Language, Cognition and Neuroscience

ISSN: 2327-3798 (Print) 2327-3801 (Online) Journal homepage: http://www.tandfonline.com/loi/plcp21

About sharing and commitment: the retrieval of biased and balanced irregular polysemes

Andreas Brocher, Jean-Pierre Koenig, Gail Mauner & Stephani Foraker

LANGUAGE, COGNITION AND NEUROSCIENCE, 2018
VOL. 33, NO. 4, 443–466
https://doi.org/10.1080/23273798.2017.1381748

Routledge
Taylor & Francis Group

REGULAR ARTICLE

Check for updates

About sharing and commitment: the retrieval of biased and balanced irregular polysemes

Andreas Brocher[a,b], Jean-Pierre Koenig[a], Gail Mauner[c] and Stephani Foraker[d]

[a]Department of Linguistics, University at Buffalo, The State University of New York, Buffalo, NY, USA; [b]SFB Prominence in Language, Universität zu Köln, Köln, Germany; [c]Department of Psychology, University at Buffalo, The State University of New York, Buffalo, NY, USA; [d]Department of Psychology, SUNY Buffalo State, Buffalo, NY, USA

ABSTRACT
We examined how the degree of semantic similarity between an ambiguous word's meanings (homonyms vs. irregular polysemes) and meaning frequency (biased vs. balanced meanings) interact during lexical access and disambiguation. In Experiment 1, which was a continuous priming experiment, and with an ITI of 50 ms, we observed exhaustive access of meanings for all ambiguous words. With an ITI of 200 ms, we found a dominance effect for biased homonyms. There was no priming for biased irregular polysemes. For balanced homonyms and polysemes, we observed strong and roughly equivalent priming for target words associated with either meaning. In Experiment 2, using sentence reading, all ambiguous words elicited longer reading times in the absence of biasing context, while only biased and balanced homonyms also led to longer reading times in subsequent subordinate-biased context. Taken together, our data support a shared features model of irregular polyseme representation and retrieval.

ARTICLE HISTORY
Received 23 December 2016
Accepted 12 September 2017

KEYWORDS
Polysemy; lexical ambiguity; priming; eye tracking

Highlights

- Meanings of irregular polysemes but not homonyms overlap in representation.
- Activation of overlapping representations allows meaning non-commitment.
- Equally frequent meanings compete for retrieval, regardless of meaning overlap.
- Availability of overlapping representations moderate resolution of between-sense competition.

Words can be ambiguous between two or more meanings, and when they are, language users often need to disambiguate them during comprehension. For example, readers need to know whether the letter string BANK is used to refer to a financial institution or the side of a river. Extensive research on the lexical access of BANK-like words in memory has revealed that the *relative frequency* of an ambiguous word's meanings can determine the order in which meanings are accessed. For example, in the absence of preceding biasing context, the more frequent financial-institution reading of BANK is retrieved before the less frequent side-of-river interpretation, consistent with separate lexical representations for the meanings (Duffy, Morris, & Rayner, 1988). We will refer to this observation as a *dominance effect*.

To date, a number of studies have also investigated the extent to which *semantic similarity*, or *semantic relatedness*, i.e. the degree to which the representation of the meanings of an ambiguous word share semantic information, affects retrieval ease and how semantic similarity may interact with the relative frequency of meanings (see Eddington & Tokowicz, 2015 for a review). Some studies (Armstrong & Plaut, 2008, 2011; Azuma & Van Orden, 1997; Beretta, Fiorentino, & Poeppel, 2005; Brocher, Foraker, & Koenig, 2016; Klepousniotou, 2002; Rodd, Gaskell, & Marslen-Wilson, 2002) have found that *polysemes*, whose senses are related (e.g. WIRE: electricity; spying), are accessed differently from *homonyms*, whose meanings are unrelated (e.g. BANK: financial institution; side of river). Other studies have failed to observe reliable effects of meaning relatedness (Klein & Murphy, 2001, 2002; Rabagliati & Snedeker, 2013).

There are several potential reasons for the lack of consensus about the retrieval of polysemes. First, some studies failed to distinguish different kinds of polysemes from each other, making it difficult to evaluate whether one type drove the observed effects more strongly than others (Foraker & Murphy, 2012; Klein & Murphy, 2001, 2002). Second, factors known to affect the retrieval

CONTACT Andreas Brocher abrocher@buffalo.edu; abrocher@uni-koeln.de
Supplemental data for this article can be accessed doi:https://doi.org/10.1080/23273798.2017.1381748

and integration of lexically ambiguous words, such as the relative frequency of meanings and the semantic similarity or relatedness between meanings, have not always been controlled by quantificational measures (Azuma & Van Orden, 1997; Beretta et al., 2005; Foraker & Murphy, 2012; Klein & Murphy, 2001, 2002; Klepousniotou, 2002; Locker, Simpson, & Yates, 2003; Rabagliati & Snedeker, 2013; Rodd et al., 2002). Third, disparate results may be at least partially due to different experimental paradigms, which can tap different aspects of representation or phases of processing. These paradigms include single-word lexical decision (Armstrong & Plaut, 2008, 2011; Azuma & Van Orden, 1997; Beretta et al., 2005; Klepousniotou & Baum, 2007; Locker et al., 2003; Rodd et al., 2002), cross-modal priming (Klepousniotou, 2002), paired-word semantic priming (Klein & Murphy, 2001; Pylkkänen, Llinas, & Murphy, 2006), and sentence reading (Brocher et al., 2016; Foraker & Murphy, 2012; Frisson, 2015; Frisson & Frazier, 2005; Frisson & Pickering, 1999; Pickering & Frisson, 2001).

In this paper, we focus on *irregular* polysemes whose *senses* bear some similarity to each other, but unlike regular polysemes, cannot be derived from one another via a rule (cf., WIRE: electricity; spying). The unpredictable, idiosyncratic relation(s) amongst an irregular polyseme's multiple senses means that these meanings must be learned separately and distinguished in some way in the mental lexicon. This semantic property separates irregular polysemes from *regular* polysemes such as CHICKEN, whose similar senses (living barnyard animal; cooked meat) can be derived from one another via a productive rule. The rule, or referring function (Nunberg, 1979), operates within one conceptual domain (Gibbs, 1993), often through metonymic relations like part-for-whole, institution-for-person, product-for-producer, place-for-event, and so on (Apresjan, 1974; Copestake & Briscoe, 1995; Frisson & Frazier, 2005; Klepousniotou, 2002; Li & Slevc, 2017; Rabagliati & Snedeker, 2013; Rabagliati, Marcus, & Pylkkänen, 2011). In other words, for regular polysemes it has been proposed that only one of the senses is stored while the other sense is derived by applying the relevant rule (Apresjan, 1974; Copestake & Briscoe, 1995; Cruse, 1986; Frisson & Frazier, 2005; Lakoff, 1987; Rabagliati et al., 2011) or that the multiple senses are filled in from an underspecified node (Bott, Rees, & Frisson, 2016; Frisson, 2009, 2015; Frisson & Pickering, 1999; Li & Slevc, 2017; Pickering & Frisson, 2001; see also Frazier & Rayner, 1990). Irregular polysemes, on the other hand, do not have this rule relation between senses, so therefore, their senses could plausibly be represented and stored differently than those of regular polysemes. Although not in the focus of the present study, another noteworthy difference between irregular and regular polysemes is that the senses of regular polysemes are typically judged to be more similar or more strongly related than the senses of irregular polysemes (e.g. Klepousniotou, Titone, & Romero, 2008). This observation makes it difficult to determine whether the two kinds of polysemy differ on a processing level (lexical rule, underspecification), a representational level (semantic overlap), or both. Our aim in this paper is to develop a detailed and explicit model of irregular polyseme representation and retrieval by comparing irregular polysemes to homonyms. This design choice highlights the point that homonyms and irregular polysemes share the characteristic of meanings/senses that are more distinct from each other, in contrast to the greater similarity or relatedness of regular polyseme senses.

There are two novel aspects of our study. First, we investigate the lexical retrieval and disambiguation of irregular polysemes while carefully manipulating the two most salient variables that have been shown to affect the lexical access of ambiguous words: relative frequency of meanings and degree of semantic similarity. Second, the majority of studies on lexical ambiguity resolution have concentrated on *biased* ambiguous words. These words have one interpretation that is more frequent (*dominant*) and one that is less frequent (*subordinate*). To our knowledge, ours is the first study to also investigate how *balanced* irregular polysemes like CONE (traffic; ice cream), whose two senses are roughly equally frequent, are accessed and disambiguated. We test whether balanced irregular polysemes are processed differently from either biased irregular polysemes, such as WIRE (electricity; spying), whose semantically related meanings differ in their relative frequencies, or from balanced homonyms, such as CALF (body; animal), whose equally frequent meanings are semantically unrelated. Differences in the relative frequencies of meanings have been shown to affect the retrieval and disambiguation of homonyms, with balanced meanings competing more than biased ones (Duffy et al., 1988; Folk & Morris, 2003; Mason & Just, 2007; Rayner & Duffy, 1986).

Retrieval of lexically ambiguous words

While a number of studies have attested robust dominance effects for biased homonyms, with the subordinate meaning taking longer to reach its retrieval threshold than the dominant meaning (e.g. Dholakia, Meade, & Coch, 2016; Duffy et al., 1988; Gottlob, Goldinger, Stone, & Orden, 1999; Gunter, Wagner, & Friederici, 2003; Simpson, 1981; Simpson & Burgess, 1985; Simpson & Krueger, 1991), considerably less data are available for biased irregular polysemes. Importantly, the data that are available are inconclusive. Some studies support a separate entries account in which a polyseme's senses are stored separately and accessed very similarly to the meanings of homonyms. Other studies have found evidence for a homonymy-polysemy distinction but, unfortunately, have often fallen short of providing a detailed model of (irregular) polyseme representation and retrieval (see below for a discussion).

Two of the earliest studies that looked at the processing of polysemes, including a number of irregular polysemes, were Klein and Murphy (2001, 2002). Based on evidence from paired-word semantic priming involving sensicality judgments, Klein and Murphy (2001) argued that polysemes are accessed like homonyms. However, the relevance of their results for how irregular polysemes are represented is unclear for several reasons. First, their materials included several kinds of polysemy. Second, their materials were not normed for sense dominance or sense similarity. Third, their materials elicited different results in other paradigms. Specifically, Pylkkänen et al. (2006) included items used by Klein and Murphy (2001) in an MEG experiment and found that their homonyms were retrieved more slowly than their irregular polysemes in the left hemisphere, while the reverse was observed in the right hemisphere. The authors concluded that homonyms lead to immediate competition between unrelated meanings whereas polysemes are accessed via a shared morphological root with the polyseme's representation subsequently filled in towards one specific interpretation.

Another set of experiments, by Foraker and Murphy (2012), used Klein and Murphy's (2001) polysemy materials in sentence reading. The authors found dominance effects for polysemes and, based on their results, argued that polyseme senses, like homonym meanings, have separate entries. Here we point out that the observed dominance effect for polysemes occurred later and more diffusely distributed in time course and in sentence region than what has been observed for homonyms in other research (e.g. Duffy et al., 1988). So, while the reported dominance effect supported separate representations, its weakness could well have been due to the mix of polyseme types, and partial confounding with end of sentence wrap-up, rather than reflecting polyseme retrieval and/or disambiguation. Second, Foraker and Murphy did not include homonym or unambiguous control conditions for comparison. This makes it difficult to evaluate the effect size of the observed longer reading times associated with polysemes. Third and finally, in post-hoc testing, Foraker and Murphy found that the strength of the dominance effect was predicted by the degree to which senses were rated to be similar. Crucially, this should not occur on a separate entries account.

Explicitly addressing differences in semantic relatedness, Klepousniotou et al. (2008) divided their ambiguous words into low, moderate, and high semantic overlap items. The authors found that high overlap ambiguous words produced a different pattern of priming than moderate or low overlap words. For high overlap words, a dominant sense target, like TENDER LAMB was equally available following either a dominant sense prime (MARINATED LAMB) or a subordinate sense prime (BABY LAMB). In contrast, subordinate targets (FRIENDLY LAMB) were responded to faster when they followed a subordinate sense prime (BABY LAMB) than when they followed a dominant sense prime (MARINATED LAMB). On the other hand, for low (e.g. PANEL) and moderate overlap words (e.g. ORANGE), stronger priming was observed when primes and targets instantiated the same meaning/sense of the ambiguous word, for both dominant and subordinate targets, than when they instantiated different meanings/senses. Taken together, this pattern of results suggests that sense dominance differentially affects sense selection depending on how similar the senses are. Klepousniotou and colleagues argued that lexical ambiguity forms a spectrum where homonyms are at one end, regular polysemes at the other, with metaphorical polysemes, which constitute a subclass of irregular polysemes, falling in between (see also Klepousniotou, 2002; Klepousniotou & Baum, 2007; Klepousniotou, Pike, Steinhauer, & Gracco, 2012).

Finally, Brocher et al. (2016) had participants read single sentences containing either a biased homonym or a biased irregular polyseme. Disambiguating context towards the subordinate meaning/sense was presented either after the ambiguous word, or before. The authors found a typical dominance effect for the homonyms, with longer reading times on the subordinate-biased disambiguation region when it followed the homonym (BANK … *FISHING*) as well as longer reading times on the homonym region when it followed a subordinate-biased context (*FISHING* … BANK). This pattern is consistent with the lexically dominant meaning competing with the contextually supported meaning. For irregular polysemes, no dominance effect was observed on the disambiguating region that followed the polyseme (WIRE … *SPY*). However, between-sense competition was registered when the polyseme followed subordinate context (*SPY* … WIRE), although this competition was somewhat weaker than for the homonyms.

Brocher and colleagues argued that the results support a shared features model of irregular polyseme representation: The multiple senses of irregular polysemes share semantic features in memory. Access of this shared information is what precludes a dominance effect. The shared semantic information accessed initially could support the following subordinate-sense or dominant-sense context equally well. Importantly, the authors posited that the degree to which features that are not shared between the senses are activated is a function of their relative frequency. Unshared features associated with the dominant sense will have higher frequencies than unshared features associated with the subordinate sense. However, activation of unshared features associated with a subordinate sense can be boosted by a previous subordinate-biased context. Thus, if subordinate-biased context occurs before encountering a polyseme, readers will more strongly activate the subordinate sense, consisting of the shared portion of meaning and the unshared (and boosted) portion of meaning. Crucially, activating the unshared semantic information leads to some resemblance of how biased homonyms that follow subordinate-biased contexts are processed. Note that this latter property of the shared feature model might explain the homonym-like, albeit not identical, processing of polysemes in Klein and Murphy's (2001) study. In this study, ambiguous words were presented together with a disambiguating modifier (e.g. SHREDDED PAPER). This could have led to full activation of the specific sense's unshared features, and when the modifier realised the subordinate sense, would lead to between-sense competition in inconsistent trials (SHREDDED PAPER – LIBERAL PAPER), with longer processing times.

Although a few studies have investigated ambiguous words with two equally frequent interpretations, as of this writing, there have been no studies that have investigated lexical access in balanced irregular polysemes. This gap is important because balanced irregular polysemes might be retrieved differently from their biased counterparts, as is the case for biased and balanced homonyms. Most of the extant data suggest that both meanings of balanced, but not biased, homonyms are activated in parallel when no biasing context is provided, leading to between-meaning competition (Duffy et al., 1988; Folk & Morris, 2003; Mason & Just, 2007; Rayner & Duffy, 1986; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979). It is thus reasonable to hypothesise that processing differences might also be observed between balanced and biased irregular polysemes.

Models and predictions

Building on past research, we contrast a *separate representation account* with an *overlapping representation account*. These two accounts make different predictions for the processing of irregular polysemes whose senses are unequal in frequency. Proponents of a separate entries account assume that the senses of polysemes are represented similarly to the meanings of homonyms, i.e. both senses have their own entry and their representations do not overlap (see Figure 1a). One consequence of a separate entries account is that the two kinds of ambiguous words should be accessed and disambiguated similarly, no matter the degree of meaning/sense similarity. For example, when ambiguous words are biased, robust dominance effects should occur, regardless of whether the ambiguous word is a homonym or a polyseme. In contrast, an overlapping representation account states that the multiple senses of irregular polysemes overlap in their lexical representations. For example, the two senses of WIRE may share the meaning features thin, cylindrical, and metal, while for BANK, an institution for housing money may not share any meaning features with the slope bordering a body of water.

These differences in representation should lead to differences in retrieval and disambiguation. In particular, the availability of shared meaning should substantially reduce dominance effects for biased irregular polysemes, as the semantic information that overlaps between senses equally characterises a common portion of the dominant and subordinate senses. Thus, comparing

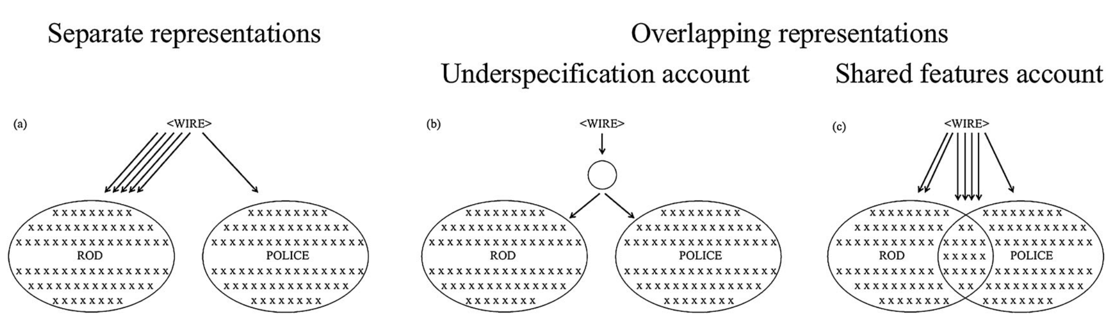

Figure 1. Illustration of representation models for irregular polysemes. A separate representations model of lexical representation is depicted in (a), while the two versions of an overlapping representations model are shown in (b) underspecification account, and (c) shared features account. Orthographic representations are represented via angled brackets while meaning representations are illustrated via the series of *x*'s in the ovals. Number of arrows represent strength of activation from orthography to meaning; *x*'s represent meaning features.

biased irregular polysemes with biased homonyms will distinguish a separate from an overlapping representation account of irregular polysemes.

The inclusion of ambiguous words whose meanings/senses are relatively equal in their frequencies makes it possible to test two variants of an overlapping representation account: an *underspecification account* and a *shared features account* (illustrated in Figure 1b and 1c, respectively). Proponents of an underspecification model (Frisson, 2009, 2015; Frisson & Pickering, 1999; Pickering & Frisson, 2001; see also Frazier & Rayner, 1990) propose that a polyseme's senses overlap via an underspecified node and that readers access this underspecified node when a polyseme is encountered. Only later, when a disambiguating context is available, do readers fill in meaning beyond the underspecified node to home-in on a specific interpretation, if needed. Importantly, the underspecified node that is shared between senses does not contain information about the relative frequencies of a polyseme's senses, as some research with regular polysemy indicates (e.g. Frisson, 2009; Frisson & Pickering, 1999). In contrast, a shared features account states that the two senses of an irregular polyseme partially overlap in memory, dividing the semantic information associated with the polyseme into shared and unshared information or features (Brocher et al., 2016; Klepousniotou, 2002; Klepousniotou et al., 2012; Klepousniotou & Baum, 2007). Initially the shared features are most quickly and highly activated during retrieval, since this information is reliable for either sense. Prima facie, there is no reason to believe that unshared features would not be activated as well, and we posit here that, if they are activated, they should be activated according to frequency strength, or dominance of that sense.

In sum, proponents of an underspecification model predict no retrieval time differences between biased and balanced irregular polysemes. In the absence of a preceding biasing context, readers access a polyseme through an underspecified node that does not carry sense frequency information. In contrast, a shared features model is compatible with the assumption that balanced polysemes are accessed differently than their biased counterparts. The overlap between the two senses of an irregular polyseme is only partial, so, the activation status of the unshared features should affect retrieval processes, in addition to the shared portion. Importantly, the unshared features of balanced words might lead to competition between the senses, similar to what has been observed for balanced homonyms (Duffy et al., 1988; Folk & Morris, 2003; Mason & Just, 2007; Rayner & Duffy, 1986). This should lead to much stronger competition between the unshared features of balanced polysemes than biased polysemes during lexical access, in particular when no conclusive context is available at the point of polyseme retrieval.

The predictions laid out above were tested in two experiments. In Experiment 1, we employed a semantic priming paradigm. Ambiguous words were used as prime words, while target words were either related to the prime's dominant or subordinate meaning/sense (for biased primes) or to the prime's meaning/sense1 or meaning/sense2 (for balanced primes). This design allowed us to test (a) for dominance effects in biased homonyms and biased polysemes and (b) for potential differences in priming between biased and balanced polysemes. In Experiment 2, we used sentence reading, in which participants read single sentences. The first clause contained either an ambiguous word or a matched control word. The second clause contained subordinate-biasing context information. This design, again, allowed us to test for dominance effects in biased homonyms and irregular polysemes and, likewise, for differences between biased and balanced polysemes, thereby providing converging evidence from an ecologically more valid paradigm.

Experiment 1: semantic priming experiment

We tested the retrieval and disambiguation of biased and balanced homonym and irregular polyseme prime words by examining how well they semantically facilitated target words related to either their dominant or subordinate meanings/senses. We employed a lexical decision task for two reasons. First, this task has been shown to be sensitive to semantic processing (Bueno & Frenck-Mestre, 2008; McRae & Boisvert, 1998; Neely, 1977; Perea & Gotor, 1997; Perea & Rosa, 2002; Tamminen & Gaskell, 2013). Second, lexical decision results make it possible to more directly compare our results to the studies we have cited in the Introduction. We used a continuous priming format (McNamara & Altarriba, 1988; McRae & Boisvert, 1998) in which participants made a lexical decision to every string presented, including both prime and target words, which were presented sequentially. Words were presented until participants made a response. Finally, we included a 50 ms intertrial interval (ITI) condition and a 200 ms ITI condition to track when in the time course of meaning/sense retrieval, the activation of shared and unshared semantic information in irregular polysemes is observable.

Method

Participants

One hundred and three students at the University at Buffalo (62 female, seven left-handed) received partial course credit for their participation. All participants were recruited from an introductory class of psychology, were monolingual native speakers of American English, and reported normal or corrected-to-normal vision. No participant contributed data to more than one on-line or norming experiment reported in this paper.

Materials

We normed over 180 words. Norming was done across successive cycles. All ambiguous words used in the experiment were normed for both meaning/sense similarity, including 12 separate norming cycles, and meaning/sense dominance, including an additional 12 norming cycles. They were either selected from published work (39% of items also appeared in Klein & Murphy, 2001; 16% of items in Klepousniotou et al., 2008 and 30% of items in Rodd et al., 2002) or agreed upon through group discussions. The major requirement for counting as *irregular* polyseme was that senses could not be derived from another by a rule, which distinguishes irregular and regular polysemes. Note that of the 32 irregular polysemes selected for the present study, 10 had a literal and a figurative interpretation and 22 had two literal interpretations.

Similarity norming. In each similarity norming cycle, 20 monolingual native speakers of American English were presented with booklets containing pairs of single sentences. Each sentence pair had one word in common. An example is provided in (1).

(a) Paul wanted to deposit all his cash but the bank was already closed.
(b) The couple went for a nice, long walk alongside the bank.

Participants were instructed to judge the semantic similarity between the two tokens of the underlined word based on the following questions (adapted from McRae, Cree, Seidenberg, & McNorgan, 2005): *Can the two meanings appear in similar contexts? Do they share physical or functional properties? Do they taste, smell, sound, or feel similarly? Do they behave similarly?* We included these questions in the instructions to elicit judgments that are based on semantic features rather than on *ad hoc* associations that are difficult to interpret (like GO THERE EVERY WEEK or RHYMES WITH TANK for the homonym BANK). The provided similarity scale ranged from 1 for "not similar at all" to 7 for "the very same meaning." Participants were encouraged to use the whole range of the scale.

Each group of 20 participants provided similarity scores for 50 sentence pairs, including 16 homonyms (BANK), 16 polysemes (WIRE), and 18 neutral filler words (ORIGAMI). Presentation was organised such that a particular ambiguous word appeared in a presentation list only once. The 16 biased and 16 balanced homonyms that were selected for the main study had a mean similarity score of 1.35 ($SD = 0.22$) and 1.32 ($SD = 0.25$), respectively. The 16 biased and 16 balanced polysemes had a mean similarity score of 3.27 ($SD = 0.61$) and 3.23 ($SD = 0.59$), respectively. A linear regression model fit to predict similarity scores revealed a main effect of ambiguity, $t = 11.90$, indicating that the meanings of homonyms were significantly less similar than the senses of polysemes. There was no main effect of bias and no Ambiguity × Bias interaction, $ts < 1$.

Dominance norming. In each dominance norming cycle, 20 monolingual native speakers of American English from the University at Buffalo received booklets that contained five tokens of single words, one token per line (e.g. WIRE ________ for five lines), followed by a blank line. For each token, participants were asked to write down whatever came to their minds. Each booklet contained eight irregular polysemes (WIRE), eight homonyms (BANK), and 32 additional filler words (ORIGAMI). Thus, in total, each participant provided $48 \times 5 = 240$ contributions.

Two different raters, who were trained in the task, decided for each contribution whether it instantiated one of the ambiguous word's targeted meanings (e.g. WIRE-CABLE or WIRE-SPY), a different/non-comprehensible meaning (e.g. BANG for WIRE), or instantiated either meaning (e.g. BEAUTIFUL for WIRE). Raters were conservative in their decisions and only assigned a particular completion to one of the ambiguous word's meanings/senses (e.g. WIRE-CABLE) when it could not, in principle, also be assigned to the competing meaning/sense (e.g. WIRE-SPY). Disagreements were resolved by discussion. After resolution, overall agreement was 92%.

For all selected items, we judged a meaning/sense to be the dominant meaning/sense if it had been produced most often across participants. We then derived the dominance score for each word relative to the other, subordinate meaning, which was the meaning/sense that had been produced the second most often. We opted for this procedure, rather than taking all produced completions into account, to reduce noise from unresolved raters' disagreements and incomprehensible productions. The 16 biased homonyms and 16 biased polysemes had a mean dominance score of .91 ($SD = .07$) and .88 ($SD = .08$), respectively. The 16 balanced homonyms and 16 balanced polysemes had a mean dominance score of .57 ($SD = .04$) and .56 ($SD = .04$), respectively. These scores are in line with previous studies that used only the first association provided by participants as a measure of dominance (Duffy et al., 1988; Folk & Morris, 2003; Klepousniotou & Baum, 2007; Rayner & Duffy, 1986; Twilley, Dixon, Taylor, & Clark, 1994, see also Armstrong & Plaut, 2011). Indeed, when comparing dominance scores for a subset of 37 items that appear in our study as well as in the Alberta Homograph Norms (Twilley et al., 1994), we found a high correlation, Pearson's $r = .68$. This suggests some robustness of dominance scores across studies. A linear regression model predicting dominance scores revealed a main effect of bias, $t = 15.71$, indicating that biased items had a larger frequency difference between the two meanings/senses than balanced items. There was no main effect of ambiguity and no Ambiguity × Bias interaction, $ts < 1.2$.

Priming stimuli. For each of the ambiguous prime words, we selected two target words. For biased ambiguous prime words, one target was related to the word's dominant interpretation and the other to the word's subordinate reading (see Table 1). For balanced ambiguous prime words, one target was selected for each of their two interpretations. Because the meanings/senses of balanced words rarely have the exact same frequencies, we will refer to the slightly more frequent meaning/sense of an ambiguous word as meaning/sense1 and to its slightly less frequent meaning/sense as meaning/sense2.

Whenever possible, target words were selected from the set of words provided in the dominance norming study to ensure that they were known to our undergraduate population and were valid examples of words that pick out one of the two targeted meanings (Gernsbacher, 1984). Although target words were always compared to themselves following ambiguous or nonword-baseline primes, care was taken to use primes and targets of comparable length and frequency across ambiguity types (e.g. BANK vs. WIRE), bias (e.g. BANK vs. CALF), and meaning/sense dominance (e.g. ROB vs. SWIM). Norming characteristics for primes and targets are provided in an Online Supplement.

We purposely chose target words with somewhat weak association scores to their ambiguous word primes because we found that strong associates predominantly refer to the dominant meaning/sense of a biased ambiguous word. The mean forward and backward association scores for homonyms and polyseme primes and their targets, based on norms developed by Nelson, McEvoy, and Schreiber (2004), are presented in Table 2. For forward association scores, a linear regression model revealed a main effect of bias, $t = 2.14$, which was driven by stronger associations between balanced items and their targets than biased items and their targets. No other effects were significant, $ts < 1$. Similarly, targets were weakly backward associated with their preceding primes. The backward association scores for dominant meaning/sense targets and their primes were numerically larger than subordinate/sense targets and their primes, but a linear regression model revealed that this difference was not statistically reliable, $t = 1.68$. No other effects reached significance, $ts < 1.6$.

An experimental trial consisted of a homonym or polyseme prime and a subsequently presented dominant-meaning/meaning1 or subordinate-meaning/meaning2 related target (see Table 1). The unrelated baseline consisted of a nonword (generated from an unrelated base word) followed by the same dominant-meaning/meaning1 or subordinate-meaning/meaning2 related target used for the ambiguous prime words. Thus, four factors were manipulated across experimental and control trials: Ambiguity (homonymous or polysemous prime), bias (biased or balanced ambiguous prime word), dominance (dominant-meaning/meaning1 or

Table 1. Example set of materials for Experiment 1.

Ambiguity	Bias	Dominance	Prime type	PRIME	TARGET
Homonymy	Biased	Dominant	Ambiguous	BANK	ROB
	Biased	Subordinate	Ambiguous	BANK	CREEK
	Biased	Dominant	Nonword	TRANSITIF	ROB
	Biased	Subordinate	Nonword	TRANSITIF	CREEK
Polysemy	Biased	Dominant	Ambiguous	WIRE	CABLE
	Biased	Subordinate	Ambiguous	WIRE	POLICE
	Biased	Dominant	Nonword	GINDER	CABLE
	Biased	Subordinate	Nonword	GINDER	POLICE
Homonymy	Balanced	Meaning 1	Ambiguous	CALF	GOAT
	Balanced	Meaning 2	Ambiguous	CALF	SHIN
	Balanced	Meaning 1	Nonword	INSTITUDE	GOAT
	Balanced	Meaning 2	Nonword	INSTITUDE	SHIN
Polysemy	Balanced	Sense 1	Ambiguous	CONE	WAFFLE
	Balanced	Sense 2	Ambiguous	CONE	CRASH
	Balanced	Sense 1	Nonword	SPACEZ	WAFFLE
	Balanced	Sense 2	Nonword	SPACEZ	CRASH

Notes: *Biased* = biased ambiguous word, *Balanced* = balanced ambiguous word, *Dominant* = dominant meaning/sense of biased ambiguous word, *Subordinate* = subordinate meaning/sense of biased ambiguous word, *Meaning1/Sense1* = first meaning/sense of balanced ambiguous word, *Meaning2/Sense2* = second meaning/sense of balanced ambiguous word.

Table 2. Mean forward and backward association strengths for dominant/meaning1 and subordinate/meaning2 targets of homonym and irregular polyseme primes in Experiment 1.

		Homonym		Polyseme	
Bias		Dominant/Meaning1	Subordinate/Meaning2	Dominant/Sense1	Subordinate/Sense2
Biased	Forward	.03	.01	.02	.01
	Backward	.01	.01	.06	.01
Balanced	Forward	.06	.06	.06	.04
	Backward	.01	.03	.02	.01

Notes: *Biased* = biased ambiguous word, *Balanced* = balanced ambiguous word, *Dominant* = dominant meaning/sense of biased ambiguous word, *Subordinate* = subordinate meaning/sense of biased ambiguous word, *Meaning1/Sense1* = first meaning/sense of balanced ambiguous word, *Meaning2/Sense2* = second meaning/sense of balanced ambiguous word, *Forward* = forward association score between prime and target, *Backward* = backward association score between prime and target.

subordinate-meaning/meaning2 target), and prime type (target preceded by an ambiguous word or an unrelated baseline nonword). Note that our first prediction involves the comparison of biased polysemes with biased homonyms to test a separate vs. overlapping representation account (i.e. Ambiguity × Dominance × Prime Type within biased words). The second prediction compares biased polysemes with balanced polysemes to investigate two versions of an overlapping representation account; namely, shared features vs. underspecification (i.e. Bias × Dominance × Prime Type within polysemes). We did not have any predictions regarding the full crossing of the four predictors in the more complex Bias × Ambiguity × Dominance × Prime Type interaction.

The 192 experimental items (64 ambiguous prime words, 64 dominant-meaning/sense and 64 subordinate meaning/sense related target words) were counterbalanced across two presentation lists, each occurring in a 50 ms ITI condition and a 200 ms ITI condition. The ITI manipulation was a between-subjects manipulation. Experimental items were interspersed amongst 384 distractor items that included words and nonwords. Nonwords were created either by changing one letter (CRAVT), adding one letter (FIEVER), or leaving out one letter (GOST). Nonwords also included 40 pseudohomophones (nonwords like GRANE) as distractors to increase participants' engagement in the task and to potentially increase the likelihood of semantic processing, based on a successful phonological look-up in the mental lexicon despite the wrong orthography (Azuma & Van Orden, 1997; Borowsky & Masson, 1996; but see Armstrong & Plaut, 2016 for recent arguments why pseudohomophones might in fact de-emphasize semantic processing). Pseudohomophones never appeared in experimental trials. The ratio between words and nonwords was 1:1. The complete set of materials is available in the Online Supplement.

Procedure

We employed a continuous priming procedure (McNamara & Altarriba, 1988; McRae & Boisvert, 1998) in which lexical decisions were made to primes and targets (see Figure 2). We used continuous rather than paired priming to obscure our experimental manipulations, and to lessen the likelihood that our results could be due to response strategies (Hutchison, 2003). Depending on the intertrial interval (ITI) (50 ms or 200 ms), a trial began either with a 50 ms or 200 ms fixation cross in the centre of the screen, which was then replaced by a letter string. Participants decided, as quickly and accurately as possible for each string (primes, targets, distractors), whether or not it was a word of English, by pressing a labelled "yes" key with their right index finger if the string was a word and a labelled "no" key with their left index finger if the string was not. Thus, a letter string was presented for as long as a participant needed to make a lexical decision on it. After each lexical decision a new trial began with a fixation cross presented for 50 ms or 200 ms. Forty-eight practice trials were first presented to familiarise participants with the task. Feedback on speed and accuracy was provided throughout the practice session but not during the experimental trials. Participants were tested individually.

All participants were debriefed after they finished the experiment. Thirty participants reported having noticed some relationship between presented stimuli. Out of these participants, 12 were able to report back one or two pairs of words that were in fact related. Of the reported pairs, 11 appeared as part of an experimental

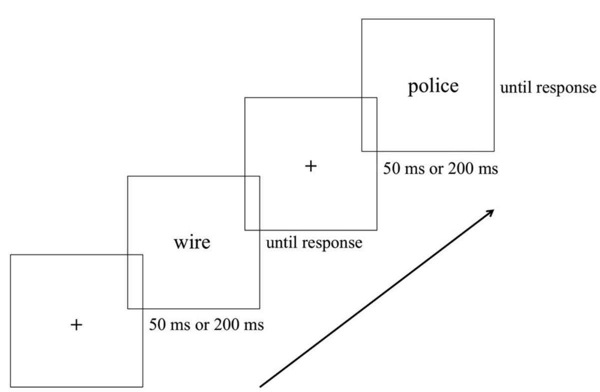

Figure 2. Illustration of the structure of trials in Experiment 1. The difference between 50 and 200 ms is due to the two different ISIs (see text).

trial (one biased polyseme, four balanced polysemes, four biased homonyms, and two balanced homonyms). Importantly, however, no participant mentioned having noticed any ambiguity in the materials.

Results

Data from three participants were excluded from analyses due to error rates (ER) higher than 20%. For analyses of lexical decision latencies of the remaining participants, which included 50 participants in the 50 ms ITI condition and another 50 participants in the 200 ms ITI condition, all incorrect responses (2.2% of the data) as well as response times (RTs) longer than 3000 ms and shorter than 200 ms (0.7% of the data) were excluded from the analyses. RTs were then log-transformed using Box–Cox power transformations before analyses to reduce skewness of the RT data. We used R (R Core Team, 2015) and lme4 (Bates, Maechler, Bolker, & Walker, 2014) to perform a linear mixed effects analysis on RTs. For the analyses of error data, we performed generalised linear regression, because error data follow a binomial distribution. Cohen's *d* effect size estimates were calculated using the lsr package (Navarro, 2014). For the 50 and 200 ms ITI conditions, each analyzed separately, ambiguity (homonym vs. polyseme), bias (biased vs. balanced), dominance (dominant vs. subordinate), and prime type (ambiguous vs. unrelated nonword-baseline prime) as well as their interactions were entered into the model as predictors and sum-coded prior to analyses. In order to fully cross all factors when comparing biased and balanced items directly and to address the fact that "truly" balanced ambiguous words (displaying a 50:50 relative frequency distribution) are rare, we treated the slightly more frequent meanings/senses of balanced items as dominant and the slightly less frequent meanings/senses as subordinate.

To account for the modest variability within association strengths between primes and targets, we also included the forward and backward association scores as well as their interaction with prime type as predictors in our models. In the same vein, to account for the variability within prime and target words, we also included length and bigram frequency of the primes as well as length, frequency, orthographic neighbourhood size, semantic neighbourhood size, and concreteness for the targets as covariates. Bigram frequencies, orthographic neighbourhood size, semantic neighbourhood size, and concreteness were assessed using the CELEX corpus (Baayen, Piepenbrock, & Gulikers, 1995). Frequency of use was assessed using the Subtlex-US corpus (Brysbaert & New, 2009). Finally, to address concerns of a reviewer, we tested for each covariate described above individually (i.e. for length, frequency, etc.), seeing whether it significantly contributed to the observed priming. To that end, we compared a model that included only the covariate in question (and the other covariates) as a linear predictor with a model that allowed that covariate to interact with soa, bias, ambiguity, dominance, and prime type, depending on the model. For example, in the 50 ms ITI condition, we compared a model containing the Bias × Ambiguity × Dominance × Prime Type interaction in addition to the main effect of length with a model that included the five-way Bias × Ambiguity × Dominance × Prime Type × Length interaction. In case a model with a five-way interaction was a better fit to the data than the model with the potentially confounding covariates as linear predictors, the five-way interaction model was used for subsequent comparisons. For all comparisons, we performed log-likelihood ratio tests and results presented below will always come from the model that best fit the data. These models are provided in the Online Supplement.

As random effects, we included intercepts for participants and items. We also included by-participant and by-item random slopes and kept them maximal (Barr, Levy, Scheepers, & Tily, 2013). However, because some of the models only converged with a random effects structure that only included prime type, all models were reduced to that random effects structure to license comparability. For all analyses reported in this paper, we adopted the decision rule that an absolute *t*-value of at least 2 indicates statistical significance (Baayen, 2008; Gelman & Hill, 2007). In what follows, we will first report response latencies of the 50 ms ITI condition and then turn to the 200 ms ITI condition. Direct comparisons between the two ITIs will be discussed in the text whenever appropriate.

50 ms ITI

Inspection of Figures 3 (biased words) and 4 (balanced words) reveals priming effects for all conditions, with numerically largest priming for the dominant meaning of biased homonyms. Error rates, priming effects, and Cohen's *d* are shown in Table 3. We computed a full model, which included all conditions, as well as separate models for homonyms, irregular polysemes, biased words, and balanced words. We fitted subset models to be able to explicitly address the various predictions laid out above, i.e. to compare biased polysemes with biased homonyms (subset model only including biased words) and to compare biased polysemes with balanced polysemes (subset model only including polysemes). Results of the regression models are provided in Table 4. As is clear in Figures 3 and 4, the only robust effect to emerge in these analyses was that of prime type, which was significant regardless of whether the ambiguous word was a balanced or biased homonym

Table 3. Error rates, priming effects, and Cohen's *d* for target words in Experiment 1.

ITI	Bias	Condition	Target ER	Baseline ER	RT Priming (ms)	*d*
50	Biased	Homonymy				
		Dominant	1.8 (1.9)	1.3 (1.6)	60	0.33
		Subordinate	0.8 (1.2)	1.5 (1.8)	31	0.22
		Polysemy				
		Dominant	1.5 (1.8)	0.8 (1.2)	29	0.16
		Subordinate	2.3 (2.1)	4.3 (2.9)	28	0.15
	Balanced	Homonymy				
		Meaning1	2.6 (2.3)	2.3 (2.2)	59	0.26
		Meaning2	3.0 (2.5)	3.3 (2.6)	42	0.19
		Polysemy				
		Sense1	1.0 (1.4)	1.5 (1.8)	33	0.31
		Sense2	4.9 (3.1)	1.5 (1.8)	23	0.19
200	Biased	Homonymy				
		Dominant	1.8 (1.9)	2.3 (2.1)	64	0.46
		Subordinate	2.0 (2.0)	2.2 (2.1)	0	0.08
		Polysemy				
		Dominant	1.7 (1.9)	1.7 (1.9)	4	0.04
		Subordinate	2.2 (2.1)	3.0 (2.4)	4	0.03
	Balanced	Homonymy				
		Meaning1	2.2 (2.1)	3.0 (2.4)	66	0.47
		Meaning2	1.5 (1.7)	3.7 (2.7)	40	0.27
		Polysemy				
		Sense1	1.5 (1.7)	2.5 (2.2)	59	0.27
		Sense2	2.7 (2.3)	3.2 (2.5)	53	0.32

Notes: *Biased* = biased ambiguous word, *Balanced* = balanced ambiguous word, *Dominant* = dominant meaning/sense of biased ambiguous word, *Subordinate* = subordinate meaning/sense of biased ambiguous word, *Meaning1/Sense1* = first meaning/sense of balanced ambiguous word, *Meaning2/Sense2* = second meaning/sense of balanced ambiguous word, *d* = Cohen's *d*, ER = mean error rate (in %).

or polyseme or whether the target word was associated with the dominant or subordinate meaning of its ambiguous word prime.

The absence of significant dominance effects for biased homonyms is in line with studies showing that, with an intermediate delay between a homonym prime and a dominant or subordinate target word, results consistent with exhaustive access is observed (Simpson & Burgess, 1985; Simpson & Krueger, 1991). For example, Simpson and Burgess (1985) found that, with an SOA

Table 4. Inferential statistics for RT data of the 50 ms ITI condition in Experiment 1.

	Main effect/Interaction	*b*	*SE*	*t*
Full model	Intercept	1.04	16.44 e-05	6347
	Prime Type	**8.47 e-05**	**2.16 e-05**	**3.92**
	Bias × Prime Type	−1.30 e-05	3.68 e-05	−0.35
	Ambiguity × Prime Type	2.85 e-05	3.54 e-05	0.81
	Dominance × Prime Type	1.52 e-05	3.54 e-05	0.43
	Bias × Ambiguity × Prime Type	2.64 e-05	6.98 e-05	0.38
	Bias × Dominance × Prime Type	−1.38 e-05	6.98 e-05	−0.20
	Ambiguity × Dominance × Prime Type	2.07 e-05	6.96 e-05	0.30
	Bias × Ambiguity × Dominance × Prime Type	7.85 e-05	13.92 e-05	0.56
Homonyms	Intercept	1.04	25.14 e-05	4150
	Prime Type	**11.57 e-05**	**3.31 e-05**	**3.80**
	Bias × Prime Type	−0.46 e-05	4.75 e-05	−0.10
	Dominance × Prime Type	1.66 e-05	5.04 e-05	0.33
	Bias × Dominance × Prime Type	2.44 e-05	9.07 e-05	0.27
Polysemes	Intercept	1.04	23.11 e-05	4514
	Prime Type	**5.98 e-05**	**2.99 e-05**	**2.00**
	Bias × Prime Type	−1.76 e-05	5.55 e-05	−0.32
	Dominance × Prime Type	1.34 e-05	5.28 e-05	0.25
	Bias × Dominance × Prime Type	−3.86 e-05	10.38 e-05	−0.37
Biased	Intercept	1.04	23.59 e-05	4422
	Prime Type	**7.78 e-05**	**3.30 e-05**	**2.36**
	Ambiguity × Prime Type	4.45 e-05	4.88 e-05	0.91
	Dominance × Prime Type	1.60 e-05	4.95 e-05	0.32
	Ambiguity × Dominance × Prime Type	7.02 e-05	9.72 e-05	0.74
Balanced	Intercept	1.04	2745 e-05	3800
	Prime Type	**8.07 e-05**	**3.05 e-05**	**2.65**
	Ambiguity × Prime Type	1.47 e-05	5.43 e-05	0.27
	Dominance × Prime Type	1.44 e-05	5.37 e-05	0.27
	Ambiguity × Dominance × Prime Type	−1.58 e-05	10.42 e-05	−0.15

Notes: *Prime Type* = ambiguous vs. unrelated prime word; *Bias* = biased vs. balanced prime word; *Ambiguity* = homonym vs. polyseme prime word; *Dominance* = dominant vs. subordinate target word. Significant effects appear in bold.

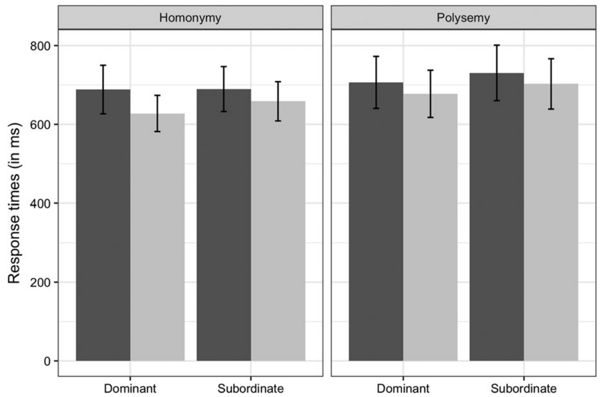

Figure 3. Lexical decision latencies and 95% confidence intervals for target words with biased primes in the 50 ms ITI condition. *Dominant* = dominant meaning/sense of biased ambiguous word, *Subordinate* = subordinate meaning/sense of biased ambiguous word; Decision latencies with nonword primes are shown in dark grey; Decision latencies with ambiguous primes are shown in light grey.

of around 400 ms in a paired-word priming task, both dominant and subordinate meanings of biased homonyms yielded priming. Given that participants could dwell on primes words *ad libitum*, it is likely that the 50 ms ITI condition of our Experiment 1 corresponded to an intermediate delay and thus tapped into this phase of exhaustive access. Finally, analyses of error rates did not reveal any reliable differences between conditions, $zs < 1.4$, $ps = .19$.

200 ms ITI

As can be seen in Figure 5 (biased words), with an ITI of 200 ms, priming was restricted to the dominant meanings of biased homonyms. In contrast, for balanced words (Figure 6), priming was observed for both homonyms and polysemes regardless of whether their target words corresponded to the more dominant or more subordinate meaning of an ambiguous prime word. The outputs of the regression models are provided in Table 5; error rates, priming effects, and Cohen's *d*s are shown in Table 3. As Table 5 reveals, the overall larger priming for balanced versus biased words led to a marginal Bias × Prime Type interaction in the full model. More importantly, the output of the full model also suggests that, while there were no dominance effects in balanced words, there were stronger dominance effects for biased homonyms than biased polysemes. This led to a significant Bias × Ambiguity × Dominance ×

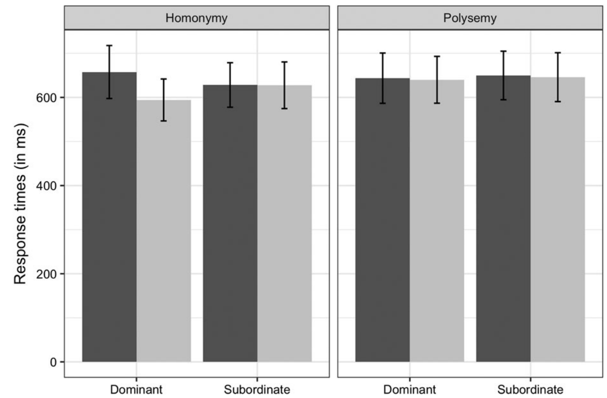

Figure 5. Lexical decision latencies and 95% confidence intervals for target words with biased primes in the 200 ms ITI condition. *Dominant* = dominant meaning/sense of biased ambiguous word, *Subordinate* = subordinate meaning/sense of biased ambiguous word; Decision latencies with nonword primes are shown in dark grey; Decision latencies with ambiguous primes are shown in light grey.

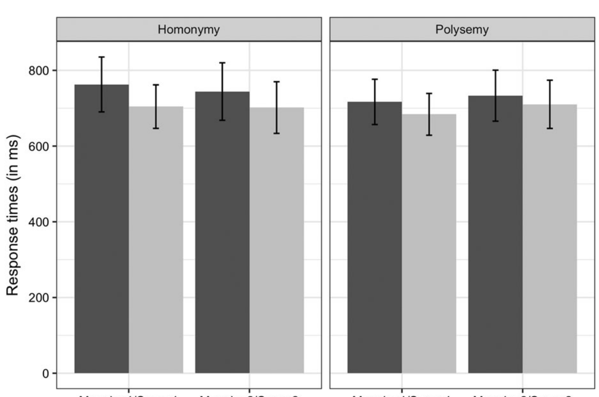

Figure 4. Lexical decision latencies and 95% confidence intervals for target words with balanced primes in the 50 ms ITI condition. *Meaning/sense1* = more frequent meaning/sense of balanced ambiguous word, *Meaning/sense2* = less frequent meaning/sense of balanced ambiguous word; Decision latencies with nonword primes are shown in dark grey; Decision latencies with ambiguous primes are shown in light grey.

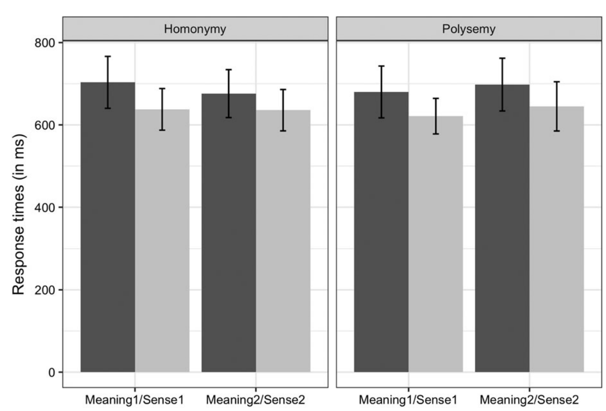

Figure 6. Lexical decision latencies and 95% confidence intervals for target words with balanced primes in the 200 ms ITI condition. *Meaning/sense1* = more frequent meaning/sense of balanced ambiguous word, *Meaning/sense2* = less frequent meaning/sense of balanced ambiguous word; Decision latencies with nonword primes are shown in dark grey; Decision latencies with ambiguous primes are shown in light grey.

Table 5. Inferential statistics for RT data of the 200 ms ITI condition in Experiment 1.

	Main effect/Interaction	*b*	*SE*	*t*
Full model	Intercept	1.04	20.96 e-05	4977
	Prime Type	**6.80 e-05**	**2.64 e-05**	**2.58**
	Bias × Prime Type	−8.03 e-05	4.40 e-05	−1.83
	Ambiguity × Prime Type	3.65 e-05	4.35 e-05	0.84
	Dominance × Prime Type	1.52 e-05	4.41 e-05	0.35
	Bias × Ambiguity × Prime Type	6.97 e-05	8.68 e-05	0.80
	Bias × Dominance × Prime Type	3.65 e-05	8.76 e-05	0.42
	Ambiguity × Dominance × Prime Type	1.12 e-05	8.69 e-05	1.29
	Bias × Ambiguity × Dominance × Prime Type	**50.76 e-05**	**17.31 e-05**	**2.93**
Homonyms	Intercept	1.04	28.10 e-05	3713
	Prime Type	17.40 e-05	18.74 e-05	0.93
	Bias × Prime Type	−7.04 e-05	4.53 e-05	−1.56
	Dominance × Prime Type	**9.16 e-05**	**4.15 e-05**	**2.21**
	Bias × Dominance × Prime Type	8.01 e-05	8.08 e-05	0.99
Polysemes	Intercept	1.04	25.13 e-05	4149
	Prime Type	4.38 e-05	2.86 e-05	1.54
	Bias × Prime Type	**−12.15 e-05**	**5.13 e-05**	**−2.37**
	Dominance × Prime Type	−3.28 e-05	5.13 e-05	−0.64
	Bias × Dominance × Prime Type	−0.86 e-05	10.09 e-05	−0.09
Biased	Intercept	1.04	29.71 e-05	3510
	Prime Type	1.73 e-05	2.86 e-05	0.60
	Ambiguity × Prime Type	**9.30 e-05**	**3.98 e-05**	**2.33**
	Dominance × Prime Type	5.14 e-05	3.99 e-05	1.29
	Ambiguity × Dominance × Prime Type	**17.13 e-05**	**7.86 e-05**	**2.18**
Balanced	Intercept	1.04	30.30 e-05	3442
	Prime Type	**11.33 e-05**	**3.62 e-05**	**3.13**
	Ambiguity × Prime Type	2.12 e-05	6.64 e-05	0.32
	Dominance × Prime Type	−0.76 e-05	6.36 e-05	−0.12
	Ambiguity × Dominance × Prime Type	−12.23 e-05	12.98 e-05	−0.94

Notes: *Prime Type* = ambiguous vs. unrelated prime word; *Bias* = biased vs. balanced prime word; *Ambiguity* = homonym vs. polyseme prime word; *Dominance* = dominant vs. subordinate target word. Significant effects of interest appear in bold.

Prime Type interaction. Focusing on the more specific subset models, Table 5 also shows that the Ambiguity × Dominance × Prime Type interaction was upheld only for biased words. For balanced ambiguous words, no effects, other than the main effect of prime type, was significant. This suggests that there were no great differences in priming between balanced homonyms and balanced irregular polysemes.

Interestingly, when only polysemes were analyzed, we obtained a significant Bias × Prime Type interaction. This interaction is due to the fact that balanced polysemes yielded strong priming effects while, as is clear from Figure 5, their biased counterparts did not elicit any priming at all. Finally, analyses of the error data showed that there were overall fewer errors in the primed than the unprimed conditions, $z = 1.9$, $p = .057$, with no other effects reaching significance, zs < 1.1, ps $> .3$.

Discussion

The RT data obtained for homonyms are consistent with existing literature that supports the claim that the two meanings of homonyms are stored separately. For balanced homonyms, separate representations would lead to strong activation of either meaning (Duffy et al., 1988). This is a plausible interpretation of the observed pattern in our data. For balanced homonym prime words, both meanings were activated and competed with each other until a disambiguating target word was encountered. The lack of difference in the priming effects on target words across the 50 and 200 ms ITI conditions suggests that competition between the equally frequent meanings was present at both intertrial intervals. Turning to the biased homonyms, the existing literature suggests that their separate representations would lead to exhaustive access at a shorter ITI and selective access at a longer ITI (Simpson & Burgess, 1985; Simpson & Krueger, 1991). This is because when a biased homonym is accessed in memory, its dominant meaning reaches its retrieval threshold much faster than its subordinate meaning. This in turn leads to larger priming for dominant than subordinate meanings with very short prime-target intervals. As intimated above, it is likely that our 50 ms ITI was not short enough to capture this early stage, since the continuous lexical decision task involves subject-paced decisions on the prime words which may have led to extended processing time for prime words in this condition. This would explain the absence of dominance effects at this intertrial interval. Thus, our 50 ms ITI condition seems to reflect a non-selective activation stage. Importantly, with a longer delay between prime and target, as with our 200 ms ITI condition, as competition settles, the subordinate meanings did appear to lose activation and return to baseline levels, in comparison with the dominant meanings, which were still facilitated relative to their controls.

Turning to the irregular polysemes, we found no evidence of dominance effects for biased or balanced words, regardless of the ITI. This finding speaks against a separate entries framework for irregular polysemes. If the two senses of biased irregular polysemes were represented separately, just like the two meanings of biased homonyms, we would have expected similarly large dominance effects for both kinds of words. This clearly was not the case. Results of biased irregular polysemes favour a shared representation account. With a short 50 ms ITI, upon reading an irregular polyseme, participants quickly and strongly activated its shared semantic information, with unshared semantic information presumably not being activated strongly enough initially to yield observable dominance effects. As a result, we found marginal response facilitation for both senses. We posit that the difference in activation strength between shared and unshared semantic information is accounted for by our proposal that shared features are more frequent than unshared features. Indeed, within a shared features framework, the semantic information that is shared by a polyseme's senses is expected to always be activated when the polyseme is encountered, irrespective of the targeted sense.

Importantly, with increasing ITI, facilitation for dominant and subordinate senses was no longer observed. This suggests that activation of the unshared semantic information decayed more than activation of the shared information, likely because the former was more weakly activated to begin with. We argue that activation of the shared semantic information does not speed the recognition of disambiguating target words because it is the non-overlapping semantic information that needs to be activated for disambiguation. Non-overlapping semantic information only begins to become fully available when a target word is encountered. Since non-overlapping information is, at best, only weakly activated at the time a disambiguating target is encountered, no response facilitation for that target word is expected.

Testing the predictions of the shared features and the underspecification model more specifically, we compared the retrieval and disambiguation of biased polysemes with that of balanced polysemes. With a short ITI, we found no differences between the two. However, with an ITI of 200 ms, we found significantly stronger response facilitation for targets of balanced polysemes than targets of biased polysemes. This difference was also evident when comparing biased and balanced polysemes across ITIs. The regression model revealed a marginally significant ITI × Bias × Prime Type interaction, $b = 10.70\text{e-}04$, $SE = 5.57\text{e-}04$, $t = 1.92$. These data suggest that the two senses of balanced polysemes competed for activation much more than the two senses of biased polysemes. This observation makes polyseme processing incompatible with both an underspecification approach to irregular polysemy and with a separate entries account. Underspecification models predict biased and balanced polysemes to behave very similarly, as frequency information associated with the various senses are not encoded in the underspecified node, which is always initially accessed (Frisson, 2009; Frisson & Pickering, 1999).

The difference in response facilitation for targets of balanced and biased polysemes and the similarity in response facilitation between balanced polysemes and balanced homonyms suggest that the initially weak activation of unshared features led to between-sense competition for balanced words, similar to the often observed between-meaning competition for balanced homonyms in previous studies (Duffy et al., 1988; Folk & Morris, 2003; Mason & Just, 2007; Rayner & Duffy, 1986). Importantly, with a longer delay between prime and disambiguating target, competition increases, as readers presumably attempt to resolve the ambiguity. In other words, equally strong activation of incompatible, unshared features requires readers to make a choice as to which sense to retrieve. Under these conditions, readers continue activating unshared features in the same fashion as they continue activating the two meanings of a balanced homonym, in order to resolve the ambiguity.

Experiment 2: sentence reading experiment

Experiment 2 employed eye tracking during sentence reading and was designed to examine whether the shared features model supported by the word recognition data of Experiment 1 extends to when biased words are encountered in sentence contexts. One important feature of sentence reading is that it allows us to investigate meaning/sense integration in addition to ambiguous word retrieval. Sentence reading is arguably a more common and natural form of language processing than lexical decisions, and likely requires deeper semantic processing of the ambiguous word, as well as interpretation and integration of the word's meaning into the sentence context.

We had two predictions for Experiment 2. The first is both empirically and theoretically grounded. Assuming that the representations of irregular polysemes share features, we predicted smaller dominance effects for biased polysemes than biased homonyms or, perhaps, even a complete absence of dominance effects for biased polysemes. This prediction is consistent with Brocher et al. (2016), and is also in line with the results of Experiment 1. More specifically, we predicted that, when subordinate-biased disambiguating information appeared after an irregular polyseme, we would not observe longer reading times for the disambiguating region relative to an unambiguous control condition. On a more theoretical basis, this prediction is based on the assumption that a shared features representation would allow readers to avoid or delay commitment to one sense, even over a clause boundary (compared to lexical decisions). To test this prediction, we created sentences in which an ambiguous word, i.e. a homonym or polyseme, appeared in a first clause, and context supporting that word's subordinate meaning/sense appeared in a second clause. Based on the results of Experiment 1, we should find stronger dominance effects for biased homonyms than for biased irregular polysemes in this condition.

Our second prediction was that if unshared features for the different senses are activated along with, or after the shared features, they should engage in between-sense competition. For balanced polysemes, this competition should lead to longer reading times for both senses, compared to an unambiguous control word. In fact, the results of Experiment 1 suggest that we should observe longer reading times for balanced polysemes and balanced homonyms relative to controls, and that these reading time differences should be roughly equal. Importantly, however, based on findings by Duffy *et al*., (1988), we should also observe longer reading times at disambiguating regions following balanced as compared to unambiguous control words, because readers may select the incorrect meaning/sense on average half of the time for balanced words. This, in turn, means that readers would need to engage in a costly reanalysis at the disambiguating region approximately half of the time. Now, our critical prediction here is that balanced polysemes should display much smaller reanalysis costs than balanced homonyms, because the availability of shared features should make it easier to switch to the competing sense than is the case when meanings are unrelated.

Method

Participants

Thirty students from SUNY at Buffalo (seven female, two left-handed) participated for course credit. All participants were recruited from an introductory psychology class and were monolingual native speakers of American English with no history of a language or learning disability, and reported normal or corrected-to-normal vision.

Materials

We assessed 64 ambiguous words, with 16 each of balanced homonyms (CALF), biased homonyms (BANK), balanced irregular polysemes (CONE), and biased irregular polysemes (WIRE), and a matched control word for each ambiguous word. The sentences consisted of two clauses. The first clause always contained the ambiguous or control word, while the second clause contained contextual information that always disambiguated the ambiguous word toward its less frequent interpretation (subordinate meaning/sense for biased ambiguous words and meaning2/sense2 for balanced ambiguous words). A set of example sentences is shown in Table 6. The full set of materials is available in the Online Supplement.

Table 6. Example materials for Experiment 2.

Frequency	Ambiguity	Sentence
Biased	Homonym	Ken decided on the |**bank** (lake) |near the | clubhouse, since the other |*beaches were too* | crowded for swimming.
Biased	Polyseme	When Mr. Jordon discovered the |**wire** (bomb) |in the |lamp, the |*FBI aborted* |the top secret mission.
Balanced	Homonym	Something seemed to be wrong with the |**calf** (pony) |that day |, because the |*animal did not* | drink nor eat.
Balanced	Polyseme	Marlene looked out for a |**cone** (barrel) |on her | way home, since a big |*pothole had been* | marked there yesterday.

Notes: For illustration purposes only, the ambiguous word appears in bold and its matched control follows in parentheses, the disambiguating region appears in italics, and the pipe symbols indicate analysis regions.

Most of the ambiguous words in our sentences were drawn from those used in Experiment 1. For the 64 ambiguous words, the dominance scores for biased homonyms ($M = .92$), biased polysemes, $M = .87$, balanced homonyms, $M = .57$, and balanced polysemes, $M = .56$, were very similar to the distribution in Experiment 1, as were the similarity scores: biased homonyms, $M = 1.37$, biased polysemes, $M = 3.29$, balanced homonyms, $M = 1.4$, balanced polysemes, $M = 3.25$. The senses of polysemes remained more similar than the meanings of homonyms. A linear regression model revealed a main effect of similarity, $t = 11.25$. Likewise, meanings/senses of biased words were more biased than meanings/senses of balanced words, $t = 14.90$. No other effect reached significance, $ts < 1.2$. All control words were matched to their ambiguous counterparts in frequency, using the contextual diversity measure in Subtlex-US (Brysbaert & New, 2009), as well as word length, and number of syllables. There were no significant differences within any set of ambiguous and unambiguous control words, all $ps > .5$.

Sentence norming. We collected plausibility judgments for each sentence, asking 95 native American English-speaking participants (at SUNY Buffalo State College). The experimental sentences were interspersed among

16 plausible distractor sentences without noticeable ambiguity, and 16 implausible ones, and then counterbalanced across four paper-and-pencil forms. Participants were asked to rate these sentences on a scale from 1 for "makes no sense at all" to 7 for "makes complete sense". Sentences containing biased homonyms and unambiguous control sentences had plausibility scores of 4.1 ($SD = 2.2$) and 4.9 ($SD = 1.9$), respectively, and those with biased polysemes and unambiguous control sentences were 4.3 ($SD = 2.1$) and 4.7 ($SD = 2.0$), respectively. Balanced homonyms and their unambiguous control sentences had a mean plausibility score of 4.8 ($SD = 2.0$) and 5.0 ($SD = 2.0$), respectively, while balanced polysemes and their controls had a mean plausibility score of 4.7 ($SD = 2.1$) and 4.7 ($SD = 2.0$). A linear regression model revealed a Bias (biased vs. balanced) × Word Type (ambiguous vs. neutral control word) interaction, $t = 3.43$, suggesting that sentences with balanced ambiguous words and their controls were judged to be about equally plausible, whereas sentences with biased ambiguous words were rated as less plausible than their unambiguous control sentences. This is expected, because, upon reading the entire sentence, readers are likely to sometimes experience reanalysis difficulty with biased words, especially for biased homonyms. Crucially, however, neither bias nor word type significantly interacted with ambiguity, $ts < 1.6$, nor was the three-way interaction reliable, $t = 1.43$.

Counterbalancing. The full set of 128 experimental sentences was counterbalanced across two presentation lists. Each list contained 32 sentences with an ambiguous word and 32 sentences with a matched control word. For a given ambiguous word, the ambiguous sentence and its control sentence always appeared in different lists. The experimental items were interspersed with 64 distractor sentences designed to contain no obvious lexical or structural ambiguities. The full set of 128 sentences was presented in a unique random order for each participant. A Yes-No comprehension question followed 32 of the distractors (half "yes", half "no"), which were the same on both presentation lists, representing 25% of the total set of trials.

Procedure

Participants were seated in front of an SR Research Eyelink 1000 eye tracker configured with the tower mount. We used SR Research's Experiment Builder programme template for reading experiments (version 1.6.121) to present stimuli and record eye movements. Each sentence was displayed all at once on one line, left-justified, about 1/3 down the monitor's screen, in 11-point Lucida Console fixed-width font. Participants' eyes were 55 cm from the screen, such that three letters subtended 1° of visual angle. They were instructed to read each sentence normally for comprehension and to press the right button of the mouse when they felt they understood it. The next screen displayed either a comprehension question, answered with either the right ("yes") or left ("no") button of the mouse, or a READY? message. Prior to beginning the experiment, participants first completed five practice trials.

Like for Experiment 1, all participants were debriefed after the experiment. During debriefing, many participants reported that some of the presented sentences failed to make sense. When asked why they felt this way, they typically responded that some words did not fit into the sentence. Two participants reported that some words had more than one meaning and that the used meaning did not make sense in the given sentence context.

Dependent measures. Four dependent measures are reported. *First Fixation* refers to the duration of the first fixation in a region. *First pass reading time* is defined as the sum of fixations that occur in a region before exiting that region to the right or left (For convenience, we use the term "first pass" for single word regions as well, rather than "gaze duration."). *Regression path duration* (also called *go-past time*) is the sum of all fixations in a region, as well as regressive fixations to earlier parts of the sentence before progressing past the region's right boundary. Regression path duration measures both early reading and some rereading. *Total reading time* is the sum of all fixations in a region and reflects a comprehensive integration of the information that was read.

Regions of analysis. We recorded and analyzed eye movements for two sentence regions, which are illustrated in Table 6. The first region consisted of an ambiguous word (e.g. WIRE) and its spillover, spanning 5–7 characters (e.g. IN THE). For this first region, we predicted longer reading times relative to controls for balanced but not biased polysemes, because results from Experiment 1 and from previous studies have shown that competition between meanings/senses leads to longer meaning/sense retrieval and selection times (Duffy et al., 1988; Folk & Morris, 2003; Rayner & Duffy, 1986). The second region of interest was the disambiguating region. It consisted of the first word that was strongly biased towards an ambiguous word's subordinate/meaning2 interpretation (e.g. FBI for WIRE) plus the subsequent 5–7 characters (ABORTED). We opted for a disambiguating region rather than a separate analysis of disambiguating word and spillover because previous studies (Brocher et al., 2016; Duffy et al., 1988; Foraker & Murphy, 2012) have shown that readers rarely disambiguate an ambiguous word on a specific word, but that meaning/sense disambiguation processes can unfold across multiple words.

Results

Data from two participants were excluded from analyses due to question-answering error rates higher than 20%. For the remaining reading time data, we excluded all blinks as well as all fixations that were either shorter than 60 ms or longer than 1000 ms. In addition, we checked that readers landed on the ambiguous/control word in computing first pass reading times. If they did not, the trial was excluded for all measures. Overall, 13.8% of the data were excluded. All reading time measures were log-transformed using Box–Cox power transformations before analyses. We used R and lme4 to perform linear mixed effects analyses on reading times for each region and reading measure of interest. Ambiguity (homonym vs. polyseme), bias (biased vs. balanced), and word type (ambiguous vs. neutral control word) together with their interactions were entered into the model as fixed effects, and sum-coded.

As in Experiment 1, we also included various word properties likely to affect reading times as covariates in the models. For analyses on ambiguous words and their controls, we included length, frequency, bigram frequency, orthographic neighbourhood size, semantic neighbourhood size, and concreteness. For analyses on ambiguous word spillover regions, we included the frequency, bigram frequency, orthographic neighbourhood size, semantic neighbourhood size, and concreteness of the ambiguous/control word as well as the length of the region. For disambiguating regions, we included the frequency, bigram frequency, orthographic neighbourhood size, semantic neighbourhood size, and concreteness of the first disambiguating word in the region (see the first word in italics in the example sentences in Table 6) and the length of the region.

In addition, and also like for Experiment 1, we tested for each covariate (e.g. length, frequency, etc.) individually, ascertaining to what extent they contributed to reading time differences. We compared a model that included only main effects of the covariates as linear predictors with models that allowed each particular covariate to interact with bias, ambiguity, and word type. We again adopted log-likelihood ratio tests to find the best-fit model, which is reported here. Finally, we again treated the somewhat more frequent meanings/senses of balanced words as dominant and the somewhat less frequent meanings/senses as subordinate. All models included random intercepts for participants and items as well as a random slope for word type. More complex random slopes led to convergence failure. All final models can be accessed in the Online Supplement. In what follows, we will first discuss reading times for the ambiguous word region and then turn to the reading times of the disambiguating region.

Ambiguous word region

Reading times for the ambiguous word and its spillover region are presented in Table 7. Outputs of the best-fit regression models are provided in Table 8. Results from reading times on ambiguous words first show that biased words elicited overall longer regression path times than their controls. This was not the case for balanced words. This bias effect, which was more strongly driven by the biased homonyms than the biased polysemes (see Table 7), is unpredicted and not in line with previous research. We are not certain why our readers regressed back to the beginning of the

Table 7. Reading time data of Experiment 2.

Measure	Ambiguity	Bias	Ambiguous word		Ambiguous spillover		Disambiguating region	
			Ambiguous	Control	Ambiguous	Control	Ambiguous	Control
First fixation	Homonym	Biased	226 (7)	215 (6)	237 (9)	235 (9)	231 (6)	219 (6)
		Balanced	215 (5)	217 (6)	243 (9)	225 (7)	234 (6)	230 (6)
	Polyseme	Biased	221 (5)	208 (5)	235 (7)	230 (7)	233 (6)	226 (6)
		Balanced	229 (6)	225 (6)	240 (8)	234 (8)	224 (6)	227 (6)
First pass	Homonym	Biased	245 (9)	229 (7)	299 (15)	287 (13)	480 (24)	414 (20)
		Balanced	249 (8)	246 (8)	330 (15)	298 (12)	466 (19)	427 (18)
	Polyseme	Biased	251 (8)	240 (9)	291 (13)	290 (12)	419 (21)	417 (16)
		Balanced	252 (9)	263 (10)	327 (16)	291 (14)	402 (17)	435 (21)
Regression path	Homonym	Biased	311 (19)	273 (14)	398 (25)	374 (32)	636 (41)	450 (35)
		Balanced	290 (13)	309 (18)	361 (21)	363 (19)	697 (49)	615 (45)
	Polyseme	Biased	298 (16)	280 (16)	360 (25)	356 (20)	526 (27)	538 (29)
		Balanced	318 (17)	326 (16)	427 (26)	408 (30)	517 (30)	606 (46)
Total time	Homonym	Biased	451 (25)	376 (19)	468 (30)	409 (26)	633 (29)	525 (24)
		Balanced	448 (21)	410 (21)	522 (30)	449 (21)	645 (25)	611 (24)
	Polyseme	Biased	382 (20)	388 (22)	406 (23)	424 (24)	594 (29)	550 (20)
		Balanced	420 (22)	444 (23)	441 (21)	486 (28)	574 (23)	630 (29)

Note: Standard error appears in parentheses following the mean, both in ms.

Table 8. Inferential statistics for reading time data of Experiment 2.

Region	Measure	Main effect/Interaction	*b*	*SE*	*t*
Ambiguous word	First Fixation	Intercept	6.83	0.221	32.67
		Word Type	0.035	0.033	1.07
		Bias × Word Type	−0.061	0.064	−0.95
		Ambiguity × Word Type	−0.021	0.064	−0.33
		Bias × Ambiguity × Word Type	−0.027	0.127	−0.21
	First Pass	Intercept	3.25	0.053	64.02
		Word Type	0.006	0.008	0.76
		Bias × Word Type	−0.021	0.015	−1.38
		Ambiguity × Word Type	0.004	0.015	0.22
		Bias × Ambiguity × Word Type	0.013	0.030	0.44
	Regression Path	Intercept	1.84	0.012	164.59
		Word Type	0.002	0.002	1.04
		Bias × Word Type	**−0.007**	**0.003**	**−2.29**
		Ambiguity × Word Type	−0.001	0.003	−0.21
		Bias × Ambiguity × Word Type	0.005	0.006	0.73
	Total Time	Intercept	2.74	0.045	60.83
		Word Type	0.005	0.007	0.68
		Bias × Word Type	**−0.026**	**0.011**	**−2.26**
		Ambiguity × Word Type	0.019	0.011	−1.69
		Bias × Ambiguity × Word Type	0.007	0.023	0.30
Ambiguous spillover	First Fixation	Intercept	3.85	0.075	51.46
		Word Type	0.019	0.016	1.17
		Bias × Word Type	0.005	0.032	0.17
		Ambiguity × Word Type	0.023	0.031	0.74
		Bias × Ambiguity × Word Type	0.069	0.062	1.11
	First Pass	Intercept	2.91	0.051	57.36
		Word Type	0.009	0.008	1.15
		Bias × Word Type	**0.029**	**0.013**	**2.14**
		Ambiguity × Word Type	−0.018	0.013	−1.34
		Bias × Ambiguity × Word Type	0.002	0.026	0.07
	Regression Path	Intercept	2.16	0.024	89.11
		Word Type	0.003	0.003	0.83
		Bias × Word Type	0.001	0.006	0.11
		Ambiguity × Word Type	−0.004	0.006	−0.64
		Bias × Ambiguity × Word Type	−0.019	0.013	−1.55
	Total Time	Intercept	4.18	0.149	27.91
		Word Type	0.003	0.021	0.16
		Bias × Word Type	−0.022	0.039	−0.55
		Ambiguity × Word Type	0.074	0.039	1.89
		Bias × Ambiguity × Word Type	0.059	0.078	0.75
Disambiguating region	First Fixation	Intercept	4.74	0.071	66.55
		Word Type	0.016	0.015	1.05
		Bias × Word Type	0.010	0.026	0.40
		Ambiguity × Word Type	0.011	0.026	0.41
		Bias × Ambiguity × Word Type	−0.065	0.052	−1.26
	First Pass	Intercept	5.15	0.229	22.47
		Word Type	0.032	0.032	0.99
		Bias × Word Type	0.054	0.060	0.90
		Ambiguity × Word Type	**0.127**	**0.060**	**2.11**
		Bias × Ambiguity × Word Type	−0.067	0.120	−0.55
	Regression Path	Intercept	3.22	0.076	42.11
		Word Type	0.014	0.010	1.40
		Bias × Word Type	−0.003	0.019	−0.15
		Ambiguity × Word Type	**0.048**	**0.019**	**2.58**
		Bias × Ambiguity × Word Type	−0.030	0.037	−0.82
	Total Time	Intercept	9.41	0.800	11.76
		Word Type	−0.155	0.102	1.52
		Bias × Word Type	0.037	0.181	0.20
		Ambiguity × Word Type	**0.418**	**0.181**	**2.31**
		Bias × Ambiguity × Word Type	−0.509	0.362	−1.41

Notes: *Word Type* = ambiguous vs. matched control word; *Bias* = biased vs. balanced ambiguous word; *Ambiguity* = homonym vs. polyseme word. Significant effects of interest appear in bold.

sentence more often after encountering a biased ambiguous word than after encountering a matched control word, or why they did so more often for biased homonyms than biased polysemes. One possibility is that, on some level and in particular for biased homonyms, readers sometimes noticed the ambiguity and, as a result, regressed back to the left to disambiguate the word. This could point to a difference in sensitivity of the current sample and the sample used by Brocher et al. (2016). These authors, who only tested biased ambiguous words, do not report longer reading times for biased ambiguous words when they preceded subordinate bias context (see also Duffy et al., 1988; Rayner & Duffy, 1986; Rayner, Pacht, & Duffy, 1994 who also found control-like reading of biased homonyms in the absence of biasing context).

Pertinent to our hypotheses, we found a marginal Ambiguity × Word type interaction and a significant Bias × Word Type interaction for total reading times. Inspection of Table 7 reveals that participants reread homonyms more often than polysemes and reread biased homonyms more than balanced homonyms. These effects suggest that participants regressed back to biased homonyms as a result of a reanalysis prompted by subordinate context in the subsequent clause.

For the ambiguous word spillover region, we observed longer reading times for balanced polysemes and homonyms relative to their controls. Table 7 reveals that there was a numerical, albeit non-significant, trend for balanced words to elicit longer first fixations than biased words, especially balanced homonyms. The Bias × Word type interaction reached significance in the first pass measure. These data are fully in line with the observation in Experiment 1 that balanced homonyms and polysemes lead to competition between meanings and senses, respectively. It is this competition that slows retrieval and selection of one of the two meanings/senses.

Disambiguating region

Reading times for disambiguating regions are shown in Table 7. Results of the regression analyses are presented in Table 8. In line with our predictions, we found significantly longer first pass, regression path, and total reading times for homonym than irregular polyseme conditions relative to their unambiguous controls. Although not significant, inspection of Table 7 also shows that the observed longer reading time differences (relative to controls) were somewhat larger for biased than balanced homonyms. This is predicted because reanalysis costs should increase when the frequency of the unselected meaning decreases. In other words, accessing a subordinate meaning after having selected the dominant meaning should be harder than accessing a second meaning of a word after its slightly more frequent meaning has been selected.

While the data for biased homonyms and polysemes in the disambiguating region fully replicate Brocher et al.'s (2016) findings, as well as data reported in the homonymy literature, it is interesting that balanced polysemes failed to show reliably longer reading times relative to their controls. This suggests that while the availability of shared semantic information does not prevent between-sense competition when the unshared information is equally frequent, it does help readers in switching to a previously unselected sense.

Discussion

Results of Experiment 2 are consistent with the data from Experiment 1 and to a large extent replicate previous research. First, in the absence of biasing contexts, balanced words elicit longer reading times than biased words. This is compatible with the observation that participants in Experiment 1 strongly activated both meanings/senses of balanced homonyms and polysemes in the 200 ms ITI condition but did not strongly and fully activate the two meanings/senses of biased words. Second, when readers encountered a subordinate-biased context after having retrieved and integrated a biased ambiguous word, they took longer to read the disambiguating region following a homonym versus a polyseme. These data suggest that, for biased words, readers selected the dominant meaning of a biased homonym, but no specific sense of a biased polyseme. For balanced words, we assumed that readers accessed either meaning about 50% of the time, which was followed by disambiguation supporting the slightly less frequent meaning/sense of balanced words. As a consequence, readers needed to reanalyse their initial interpretation of the ambiguous word roughly half of the time. What is important is that the reanalysis was only costly in the case of balanced homonyms, not in the case of balanced polysemes. This result nicely lines up with the finding that less effort is required to overcome between-sense competition when context and frequency bias support different senses of an irregular polyseme than when they support different meanings of a homonym (Brocher et al., 2016).

General discussion

In this paper, we investigated how speakers resolve the ambiguity of irregular polysemes, i.e. words which have semantically similar senses, but whose senses are idiosyncratically related rather than falling into a general pattern resembling a rule. Our findings support the claim that readers retrieve partially overlapping meaning representations rather than fully specified senses. That is, when readers access an irregular polyseme like WIRE, they quickly and strongly activate features that are part of both the cable-sense and the spy-sense, with unshared features being activated more weakly. This idea follows naturally from the observation that high-frequency words are accessed more quickly than low-frequency words (Forster & Chambers, 1973; Whaley, 1978). Because shared features would be activated irrespective of sense, these features can be considered more frequent and therefore are more quickly activated than unshared features, which are only strongly activated when one specific sense is accessed. Our data support an overlapping representation account and are incompatible with the view that senses of irregular polysemes are represented separately in the mental lexicon (Foraker & Murphy, 2012; Klein & Murphy, 2001, 2002; Rabagliati & Snedeker, 2013). If this were the case, biased irregular polysemes should have patterned with biased homonyms, which is not what we observed.

Data from our continuous priming experiment suggest that the initial strong activation of the shared semantic information leads to an absence of dominance effects in the case of biased polysemes. While readers are likely to have activated unshared semantic information in addition to the shared information, activation of the unshared features presumably decayed once the word had successfully been accessed. This explains why marginal response facilitation for the dominant and subordinate sense was observed with a shorter ITI but not with a longer ITI. For balanced polysemes, readers also initially activated some unshared semantic information in addition to shared information, which is why with a 50 ms prime-target delay, biased and balanced polysemes patterned together. In accounting for the differences in retrieval with a 200 ms ITI, we propose that the initially weaker activation of the less frequent, unshared features of a polyseme led to stronger between-sense competition when senses were roughly equally frequent than when senses were of unequal frequency. This difference in activation of unshared features, we surmise, explains why balanced polysemes patterned with balanced homonyms and not biased polysemes in the 200 ms ITI condition. Between-sense competition in the case of balanced polysemes occurred because the equally strong activation of incompatible, unshared features required readers to make a choice about which sense to retrieve. This is consistent with the longer reading times observed for balanced homonyms compared to unambiguous control words in sentence reading (Duffy et al., 1988; Folk & Morris, 2003; Rayner & Duffy, 1986; see also Armstrong & Plaut, 2011; MacGregor, Bouwsema, & Klepousniotou, 2015; Pylkkänen et al., 2006, our Experiment 2). Senses remained activated in an attempt to resolve the ambiguity, until a disambiguating context was encountered (which was at target in our experiment).

Our account of the activation pattern found in Experiment 1 is consistent with a battery of studies conducted by Gernsbacher and colleagues (Gernsbacher & Faust, 1991; Gernsbacher & St. John, 2001). These authors found that the two meanings of (balanced) homonyms became initially equally strongly activated and remained activated until a disambiguating clue was provided. Context then triggered a *signal* that initiated the suppression of the meaning that was not supported by the context. In contrast, in our results, the near absence of competition between senses of biased polysemes (due to the imbalance in the relative frequency of the senses) led to weaker activation of the unshared features (and possibly decay of the activation of unshared features soon after the polyseme had been retrieved; see Simpson & Burgess, 1985; Simpson & Krueger, 1991). This weaker activation of unshared features, in turn, led to the absence of observable effects of polyseme primes on the retrieval of related target words.

The assumption that, in the absence of a biasing context, polyseme retrieval comes with strong activation of a shared subset of meaning features explains why polysemes have often been found to be accessed more quickly than unambiguous words (Armstrong & Plaut, 2008, 2011; Azuma & Van Orden, 1997; Beretta et al., 2005; Klepousniotou & Baum, 2007; Locker et al., 2003; Rodd et al., 2002; Taler, Kousaie, & Zunini, 2013). It also explains why, for biased polysemes, disambiguating target words showed no priming with a 200 ms ITI. When polyseme retrieval is based on strong activation of shared features and decaying or weaker activation of unshared features, and when retrieval of a subsequent disambiguating target word requires strong activation of unshared features (to disambiguate the preceding prime), response facilitation associated with that target will be weak at best.

It is important to clarify the possibly counterintuitive observation that between-sense competition for balanced polysemes led to *faster* and not *slower* decision times for subsequent target words in Experiment 1. To understand this result, it is crucial to keep separate the retrieval processes involved in lexical access of the prime word and the retrieval processes involved in lexical access of the target word. While lexical access of balanced ambiguous words is predicted to be slowed down due to competition (Duffy et al., 1988; Folk & Morris, 2003; Rayner & Duffy, 1986), stronger competition leads to stronger activation of the meanings that caused the competition. Thus, when a word is encountered that is related to one of the competing meanings, response facilitation, and not inhibition, of that word is expected.

Turning to the data from sentence reading (Experiment 2), we first note that they largely replicate the results presented in Brocher et al. (2016) and other studies (Duffy et al., 1988; Folk & Morris, 2003; Rayner et al., 1994; Rayner & Duffy, 1986). For biased words, we found typical dominance effects for the disambiguation of homonyms and no difference between the ambiguous and neutral control condition for the disambiguation of irregular polysemes. These data nicely converge on the results from our priming study. Upon reading a biased homonym in a neutral context, readers committed to its dominant meaning. This then led to the need for a costly reanalysis when a subsequent context supported its less frequent meaning. For biased irregular polysemes, in contrast, readers did not commit to a specific sense and, when reading subordinate-biased contexts, they only needed to fill in the subordinate sense, without the need for reanalysis.

Data from balanced polysemes also align with the results from Experiment 1. These words again patterned with balanced homonyms: Readers slowed down for both kinds of balanced words. Importantly however, readers only took longer to read subsequent subordinate-biased context regions when they had retrieved a balanced homonym in the first clause and not when they had retrieved a balanced polyseme. This finding is compatible with the observation that readers more quickly overcome between-sense competition associated with biased polysemes than competition associated with biased homonyms (Brocher et al., 2016). In other words, although readers have presumably resolved the ambiguity of a balanced polyseme by selecting one sense, when that sense turned out to be incorrect, they were able to quickly compute the competing sense because of the activation of the shared features.

A shared features model of irregular polyseme representation and retrieval

Armstrong and Plaut (2008, 2011) have proposed a model of word retrieval that involves both co-operative and competitive settling dynamics and that, we believe, can help explain our results. In their model, during the initial stages of word retrieval, activation of semantic information engages mainly co-operative processes: Semantic features of ambiguous words that are compatible with one another, such as shared features, collaborate towards retrieval. At later stages of retrieval, competitive dynamics accrue, in that semantic features that are activated and incompatible with one another (e.g. unshared features) compete. The model put forward by Armstrong and Plaut can explain why words with related meanings are typically accessed more quickly than words with unrelated meanings (Armstrong & Plaut, 2008, 2011; Azuma & Van Orden, 1997; Beretta et al., 2005; Klepousniotou & Baum, 2007; Locker et al., 2003; Rodd et al., 2002; Taler et al., 2013) by alluding to the time course of different retrieval dynamics involved in lexical access. Importantly, and as noted by Armstrong and Plaut (2011), competitive processes should be strongest for balanced ambiguous words. Indeed, for biased ambiguous words, competitive processes are often weak or not observed at all (Armstrong & Plaut, 2008). Our own materials were not controlled in such a way that allows us to directly compare RTs of biased and balanced ambiguous words, but we did find compatible effects for biased and balanced ambiguous words in comparison to control word conditions.

If Armstrong and Plaut's model and the representation we propose for irregular polysemes are on the right track, at early stages of lexical access, co-operative settling dynamics, involving strong facilitatory connections, lead to strong activation of the shared features. At this stage, competitive settling dynamics are rather weak, leading to weak activation of the unshared features. Indeed, because shared features are most frequent they should contribute most strongly towards co-operative retrieval processes, for both biased and balanced polysemes. The crucial difference between the two kinds of irregular polysemy is that, even at early stages of retrieval, competitive retrieval processes should be stronger for balanced than biased polysemes. This is predicted because the closer two senses are to being equal in frequency, the stronger the competition between the unshared semantic information of the two senses should be (Armstrong & Plaut, 2013). As a consequence, lexical access should be slower for balanced compared to biased polysemes because of between-sense competition (see Armstrong & Plaut, 2008, 2013; Beretta et al., 2005; Locker et al., 2003; Rodd et al., 2002). Slower lexical access for balanced polysemes leaves more time for competitive processes to accrue and leads to further activation of the unshared features (because competition has not yet been resolved).

In sum, given the time course of co-operative and competitive retrieval processes and the claim that competition works as a function of sense frequency, the two senses of balanced irregular polysemes are expected to be much more strongly activated than the two senses of biased irregular polysemes. The difference in target facilitation between biased and balanced polysemes that we found in Experiment 1 therefore stems from the strength of competitive rather than co-operative retrieval dynamics. For polysemes with roughly equally frequent senses, between-sense competition leads to an increase in activation of the unshared features, which then begins to resolve or decay when a disambiguating context is provided. For polysemes with senses of unequal frequencies, there is, at best, weak competition between the two senses at first, leading to a decay of activation of the unshared features (Simpson & Burgess, 1985; Simpson & Krueger, 1991).

On a representational level, we take the data from balanced irregular polysemes as evidence that the relative frequencies of a polyseme's senses are associated with the ambiguous word's lexical representation. This aspect of our data speaks against an underspecification model and in favour of a shared features model. That is, while the relative frequencies of senses are associated with the unshared features according to a shared features model, they are not associated with the underspecified node, according to an underspecification model (see Figure 1). It bears repeating, though, that the underspecification model has been designed to account for the processing of *regular* polysemes like CHICKEN and BOOK and has not yet been tested on irregular polysemes. But, as it stands, it is not clear how an underspecification model could account for the observed differences between biased and balanced irregular polysemes. If all polysemes are accessed via an underspecified node, there should be no retrieval differences between biased and balanced (irregular) polysemes.

Our conception of shared features overlap is generally consistent with the concept of sharing a core meaning (as proposed by Klepousniotou, 2002; Klepousniotou & Baum, 2007; Klepousniotou et al., 2012; Rodd et al., 2002, 2004; Williams, 1992). However, we stress that we have focused specifically on irregular polysemy in this investigation and in our model. The assumption that senses are overlapping, an assumption that is crucial to our model, does not necessarily apply to regular polysemes. It is still a matter of discussion whether the senses of regular polysemes, which can be *related* by rules, share a significant number of salient *features*. For example, it is not clear to what extent NEWSPAPER, as a concrete object with sheets of newsprint, and NEWSPAPER, as the abstract content communicated in the publication, do in fact share many features (Klein & Murphy, 2001, 2002; Murphy, 2007; Rice, 1992; Taylor, 2003).

Finally, we agree with Klepousniotou (2002), Klepousniotou and Baum (2007), and Klepousniotou et al. (2012) that lexical ambiguity is best understood as a continuum. However, we view relatedness, similarity, and type of lexical ambiguity (homonyms, regular polysemes, and irregular polysemes) as three distinct dimensions of classification, with different representational and processing impacts. Although it is possible that the multiple senses of regular polysemes are on average conceived to be more related than the multiple senses of irregular polysemes, our similarity rating does not speak to this issue. An important question for future research is to discover the point at which the two readings of an ambiguous word share a sufficient number of features (or sufficiently salient features) such that the relative frequency of senses does not noticeably interfere with retrieval. Indeed, the homonym/irregular polyseme distinction is only a theoretical one. Semantic similarity strongly varies across lexically ambiguous items (as is supported by our local norming studies as well as Klepousniotou et al., 2008). For irregular polysemes with weakly related senses, activation of the shared features might not be sufficiently strong to greatly contribute to lexical retrieval. Under these conditions, unshared features should also quickly and strongly be activated, making these irregular polysemes behave more like homonyms.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Apresjan, J. (1974). Regular polysemy. *Linguistics*, *12*, 5–32. doi:10.1515/ling.1974.12.142.5

Armstrong, B. C., & Plaut, D. C. (2008). Settling dynamics in distributed networks explain task differences in semantic ambiguity effects: Computational and behavioral evidence. In *Proceedings of the 30th annual conference of the cognitive science society* (pp. 273–278). Hillsdale, NJ: Cognitive Science Society.

Armstrong, B. C., & Plaut, D. C. (2011). Inducing homonymy effects via stimulus quality and (not) nonword difficulty: Implications for models of semantic ambiguity and word recognition. In *Proceedings of the 33rd annual conference of the cognitive science society*. Hillsdale, NJ: Cognitive Science Society (pp. 2223–2228).

Armstrong, B. C., & Plaut, D. C. (2013). Simulating overall and trial-by-trial effects in response selection with a biologically plausible connectionist network. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 139–144). Austin, TX: Cognitive Science Society.

Armstrong, B. C., & Plaut, D. C. (2016). Disparate semantic ambiguity effects from semantic processing dynamics rather than qualitative task differences. *Language, Cognition and Neuroscience*, *31*, 940–966. doi:10.1080/23273798.2016.1171366

Azuma, T., & Van Orden, G. C. (1997). Why safe is better than fast: The relatedness of a word's meanings affect lexical decision times. *Journal of Memory and Language*, *36*, 484–504. doi:10.1006/jmla.1997.2502

Baayen, R. H. (2008). *Analyzing linguistic data. A practical Introduction to statistics using R*. Cambridge University Press. doi:10.1017/CBO9780511801686

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (release 2, CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–6*. Retrieved from http://CRAN.R-project.org/package=lme4

Beretta, A., Fiorentino, R., & Poeppel, D. (2005). The effects of homonymy and polysemy on lexical access: An MEG study. *Cognitive Brain Research*, *24*, 57–65. doi:10.1016/j.cogbrainres.2004.12.006

Borowsky, R., & Masson, M. E. J. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 63–85. doi:10.1037/0278-7393.22.1.63

Bott, L., Rees, A., & Frisson, S. (2016). The time course of familiar metonymy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 1160–1170. doi:10.1037/xlm0000218

Brocher, A., Foraker, S., & Koenig, J. P. (2016). Processing of irregular polysemes in sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 1798–1813. doi:10.1037/xlm0000271

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*, 977–990. doi:10.3758/BRM.41.4.977

Bueno, S., & Frenck-Mestre, C. (2008). The activation of semantic memory: Effects of prime exposure, prime-target relationship, and task demands. *Memory & Cognition*, *36*, 882–898. doi:10.3758/MC.36.4.882

Copestake, A., & Briscoe, E. J. (1995). Semi-productive polysemy and sense extension. *Journal of Semantics*, *12*, 15–67. doi:10.1093/jos/12.1.15

Cruse, D. A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press. doi:10.1016/B0-08-043076-7/02990-9

Dholakia, A., Meade, G., & Coch, D. (2016). The N400 elicited by homonyms in puns: Two primes are not better than one. *Psychophysiology*, *53*, 1799–1810. doi:10.1111/psyp.12762

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, *27*, 429–446. doi:10.1016/0749-596X(88)90066-6

Eddington, C. M., & Tokowicz, N. (2015). How meaning similarity influences ambiguous word processing: The current state of the literature. *Psychonomic Bulletin & Review*, *22*, 13–37. doi:10.3758/s13423-014-0665-7

Folk, J. R., & Morris, R. K. (2003). Effects of syntactic category assignment on lexical ambiguity resolution in reading: An eye movement analysis. *Memory & Cognition*, *31*, 87–99. doi:10.3758/BF03196085

Foraker, S., & Murphy, G. L. (2012). Polysemy in sentence comprehension: Effects of meaning dominance. *Journal of Memory and Language*, *67*, 407–425. doi:10.1016/j.jml.2012.07.010

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*, 627–635. doi:10.1016/S0022-5371(73)80042-8

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language*, *29*, 181–200. doi:10.1016/0749-596X(90)90071-7

Frisson, S. (2009). Semantic underspecification in language processing. *Language and Linguistics Compass*, *3*, 111–127. doi:10.1111/j.1749-818X.2008.00104.x

Frisson, S. (2015). About bound and scary books: The processing of *book* polysemies. *Lingua*, *157*, 17–35.

Frisson, S., & Frazier, L. (2005). Carving up word meaning: Portioning and grinding. *Journal of Memory and Language*, *53*, 277–291. doi:10.1016/j.jml.2005.03.004

Frisson, S., & Pickering, M. J. (1999). The processing of metonymy: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1366–1383. doi:10.1037/0278-7393.25.6.1366

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York, NY: Cambridge University Press. doi:10.1017/CBO9780511790942

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*, 256–281. doi:10.1037/0096-3445.113.2.256

Gernsbacher, M. A., & Faust, M. (1991). The role of suppression in sentence comprehension. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 245–262). North-Holland: Elsevier Science Publishers B.V.

Gernsbacher, M. A., & St. John, M. F. (2001). Modeling suppression in lexical access. In D. S. Gorfein (Ed.), *Perspectives on resolving lexical ambiguity*. Washington, DC: American Psychological Association. doi:10.1037/10459-003

Gibbs Jr, R. W. (1993). Process and products in making sense of tropes. In A. Ortony (Ed.), *Metaphor and thought* (pp. 258–276), Cambridge: Cambridge University Press. doi:10.1017/CBO9781139173865.014

Gottlob, L. R., Goldinger, S. D., Stone, G. O., & Orden, G. C. V. (1999). Reading homographs: Orthographic, phonologic, and semantic dynamics. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 561–574. doi:10.1037/0096-1523.25.2.561

Gunter, T. C., Wagner, S., & Friederici, A. D. (2003). Working memory and lexical ambiguity resolution as revealed by ERPs: A difficult case for activation theories. *Journal of Cognitive Neuroscience*, *15*, 643–657. doi:10.1162/jocn.2003.15.5.643

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review*, *10*, 785–813. doi:10.3758/BF03196544

Klein, D. E., & Murphy, G. L. (2001). The representation of polysemous words. *Journal of Memory and Language*, *45*, 259–282. doi:10.1006/jmla.2001.2779

Klein, D. E., & Murphy, G. L. (2002). Paper has been my ruin: Conceptual relations of polysemous senses. *Journal of Memory and Language*, *47*, 548–570. doi:10.1016/S0749-596X(02)00020-7

Klepousniotou, E. (2002). The processing of lexical ambiguity: Homonymy and polysemy in the mental lexicon. *Brain and Language*, *81*, 205–223. doi:10.1006/brln.2001.2518

Klepousniotou, E., & Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics*, *20*, 1–24. doi:10.1016/j.jneuroling.2006.02.001

Klepousniotou, E., Pike, G. B., Steinhauer, K., & Gracco, V. (2012). Not all ambiguous words are created equal: An EEG investigation of homonymy and polysemy. *Brain & Language*, *123*, 11–21. doi:10.1016/j.bandl.2012.06.007

Klepousniotou, E., Titone, D., & Romero, C. (2008). Making sense of word senses: The comprehension of polysemy depends on sense overlap. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1534–1543. doi:10.1037/a0013012

Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago: University of Chicago Press. doi:10.7208/chicago/9780226471013.001.0001

Li, L., & Slevc, R. (2017). Of papers and pens: Polysemes and homophones in lexical (mis)selection. *Cognitive Science*, *41*, 1532–1548. doi:10.1111/cogs.12402

Locker, L., Simpson, G. B., & Yates, M. (2003). Semantic neighborhood effects on the recognition of ambiguous words. *Memory & Cognition*, *31*, 505–515. doi:10.3758/BF03196092

MacGregor, L. J., Bouwsema, J., & Klepousniotou, E. (2015). Sustained meaning activation for polysemous but not homonymous words: Evidence from EEG. *Neuropsychologia*, *68*, 126–138. doi:10.1016/j.neuropsychologia.2015.01.008

Mason, R. A., & Just, M. A. (2007). Lexical ambiguity in sentence comprehension. *Brain Research*, *1146*, 115–127. doi:10.1016/j.brainres.2007.02.076

McNamara, T. P., & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language*, *27*, 545–559. doi:10.1016/0749-596X(88)90025-3

McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 558–572. doi:10.1037/0278-7393.24.3.558

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*, 547–559. doi:10.3758/BF03192726

Murphy, G. L. (2007). Parsimony and the psychological representation of polysemous words. In M. Rakova, G. Petho, & C. Rákosi (Eds.), *The cognitive basis of polysemy*. Frankfurt am Main, Germany: Peter Lang Verlag.

Navarro, D. J. (2014). *Learning statistics with R: A tutorial for psychology students and other beginners (Version 0.4)*. Adelaide: University of Adelaide.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254. doi:10.1037/0096-3445.106.3.226

Nelson, D., McEvoy, C. L., & Schreiber, T. (2004). The University of South Florida free association rhyme and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*, 402–407. doi:10.3758/BF03195588

Nunberg, G. (1979). The non-uniqueness of semantic solutions: polysemy. *Linguistics and Philosophy*, *3*, 143–184. doi:10.1007/BF00126509

Perea, M., & Gotor, A. (1997). Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. *Cognition*, *62*, 223–240. doi:10.1016/S0010-0277(96)00782-2

Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological Research*, *66*, 180–194. doi:10.1007/s00426-002-0086-5

Pickering, M. J., & Frisson, S. (2001). Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 556–573. doi:10.1037/0278-7393.27.2.556

Pylkkänen, L., Llinas, R., & Murphy, G. L. (2006). The representation of polysemy: Meg evidence. *Journal of Cognitive Neuroscience*, *18*, 97–109. doi:10.1162/089892906775250003

Rabagliati, H., Marcus, G. F., & Pylkkänen, L. (2011). Rules, radical pragmatics and restrictions on regular polysemy. *Journal of Semantics*, *28*, 485–512. doi:10.1093/jos/ffr005

Rabagliati, H., & Snedeker, J. (2013). The truth about chickens and bats: Ambiguity avoidance distinguishes types of polysemy. *Psychological Science*, *24*, 1354–1360. doi:10.1177/0956797612472205

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*, 191–201. doi:10.3758/BF03197692

Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory and Language*, *33*, 527–544. doi:10.1006/jmla.1994.1025

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

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.

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, *46*, 245–266. doi:10.1006/jmla.2001.2810

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science*, *28*, 89–104. doi:10.1207/s15516709cog2801_4

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, *14*, 489–537. doi:10.1016/0010-0285(82)90017-2

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior*, *20*, 120–136. doi:10.1016/S0022-5371(81)90356-X

Simpson, G. B., & Burgess, C. (1985). Activation and selection processes in the recognition of ambiguous words. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 28–39. doi:10.1016/S0022-5371(81)90356-X

Simpson, G. B., & Krueger, M. A. (1991). Selective access of homograph meanings in sentence context. *Journal of Memory and Language*, *30*, 627–643. doi:10.1016/0749-596X(91)90029-J

Swinney, D. A. (1979). Lexical access during sentence comprehension: (re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645–659. doi:10.1016/S0022-5371(79)90355-4

Taler, V., Kousaie, S., & Zunini, R. L. (2013). ERP measures of semantic richness: The case of multiple senses. *Frontiers in Human Neuroscience*, *7*, 1–6. doi:10.3389/fnhum.2013.00005

Tamminen, J., & Gaskell, M. G. (2013). Novel word integration in the mental lexicon: Evidence from unmasked and masked semantic priming. *The Quarterly Journal of Experimental Psychology*, *66*, 1001–1025. doi:10.1080/17470218.2012.724694

Taylor, J. (2003). Cognitive models of polysemy. In B. Nerlich, Z. Todd, V. Herman, & D. D. Clarke (Eds.), *Polysemy – Flexible patterns of meaning in mind and language* (pp. 31–48), Berlin: Mouton de Gruyter. doi:10.1515/9783110895698.31

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory & Cognition*, *22*, 111–126. doi:10.3758/BF03202766

Whaley, C. P. (1978). Word—nonword classification time. *Journal of Verbal Learning and Verbal Behavior*, *17*, 143–154. doi:10.1016/S0022-5371(78)90110-X

Williams, J. (1992). Processing polysemous words in context: Evidence for interrelated meanings. *Journal of Psycholinguistic Research*, *21*, 193–218. doi:10.1007/BF01068072