Available online at www.sciencedirect.com
SCIENCE DIRECT
Cognitive Brain Research 24 (2005) 57–65
COGNITIVE BRAIN RESEARCH
www.elsevier.com/locate/cogbrainres

Research report

The effects of homonymy and polysemy on lexical access: an MEG study

Alan Beretta[a,*], Robert Fiorentino[b,c], David Poeppel[b,c,d]

[a]*Department of Linguistics, Michigan State University, A-628 Wells Hall, East Lansing, MI 48824, USA*
[b]*Cognitive Neuroscience of Language Laboratory, University of Maryland, College Park, MD 20742, USA*
[c]*Department of Linguistics, University of Maryland, College Park, MD 20742, USA*
[d]*Department of Biology, University of Maryland, College Park, MD 20742, USA*

Accepted 3 December 2004
Available online 4 February 2005

Abstract

We examined the neural correlates of semantic ambiguity by measuring changes in MEG recordings during a visual lexical decision task in which the properties of ambiguous words were manipulated. Words that are ambiguous between unrelated meanings (like *bark*, which can refer to a tree or to a dog) were accessed more slowly than words that have no unrelated meanings (such as *cage*). In addition, words that have many related senses (e.g., *belt*, which can be an article of clothing or, closely related in sense, a fan belt used in machinery) were accessed faster than words that have few related senses (e.g., *ant*). The findings are inconsistent with accounts that posit that both kinds of ambiguity involve separate lexical entries, but instead offer both behavioral and neurophysiological support for separate entry accounts only for homonymy, and a single-entry model of polysemy. The findings also provide neural correlates for a behavioral study of lexical ambiguity that demonstrated that the frequently reported ambiguity advantage in lexical decision tasks is not due to the presence of many unrelated meanings but to the presence of many related senses.
© 2004 Elsevier B.V. All rights reserved.

Theme: Neural basis of behavior
Topic: Cognition

Keywords: MEG; Language; Ambiguity; Lexical access; Lexical decision

1. Introduction

The issue of lexical ambiguity has been of great interest because it addresses foundational issues regarding the nature of the mental lexicon and lexical access. There exist rich behavioral and theoretical linguistic literatures on ambiguity and the nature of the lexicon. This extensive work, in combination with recent electrophysiological (EEG, MEG) studies that have identified components for visual word recognition, makes possible, and forms the basis for, a range of behavioral and neurological predictions about lexical access, in general, and lexical ambiguity in particular. At issue is the representation of lexical knowledge in the human brain; in this regard, the nature of the processing of ambiguous words becomes critical for adjudicating among theories of lexical representation.

A frequently reported and standardly accepted phenomenon in the behavioral literature on lexical access is the so-called ambiguity advantage. It has been found in a number of studies [1,3,14,15,25] that in visual lexical decision tasks, ambiguous words yield faster reaction times than unambiguous words. In attempts to account for these findings, it has been typically assumed that what needs to be explained is why words which have unrelated meanings should confer a processing benefit, for example Ref. [3]. But this assumption is not warranted. Ambiguity can arise in different ways, and

* Corresponding author. Fax: +1 517 432 2736.
E-mail address: beretta@msu.edu (A. Beretta).

0926-6410/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.cogbrainres.2004.12.006

by far, the least common type of ambiguity is the type that is based on unrelated meanings. Far more pervasive is ambiguity that is due to multiple related senses. (For a count that demonstrates this disparity, see Ref. [31]; also, see Ref. [17] for a detailed explanation as to why different related senses should so outnumber unrelated meanings; and Ref. [4] for a count of the number of senses available for a single word—26 for the word line.) We will see that it is this type of ambiguity that actually provides the processing advantage.

Traditionally, ambiguity between unrelated meanings is known as *homonymy*, where two words happen to share the same orthography and phonology. An example is *bank*, which can be the side of a river or a financial institution. How homonymy comes about (historical relations or accidental coincidence of orthography and phonology) is of no consequence to models of lexical access. Ambiguity between related senses is known as *polysemy*. Consider the word *door* in the following sentences.

(1) The door fell off its hinges.
(2) The child ran through the door.

The sense of *door* in (1) is clearly different from the sense of *door* in (2). In (1), *door* is a physical object, whereas in (2), it is an aperture. As a further example, let us take another look at the homonymous word *bank*. One of its meanings (financial institution) can be seen to possess different polysemous senses:

(3) The bank apologized to its customers.
(4) The bank was destroyed in an earthquake.

In (3), *bank* is an institution; in (4), a building. (For many examples of different kinds of polysemy and a well-known theoretical account, see Pustejovsky [27]).

So ambiguous words can be either homonymous or polysemous, and it is also possible for one or more meanings of homonyms to be polysemous. In addition, the number of polysemous senses a word may have can vary a great deal. These simple facts are directly relevant to the design of the present study.

A recent lexical decision study (Rodd et al. [31]) addressed both kinds of ambiguity. This study manipulated homonymy and polysemy in 2 $\times$ 2 factorial design and found that the familiar ambiguity advantage was entirely due to polysemy, and that homonymous words, far from conferring any processing benefit, actually delay access. Many related senses help, but many unrelated meanings hurt. This was a rather dramatic finding because the homonymous word disadvantage was precisely the opposite of what most accounts of lexical access assumed. The fact that having many senses was responsible for the well-known ambiguity advantage had not been properly recognized before.

This study raises anew fundamental questions with respect to the two classes of ambiguity that have been debated for many years. While it is standardly assumed that unrelated homonymous meanings must be separately represented in separate lexical entries, it is not obvious that the same must apply to related polysemous senses. A central question that arises is therefore: do polysemous senses have separate entries just as homonymous meanings do? Pustejovsky [27] argues that listing senses in separate entries is an entirely inadequate means of accounting for polysemy, but he observes that other theorists nevertheless do assume enumerated senses both for nouns and for variation in verb complementation. There are, however, several critical disadvantages in assuming a multiple separate entry model for polysemy, according to Nunberg [23], Pustejovsky [27], and others, among them their failure to capture the logical relationship between senses, the fact that words can take on new senses more or less on the fly, and the fact that a single sense can have many syntactic realizations [27].

The alternative to a lexicon in which different senses are stored as separate words is one in which each polysemous word is stored as a single lexical entry. One instantiation of this idea that has been proposed in a variety of guises by numerous researchers is that polysemous words are stored as a single 'core' meaning. For example, Nunberg [23] argued that some senses can be derived from others; thus, it is possible to derive *chair* as a token (the chair was damaged) from *chair* as a type (the chair was an unknown fixture in Smith's house), or *chicken* as a meat from *chicken* as a bird. So, a polysemous word like *window* would be assumed to have "only one conventional use, with the other normal uses generated pragmatically" (p. 153). Nunberg's linguistic perspective on polysemous sense extension is clearly consistent with the idea that all that is represented is a core meaning. A similar proposal was made by Lehrer [22], who shared the view that polysemy can be largely explained through rules of meaning extension. Caramazza and Grober [4] make the direct claim that a core meaning is what is stored in a psychological representation of a word with many senses.

An account that is rather different in specifics but which is still consistent with the idea that a single entry suffices for polysemous words is offered by Chomsky [5]. On this account, two senses for *book*, which can denote a physical object or its contents, are represented (disjunctively as either $\pm$ physical object, according to Nunberg) within a single lexical entry. The advantage of this kind of proposal, according to Nunberg, is that, among other things, it preserves the requirement that syntactic rules governing anaphoric coreference operate on the 'same lexical item' ([23], p. 151). This matters when one considers sentences such as (5):

(5) The newspaper has decided to change its format.

In (5), syntactically, *the newspaper* and *its* share identical reference. However, *the newspaper* here denotes a publisher, whereas *its* denotes a publication, two different senses. By assuming single lexical entries for polysemous words, the notion of syntactic identity need not be compromised.

Yet another account that is consistent with a single-entry model of polysemy can be found in the work of Frisson and Pickering [9,10]. They showed that different senses were equally easy to process. They proposed that readers initially adopt an underspecified interpretation, choosing between senses later in processing. This notion of sense underspecification implicates a single-entry.

Suffice to say, there are a number of proposals that are more or less compatible with the idea of a single-entry model of polysemy. But are there some polysemous senses that might require separate entries? As Nunberg ([23], p. 142) and others have noted, "polysemy is a gradient phenomenon." Klein and Murphy ([17], p. 569) elaborate: "It is widely recognized that polysemous senses range from nearly identical to nearly unrelated." In a series of experiments [16,17], it was demonstrated that when subjects were presented with a word like *paper* in one sense (in the phrase *wrapping paper*), this slowed down later processing of the same word in a 'nearly unrelated' sense (in the phrase *daily paper*). Further, subjects did not judge that these 'nearly unrelated' senses referred to the same sorts of things. So, on an extreme of what could be considered polysemy, subjects treat senses as they do unrelated homonyms, which were also slowed relative to unambiguous words.

Do these findings mean that single-entry models of polysemy must be abandoned? There would seem to be little motivation for this. After all, as Klein and Murphy ([17], p. 568) themselves point out, "It is important not to exaggerate the separation of polysemous senses in our results. For example, we found that more similar senses were sorted together…In addition, we chose polysemous senses that were clearly distinct in meaning…did not use type-token polysemy which naïve subjects might not even identify as being different senses. Nor did we use subtle differences…in which different aspects of the same word are emphasized depending on the perspective of the speaker." So, by their own assessment, most kinds of polysemy either were not tested or behaved in a way consistent with a single-entry model. What their findings do highlight, however, is that polysemous senses can be unrelated to the extent that they are not obviously distinct from homonymy and it is possible that, in this gray area alone, a single-entry model cannot be posited.

In this connection, it is worth noting that, in Nunberg's [23] estimation, many forms of polysemy in English exhibit the same patterns in other languages; so, for example, he can think of no language in which *window* does not occur in both its aperture and physical object senses. By contrast, *paper* does not have the sense of *newspaper* in any language other than English (so far as we have yet been able to determine). This suggests that the kinds of nearly unrelated senses investigated by Klein and Murphy [16,17] may be qualitatively different from other kinds of polysemy.

It is reasonable to conclude, then, that a single-entry model of polysemy is a viable possible alternative to positing separate lexical entries for each sense. We will not seek to distinguish all of the different single-entry proposals we have discussed. In attempting to address an area of inquiry as complex as polysemy is, it is impossible to tease apart all of the relevant factors in a single experiment. What we wish to establish first is whether a single-entry proposal is a good characterization of polysemy at all.

Let us return now to the Rodd et al. [31] study and consider its implications for single-entry and separate-entry models of polysemy. In addition to the surprising finding that the ambiguity advantage is due to many related senses, the Rodd et al. finding appears, at first, to pose a problem for accounts that assume that each related polysemous sense entails a separate representation allowed by a separate lexical entry. If polysemous senses did involve separate entries of some sort, it would be expected, other things being equal, that they would be processed in a manner very similar to homonyms, which uncontroversially involve separate entries for each of their unrelated meanings. Yet Rodd et al. demonstrate that polysemy and homonymy produce very different processing responses, inconsistent with the idea of a separate-entry model of polysemy, but consistent with a single-entry account.

However, the evidence this study offers for a single-entry account could have originated in later stages of (postlexical) processing. Lexical decision can tap into a post-lexical-access stage of processing but is not necessarily informative about initial stages of lexical processing, such as spreading activation. Considering the lexical decision results alone, then, it is possible that at an early stage of lexical access, there is no processing difference between homonymy and polysemy, which would be consistent with a separate-entry account, and that the difference reported by Rodd et al. is a late-emerging difference (perhaps due to semantic competition in the case of homonymous words but not in the case of polysemous words). A way of pursuing the possible early-late distinction in lexical processing is suggested by the recent MEG literature, to which we now turn.

An important MEG lexical access experiment by Embick et al. [7] found that the first neuromagnetic evoked component whose peak latency shifted systematically with the frequency of the lexical stimulus was at 350 ms (the M350). In a lexical decision task, the M350 was earlier for frequent words than for infrequent words. In this study, the M350 frequency advantage correlated with the RT frequency advantage. Here, even though the M350 occurs several hundred milliseconds earlier than the RTs, there was no distinction between the two in the nature of the effect. What was true of the earlier M350 stage was also true of the later RT stage. Thus, with this result by itself, it was not possible to tell if the M350 component reflects initial lexical activation or later processing. Since this initial study, however, the MEG literature has reported a striking dissociation between M350s and RTs [28]. In a lexical decision study, Pylkkänen et al. demonstrated that "M350 latencies vary independently from reaction times when stimuli are simultaneously varied along a dimension that affects lexical activation and a dimension that affects selection/decision" ([28], p. 11). The dissociations observed between M350s and RTs with these manipulations supported the hypothesis that the M350 is an index of initial lexical activation and not of postlexical processing. Given the demonstration in the study by Pylkkänen et al. [28] (see also [29]) that a dissociation between M350 and RT is measurable, it becomes possible to test other hypotheses which predict a dissociation between MEG responses and RTs.

Returning to the separate-entry account, if it is correct, it should be the case that at the initial stage of lexical access indexed by the M350, polysemy and homonymy will behave in the same way, unlike the RT behavior in which polysemy and homonymy are different. That is, there should be a dissociation between M350 and RT. On the other hand, if separate-entry accounts of polysemy are mistaken, then homonymy should yield distinct processing behavior from polysemy, and the M350 results should parallel the RT results. Thus, the present study aims primarily to test the competing separate-entry and single-entry hypotheses. To this end, the design and stimulus materials of Rodd et al. [31] were chosen for our MEG study.

2. Methods

2.1. Participants

Subjects were 19 right-handed, English-speaking volunteers with normal or corrected-to-normal vision, 13 females, aged 18–31 years. All subjects were students at the University of Maryland and gave their written informed consent to take part in the experiment. They were paid $20 for their participation.

2.2. Stimuli

The experimental stimuli and design that were used were those in Appendix B and Experiment 2 of Rodd et al. [31]. 128 lexical items were included in a 2×2 factorial design, the two factors being homonymy (single meaning (non-homonyms) vs. more than one meaning (homonyms)) and polysemy (many senses vs. few senses). Words were considered to have more than one meaning (homonyms) if they had two or more entries in the Wordsmyth dictionary [24], and to have a single meaning (non-homonyms) if they had only a single Wordsmyth entry. Measures of the number of senses were based on the total number of senses for all entries of each word in the Wordsmyth dictionary and the total number of senses for each word in the WordNet lexical database [8]. This yielded 32 items in each of the four cells of the design: (i) single-meaning (non-homonym), few senses, for example, *ant* (one Wordsmyth entry and one sense) (ii) single-meaning (non-homonym), many senses, for example, *mask* (one Wordsmyth entry and 11 senses) (iii) more than one meaning (homonym), few senses, for example, *calf* (2 Wordsmyth entries and 3 senses, and (iv) more than one meaning (homonym), many senses, for example, *bark* (2 Wordsmyth entries and 10 senses). (To avoid confusion, it is worth stressing that using a dictionary merely provides us with an independent numerical measure of homonymy/non-homonymy and many/few senses. The fact that many dictionaries happen to list homonyms as separate entries but list different senses within the same entries simply reflects the preferences of lexicographers, but of course this in no way prejudges the issue the present study seeks to investigate, that is, whether different senses also involve separate entries).

Each of the four cells was matched for frequency using the CELEX database [2], number of syllables, concreteness ratings, and familiarity ratings. Words that were different from each other by a single letter were not significantly different between groups.

Words were pseudorandomly split into four lists, each with the same number of items from each of the four stimulus conditions. The order of the four lists was randomized across subjects, and for each subject, the order of items within lists was also randomized.

Nonwords were pseudohomophones matched for length with the word stimuli. The ratio of words to nonwords was 1:1.

2.3. Procedure

Subjects were placed horizontally in a dimly lit magnetically shielded room (Yokogawa Electric Corporation, Tokyo, Japan). Stimulus presentation was carried out using Psyscope [6]. For each trial of word or nonword stimuli, subjects were presented with a fixation point projected onto the center of a rear-projection screen for 500 ms. This was followed by presentation of the stimulus. Presented stimuli subtended 1.4° of visual angle vertically and 3.5° horizontally, based on an average of the shortest words (3 letters) and the longest words (7 letters), range 2.3° to 4.6°. Subjects were instructed to decide whether each stimulus item was a real word or not (lexical decision), and to respond as quickly and accurately as possible. Word decisions were made by button press with the right hand, nonword decisions with the left hand. Once the subject responded, the fixation point returned following an intertrial interval which varied pseudorandomly from 500 to 1000 ms at 50 ms intervals. Accuracy and reaction times were recorded.

Subjects were first given a practice session of 64 items to help familiarize them with the task. Also, the four lists were presented in separate blocks, and each block commenced with 10 stimulus items not included in analysis. There was a brief break for subjects between each stimulus block.

2.4. Recording and analysis

MEG recordings were conducted using a 160-channel axial gradiometer whole-head system (Kanazawa Institute of Technology, Kanazawa, Japan). The sampling rate was 1000 Hz, and data were acquired continuously with a band width of 1.0 Hz to 200 Hz. Data were noise-reduced prior to analysis to remove external sources of noise artifacts. Epochs with artifacts exceeding ± 2 pT in amplitude were removed before averaging; more than 97% of epochs survived this procedure. Signals for each stimulus condition were averaged. Following averaging, data were baseline corrected using a 100-ms prestimulus interval and were low-pass filtered at 20 Hz.

Reaction times were measured from the onset of the visual stimulus until button press. Incorrect responses and responses longer than 1200 ms were removed from analysis of both behavioral and MEG data. Data from one subject were not included in analysis because of an error rate of more than 10%. Of the remaining 18 subjects for reaction time analysis, the overall error rate for responses was 4.58%, ranging from 1.17% to 8.98% for each subject. Responses longer than 1200 ms resulted in 2.39% of the data points being removed from the analysis. In total, incorrect and overly slow responses resulted in the loss of 6.97% of the trials.

In the analysis of MEG data, the averaged signals were examined to find the dipolar field distributions which characteristically appear in the following time ranges: 100–220 ms (M170), 200–300 ms (M250), and 300–420 ms, (M350). Components in these time ranges have been reported in a number of studies of visual word or character recognition (see [12,13,18–21,32]). Based on the previous literature concerning the functional significance of these components, and on their presence across conditions in the current data set, our analytical method involved evaluating the differences in these components due to the stimulus manipulation. For this reason, we analyzed the subjects' data only if these components appeared in the characteristic temporal range and source distribution, and only if these components were clearly identifiable across conditions within a subject. Three participants lacked a clearly identifiable M350 component in one or more of the four target conditions, and three additional participants lacked one or two of the early components in one or more of the four target conditions. Thus, the six subjects who did not meet our strict inclusion criterion were eliminated from analysis (along with the subject who was eliminated for an excessive lexical decision error rate). The data from 12 subjects was carried forward for statistical analysis. For one subject, two M170 latency/amplitude values were replaced with the condition mean in the statistical analysis.

For each subject, five sensors in the source (outgoing magnetic field) and five sensors in the sink (ingoing magnetic field) were chosen for the M350 component. The selection was made based on which sensors best captured the left hemisphere dipolar fields for all four stimulus conditions. The latency and amplitude of the first root mean square (RMS) peak for each stimulus condition across the 10 chosen sensors were recorded and used in data analysis. Without altering the sensors, the M250 and M170 latencies and amplitudes were recorded and also entered into data analysis (see Fig. 1 for a comparison of two conditions contrasting in number of senses, for one participant). Our channel selection, optimized to quantify the M350, does not capture the M170 and M250 evoked fields in the best way;

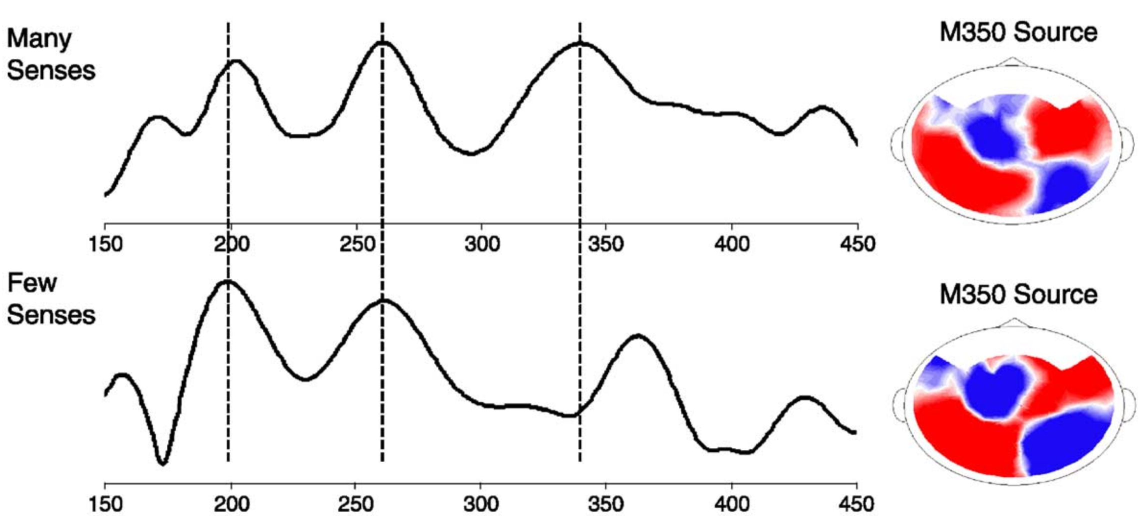

Fig. 1. M350 latency difference by senses: waveforms and M350 field distributions. The figure shows the latency difference at M350 for the contrast among many senses and few senses for one participant (many senses vs. few senses among words with more than one meaning (homonyms)). The waveforms are RMS-averaged waveforms from the 10 left-hemisphere channels selected for the M350 component, held constant across the two conditions, and plotted from 150 to 450 ms post-onset of the visual word. The vertical dashed lines show the peak latency for the M170, M250, and M350 for the condition in which it appears first. As can be seen, the M350 peaks earlier for many than few senses. Right: the source distribution at M350 for each condition. The red areas represent the outgoing portions of the source distribution, and the blue areas represent the ingoing portions of the source distribution.

however, these earlier responses were clearly evident and did permit unambiguous quantification. Because the hypothesis concerns the M350, a coarse assessment of these earlier components sufficed and we focus the detailed analysis on the M350.

3. Results

3.1. Behavioral results

Mean values of response time and accuracy were calculated across participants and items. The mean response time for each of the four stimulus conditions, and for the main effects of the two factors polysemy and homonymy, is given in Table 1.

3.2. Response time

ANOVA (2 Homonymy × 2 Polysemy) for response time showed a significant effect of polysemy by participants ($F(1,17) = 15.616$, $P < 0.002$), and a significant effect of polysemy by items ($F(1,31) = 4.325$, $P < 0.046$). Words with multiple senses (mean 617 ms, Standard Error 17) were responded to more *quickly* than words with few senses (mean 637 ms, SE 19). The effect of homonymy was also significant by participants ($F(1,17) = 7.832$, $P < 0.013$) and marginal by items ($F(1,31) = 3.508$, $P < 0.089$). Words with more than one meaning (homonyms) were responded to more *slowly* (mean 635 ms, SE 17) than words with only one meaning (non-homonyms) (mean 619 ms, SE 19). The interaction of homonymy and polysemy was not significant by participants or items ($F < 1$).

3.3. Accuracy

ANOVA (2 Homonymy × 2 Polysemy) for accuracy (proportion correct) showed a significant effect of polysemy by participants ($F(1,17) = 39.085$, $P < 0.001$) and by items ($F(1,31) = 11.332$, $P < 0.003$). Words with multiple senses were responded to more accurately (mean 0.986, SE 0.004) than words with few senses (mean proportion correct 0.938, SE 0.011). The effect of homonymy was not significant by participants ($F(1,17) = 1.\ 946$, $P < 0.182$) or by items ($F < 1$). Words with more than one meaning (homonyms) (mean 0.956, SE 0.011) were responded to with approximately equal accuracy to words with only one meaning (non-homonyms) (mean 0.968, SE 0.009). The interaction of homonymy and polysemy was not significant for accuracy by participants ($F(1,17) = 1.068$, $P < 0.317$) or items ($F < 1$).

The behavioral results for polysemy are comparable to the results reported in Rodd et al. [31]. Both Rodd et al. and the current study found a significant advantage for many senses in both response time and in accuracy. The current study and Rodd et al. did not find a response time or accuracy interaction among homonymy and polysemy, and like Rodd et al., the current study did not find a significant accuracy difference by homonymy.

The current study did find a significant response time difference by homonymy, however, whereas Rodd et al. report only a numerical difference (6 ms slowdown) for words with more than one meaning (homonyms) vs. words with one meaning (non-homonyms), which reached marginal significance only in an analysis of inverse response times by participants. Rodd et al.'s Experiment II (auditory task on a similar stimulus set) does report a slowdown (29 ms) for words with more than one meaning (homonyms) vs. words with one meaning (non-homonyms) which was significant by subjects and items, and a homonymy by polysemy interaction significant by subjects but not items.

Table 1
Mean response time (ms): factorial design and main effects

Homonymy	Polysemy		
	Few senses	Many senses	Mean
Single meaning (non-homonyms)	626	611	619
More than one meaning (homonyms)	648	622	635
Mean	637	617	

3.4. MEG results

Mean values of peak latency for the M350 RMS peaks were calculated across subjects. The peak latency differences at M350 are shown in Table 2.

3.5. M350 Latency

Words with multiple senses elicited an *earlier* peak (Mean Peak Latency 338, SE 5.76) than words with few senses (mean 352 ms, SE 6.12); ANOVA (2 Homonymy × 2 Polysemy) for peak latency of the RMS averaged M350 waveform ($F(1,11) = 4.018$, $P < 0.071$) by participants. Words with more than one meaning (homonyms) elicited *later* M350 peak latency (mean 354 ms, SE 5.96) than words with a single meaning (non-homonyms) (mean 336 ms, SE 5.7); ($F(1,11) = 3.514$, $P < 0.089$) by participants. The interaction of homonymy and polysemy was not significant in this analysis by participants. Planned direct comparisons of the main effects of Homonymy and of Polysemy yielded significant results on paired t tests (Polysemy: $t(23) = 2.071$, $P < 0.05$, two-tailed; Homonymy: $t(23) = 2.209$, $P < 0.038$, two-tailed) and on the Wilcoxon Signed Ranks test, a non-parametric test of mean differences suited for comparisons involving a small N (Polysemy: $P < 0.013$; Homonymy: $P < 0.033$).

We note that there was also a marginal effect of polysemy in M350 amplitude, in ANOVA by participants ($F(1,11) = 4.037$, $P < 0.071$). There was no effect of homonymy, ($F < 1$). The interaction was also marginal $F(1,11) = 4.12$, $P < 0.068$).

Table 2
Mean M350 latency (ms): factorial design and main effects

Homonymy	Polysemy		
	Few senses	Many senses	Mean
Single meaning (non-homonyms)	345	328	336
More than one meaning (homonyms)	359	349	354
Mean	352	338	

Words with a single meaning (non-homonyms) and few senses had a mean amplitude of 66fT (SE = 5.78); words with a single meaning (non-homonyms) and many senses had a mean amplitude of 94fT (SE = 10.1). Words with more than one meaning (homonyms) and few senses had a mean amplitude of 80fT (SE = 7.22) and words with more than one meaning (homonyms) and multiple senses had a mean amplitude of 76fT (SE = 6.64).

4. Discussion

The study reported in this paper sought to test competing accounts of the nature of the lexicon. One account, the separate-entry account, assumes that different polysemous senses are on a par with different homonymous meanings insofar as they involve separate lexical entries. Thus, other things being equal, processing of both kinds of ambiguity should be the same. The alternative, single-entry account, maintains that polysemous senses are different from homonymous meanings in that only the latter involve separate lexical entries. On this account, other things being equal, processing of both kinds of ambiguity should be distinct. An earlier RT study [31] had shown that homonymy and polysemy produced quite distinct responses, supporting a single-entry account. However, it was possible that sensitivity to ambiguity type was a late-occurring response and that, at an earlier stage of processing, homonymy and polysemy would behave in the same way, as would be consistent with separate-entry accounts. MEG permits recording of neural responses at earlier stages of lexical processing, and the MEG literature has shown that dissociations between neuromagnetic responses and RTs was possible [28,29]. Hence, the importance of MEG in addressing this possibility.

In the present study, the latency of RTs was significantly later (16 ms) for words with more than one meaning (homonyms) relative to words with one meaning (non-homonyms), and significantly earlier (20 ms) for words with many senses than for words with few senses. Thus, the behavioral results replicate the results of Rodd et al. [31].

The latency of the M350 component was significantly slower (14 ms) for words with more than one meaning (homonyms) than for words with one meaning (non-homonyms), and also significantly faster (18 ms) for words with many senses than for words with few senses. Thus, the MEG results mirror the behavioral results, but show that the effect holds hundreds of milliseconds prior to the reaction time judgments. It was possible that the MEG results and the RT results would show a dissociation but, in the event, both the behavioral and the neural results correlated. Polysemy and homonymy yield distinct processing profiles not only in behavioral responses occurring around 600–650 ms, but also in neural M350 responses occurring approximately 300 ms earlier. These results support a single-entry account of polysemy and, conversely, provide no support for a separate-entry account, that is, for an account of lexical ambiguity which claims that both homonymy and polysemy involve multiple lexical entries at some stage of processing.

The results offer an answer to the question that motivated the experiment reported here, but they raise another intriguing question, namely, why do homonymy and polysemy show the *direction* of effects that they do? That is, while there is a straightforward theoretical interpretation of the fact that homonymy and polysemy manifest distinct processing profiles, why should words with more than one meaning (homonyms) slow access relative to words with one meaning (non-homonyms), and why should words with many polysemous senses speed access relative to words with few polysemous senses.

In addressing these problems, we are mindful that linguistic approaches to polysemy typically never discuss processing issues, and Rodd et al, have shown that the psychological literature has typically conflated polysemy with homonymy. However, by conceiving of linguistic models of polysemy in terms of single-entry processing accounts, as we have done, it is possible to consider the relative effects of homonymy and polysemy in the time course of lexical processing.

First, let us consider the problem posed by the homonymy disadvantage. Why should there be longer latencies for words with more than one meaning (homonyms) compared to words with one meaning (non-homonyms)? Network models of word recognition (e.g., [11,26]) that implicate competition between words to activate meaning representations may be informative here since interference between the different meanings of homonyms ought to result in slower recognition than for single-meaning non-homonymous words. In this sort of model, "each word is represented as a unique pattern of activation across a set of orthographic/ phonological and semantic units" ([31], p. 247). Orthographic patterns of words are linked to more than one semantic pattern if a word is homonymous. When the network encounters an orthographic pattern of a homonymous word, both of its meaning representations will compete with each other. The consequence of this competition is that it will take longer to arrive at a stable activation pattern.

The competition explanation of the homonymy disadvantage is not confined to network models, but is raised in other accounts too. For example, Traxler et al. ([34], p. 542) comment that, "lexically ambiguous words can lead to difficulty in comparison to unambiguous words (...). Presumably, readers are at some level comparing alternative interpretations, and this process of comparison is in itself costly." (See also [30,33]).

While competition is certainly a possibility with regard to the processing of homonymy, there is another candidate given the design of the present study. Assuming words with more than one meaning (homonyms) do have separate entries, frequency alone could constitute an explanation of the homonymy disadvantage. The stimulus items were matched for form frequency, so each entry for a homonym must be less frequent than a single-meaning non-homonym entry. Since frequency is known to affect both RTs and M350 responses, it should be expected that homonyms matched for form frequency will slow access relative to non-homonyms.

However, when we turn to the problem of the many senses advantage (why words with more senses have a processing advantage over words with few senses), it becomes evident that frequency makes precisely the wrong predictions. If a word has two senses, and another word has 12 senses, on items matched for form frequency, the frequency of each sense of the word with 12 senses ought, on average, to be less frequent than each sense of the word with two senses. Following the logic used in the case of homonymous word disadvantage, it should be expected that words with more senses ought to elicit slower RTs and M350 responses than words with fewer senses, contrary to fact. In the case of the many senses advantage, then, frequency does not explain the results.

Why might frequency provide a possible explanation for the homonymous word disadvantage and yet predict a counterfactual few senses advantage? Note that frequency differs between words with more than one meaning (homonyms) and single-meaning non-homonymous words only if it is true that homonyms involve separate lexical entries. The corollary is that frequency explains nothing if there are not separate lexical entries (because form matching would not be compromised). If many senses do not involve separate entries, it is not at all surprising that frequency should have no effect. The present study has yielded data that are consistent with the claim that while homonyms have separate entries, polysemous words do not. The fact that frequency can explain the homonymy disadvantage but cannot explain the many senses advantage may be seen as further confirmation of this claim.

While it is true that either competition or frequency differences attendant on separate homonymous entries may explain the homonymy disadvantage, competition or frequency do *not* explain the many senses advantage.

What might explain the many senses advantage? As argued in [31], if we make certain plausible assumptions about the network models mentioned above, a partial explanation follows. So, if it is assumed that different senses implicate different word nodes (comparable to separate lexical entries), then a polysemous word with many senses ought to slow recognition just as the different meanings of homonyms do. But our results demonstrate that this cannot be the case. The alternative assumption, that many polysemous senses implicate a single word node (comparable to a single entry), the consequence should be that they are recognized as fast as words that have few polysemous senses, *but not faster*. The single-node (single-entry) assumption does some work for us. It explains why polysemous words are recognized faster than homonyms, but it falls short of explaining why words with many senses should elicit faster RT and M350 responses than words with few senses.

To attempt to explain the many senses vs. few senses advantage, Rodd et al. [31] consider several intuitive candidates, including the possibility that words with many senses may be semantically richer than words with fewer senses, or that words with many senses are used in a wider range of contexts than words with few senses and so develop context independent representations.

To explore these intuitions, it would be useful to link them to one of the theories that involves single entries for polysemous words in very specific ways, and to examine very particular kinds of polysemy. Since the present study was not designed to tease apart the various theoretical models of single-entry polysemy, further experimentation is indicated. However, what the present study *has* addressed is the prior question of whether or not a single-entry lexical model of polysemy is a viable proposition. The findings in this respect are rather clear: this study provides firm behavioral and neural support for a single-entry model of polysemy.

Acknowledgments

We would like to thank Jeff Walker for excellent lab assistance, three reviewers for detailed constructive comments, Fernanda Ferreira for a detailed and valuable reading of an earlier draft, and Alan Munn and Cristina Schmitt for helpful discussions of theoretical issues surrounding the lexicon. RF and DP were supported by NIH DC R01 05660 to DP. During the preparation of this manuscript, DP was a Fellow at the Wissenschaftskolleg zu Berlin.

References

[1] T. Azuma, G.C. Van Orden, Why safe is better than fast: the relatedness of a word's meanings affects lexical decision times, J. Mem. Lang. 36 (1997) 484–504.

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

[3] R. Borowsky, M.E.J. Masson, Semantic ambiguity effects in word identification, J. Exp. Psychol. Learn. 22 (1996) 63–85.

[4] A. Caramazza, E. Grober, Polysemy and the structure of the subjective lexicon, in: C. Rameth (Ed.), Georgetown University Roundtable on Languages and Linguistics, Georgetown Univ. Press, Washington, DC, 1976, pp. 181–206.

[5] N. Chomsky, Remarks on nominalization, in: R.A. Jacobs, P.S. Rosenbaum (Eds.), Ginn and Company, Waltham, MA, 1970, pp. 184–221.
[6] J.D. Cohen, B. MacWhinney, M. Flatt, J. Provost, Psyscope: a new graphic interactive environment for designing psychology experiments, Behav. Res. Meth. Instr. 25 (1993) 257–271.
[7] D. Embick, M. Hackl, J. Schaeffer, M. Kelepir, A. Marantz, A magnetoencephalographic component whose latency reflects lexical frequency, Cogn. Brain Res. 10 (2001) 345–348.
[8] C. Fellbaum (Ed.), WordNet: An Electronic Lexical Database, MIT Press, Cambridge, 1998.
[9] S. Frisson, M.J. Pickering, The processing of metonymy: evidence from eye movements, J. Exp. Psychol. Learn. 25 (1999) 1366–1383.
[10] S. Frisson, M.J. Pickering, Obtaining a figurative interpretation of a word: support for underspecification, Metaphor Symb. 16 (2001) 149–171.
[11] M.G. Gaskell, W.D. Marslen-Wilson, Integrating form and meaning: a distributed model of speech perception, Lang. Cogn. Processes 12 (1997) 613–656.
[12] P. Helenius, R. Salmelin, E. Service, J.F. Connolly, Distinct time courses of word and context comprehension in the left temporal cortex, Brain 121 (1998) 1133–1142.
[13] P. Helenius, R. Salmelin, E. Service, J.F. Connolly, Semantic cortical activation in dyslexic readers, J. Cogn. Neurosci. 11 (1999) 535–550.
[14] Y. Hino, S.J. Lupker, Effects of polysemy in lexical decision and naming—An alternative to lexical access accounts, J. Exp. Psychol. Hum. 22 (1996) 1331–1356.
[15] G. Kellas, F.R. Ferraro, G.B. Simpson, Lexical ambiguity and the timecourse of attentional allocation in word recognition, J. Exp. Psychol. Hum. 14 (1988) 601–609.
[16] D.E. Klein, G.L. Murphy, The representation of polysemous words, J. Mem. Lang. (2001) 259–282.
[17] D.E. Klein, G.L. Murphy, Paper has been my ruin: conceptual relations of polysemous senses, J. Mem. Lang. 47 (2002) 548–570.
[18] S. Koyama, R. Kakigi, M. Hoshiyama, Y. Kitamura, Reading of Japanese Kanji (morphograms) and Kana (syllabograms): a magnetoencephalographic study, Neuropsychologia 36 (1998) 83–98.
[19] S. Koyama, D. Naka, R. Kakigi, Evoked magnetic responses during a word completion task, Electroencephalogr. Clin. Neurophysiol. 49 (1999) 174–178 (Suppl.).
[20] S. Kuriki, Y. Hirata, N. Fujimaki, T. Kobayashi, Magnetoencephalographic study on the cerebral neural activities related to the processing of visually presented characters, Cogn. Brain Res. 4 (1996) 185–189.
[21] S. Kuriki, F. Takeuchi, Y. Hirata, Neural processing of words in human extrastriate visual cortex, Cogn. Brain Res. 6 (1998) 193–202.
[22] A. Lehrer, Polysemy, conventionality, and the structure of the lexicon, Cogn. Linguist. 1 (1990) 207–246.
[23] G. Nunberg, The non-uniqueness of semantic solutions: polysemy, Linguist. Philos. 3 (1979) 143–184.
[24] R. Parks, J. Ray, S. Bland. Wordsmyth English dictionary—Thesaurus [ONLINE], Available: http://www.wordsmyth.net/ (2003, April).
[25] P.M. Pexman, S.J. Lupker, Ambiguity and visual word recognition: can feedback explain both homophone and polysemy effects? Can. J. Exp. Psychol. 53 (1999) 323–334.
[26] D.C. Plaut, Structure and function in the lexical system: insights from distributed models of word reading and lexical decision, Lang. Cogn. Processes 12 (1997) 765–805.
[27] J. Pustejovsky, The Generative Lexicon, MIT Press, Cambridge, 1995.
[28] L. Pylkkänen, A. Stringfellow, A. Marantz, Neuromagnetic evidence for the timing of lexical activation: an MEG component sensitive to phonotactic probability but not to neighborhood density, Brain Lang. 81 (2002) 666–678.
[29] L. Pylkkänen, S. Feintuch, E. Hopkins, A. Marantz, Neural correlates of the effects of morphological family frequency and size: an MEG study, Cognition 91 (2004) B35–B45.
[30] K. Rayner, Eye movements in reading and information processing: 20 years of research, Psych. Bulletin 124 (1998) 372–422.
[31] J. Rodd, G. Gaskell, W. Marslen-Wilson, Making sense of semantic ambiguity: semantic competition in lexical access, J. Mem. Lang. 46 (2002) 245–266.
[32] T. Sekiguchi, S. Koyama, R. Kakigi, The effect of word repetition on evoked magnetic responses in the human brain, Jpn. Psychol. Res. 42 (2000) 3–14.
[33] D.A. Swinney, Lexical access during sentence comprehension: (Re)consideration of context effects, J. Verb. Learn. Verb. Behav. 18 (1979) 646–659.
[34] M.J. Traxler, M.J. Pickering, B. McElree, Coercion in sentence processing: evidence from eye-movements and self-paced reading, J. Mem. Lang. 47 (2002) 530–547.