uses draw on a single core meaning, recall might be quite high even when the sense is changed. If the senses rely on separate representations, recall would be expected to be lower when different senses are used in learning and test. Because homonyms do not share a common core meaning, Light and Carter-Sobell's results do not speak to this possibility.

A related experiment was Perfetti and Goodman's (1970) study of semantic constraints on disambiguation. They found that when ambiguous words were presented in a sentence context, subjects later false-alarmed in a recognition memory test to words associated with the relevant sense of the word. For example, after reading a sentence like "Many families rent a house in the country for the summer months" for the critical word *country*, subjects later were more likely to say that the word *city* (related to the

rural sense of *country*) but not the word *nation* (related to the political sense of country) had appeared in the list. In another experiment, Perfetti and Goodman presented a list of words for subjects to learn. They did not find analogous context effects when the context was the word on the list just prior to the ambiguous word. Thus, their results suggest that the context of a single word may not be sufficient to distinguish the senses of a polysemous word in the present experiment.

However, the Perfetti and Goodman (1970) study is not entirely appropriate for investigating the issues raised in the present enterprise. They did not look at memory for the ambiguous word itself but instead at false alarms to associates. In Experiment 1, we looked at recognition memory for the polysemous word itself, which may be a more sensitive measure. Furthermore, the Perfetti and Goodman stimuli seem to have included both polysemous items (like *country*) and homonyms. They argued (p. 427) that only 4 of the 30 items used in most of their experiments were polysemous. So, again, it is unclear whether evidence of distinct memory representation will be found for words that have highly related semantic representations. The present experiment attempts to answer that question.

The experiment had two parts. In the learning phase, subjects read phrases and were told to study them for a later memory test. The phrases contained polysemous words, biased in interpretation toward one of two senses. In the test phase, subjects viewed similar phrases in which one of the two words was capitalized. Their task was to decide if they had seen the capitalized word before. In experimental conditions, the word that was capitalized was the polysemous word, which could be in the same phrase as previously seen, in a different phrase that used the same sense, or in a phrase that used the alternative sense. However, the capitalized word was always the repeated word. Therefore, in the experimental conditions, the response should always be YES. In foils, the noncapitalized word was repeated.

If polysemous words are represented with separate sense representations, memory should be better if a word is used in the same sense than if it is in a different sense. If the core meaning view is correct, there should be no such difference, as subjects will access the single sense of a polysemous word every time it is encountered.

Method

Materials. Twenty-four polysemous words were used (see the Appendix). We started with Durkin and Manning's (1989) list of 175 polysemous words and their most common different senses and selected senses that were clearly related. Polysemy of senses was also ensured by consulting the words' listings in the *Oxford English Dictionary*. The meanings of homonyms in this dictionary are given separate entries, whereas the related senses of our polysemous items were all listed under a single entry.

These potential stimuli were then normed on our subject population. Thirty-six subjects produced definitions of polysemous words, and the two most common senses produced were chosen for use in the first experiment. After the two senses were chosen, two phrases using the word were constructed for each sense so that there were four phrases for each polysemous word. For example, *paper* is a polysemous word that can mean sheets of material made from wood pulp or a newspaper. (Note that our paraphrases of these senses are simply expository conveniences. Since the descriptions were never used in the experiments, they have no bearing on the results.) There were two phrases created for each of these senses, as in the examples given below.

Sheets of a material	Newspaper
wrapping paper	daily paper
shredded paper	liberal paper

The phrases were selected so that there would be minimal semantic overlap in the modifiers of each sense. For example, *daily* and *liberal* are not from the same semantic field, even though they both serve to mark *paper* as indicating a newspaper. In addition, there was minimal morphological overlap between the modifiers of an item (e.g., two modifiers ending in *-ing*), and consistent pairs did not share morphological structure more than did inconsistent pairs. In the learning phase, subjects might see a phrase using one of the senses of paper, either *wrap-*

ping paper or *daily paper*. In the test phase they would see either the same phrase repeated (the *same phrase condition*), a new phrase using the same sense (the *consistent sense condition*), or a new phrase using a different sense (the *inconsistent sense condition*). So, the test item *shredded PAPER* was in the consistent condition when *wrapping paper* was studied and was in the inconsistent condition when *daily paper* was studied. Note that the critical comparison (of consistent vs inconsistent conditions) always involved a new modifier.

Counterbalancing items across conditions required six lists (each of two test phrases in a quadruple was preceded by a phrase that was identical, consistent, or inconsistent). Thus, none of the effects could be attributed to one sense being more familiar or memorable than the other, since both senses appeared in both consistent and inconsistent conditions.

Foils also consisted of pairs of phrases, one appearing in the study portion of the experiment and the second appearing in test. As in the experimental items, one word of the phrase was repeated, but unlike in the experimental items, the repeated word was not the word subjects were asked to judge. For example, the study phrase might be *tennis ball*, which would be followed in the test phase by *tennis SHIRT*, where subjects were asked if they had seen *shirt* before. These foils were used to force subjects to focus on the critical word, rather than allowing them to respond positively if one of the words seemed familiar. There were 24 foils and 24 experimental items, for a total of 48 test trials.

Procedure. Subjects viewed the materials on a Macintosh Quadra 630 computer, which was connected to a PsyScope button box (Cohen, MacWhinney, Flatt, & Provost, 1993). Their dominant hands were assigned to the YES response and their nondominant hands to the NO response. For the first phase, subjects were instructed to study the phrases that would be viewed for a later memory test. Phrases appeared on the computer screen for 2 s apiece, with 2 s between trials. Each phrase was viewed once. At the end of this phase, instructions appeared telling subjects about the recognition memory test. They also received verbal instructions from the experimenter and were allowed to ask any questions about the test. The test phase began immediately afterward. In this second part, subjects were instructed to decide whether the capitalized word in each phrase had appeared in the first part of the study. If it had, they pressed the YES button; if it had not, they pressed the NO button. Subjects were told to read the whole phrase, because the other words might serve as a memory aid. They were also told to go as fast as possible, without making mistakes, although no explicit feedback about errors was given.

Subjects. Subjects were 61 introductory psychology students who participated in the experiment for partial fulfillment of course credit. They were all native speakers of English.

Results and Discussion

One subject who had over 50% errors was dropped from the study. Reaction times (RTs) were not analyzed, given the large number of missing RTs (due to memory errors).

Same phrase items were the most accurately evaluated (79% correct) (SD = 19%), followed by the consistent sense items (64%) (SD = 25%), with the inconsistent sense items being the most error prone (56%) (SD = 24%), $F_1(2,118) = 23.14$, $p < .001$; $F_2(2,46) = 25.75$, $p < .001$. For example, when *paper* was seen initially in a phrase like *wrapping paper*, which supported the "sheets of a material" sense, it was easier to verify having seen *paper* when *shredded PAPER* was presented at test than when *liberal PAPER* was presented at test. Not surprisingly, seeing a word in an identical context was the most helpful. These results indicate that the way a polysemous word is processed initially affects later memory access.

The repeated items were significantly more accurate than the consistent sense phrases, $t(59) = 4.25$, $p < .001$; $t(23) = 5.80$, $p < .001$. Most importantly, the consistent sense phrases were reliably more accurate than the inconsistent sense phrases, $t(59) = 2.75$, $p < .01$; $t(23) = 2.23$, $p < .05$. In fact, the inconsistent condition was not reliably different from chance (p's $> .05$).

These results lend support to the hypothesis that senses are stored separately and are evidence against a single core meaning hypothesis. If the words were interpreted in terms of a common core meaning (as suggested by a monosemy view such as Ruhl's, 1989), memory would have been equivalent for the consistent and inconsistent senses. Thus, these results suggest that polysemous senses may be stored separately. At the very least, they suggest that the senses are functionally distinct in that same-sense uses are more related than are different-senses uses of the same word. The fact that the sense-inconsistent condition was not reliably different from chance indicates that if there is a core sense, it is not at all strong, since it did not provide a basis for memory. Semantic overlap across senses is less, then, than might be expected, a conclusion that is examined further in Experiment 3.

This conclusion contrasts with that of Perfetti and Goodman's (1970) study described above. In their experiment, a single context word was not sufficient to force subjects to distinguish the senses of the learned words, whereas a sentence context was. Although there are many differences between their study and ours, perhaps the most important one is that our subjects read the words as a phrase and therefore presumably interpreted them as a coherent concept (Gerrig & Murphy, 1992). For example, in order to understand *liberal paper*, a subject would have had to determine that *paper* referred to a newspaper, rather than to a sheet of blank paper, say, or else the phrase would not have made sense. In contrast, Perfetti and Goodman's context was a preceding item on a list of to-be-remembered words, so subjects were not encouraged to integrate the stimuli. Also, as mentioned earlier, our dependent measure was memory for the target word itself, whereas their measure was the less direct measure of false alarms to another word related to one sense of the target word. As the effect size was only 8% in our experiment, it is perhaps not surprising that a less direct measure would not obtain a significant difference between consistent and inconsistent senses.

EXPERIMENT 2

Our main interest in Experiment 1 was not memory for senses per se, but rather the issue of how polysemous words are represented and processed. The results did suggest that different senses are stored separately. This result is surprising enough to follow up in a task that involves comprehension and semantic processing rather than memory. In this experiment, the task was to make a sense/nonsense judgment on phrases similar to those used in the first memory experiment. The RT and accuracy of this judgment were the dependent measures. We again manipulated sense consistency by presenting phrases that involved the same or a different sense of a word and by looking for priming of a consistent use of a word compared to the inconsistent use.

For example, subjects might see *liberal paper* and have to say whether or not it made sense. On the previous trial, they would have seen either *daily paper* (consistent sense) or *wrapping paper* (inconsistent sense). The question, then, is whether the difficulty in deciding the sensicality of *liberal paper* depends on the consistency in sense of the prior use of *paper*. (Foils were used so that a sensical first phrase did not predict the sensicality of the second phrase.) Again, if there is only a core meaning, consistency should not make a difference, since all phrases would require access to the core concept of *paper*. We used a sensicality judgment because it requires subjects to access the meaning of the word, unlike lexical decision tasks, which only require subjects to verify that a string is a word. Although meaning may be used in such a task, it may not be accessed at the level of detail that would distinguish different senses. Furthermore, judging phrase sensicality has been shown to be a sensitive measure of conceptual processing in previous work (e.g., Murphy, 1991).

Bainbridge, Lewandowsky, and Kirsner (1993) performed a similar study on polysemous words using sentence contexts and a lexical decision task. Subjects made lexical decisions on words twice—preceded by contexts that evoked either the same or different senses across trials.

They found that polysemous words were judged faster in the second trial when they were preceded by the same-sense context. Bainbridge et al. concluded (p. 624) that the different senses are represented separately, and that priming is to a large degree dependent on activating the same sense of a word. However, their study did not separate repetition of a sense from repetition of the exact sentence context. That is, their condition with different senses (necessarily) had different sentence contexts, whereas their same-sense condition had the exact same sentence context on both trials. As our Experiment 1 showed, seeing the exact same context provides more priming than simply activating the same sense. (In fact, the repetition effect there was larger than the sense consistency effect.) In order to fairly compare the same-sense and different-sense conditions, one needs to use different contexts in both cases, as in the present experiment.

Method

Materials. The same phrases were used in the critical trials as in Experiment 1. These phrases were counterbalanced, so that a subject might see the prime phrase *wrapping paper* followed by either *shredded paper* (consistent) or *liberal paper* (inconsistent) as targets, and a different subject might see *daily paper* as a prime, again followed by either *liberal paper* (now consistent) or *shredded paper* (inconsistent), requiring four counterbalancing lists. Within each list, half the critical items were consistent and half inconsistent. Each subject saw 24 total experimental pairs of phrases and 72 foil pairs. The foils were also two phrases sharing a word, with at least one phrase that did not make sense. In one-third of the foils, the first phrase (*history lecture*) made sense while the second (*yellow lecture*) did not. One-third of the pairs had the reverse pattern, and the final third had two nonsensical phrases. As a result of these foils, the sensicality of the first phrase did not predict sensicality of the second across the experiment. The same equipment was used as in Experiment 1.

Procedure. To ensure that subjects understood the instructions, there was a set of eight practice trials. Subjects were asked to judge the sensicality of each phrase. They were explicitly instructed that sets of two phrases in a row would have a word in common. They were told to respond as quickly as they could without making errors. Each phrase appeared by itself on the screen until subjects responded. After every trial, subjects received feedback: Feedback telling them they were correct was on the screen for 1 s, while feedback after an error was on the screen for 2 s. There was a pause of 250 ms between the end of the feedback and the beginning of the subsequent item, and there was no particular marking of the phrase pairs that shared a word. The entire experiment took less than 20 min.

Subjects. Subjects were 27 introductory psychology students who participated in the experiment for partial fulfillment of course credit. They were all native speakers of English.

Results and Discussion

In order to be able to analyze RTs, in this and subsequent experiments we eliminated subjects who made too many errors and therefore had few trials per cell. Our criterion for dropping subjects was making more than 20% errors overall (which suggested a general lack of attention) or making errors on more than 20% of the experimental trials. Seven subjects were dropped from Experiment 2 on this basis. For the remaining 20 subjects, experimental RTs larger than 3 *SD*s above each subject's mean were omitted. Trials in which subjects made an error on the prime were omitted from both analyses, since we could not be sure that subjects had correctly processed the prime phrase, which constituted the experimental manipulation. This resulted in the exclusion of 1% of the data.

The error analysis of the critical phrases showed that it was easier to understand the phrase when the repeated word was used in the same sense than in a different sense, $F_1(1,19) = 14.77$, $p < .005$; $F_2(1,23) = 10.12$, $p < .005$. When subjects received consistent phrases, they were correct 96% of the time ($SD = 5\%$), whereas when they received a prime and target

using different senses of the polysemous word, they were correct 87% of the time ($SD = 10\%$).

The RT analysis also showed the consistency effect, although it was only reliable in the item analysis. When the prime and target were consistent, sensicality was judged more rapidly (792 ms; $SD = 96$) than when prime and target were inconsistent (859 ms; $SD = 118$), $F_1(1,19) = 1.82$, $p > .10$; $F_2(1,23) = 9.64$, $p = .005$. The lack of significance in the subject analysis is probably due to the small number of items per subject, once trials with errors (on primes or targets) were removed.

When the second occurrence of a word was consistent with the first one, subjects were more likely to judge the phrase as making sense, and they tended to make this decision faster. This constitutes further evidence against a single core meaning, using an online task. Our results confirm those of Bainbridge et al. (1993), who had a similar design using a lexical decision task. But, as discussed earlier, Bainbridge et al. used the exact same contexts in their consistent condition, and so their results could have been an effect of receiving the same context on both trials. In our experiment, the contexts were different in both conditions, and we still found an influence of same vs different sense.

The results of Experiments 1 and 2 use very different techniques to converge on the conclusion that the different senses of a polysemous word are functionally distinct. In Experiment 1, we argued that there was little absolute commonality among the different senses of the same word, because the inconsistent condition was not different from chance. Such a claim could not be made in Experiment 2, however, because there was no lower baseline against which the inconsistent sense condition could be compared. The next experiment provides such a baseline.

EXPERIMENT 3

Experiment 3 served both to replicate Experiment 2 and to gauge the size of the consistency effect. The consistency effect arises because different uses of a word involve different semantic properties. The stronger such an effect, the less significant a core meaning could be. Homonyms provide a useful benchmark for estimating the size of the effect, because the different meanings of a homonym are essentially independent. While a sense/nonsense judgment task has not been done, to our knowledge, with homonyms, one would expect that since their meanings are far more distinct, the consistency effect would be far larger. For example, reading the word *bank* to mean a financial institution will have a severe cost when the word was previously used to refer to a mound of earth (Swinney & Hakes, 1976; Tabossi, 1988; Tabossi & Zardon, 1993; Van Petten & Kutas, 1987). The prime will activate one lemma for this word, which will then be very strongly activated when the word occurs again. When this is the incorrect lemma (in the inconsistent condition), it must be suppressed and another one must be selected. If the senses of polysemous words overlap, then they should reveal a smaller consistency effect. Furthermore, the linguistic assumption is that these senses are connected to the same lemma, so selection of a new one would not be necessary.

In sum, homonyms provide a way to scale the size of the consistency effect in polysemous words, since they represent the case in which different meanings are completely unrelated. This experiment also serves as an attempt to replicate the consistency effect found with polysemous words in Experiment 2.

Method

The method was identical to that used in Experiment 2, with the one change that both homonymous and polysemous stimuli were used. For the experimental conditions in which the homonyms appeared, a subject might be asked to judge the sensicality of *commercial bank* and then *savings bank*, which use the same meaning of *bank*. However, another subject might see *creek bank* and then *savings bank*, which use different meanings of *bank*. Polysemous and homonymous items were used in the same list. Half of the experimental items a subject saw were consistent and half were inconsistent. The subjects were 34 introductory psychology students who participated

in the experiment for partial fulfillment of course credit. They were all native speakers of English.

A slightly different set of polysemous phrases was used in Experiment 3; modifiers were replaced in some of the less successful items from the prior experiments (e.g., those with many errors). There were a total of 24 polysemous words and 24 homonymous words (see the Appendix). Each subject saw 48 polysemous phrases, 48 homonymous phrases, and 288 control phrases, 96 in each of the control conditions. The foils from Experiment 2 were augmented by new items to make up the full complement. Four counterbalancing lists were used, as in the previous experiment, so that each critical phrase would appear in all conditions across subjects. The procedure was identical to that used in Experiment 2.

Results and Discussion

Two subjects were dropped from the study for making errors on more than 20% of all the phrases or of the experimental trials. For the remaining 32 subjects, RTs larger than 3 *SD* above each subject's mean were omitted, as were trials in which subjects made an error on the prime phrase. This resulted in the exclusion of 2% of the data.

One assumption was that homonyms and polysemous words would differ in the size of the consistency effect, with consistency being stronger for homonyms, due to the completely separate representations. Surprisingly, this was not found, as shown in Table 1: There were no interactions between consistency and word type, either in RTs, $F_1(1,31) = 1.69$, $p > .20$; $F_2(1,46) < 1$; or errors, $F_1(1,31) = 2.56$, $p > .10$; $F_2(1,46) = 2.62$, $p > .10$.

Importantly, although there were no interactions, consistency was still a reliable and important factor: when the prime and target were consistent in the sense they referred to, the target phrase was evaluated 85 ms more quickly and 12% more accurately than when the prime and target were inconsistent in sense, $F_1(1,31) = 46.01$, $p < .001$; $F_2(1,46) = 35.76$, $p < .001$ for RTs; and $F_1(1,31) = 77.79$, $p < .001$; $F_2(1,46) = 43.81$, $p < .001$ for errors. There was no difference overall between the homonyms and polysemous words, all F's < 1.

When results were analyzed separately for the polysemous and homonymous words, both showed the pattern found in prior experiments, with an advantage for consistency in errors and RTs (see Table 1 for the means), all p's $< .005$. If anything, the improvement of the stimuli appears to have strengthened the results for polysemous items.

This pattern of results is surprising, since a larger effect was expected for the homonyms. As discussed earlier, the different meanings of a homonym are generally thought to be represented distinctly, so there should have been larger consistency effects in these stimuli. Priming one meaning that is totally separate from the alternate meaning should have more strongly aided understanding of a phrase using the same meaning and hurt understanding of a phrase using the other meaning. Although the effect size for the homonyms was slightly larger (107 ms) than for the polysemous words (64 ms), there was no interaction by word type, so this difference is at most suggestive.

Because this finding is somewhat surprising and because there is a hint of the expected difference, we replicated the experiment. In addition to the main experiment just reported, 24 subjects performed in an essentially identical study, with the exception that some of the homonym items were different (as described in Experiment 4A). In this version of the experiment, which we will call Experiment 3A, we again found no important difference between the homonym and polysemy conditions, the interaction p's all $>.25$. In fact, in this experiment, the priming effect found for homonyms (25 ms, 12% errors) was actually *smaller* than that found for polysemous items (80 ms, 12% errors). The priming effect was significant in separate analyses of both classes of words, in RTs and error analyses, with the exception of the item analysis of RTs for homonyms, $F(1,23) < 1$. Thus, across Experiments 3 and 3A, the size of the priming effect was very close to being equal for homonyms and polysemous words. In one, the effect was slightly larger for homonyms, and in the other, the effect was slightly larger for poly-

TABLE 1

Mean Reaction Times (ms) with Standard Deviation in Parentheses and Percentage Correct for Experiments 3 and 4

	Sense			
	Consistent		Inconsistent	
	RT	% correct	RT	% correct
Experiment 3				
Polysemous words	774 (118)	95	838 (156)	86
Homonyms	743 (130)	97	850 (187)	83
M	759	96	844	85
Experiment 4				
Polysemous modifiers	593 (108)	93	583 (112)	91
Homonym modifiers	565 (111)	96	574 (115)	97
M	579	95	579	94

semous words, but in neither case did the interaction of word class and priming approach significance.

This pattern of results is consistent with the finding in Experiment 1 that memory was not cued by using the word in a different sense. In both cases, cross-sense performance was about the same as a baseline. That is, counterintuitively, the overlap of different senses of the same word is minimal—about the same as the overlap of homonym meanings. In functional terms, this means that any core meaning shared by these senses is also minimal. Such results rule out not only a full single-sense view but also any separate representation account in which a core meaning plays a significant role. We discuss possible reasons for the minimal overlap of polysemous senses in the General Discussion.

This finding is in direct contrast to that of Williams (1992), who argued for representational differences between homonyms and polysemous words. In particular, he found that contextually inappropriate senses of polysemous words were still activated for some time after presentation, whereas the literature on homonyms generally shows suppression of the incorrect meaning (see introduction to Experiment 5 below). However, Williams did not actually include homonyms in his study, so he did not directly reveal any differences between them and polysemous words. Also, his experiments did not compare priming of relevant and irrelevant senses, as the present experiments did. It is possible that the priming he found for the contextually irrelevant sense of a word would be significantly less than what would be found for the relevant sense (analogous to our consistency effect). Thus, the present experiments provide a more complete comparison of polysemy and homonymy than did Williams's study.

EXPERIMENT 4

The first three experiments looked at the relation between phrases that used words in a consistent or inconsistent sense, finding a difference between these conditions both in memory and timed semantic judgments. We have attributed these effects to the polysemous word being used in the same or a different sense across trials. However, the way in which sense consistency was manipulated was via a modifying word, so it is possible that the modifiers themselves were partly responsible for these effects. To illustrate this possibility, consider the consistent phrase pairs, *wrapping paper* and *shredded paper*. It is possible that the word *wrapping* was priming *shredded*, rather than the consistency of the use of *paper* causing the effect. In this case, "sense consistency" results would be obtained, but for the wrong reasons—having nothing to do with the noun, which was the word of interest. Thus, Experiment 4 was a control experiment that investigated whether priming of the modifiers themselves might be responsible for the consistency effects obtained.

Experiment 4 also raised the possibility of explaining the surprising findings of Experiments 3 and 3A. One hypothesis was that the homonyms in that experiment would show stronger effects of sense consistency than the polysemous words, since their meanings have little or no semantic overlap. In fact, there was no reliable difference between the two conditions. One possible reason for these findings could be uncontrolled item differences—in particular, the modifier–modifier priming described above. Perhaps the relations between the modifiers was systematically different in the polysemous and homonymous stimuli, explaining why the predicted larger effect in homonyms was not found.

To address these possibilities, it was necessary to perform a control experiment to ensure that sense priming was not being driven by the modifiers—to test for any possible modifier–modifier priming and compare such priming across word types. To this end, Experiment 4 was run on the modifiers alone, using a lexical decision task. (The sensicality judgment task used in the previous experiments could not be used with single-word stimuli.)

Method

Materials. The prior experiments used phrases that included the polysemous word and a modifier. For example, *paper* was a polysemous word, and it was paired with the modifiers *wrapping*, *shredded*, *liberal*, and *daily*. In this experiment, only the modifiers from Experiment 3 (from both polysemous and homonymous items) were used. The comparison of interest was modifiers taken from consistent pairs (like *wrapping* and *shredded*) vs modifiers taken from inconsistent pairs (like *liberal* and *shredded*) in the previous experiments. Foils consisted of words followed by nonwords. Pronounceable nonwords were constructed from words matched for frequency (Francis & Kucera, 1982) and length to the actual words used, with minimal changes (1–2 letters) to transform them into nonwords.

Procedure. Subjects viewed the materials on a Macintosh Quadra 630 computer, which was connected to a PsyScope button box. Their dominant hands were assigned to "WORD" responses and their nondominant hands to "NONWORD" responses. Subjects were told that letter strings would appear on the screen one at a time, in pairs. The first string in a pair would be a word, and it would be up for a short while (500 ms) and then disappear. Immediately, a string of letters would appear on the screen. Their task was to decide as quickly as possible, without making errors, if it was a word or not. Feedback for correct responses was on the screen for 1 s, while feedback for incorrect responses was up for 2 s, followed by a gap of 250 ms before the next pair. Subjects were also warned that there would be a memory test at the end on what they saw, to ensure they paid close attention to both primes and targets. The memory test instructed them to write down all the words they remembered seeing during the experiment.

Subjects. Subjects were 37 introductory psychology students who participated in the experiment for partial fulfillment of course credit. They were all native speakers of English.

Results and Discussion

Five subjects who made errors on over 20% of the strings were omitted from analysis. For the remaining 32 subjects, RTs larger than 3 *SD* above each subject's mean were omitted. This resulted in the exclusion of 3% of the data.

Analyses performed on the RTs found that the modifiers of the homonyms were evaluated faster (569 ms) than the modifiers of the polysemous words (588 ms), but this was reliable only in the subject analysis, $F_1(1,31) = 4.25$, $p < .05$; $F_2(1,46) = 1.22$, $p > .25$. There was also a reliable difference in accuracy, with the homonym modifiers more accurate (96% correct) than the polysemous modifiers (92% correct), $F_1(1,31) = 10.45$, $p < .005$; $F_2(1,46) = 5.31$, $p < .05$. One possible reason for this effect may be that the senses of the polysemous words required more complex modifiers to be distinguished. However, as simple differences between homonyms and polysemous phrases were not of interest in previous experiments, this result is not very revealing.

More importantly, there was no reliable priming of consistent items (we will refer to the modifiers by the names of the conditions they served in in Experiment 3), as would be expected if the modifiers were responsible for the priming of the previous experiments, all F's < 1. However, the interaction between word type and consistency, while not reliable in the RT analysis, $F_1(1,31) = 1.40$, $p > .20$; $F_2(1,46) = 1.35$, $p > .25$, was reliable in the error analysis $F_1(1,31) = 4.96$, $p < .05$; $F_2(1,46) = 4.16$, $p < .05$.

Because of this reliable interaction and because there was a hint of a difference between the results for homonyms and polysemous items in Experiment 3, the modifiers of homonyms and polysemous items were examined separately (see Table 1). The polysemous items showed no differences between the consistent and inconsistent modifiers in RTs, F's < 1, while in the error analysis, there was a nonsignificant trend toward the consistent items being more accurate, $F_1(1,31) = 3.14$, $p = .09$; $F_2(1,23) = 4.02$, $p = .06$. The homonyms showed no differences between the consistent and inconsistent modifiers in either RTs, F's < 1, or errors, $F_1(1,31) = 1.00$, $p > .30$; $F_2(1,23) < 1$. The lack of significant effects and the small size of the trends make it hard to attribute the reliable consistency results in the previous experiments to modifier priming.

A similar experiment was also carried out on the modifiers used in Experiment 3A. (We will refer to this as Experiment 4A.) In the complete analysis, we found no consistency effect, p's $> .10$, and no interaction of consistency and word type, p's $> .25$. However, visual examination of the results did seem to suggest that there was a possible difference between the word types. Separate analyses revealed no priming whatsoever for polysemous items. Therefore, the consistency effect found in Experiments 2 and 3A cannot be attributed to modifier effects. Surprisingly, a sort of reverse priming effect was found in the homonym stimuli, in which consistent items were verified 24 ms more slowly than inconsistent items, $F_1(1,27) = 5.01$, $p < .05$; $F_2(1,23) = 4.66$, $p < .05$, although there was no effect in errors.

In order to eliminate modifier priming as a possible explanation for the results, we removed eight items that contributed to this priming difference, so that the amount of modifier priming was equated in the homonyms and polysemous items. (These phrases were not used in Experiments 3 and 4.) We then reanalyzed the results of Experiment 3A with those items removed, to reevaluate the effect of consistency for these equated modifiers. We again found a reliable effect of consistency, with consistent phrases evaluated 60 ms faster than their inconsistent counterparts, $F_1(1,23) = 7.38$, $p < .05$; $F_2(1,38) = 6.06$, $p < .05$. The consistent items were also 12% more accurately judged than the inconsistent ones, $F_1(1,23) = 32.85$, $p < .001$; $F_2(1,38) = 22.39$, $p < .001$. There was again no interaction between consistency and word type (all F's < 1).

In short, although there were some differences between the modifiers of the two word types, they could not explain the consistency effects found earlier. Consistency effects were found for stimuli that had no modifier priming (Experiment 3), and the effects did not differ when the modifier priming was equated for the two word types (the reanalysis of Experiment 3A). Note that the interaction of priming and word type was not reliable in any experiment, so the reanalysis of the results of Experiment 3A was a conservative step. Nonetheless, it also found strong consistency effects for both word types. Thus, the results show very similar findings for homonyms and polysemous words.

EXPERIMENT 5

This experiment was conducted to examine whether the priming of polysemous word senses is due to inhibitory or facilitory processes. Facilitation could come about through a number of means. One obvious way is that if different word senses have separate representations, interpreting a word in one sense might activate that sense for later trials. When the word appears again, the correct sense is already selected and ready to be used in interpreting the new phrase. Thus, continued activation of a given sense could speed processing. One could imagine a rather different facilitative process that

would also lead to priming, however. It may be simply that the consistency of the interpretation of the entire phrase is responsible for the priming. For example, *wrapping paper* and *shredded paper* refer to similar kinds of things (i.e., have shared semantic properties), whereas *wrapping paper* and *daily paper* refer to different kinds of things. It could be the shared semantic components of the whole phrase that speed subjects' decision that the consistent phrase makes sense.

Inhibition of incorrect senses is another possible cause of the priming results. Gernsbacher (1990) has emphasized the importance of suppressing incorrect meanings of words and sentences as part of a fluent comprehension process. However, if senses inhibit one another, they must have separate representations. The newspaper sense of *paper* cannot be suppressed while the sheets of material sense is spared unless the two senses are functionally distinct. If judging *wrapping paper* actually made it more difficult to interpret *daily paper* (relative to a neutral condition), this would suggest that two different senses are stored. The second account given above for facilitation would not predict inhibition of senses. That is, if consistent phrases are easier to judge because they share semantic components, it does not follow that phrases using different word senses would be mutually inhibitory, merely that there would be less facilitation, because they do not share as many semantic properties.

The prior experiments do not distinguish facilitation or inhibition of senses. Because only two conditions were tested, it is not possible to say whether consistent phrases received an advantage, inconsistent phrases suffered a disadvantage, or some combination of the two. Such a conclusion requires a neutral condition that the others can be compared to: Faster responses than the neutral condition would indicate facilitation, and slower responses would indicate inhibition. In studies of homonyms, Gernsbacher and Faust (1991; also see discussion in Gernsbacher, 1990) showed that there was both facilitation and inhibition involved in the comprehension of ambiguous words in sentence contexts. When reading words like *bank*, subjects showed increased activation to words related to the intended meaning as well as to words related to the unintended meaning. However, after a delay, the activation to the unintended meaning decreased, whereas activation to intended meanings did not. Significantly, when the context did not pick out the correct meaning of an ambiguous word, both meanings remained active after the delay. Gernsbacher and Faust explained this pattern of results by arguing that both meanings initially receive activation during the comprehension process, and the incorrect meaning is then suppressed by contextual processes. Such a pattern might be expected when a single word is associated to two very different semantic representations. Inhibition is possible and indeed desirable in such a case, because one can fish from only one kind of bank and withdraw money from only the other kind: identifying one of the meanings as correct indicates that the other meaning is completely incorrect and so should be inhibited.

This situation is not exactly the same for polysemy, since the two senses are related. Indeed, in some cases, the same word can be used in two senses at once, as in *Your book is not only badly written, it is too heavy*; or *The factory fired its workers and then was burned to the ground* (Fauconnier, 1985; Cruse, 1986; Geeraerts, 1993). Thus, it is not clear that inhibition would be as desirable as it is for homonymy—or even possible. To the degree that polysemous senses share some meaning, it should be harder to facilitate one sense while suppressing the other. Thus, finding results parallel to those found for homonyms would indicate fairly distinct representations.

The present experiment therefore aimed to discover whether the priming effects observed for polysemous senses were due to inhibition of the inconsistent sense, facilitation of the consistent sense, or both. Since inhibition and facilitation are both measured relative to a neutral condition, the selection of an appropriate neutral item is obviously crucial. It would not be appropriate to use an unrepeated condition as neutral: Once one has read the word *paper*, it would be easier to read it again in the next trial completely independent of any issues of semantic representation, so an unrepeated condition could

not meaningfully be compared to the consistent and inconsistent trials in which the word was repeated.

Our neutral baseline used a prime in which a blank line took the place of the modifier. For example, _____ *paper* might be followed by *liberal paper*. This neutral condition primes the word *paper* but does not select any particular sense. This was to be compared to the conditions in which the prime selected the consistent or inconsistent sense of the noun, as in prior experiments. Because of the form of the target stimuli, which consisted of multiple words, a blank line was deemed more appropriate than simply presenting the single word. Other possible neutral modifiers would have been the use of a string of letters (XXXX) or an unrelated word (such as BLANK). Den Heyer, Taylor, and Abate (1986) found that unrelated words are more neutral than the strings of X's. However, in the current paradigm, it was impossible to find completely unrelated words that would not influence the task. For example, *BLANK paper* is inappropriate because it suggests a particular sense of *paper*, so it would hardly be neutral. This sort of problem seems likely for any neutral word that would be used. For this reason, we used a blank line preceding the word in the neutral primes.

Method

Design. As in Experiments 2 and 3, subjects were asked to read phrases containing the polysemous word and then make judgments as to their sensicality. Phrases occurred in sequential pairs. The prime could select one of the two senses of the polysemous word or be in the neutral, blank line condition. The targets were the same as before. Therefore, each polysemous word had associated with it three prime phrases (e.g., *wrapping paper*, *daily paper*, and _____ *paper*) and two target phrases (*shredded paper* and *liberal paper*), so six counterbalanced lists were necessary. One-third of the experimental items in each list were sense consistent, one-third sense inconsistent, and one-third neutral. Subjects saw each polysemous word in only one pair.

As before, we constructed foils in which the first, second, or both phrases did not make sense. Then we replaced one-third of each of the foil types with pairs of phrases in which the prime was neutral, to match the frequencies in the experimental stimuli. Although this resulted in some foil primes that now did make sense (when the initial nonsensical phrase was replaced by the neutral line modifier), it maintained the pattern that the sensicality of the target phrase was statistically unrelated to the condition of the initial phrase.

Materials. The polysemous materials from Experiment 3 were used here as well, but due to the addition of a new condition (neutral prime), 6 more items were created so that the complete set consisted of 30 polysemous word quads (see the Appendix). Each subject saw 60 experimental polysemous phrases and 180 foil phrases, 60 of each of the types of foils.

Procedure. Using the same equipment and procedure as previously, subjects were asked to judge the sensicality of each phrase. They were explicitly instructed that all the blank, neutral phrases made sense (since they were just one word).

Subjects. Subjects were 60 introductory psychology students who participated in the experiment for partial fulfillment of course credit. They were all native speakers of English.

Results and Discussion

Performance in this experiment was very accurate, and no subjects were dropped from the study. However, RTs greater than 3 *SD* above each subject's mean were omitted. Trials were omitted in which subjects made an error on either the prime or the target phrase. This resulted in the exclusion of less than 1% of the data.

An ANOVA found a clear difference between the consistency conditions, with the neutral condition (879 ms) falling between the consistent (832 ms) and inconsistent (938 ms) conditions in RTs, $F_1(2,118) = 6.33$, $p < .005$; $F_2(2,58) = 11.16$, $p < .001$. Error means followed the same pattern as the RTs, with the neutral items (89% correct) less accurate than the consistent items (96% correct) and more accurate than the inconsistent items (84% correct), $F_1(2,118) = 19.21$, $p < .001$; $F_2(2,58) = 14.26$, $p < .001$. More detailed analyses found evidence for both facilita-

tion of consistent items and inhibition of inconsistent items. The consistent condition was reliably better than the neutral condition in both RTs (a 47-ms difference), $t(59) = 2.26$, $p < .05$; $t(29) = 2.30$, $p < .05$; and errors, $t(59) = 4.39$, $p < .001$; $t(29) = 3.26$, $p < .01$. The comparison of the inconsistent and neutral conditions was fully reliable over errors, $t(59) = 2.12$, $p < .05$; $t(29) = 2.66$, $p < .05$, and the 59-ms RT difference was reliable in the item analysis, $t(59) = 1.77$, $p = .08$; $t(29) = 2.32$, $p < .05$. This suggests that for polysemous words, consistency facilitates comprehension and inconsistency inhibits comprehension.

These data indicate that multiple processes combine to give sense-consistent phrases an advantage. First, it appears that the consistent sense is being activated, which effectively lowers the threshold for subsequent activation. It is not clear from these data how long this effect lasts, but it is present at least long enough to have an effect on the next phrase read. Second, the activation of one sense appears to cause a dampening of activation for different senses.

The facilitation of consistent uses of a word is easily explained as repetition priming of semantic features or an entire word sense. However, this mechanism cannot explain the inhibition of the inconsistent sense. If inconsistency were worse than consistency only because the consistent sense was primed, then the inconsistent condition would be about as fast as the neutral condition, since neither one would have the correct sense primed (but both would have had general priming from reading the word, etc.). Instead, we found that the inconsistent condition was reliably harder than the neutral one, which is analogous to Gernsbacher and Faust's (1991) finding that the inappropriate meaning of the homonym showed less priming relative to the neutral condition after a delay. This finding could be caused by an active process of inhibition (as Gernsbacher and Faust propose), although other explanations may also be possible (see the General Discussion).

These results give even stronger evidence about the separate representation of polysemous senses. If one sense can be activated and the other inhibited as a result of an earlier encounter with the word, these senses must be functionally and therefore representationally distinct. Clearly, the notion that the word has a single core meaning that is the basis for every use cannot accommodate such results. The implications of these results go beyond that conclusion, however, in that they suggest that the senses of polysemous words are quite distinct in spite of their apparent relatedness. For example, one might have thought that the fact that newspapers are printed on paper would make it difficult to activate the newspaper sense of *paper* while simultaneously inhibiting the sheets-of-paper sense. However, this was not the case.

We argued earlier that such inhibition ruled out one explanation of the earlier priming effect. Perhaps a liberal paper is just generally more similar to a daily paper than it is to some wrapping paper. This kind of semantic consistency of the whole phrases—without any mention of polysemy—might explain the advantage of consistent over inconsistent word uses in the first three experiments, as similar concepts would have some of the same properties, which might aid in making a sensicality judgment to the target phrase. (However, we should note that, in fact, wrapping paper is not that similar to shredded paper: One cannot wrap something with shredded paper, and wrapping paper is not the most typical shredded paper. Similarly, *daily paper* emphasizes a very different aspect of newspapers than does *liberal paper*. We selected modifiers just so that the same-sense phrases would not describe very similar things, avoiding phrase pairs like *weekly paper* and *daily paper*.) This explanation would not predict suppression of an inconsistent use of the word, however, which is what we found in the present experiment. Such suppression would seem to require that a word like *paper* have different representations for different senses, one of which could be suppressed while the other was simultaneously facilitated.

We should emphasize that the present study was designed to investigate the representational issues involving polysemous senses, and in particular it did not address the processing issue that has been of so much interest in ambiguity research more generally, namely, the question of

whether the initial activation process is context-sensitive from the start or only after initial context-free activation of all meanings (i.e., whether the selection process is or is not modular). We mention this because such studies often use a priming technique similar to ours (though usually with a lexical task rather than a semantic one). Given the difficulty in answering this question in the homonymy literature (see Simpson, 1994, for a review), it seems likely that the answer will be no easier to find for polysemy. Our results suggest that selective access is *possible*, since different senses seem to be represented separately. If the senses shared a core representation that included most of each sense's meaning, selective access could not be achieved. Whether selective access in fact occurs is a question we leave to future research.

GENERAL DISCUSSION

Taken as a whole, these experiments provide evidence for separate representations of the senses of polysemous words. The studies found evidence for sense priming in polysemy. Using a word in a specific sense facilitated comprehension for a phrase that used the word in the same sense and inhibited comprehension for a phrase that used the word in a different sense. In addition, we found that the effects of sense consistency in polysemy are similar to the effects of meaning consistency in homonyms. Finally, using a word in the same sense was a good memory cue, but using the word in a different sense was not.

A straightforward core meaning view is easily disconfirmed by our data, because it would not predict any difference between using a word in the same or a different sense. And this view would certainly predict that there should be a qualitative difference between homonyms (which have no core meaning) and polysemous words (which do), yet we found no such difference. In short, words like *paper* cannot be represented by a single semantic description that is accessed every time ("flat sheets made of wood pulp" or the like).

One can elaborate a core meaning view in a number of ways that might make it more consistent with the present results. Caramazza and Grober (1976) argued for a core meaning that represents polysemous words but also suggested that the lexicon contains a set of "instruction rules," analogous to derivation rules in the grammar, that could produce more specific senses from the core. If one interprets this proposal to mean that every distinct sense has an instruction rule that is stored in order to generate that sense from the core meaning, then this proposal is actually not a core meaning account any more, since the instructions would be separate representations of each sense. Such rules would have to be quite specific in order to derive the senses they found for *line*, such as "rope" or "line of business," from the core geometric meaning, so they could not be rules that apply widely across the lexicon.

An alternative more consistent with the core meaning view might instead make a distinction between the permanent representation in lexical memory and more temporary, episodic constructions of word senses (Anderson & Ortony, 1975). That is, perhaps a word like *paper* has a single (core) representation in the lexicon, but when it appears in a specific context, a more detailed context-appropriate sense is constructed in the sentence representation. One could then argue that this sense is what is influencing the interpretation of later uses of the word. For example, once one has used the context to interpret *paper* to mean a daily publication of news, one can then interpret the subsequent appearance of *paper* in the same way much more easily. There are four major problems with this proposal. First, it is not clear how this view could explain the inhibition of a different sense, as was found in Experiment 5, since specific senses must be represented in order to be suppressed. Second, this view cannot explain why the consistency effect was at least as strong for polysemous and homonymous words. Since it assumes that the specific senses of a polysemous word are created on the fly from a shared core meaning, but different meanings of homonyms are stored separately (since they cannot be derived), it should predict a greater consistency effect for homonyms. Experiments 3 and 3A did not find this result. Third, the sole representation of a core meaning seems psycho-

logically implausible if specific senses are frequent. That is, it would be very surprising if people often used the word *paper* to refer to newspapers and to published articles, yet did not represent this fact in the lexicon, but derived it from general principles on every occurrence. Finally, linguists have argued that the details of word usage are not in fact derivable from a core meaning and that conventional senses must therefore be represented in the lexicon (Lehrer, 1990; Rice, 1992). In sum, viewing senses as episodic constructions does not save the core-meaning theory.

As mentioned in the introduction, there are no detailed psychological models that attempt to explain how polysemous words are represented and processed. The present results speak to a number of the possible models one could develop. Perhaps the main empirical result is the finding that different senses have little functional overlap—about the same as the unrelated meanings of homonyms. (We attempt to explain this finding in the next section.) Obviously, this creates problems for a core view. However, it also greatly limits all the possible core-plus-senses views. If one believes that there is a core meaning that is common to all uses of a word, or at least a prototypical meaning that is accessed first, then processing the word in one sense should provide some benefit to processing it in another sense. However, in both a memory task (Experiment 1) and a semantic judgment task (Experiments 3 and 3A), we found that cross-sense priming was minimal. In the memory task, the different-sense use of a word did not provide a reliable memory cue. In the sensicality task, the effect of switching senses was virtually the same as the effect of switching meanings of homonyms. These results argue that if there is a common core to words, it has little content and little effect on processing. The finding that cross-sense priming is inhibitory (Experiment 5) also provides evidence that sharing a core meaning is not sufficient to benefit comprehension. As a matter of parsimony, the present results do not give us any reason to believe that there is a core meaning, so separate representations of each sense appear to be a better bet than the core-plus-senses view.

Reasons for Minimal Sense Overlap

As we noted in the Introduction, polysemous senses are (by definition) related, whereas homonyms are not. Why, then, did we find no difference in the consistency effect between polysemous words and homonyms? Or, put more generally, why do our results show so little overlap in polysemous senses if the senses are related, often by productive relations? One might wonder whether there is some error in the methods that have given this counterintuitive result.

In her thesis, Klein (2000) explored the relations between polysemous senses in a variety of conceptual tasks (rather than online comprehension tasks) in order to better understand the relations among the senses of polysemous words. Her results confirmed the low overlap of different senses in a number of different tests of conceptual coherence. For example, subjects strongly preferred to sort together different words from the same superordinate category over the same word used in two different senses, and the induction of a property from one sense of a word to a different sense was very low. Interestingly, the scores for homonyms were slightly, but significantly, lower in the same tasks. For example, she found that subjects sorted different senses of a polysemous word together only 14% of the time in one experiment, but this was reliably higher than the 6.6% for the different meanings of homonyms.

These results show that naive subjects see little conceptual commonality in the different senses of a word, although it is very slightly more than that seen in the completely unrelated homonym meanings. How can such results be reconciled with the fact that senses are semantically and historically related? To explain this, we need to distinguish semantic *relatedness* from semantic *overlap*, which we have been using somewhat interchangeably prior to this point. The different senses of polysemous words are clearly related. In some cases, a strong argument can be made for how one sense developed historically from another one (Clark & Clark, 1979; Malt et al., 1999; Murphy, 1997; Sweetser, 1990). Although such cases must necessarily be "related," a relation does not make

senses similar. For example, if one uses the word *paper* to refer to sheets of writing material, one might find it very natural to extend the word to refer to the content written on the paper (*His paper was boring*.). However, the content is not similar to the sheets: One is information, the other is made of wood pulp; one has semantic content, the other has molecular structure; one has an author, the other has a manufacturer; and so on. As a result, wrapping paper and a liberal paper do not have much semantic overlap, even if the relation between the two is easily understood. When the word is extended even further (e.g., *paper* meaning the publisher of a newspaper, as in *The paper fired half its reporters*), the overlap is even less. If priming and memory cueing in these tasks (and sorting and induction in Klein, 2000) primarily result from shared semantic features, then even obviously related senses may not overlap in meaning enough to produce performance advantages.

If this argument is correct, then it converges on our previous conclusions about what a core meaning might be, for those views that include some kind of core representation. Clearly, the core cannot be a basic component of meaning that is shared by all or most of the senses. If *paper* can refer to a kind of a material but also the content printed on that material—and, indeed, the content minus the material as when one presents a paper verbally, or perhaps using a computer projection system—then the senses are related without having overlapping representations. So, if there is a core sense, it would probably be a prototypical or perhaps historically early sense that is not *shared* by the other senses but instead is the original basis from which the other senses were historically derived (although it seems likely that some senses are in turn derived from other derived senses; Malt et al., 1999; Murphy, 1997). Such a core would not play a critical part in the online processing of the word, although it might be important in linguistic analysis.

To summarize, our proposal is that linguistic analyses arguing that polysemic senses are related (e.g., Nunberg, 1979) are correct. Nonetheless, the different senses of a word may not be very similar, and it is this variable that determines priming in the present tasks. If this is correct, it is not surprising that the senses are stored separately in the mental lexicon.

Limitations on the Present Results

One question about these results is how far discriminations between senses can be drawn. We chose senses of words that were fairly distinct, so that we could select each sense with a single-word modifier. However, some forms of polysemy involve very subtle distinctions. For example, Nunberg (1979) discusses the type/token polysemy found in almost all nouns in which a word can refer to an individual or a class. The word *dog* can refer to an individual (*The dog dug up my flowers*) or to the entire class (*The dog has been domesticated for centuries*). Since the individual is a member of the class, generally having all the properties typical of the class, these two senses have very similar content. Although there may be an important ontological difference between an individual and the class it is in, when describing dogs as a class (domesticated four-legged mammals, usually with fur, etc.), the description also applies to a typical individual in that class (a particular domesticated four-legged mammal, with fur, etc.). Given this overlap in meaning, it might be quite difficult, if not impossible, to facilitate one of these senses while suppressing the other, as we found in Experiment 5. In contrast, the use of *dog* as a verb, meaning to follow and harass, seems quite different from the individual canine sense, so it should be possible to facilitate one while suppressing the other. This issue does not arise for homonyms, because their different meanings are unrelated. But a complete model of polysemy will have to explain which senses are enhanced and which are suppressed when a given sense is selected.

This issue leads to another critical question about polysemy, namely, how to tell if two different senses are indeed represented separately. In the limit, "the meaning of any word form is in some sense different in every distinct context in which it occurs" (Cruse, 1986, p. 51). Clearly, it is impossible (and probably undesirable) to represent each such distinction in the lexicon; yet the present data suggest that at least some mean-

ingful distinctions are indeed represented. Thus, it is likely that the mental lexicon represents some reasonable number of senses, rather than the extremes of one core sense or every possible semantic distinction. If two senses are only very subtly different, it seems unlikely that speakers will develop separate entries for them, since a single entry will suffice to specify most of the meaning for both. If two senses are strikingly different, then a single entry will probably be unsuccessful at representing both meanings, which will presumably lead to the formation of separate entries. Although this "in-between" proposal seems eminently reasonable, it is also rather vague. What is needed is a more specific model of what causes a sense to be separately represented, from which one could derive predictions about which uses would involve the same senses and which would involve different senses. Such a proposal is not yet to be found in the literature (though see Murphy, 1997, for some discussion) and is beyond the scope of the present article. However, one empirical demonstration that diverse uses could be reduced to a manageable number of senses can be found in Gibbs, Beitel, Harrington, and Sanders (1994), who examined 35 different senses of the word *stand*. Using hierarchical clustering, they found that three main categories could adequately represent the similarity among the 35 uses (i.e., account for substantial variance of the similarities). Thus, one might argue that people would be likely to represent three general senses of *stand* in order to make the distinctions among different uses. It would be interesting to test that grouping with the priming technique used here or the conceptual methods used by Klein (2000).

Some linguistic analyses of polysemy have proposed that there is no firm line to be drawn between polysemy and more usual ambiguity (see Geeraerts, 1993, and Tuggy, 1993, for one such discussion). Furthermore, some linguists have argued against the more basic assumption of separable senses. For example, Cruse (1986, pp. 71 ff.) argued that some words had ranges of senses that shaded imperceptibly from one to another, forming a *sense spectrum*. The difficult aspect of such spectra, he claimed, was that sometimes the endpoints of a spectrum had very different meanings and could not plausibly be counted as being part of the same sense, even though intermediate "adjacent" senses were so close as to not require separate entries. Thus, one is faced with arbitrarily dividing up the spectrum into distinct senses, in spite of the continuum of meanings, or conversely treating the very diverse meanings as all falling under one sense.

This awkward situation is a methodological puzzle, but it may not be so much of a puzzle for speakers to represent. For example, color is physically instantiated through a continuous set of dimensions, with no natural break between the colors we perceive. In order to represent color differences, the human visual system uses a few landmarks—three photoreceptors that differ in the frequencies they are most sensitive to—and any given color is identified in relation to those landmarks. The particular location of those landmarks (i.e., the frequencies the receptors are most sensitive to) is to some degree arbitrary, so long as they can triangulate the space that needs to be covered. Similarly, in order to represent a word like *paper* that has a wide range of possible senses (ranging from things made out of wood pulp to the management of a newspaper), one needs to establish enough specific senses so that most common meanings are covered—the specific senses that are represented may not be critical (see also Nunberg, 1979). If one first learns *paper* to refer to unlined sheets for writing, one may not need to develop a new sense referring to lined sheets, since such uses are easily understood by the pre-existing sense, and contextual addition of the lines is simple. But the reverse is also true: If one first learned *paper* to refer to lined sheets, one would probably not need to form a new sense to refer to the unlined sheets. Thus, the specific senses that are represented may be somewhat arbitrary, dependent on an individual speaker's learning history. What is principled is the process by which new senses are formed, based on the range of senses that a word expresses.

The development of a model of sense creating would require more detailed description of

the principles involved in establishing new senses. It is possible that the connectionist concept of local attractors may be helpful in developing such principles. That is, so long as a given use is sufficiently close enough to a local (semantic) attractor, there will be no need to form a new sense. But if a given sense is not semantically close enough to an established attractor, a new set of distinctions will have to be learned with experience. Indeed, in his article primarily about homonymy, Kawamoto (1993, pp. 511–512) makes just such a proposal for polysemy. On his account, a sense will become separately represented when it is frequent enough or different enough from existing senses to develop its own local minimum in the semantic space. If there are a number of about equally frequent senses that are only slightly different from one another, this will not result in separate representations but rather a fairly wide "basin" in semantic space. All such uses would be quickly understood, as listeners would not be attempting to distinguish different senses.

Kawamoto's (1993) proposal does not have explicit inhibition of meanings, so how could it explain the interference found in our Experiment 5? The negative effects of using a word in a different sense could be attributed to the difficulty of having to climb out of the old (incorrect) local minimum and into the correct one. That is, if one reads the word *paper* used to refer to a newspaper, connection weights would be modified so that subsequent encounters of the word would be more likely to go to that location in semantic space. When the word is read again with a different sense, contextual information will have to overcome this tendency, thereby resulting in longer comprehension times (i.e., longer times to reach the correct local minimum). Such an explanation implies that the senses used in our studies were in fact represented as separate local minima rather than in a wide basin, so it is also consistent with our conclusion that separate senses are represented in the lexicon. However, one advantage of this view, which Kawamoto points out, is that it does not require us to specify in advance which senses are separate, as the system discovers this itself by virtue of the semantic similarity structure of the entire set of uses of the word. It would be particularly interesting to see if such a model could account for the similar results for polysemous senses and homonym meaning we found.

In conclusion, polysemous senses are semantically related but are not very similar, resulting in same-sense facilitation and cross-sense inhibition. If polysemous words do have a core meaning, it cannot be a substantial semantic component that is common to all the senses of a word. It remains for a future paper to illuminate the exact form of the sense representations, to explain how new senses are added to a lexical representation, and to construct a process model of understanding polysemy. We look forward to that paper, be it a talk, a journal article, or a news story.

APPENDIX: STIMULI

Polysemous items in all experiments	Homonym items in Experiments 3 and 4	Polysemous items added in Experiment 5
book	arms	glasses
tin	ball	class
run	bat	box
see	bridge	sign
paper	calf	trunk
chicken	date	drive
television	poker	
sheet	fan	
corn	pool*	
oak	post*	
shower	bank	
drinker	pen	
atmosphere	ring	
coat	band	
cold	mint	
fortune	mold	
hall	pitcher	
letter	plot	
nail	pit	
navy	match	
orange	seal	
production	second	
cotton	tick	
filling	toast	
	bail**	
	net**	

Single asterisks indicate items used in Experiments 3 and 4 only. Double asterisks indicate items used in Experiments 3A and 4A only.

REFERENCES

Anderson, R. C., & Ortony, A. (1975). On putting apples into bottles—A problem of polysemy. *Cognitive Psychology*, **7**, 167–180.

Bainbridge, J. V., Lewandowsky, S., & Kirsner, K. (1993). Context effects in repetition priming are sense effects. *Memory & Cognition*, **21**, 619–626.

Caramazza, A., & Grober, E. (1976). Polysemy and the structure of the subjective lexicon. In C. Rameh (Ed.), *Georgetown University roundtable on languages and linguistics. Semantics: Theory and application* (pp. 181–206). Washington, DC: Georgetown Univ. Press.

Clark, E. V., & Clark, H. H. (1979). When nouns surface as verbs. *Language*, **55**, 767–811.

Clark, H. H., & Gerrig, R. J. (1983). Understanding old words with new meanings. *Journal of Verbal Learning and Verbal Behavior*, **22**, 591–608.

Cohen J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, **25**, 257–271.

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

den Heyer, K., Taylor, M. A., & Abate, A. (1986). On the nature of neutral primes in a lexical decision task. *Psychological Research*, **48**, 161–168.

Deane, P. D. (1988). Polysemy and cognition. *Lingua*, **75**, 325–361.

Durkin, K., & Manning, J. (1989). Polysemy and the subjective lexicon: Semantic relatedness and the salience of intraword senses. *Journal of Psycholinguistic Research*, **18**, 577–612.

Fauconnier, G. (1985). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge, MA: MIT Press.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Geeraerts, D. (1993). Vagueness's puzzles, polysemy's vagaries. *Cognitive Linguistics*, **4**, 223–272.

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.

Gernsbacher, M. A., & Faust, M. (1991). The role of suppression in sentence comprehension. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 97–128). Amsterdam: North-Holland.

Gerrig, R. J. (1986). Process and products of lexical access. *Language and Cognitive Processes*, **1**, 187–195.

Gerrig, R. J. (1989). The time course of sense creation. *Memory & Cognition*, **15**, 55–71.

Gerrig, R. J., & Murphy, G. L. (1992). Contextual influences on the comprehension of complex concepts. *Language and Cognitive Processes*, **7**, 205–230.

Gibbs, R. W., Beitel, D. A., Harrington, M., & Sanders, P. E. (1994). Taking a stand on the meanings of "stand": Bodily experience as a motivation for polysemy. *Journal of Semantics*, **11**, 231–251.

Heine, B. (1992). Grammaticalization chains. *Studies in Language*, **16**, 335–368.

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.

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

Klein, D. E. (2000). *"Paper has been my ruin": On the organization of polysemous senses*. Unpublished Ph.D. dissertation, Psychology Department, University of Illinois at Urbana-Champaign.

Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago: Univ. of Chicago Press.

Langacker, R. W. (1987). *Foundations of cognitive grammar* (Vols. 1 & 2). Stanford: Stanford Univ. Press.

Lehrer, A. (1990). Polysemy, conventionality, and the structure of the lexicon. *Cognitive Linguistics*, **1**, 207–246.

Levelt, W. J. M. (1989). *Speaking*. Cambridge, MA: MIT Press.

Light, L. L., & Carter-Sobell, L. (1970). Effects of changed semantic context on recognition behavior. *Journal of Verbal Learning and Verbal Behavior*, **9**, 1–11.

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language*, **40**, 230–262.

Murphy, G. L. (1991). Meanings and concepts. In P. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 11–35). Hillsdale, NJ: Erlbaum.

Murphy, G. L. (1997). Polysemy and the creation of novel word meanings. In T. B. Ward, S. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 235–265). Washington, DC: American Psychological Association.

Nunberg, G. (1979). The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy*, **3**, 143–184.

Perfetti, C. A., & Goodman, D. (1970). Semantic constraint on the decoding of ambiguous words. *Journal of Experimental Psychology*, **86**, 420–427.

Rice, S. A. (1992). Polysemy and lexical representation: The case of three English prepositions. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 89–94). Hillsdale, NJ: Erlbaum.

Ruhl, C. (1989). *On monosemy: A study in linguistic semantics*. Albany: State Univ. of New York Press.

Simpson, G. B. (1994). Context and the processing of ambiguous words. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 359–374). San Diego: Academic Press.

Sweetser, E. E. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge Univ. Press.

Swinney, D. A., & Hakes, D. T. (1976). Effects of prior context upon lexical access during sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, **15**, 681–689.

Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential context. *Journal of Memory and Language*, **27**, 324–340.

Tabossi, P., & Zardon, F. (1993). Processing ambiguous words in context. *Journal of Memory and Language*, **32**, 359–372.

Tuggy, D. (1993). Ambiguity, polysemy, and vagueness. *Cognitive Linguistics*, **4**, 273–290.

Van Petten, C., & Kutas, M. (1987). Ambiguous words in context: An event-related potential analysis of the time course of meaning activation. *Journal of Memory and Language*, **26**, 188–208.

Williams, J. N. (1992). Processing polysemous words in context: Evidence for interrelated meanings. *Journal of Psycholinguistic Research*, **21**, 193–218.

Zgusta, L. (1971). *Manual of lexicography*. The Hague: Mouton.

(Received July 28, 1999)
(Revision received November 25, 2000)
Published Online June 21, 2001