available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Understanding words in context: The role of Broca's area in word comprehension

Marina Bedny[a,b,*], *Justin C. Hulbert*[c], *Sharon L. Thompson-Schill*[a]

[a]*Center for Cognitive Neuroscience, Department of Psychology, University of Pennsylvania, 3720 Walnut Street, Philadelphia, PA 19104-6241, USA*

[b]*Center for Non-Invasive Brain Stimulation, Beth Israel Deaconess Medical Center, Harvard Medical School, 330 Brookline Ave. Kirstein Building KS-430, Boston, MA 02215, USA*

[c]*Psychology Department, 1227 University of Oregon, Eugene, OR 97403-1227, USA*

ARTICLE INFO

Article history:
Accepted 5 October 2006
Available online 22 November 2006

Keywords:
Left inferior frontal grgyrus
Word comprehension
Meaning selection
Frontal lobe
Homonym
Polysemy
Ambiguity

ABSTRACT

What role does meaning selection play in word comprehension, and what neural systems support this selection process? Most words have multiple meanings and are therefore ambiguous. This is true of both homonymous words (words that have multiple unrelated meanings) and polysemous words (words that have multiple related meanings). The extant evidence indicates that meaning selection is an integral part of homonym comprehension. However, it is not known whether meaning selection extends to polysemous words, or what neural systems support meaning selection during comprehension. Prior neuroimaging and neuropsychological evidence suggest that the left inferior frontal gyrus (LIFG) may play a role in resolving competition during language processing. We therefore sought to test the hypotheses that meaning selection is part of polysemous word comprehension, and that the LIFG resolves meaning competition during word comprehension. We tested healthy participants on a version of the triplet lexical decision task, with polysemous and homonymous stimuli. Results suggest that the meanings of polysemous words, like the meanings of homonyms, are selected based on context. However, homonymous and polysemous words differed in how meaning frequency affected meaning selection. We then administered the triplet lexical decision task to patients with LIFG damage to examine whether this region plays a role in context-dependent meaning selection. Results support the hypothesis that the LIFG serves as a top-down biasing mechanism that facilitates rapid meaning selection during word comprehension. We conclude that context-dependent meaning selection is an integral part of word comprehension for both homonyms and polysemous words, and that the LIFG facilitates this selection process.

1. Introduction

Virginia Woolf once described words as "…the wildest, freest, most irresponsible, most un-teachable of all things … because the truth they try to catch is many sided. And they convey it by being many sided, dashing first this way than that…" (Woolf, 1937). In psycholinguistics, this fluid property of words has been described as lexical-semantic ambiguity. Two major

* *Corresponding author*. Center for Non-Invasive Brain Stimulation, Beth Israel Deaconess Medical Center, Harvard Medical School, 330 Brookline Ave. Kirstein Building KS-430, Boston, MA 02215, USA. Fax: +1 617 975 5322.
E-mail address: mbedny@bidmc.harvard.edu (M. Bedny).

0006-8993/$ – see front matter
doi:10.1016/j.brainres.2006.10.012

types of lexical-semantic ambiguity have traditionally been distinguished: homonymy and polysemy. Homonyms have two or more unrelated meanings that, by historical accident, share a single form (e.g., organ: musical instrument or body part); on the other hand, polysemous words have two or more related meanings called senses (e.g., chicken: food or bird). In contrast to the meanings of homonyms, the sense extensions of polysemous words can be systematic (e.g., the names of animals can also refer to their meat) (Cruse, 1986; Pustejovsky, 1995; Weinreich, 1964).

While in linguistics, homonymy and polysemy are treated as categorically different phenomena, connectionist models of word recognition represent the homonymy/polysemy distinction as a continuum. The meanings of homonyms are instantiated as non-overlapping representations over a set of semantic units, while the sense representation of polysemous words partially overlaps (Kawamoto, 1993, 1994; Rodd et al., 2004). Word senses can be thought of as local minima in a large shallow attractor basin, while meanings of homonyms are global minima in competition with each other during comprehension (Rodd et al., 2004). The meanings of homonyms and frequent meanings of polysemous words are retrieved from long term memory (Klein and Murphy, 2001, 2002), while novel senses are generated during comprehension (Klepousniotou et al., 2006). In a connectionist framework, the retrieval of meanings and senses from long-term memory, as well as sense generation, involves activating representations in a distributed network. The selection of meanings and senses involves activating one of several preexisting representations that are associated with a given word form.[1] In the case of sense generation, a representation that has not previously been activated is carved out of the semantic space (Kawamoto, 1993).

During online word comprehension, two major factors shape meaning retrieval: meaning frequency (a measure of the association between a word-form and a meaning in long-term memory), and the current context. Meaning frequency has been repeatedly shown to affect behavior: more frequent meanings are accessed faster than less frequent meanings (Duffy et al., 1988, 2001). On the other hand, context can provide unequal support for the different meanings of a single form Martin et al. (1999). Thus, the interaction of meaning frequency and context shapes lexical ambiguity resolution.

Data from behavioral and ERP experiments have demonstrated that under most circumstances, multiple meanings of homonyms are initially activated (e.g., Swaab et al., 2003; Swinney, 1979). For example, immediately following the presentation of an ambiguous word, subjects make faster lexical decisions to words related to either one of its meanings (as compared to words unrelated to either meaning of the ambiguous word) (Swinney, 1979). While multiple meanings are initially retrieved, context is rapidly used to select the appropriate meaning (Swinney, 1979; Tanenhaus et al., 1979). Four syllables following an ambiguous word (but not before this time) subjects are faster to make a lexical decision only to target words that are related to the contextually appropriate meaning of the ambiguous word (Swinney, 1979). Although these early findings suggested that context affects only post access processes (Swinney, 1979; Tanenhaus et al., 1979), eye-tracking and ERP data have provided evidence that context can influence lexical access much earlier in processing that previously thought (Duffy et al., 1988; Martin et al., 1999; Sereno et al., 2003).

The mechanism by which context influences meaning selection continues to be debated. Most authors agree that context directly boosts the activation of the appropriate meaning (Duffy et al., 2001; Gernsbacher, 1990; Gorfein, 2001; McNamara and McDaniel, 2004). However, there is some disagreement about how (or whether) context dampens the activation of inappropriate meanings. According to one class of hypotheses, in the absence of contextual support, the inappropriate meaning passively decays over time (Duffy et al., 2001). Alternatively, the context inappropriate meaning may be dampened through lateral inhibition by the context-appropriate meaning. Finally, Gernsbacher and colleagues have argued for a direct suppression mechanism, whereby context (in the form of a higher order, sentence gestalt representation) actively suppress the inappropriate lexical-semantic candidate (Gernsbacher, 1990; Shivde and Anderson, 2001; Simpson, 2001).

The latter form of suppression is implemented in a mode of context-dependent meaning selection (Gernsbacher and St. John, 2001). This model contains four layers of representation: an orthographic input layer, lexical layer, a concept layer, and a sentence gestalt layer. The representation of the "sentence gestalt" exerts a top-down influence to suppress irrelevant conceptual representations, and increases the activation of context appropriate representations (Gernsbacher and St. John, 2001).

A fundamentally important question is how context-based meaning retrieval is implemented in the neural systems that support language. Recently, a network of brain regions that respond to lexical-semantic ambiguity was identified using functional neuroimaging (Rodd et al., 2005). Participants heard sentences in high- and low-ambiguity conditions. The high-ambiguity sentences contained two or more homonyms, while the low-ambiguity condition sentences did not contain homonyms. The two sentence types were matched on syntactic complexity and naturalness ratings. Compared to low ambiguity sentences, high ambiguity sentences produced greater activity in the left and right inferior frontal gyri (IFG), left inferior, middle and superior temporal gyri, as well as the anterior cingulate. The distinct contributions of the left and right inferior frontal, lateral temporal, and anterior cingulate regions to context-dependent meaning selection are not known. Moreover, due to limits in the inference afforded by fMRI data, activity in some of these regions such as the RIFG may be epiphenomenal, and may not play a necessary role in ambiguity resolution. For example, there is some evidence that meaning selection is restricted to the left hemisphere, suggesting that RIFG activity observed in the high ambiguity condition does not play a functional role in meaning selection (Burgess and Simpson, 1988). In contrast, indirect evidence suggests that

[1] In the present paper, we use the term "selection" to refer to the activation of one meaning among the multiple meaning-representations that are associated with a single form. As such, selection need not correspond to a discrete processing stage and may occur in parallel with meaning activation and integration.

the LIFG in particular may act as a biasing mechanism that facilitates rapid meaning selection during word comprehension (Thompson-Schill et al., 2005).

Indirect evidence for the idea that the LIFG is important for meaning selection comes from neuropsychological studies. Patients with bilateral (but not right-lateralized) frontal lobe damage do not select the context-appropriate meaning of ambiguous words in a naming task (Metzler, 2001). Furthermore, there is some evidence that Broca's aphasics are impaired in context-dependent meaning selection. The ability to make inferences regarding the function of the LIFG from studies of Broca's aphasia is limited by the fact that not all Broca's aphasics have damage to the LIFG (Dronkers, 2000). However, a number of studies of Broca's aphasia provide sufficient lesion information to conjecture that damage to the LIFG is associated with deficits in word comprehension (Dronkers, 2000). Milberg et al. (1987) were the first to show that Broca's aphasics do not show normal semantic priming in some paradigms. They used the triplet lexical decision task (Schvaneveldt and Meyer, 1976), in which participants see or hear words, one at a time. Words are grouped into triplets, with a longer inter-stimulus-interval between triplets than between individual words within a triplet. On critical trials, the first word of each triplet serves as a cue, the second word is ambiguous, and the third word is the target. The measure of interest is reaction time to the target word, as a function of condition. Most pertinent to our discussion are three conditions: consistent, inconsistent, and unrelated. In the consistent condition, the first and third words are related to the same meaning of the ambiguous word; in the inconsistent condition the first and third words are related to different meanings of the middle-ambiguous word; in the unrelated condition the first two words are not related to the target word.

In this paradigm, control subjects are significantly faster to respond to the target word in the consistent condition, relative to the inconsistent condition. The reaction time in the neutral condition is significantly slower than the consistent condition, and not different from the inconsistent condition (Schvaneveldt and Meyer, 1976). Slower response times in the inconsistent relative to the consistent condition illustrate that the priming effect of an ambiguous word depends on the context in which it appears. Said differently, control subjects are able to select the context appropriate meanings of ambiguous words (Schvaneveldt and Meyer, 1976). In contrast, Milberg et al. (1987) found that for patients with Broca's aphasia, reaction times in the consistent, inconsistent, or neutral conditions did not different from each other. Based on this finding of abnormal semantic priming, Milberg et al. argued that single word comprehension is impaired in Broca's aphasics.

The failure of Broca's aphasics to show normal semantic priming in triplet lexical decision paradigm has led some authors to hypothesize that Broca's aphasics have a deficit in automatically accessing word meanings (Milberg and Blumstein, 1989; Milberg et al., 1987; Prather et al., 1992, 1997). However, the role of the LIFG in word comprehension has remained controversial because Broca's aphasics show normal semantic priming on some paradigms but not others: for instance, Broca's aphasics show a deficit in lexical-semantic priming in the triplet paradigm described above, but they exhibit normal semantic priming when tested with word pairs (Katz, 1988). Such findings have led to the hypothesis that the LIFG is not critical for accessing word meanings, but instead plays a role in integrating words into the sentence context (Hagoort, 1989, 1993; Swaab et al., 1998). In support of this integration hypothesis, ERP data have shown that while control subjects select the context appropriate meaning of an ambiguous word within 100 ms of its presentation, Broca's aphasics do not show evidence of selection until 1250 ms. These data suggest that damage to the LIFG may delay contextual selection, a component of the integration process (Swaab et al., 1998).

While there is convincing evidence that Broca's aphasics have deficits in lexical integration, and possibly lexical access, a characterization of the LIFG as exclusively a lexical integration or access mechanism is incomplete in that it is isolated from the broader literature on the function of the LIFG. Furthermore, an understanding of the broader picture of LIFG function can inform us regarding what the LIFG contributes to the processes of lexical integration, and access. A large body of evidence suggests that the LIFG is involved in a range of language tasks including production and comprehension, as well as word and sentence level processing (Broca, 1861; Caramazza and Zurif, 1976; Novick et al., 2005; Thompson-Schill et al., 2005).

The role of the LIFG in these different tasks is controversial, but a number of studies have demonstrated that activity in the LIFG increases under conditions of competition between semantic representations (Barde and Thompson-Schill, 2002; Kan and Thompson-Schill, 2004; Thompson-Schill et al., 1997, 1999, 2005). For example, in fMRI studies of the verb-generation task, there is greater LIFG activity for nouns that have many associated verbs (high competition), than for nouns that have one strongly associated verb (low competition) (Thompson-Schill et al., 1997). Moreover, patients with LIFG damage make errors generating verbs to high but not low competition nouns (Thompson-Schill et al., 1998). Such data have led to the hypothesis that the LIFG is important for resolving competition between semantic representations.

Combining the broader literature on the function of the LIFG, and evidence from the performance of Broca's aphasics, we hypothesize that the LIFG plays a role in selecting the context-appropriate meaning of ambiguous words. In the present study, we test this hypothesis by comparing the performance of patients with LIFG damage to the performance of patients with damage to the frontal lobe that spares the LIFG. We further hypothesize that meaning selection is not restricted to the comprehension of homonyms, but rather extends to the comprehension of polysemous words. To test these hypotheses, we administered the word triplet, lexical decision task (Schvaneveldt and Meyer, 1976), to patients with LIFG damage, patients with non-LIFG frontal lobe damage, and age- and education-matched healthy controls. Critical word triplets occurred in two conditions: consistent and inconsistent. In the consistent condition, the first and third words are related to the same meaning of the middle word. In the inconsistent condition, the first and third words are related to different meanings of the middle word. The critical measure of interest is the reaction time to the third word (i.e., the target). In addition to meaning consistency, we manipulated ambiguity type. For half of the word triplets, the middle word was a homonym; for the other half of the triplets, the middle word was polysemous. We predicted that both age-matched controls and patients with non-LIFG frontal lobe damage would show evidence of meaning selection by responding more quickly to the target word in the consistent than inconsistent condition, whereas patients with damage to the LIFG would show an absent, or reduced effect of consistency.

A further goal of the present study was to compare context-dependent meaning selection for homonymous and polysemous words, and to examine the effect of meaning frequency on the selection of homonymous and polysemous word-meanings. There is currently conflicting evidence as to whether the meanings of polysemous words can be selected based on context. Can one sense of a polysemous word be more active than another, or are the senses of polysemous words inextricably linked? Initial studies on this topic indicated that meanings of polysemous words, unlike homonyms, cannot be activated independently of each other (Frazier and Rayner, 1990; Williams, 1992). However, recent data from a relatedness judgment task provide support for the idea that senses of polysemous words that are related, but saliently distinct can be retrieved independent of each other (Klein and Murphy, 2001, 2002). Moreover, based on a recent distributed model of meaning representation, it was proposed that the senses of polysemous words compete when only one sense is context appropriate (Rodd et al., 2004). Based on this framework of distributed meaning representation, we hypothesized that meanings of polysemous words would be selected in the current version of the lexical decision task, where a biasing context is provided.

2. Experiment 1

2.1. Results

Participants saw three words appear on the screen, one at a time and made a lexical decision to each word in the triplet. The dependent measure of interest was the reaction time and accuracy to the third word in the triplet (i.e., the target) as a function of condition (consistent or inconsistent) and ambiguity type (homonymous or polysemous). Overall reaction time and accuracy measures for each group are summarized in Table 1. Accuracy data, as a function of condition, are summarized in Table 2. We report only analyses on the critical trials.

Table 1 – Task performance summary (Experiment 1)

	Critical trials		Filler trials	
	Accuracy	Reaction time	Accuracy	Reaction time
LIFG patients	97.3(2)	760(115)	80.9(7)	996(156)
Frontal controls	97.4(3)	714(104)	85.2(1)	954(160)
Age-matched controls	99.0(2)	599(97)	91.8(7)	760(103)

Mean reaction times and accuracy for target trial responses. Critical trials include the consistent and inconsistent conditions.

2.1.1. *Age matched control subjects*

Participants were faster to respond to target words in the consistent than inconsistent condition as evidenced by a main effect of consistency [$M_{con}=592$, SD=96, $M_{inc}=606$, SD=99; $F_1(1,56)=26.69$, $p<0.0001$; $F_2(1,96)=17.76$, $p<0.0001$] (see Fig. 1). There was also a main effect of ambiguity type: participants were faster to respond to target words in the homonymous triplets than in the polysemous triplets [$M_{hom}=589$, SD=95, $M_{poly}=609$, SD=99; $F_1(1,56)=52.15$, $p<0.0001$, $F_2(1,96)=5.50$, $p<0.05$]. The effect of ambiguity type is most likely due to the fact that the target words in the homonymous triplets were slightly shorter and more frequent than the target words in the polysemous triplets. The effect of consistency was smaller for polysemous than for homonymous triplets, as evidenced by a marginal ambiguity-type-by-consistency interaction [$F_1(1,56)=3.50$, $p=0.07$; $F_2(1,96)=4.89$, $p<0.05$]. Planned comparisons revealed an effect of consistency for both homonymous trials [$M_{con}=579$, SD=94, $M_{inc}=598$, SD=98; $t_1(19)=5.08$, $p<0.0001$; $t_2(47)=4.47$, $p<0.0001$], and polysemous trials [$M_{con}=604$, SD=98, $M_{inc}=614$, SD=101; $t_1(19)=2.16$, $p<0.05$; $t_2(49)=1.43$, $p=0.16$]. However, the effect of consistency for polysemous words did not reach significance in the item-wise analysis.

We also examined reaction time as a function of meaning frequency, consistency, and ambiguity type. Triplets were grouped based whether the target was related to the most frequent meaning of the ambiguous word (M1 triplets), or a less frequent meaning of the ambiguous word (M2 triplets). There was no main effect of meaning frequency in the reaction time data ($F<1$). However, there was a significant meaning frequency by consistency interaction. The consistency effect was larger for M2 triplets than for M1 triplets [$M_2=20$ ms, SD=13, $M_1=9$ ms, SD=18; $F_1(1,133)=4.05$, $p<0.05$]. This was qualified by a marginal three-way interaction between meaning frequency, consistency, and ambiguity type [$F_1(1,133)=3.57$, $p=0.06$]: the interaction of meaning frequency and consistency was present for homonymous, but not polysemous triplets (see Fig. 2).

Participants were slightly more accurate in the consistent than inconsistent condition [$M_{con}=99.3\%$, SD=3%, $M_{inc}=98.7$, SD=3%; $F_1(1,56)=4.90$, $p<0.05$; $F_2(1,96)=4.43$, $p<.05$]. No other effects approached significance in the accuracy analyses ($F<2$). In the planned comparisons, the effect of consistency on accuracy did not reach significance for homonyms [$t_1(16)=1.27$, $p>0.20$; $t_2(47)=1.60$, $p=0.12$], or polysemous words [$t_1(16)=1.27$, $p>0.20$; $t_2(49)=0.95$, $p=0.35$].

2.1.2. *Patient groups*

2.1.2.1. *Frontal controls.* Patients with non-LIFG frontal lobe damage were significantly faster to respond to target words in the consistent than inconsistent condition [$M_{con}=702$, SD=103, $M_{inc}=726$, SD=107; $F_1(1,18)=6.63$, $p<0.05$; $F_2(1,96)=7.06$, $p<0.01$]. Participants were also faster to respond to target words in the homonymous triplets. The effect of ambiguity-type was significant in the subject-wise analyses, and marginal in the item analyses [$M_{hom}=700$, SD=93, $M_{poly}=729$, SD=117; $F_1(1,18)=9.85$,

Table 2 – Target response accuracy and reaction time as a function of condition (Experiment 1)

Group	Homonym				Polysemous			
	Consistent		Inconsistent		Consistent		Inconsistent	
	% Correct	RT	% Correct	RT	% Correct	RT	% Correct	RT
LIFG patients	97.7(1.9)	728(115)	96.9(2.8)	748(109)	97.5(2.0)	781(125)	96.9(3.0)	782(123)
Frontal controls	98.4(2.7)	686(83)	98.0(1.7)	714(105)	97.6(4.7)	720(123)	95.7(4.4)	739(112)
Age controls	99.5(1.2)	579(94)	98.7(2.9)	598(98)	99.1(1.6)	604(98)	98.6(2.4)	614(101)

Mean percent correct and reaction time for target words as a function of condition, for each group. Standard deviations of the mean appear in parenthesis.

$p<0.01$, $F_2(1,96)=3.00$, $p<0.09$]. There was no ambiguity-type-by-consistency interaction ($F<1$). For homonyms, the effect of consistency was marginal in the subject-wise analyses, and significant in the item-wise analysis [$M_{con}=686$, SD=83, $M_{inc}=714$, SD=105; $t_1(6)=2.06$, $p<0.09$; $t_2(47)=2.69$, $p<0.01$]. For polysemous words, the effect of consistency reached significance in the subject-wise analysis, but not in the item-wise analysis [$M_{con}=720$, SD=123, $M_{inc}=739$, SD=112; $t_1(6)=2.66$, $p<0.05$; $t_2(49)=1.27$, $p=0.21$]. The only effect to reach significance in the accuracy data was the item-wise effect of consistency for polysemous words [$t_2(49)=2.04$, $p=0.047$]. The effect of ambiguity type was marginal in the accuracy analyses [$F_1(1,18)=3.32$, $p=0.09$, $F_2(1,96)=2.72$, $p=0.10$]. The interaction of meaning frequency and consistency did not reach significance for homonymous or polysemous words.

2.1.2.2. LIFG patients. Patients with LIFG damage were not significantly faster to respond to target words in the consistent condition in the subject-wise analysis [$M_{con}=755$, SD=118, $M_{inc}=765$, SD=113; $F_1(1,18)=0.82$, $p=0.38$]; and the effect of condition was marginal in the item-wise analysis [$F_2(1,95)=3.17$, $p=0.08$]. Similarly, the consistency-by-ambiguity-type interaction did not reach significance in the subject-wise analysis [$F_1(1,18)=0.71$, $p<0.41$], and was marginal in the item-wise analysis [$F_2(1,95)=3.07$, $p=0.08$]. There was, however, a robust effect of ambiguity type [$M_{hom}=738$, SD=108, $M_{poly}=782$, SD=119; $F_1(1,18)=14.50$, $p=0.001$; $F_2(1,95)=9.36$, $p<0.01$]. For homonyms, the effect of consistency did not reach significance across participants, but was significant across items [$M_{con}=728$, SD=115, $M_{inc}=748$, SD=109; $t_1(6)=1.12$, $p=0.30$; $t_2(47)=2.63$, $p<0.05$]. There was no effect of consistency for polysemous words [$M_{con}=781$, SD=125, $M_{inc}=782$, SD=123; $t_1(6)=0.05$, $p=0.97$; $t_2(47)=0.02$ $p=0.99$]. The interaction of meaning frequency and consistency did not reach significance for either homonymous or polysemous words.

2.1.2.3. Comparison of patient groups. When reaction time data from both groups of patients were analyzed together, there was a main effect of consistency [$F_1(1,38)=5.55$, $p<0.05$;

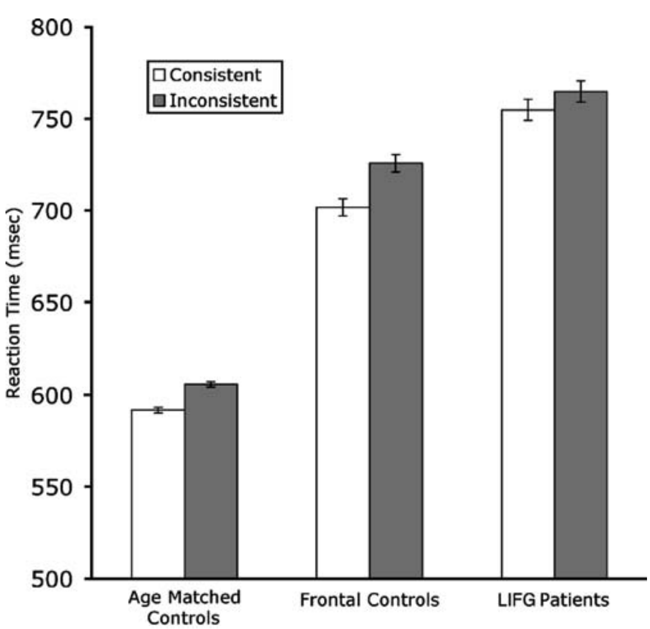

Fig. 1 – The effect of consistence on the performance of control subjects and LIFG patients (Experiment 1). Average reaction times (in milliseconds) in the consistent, and inconsistent conditions. Standard error bars represent the standard errors of the difference between the consistent and inconsistent conditions.

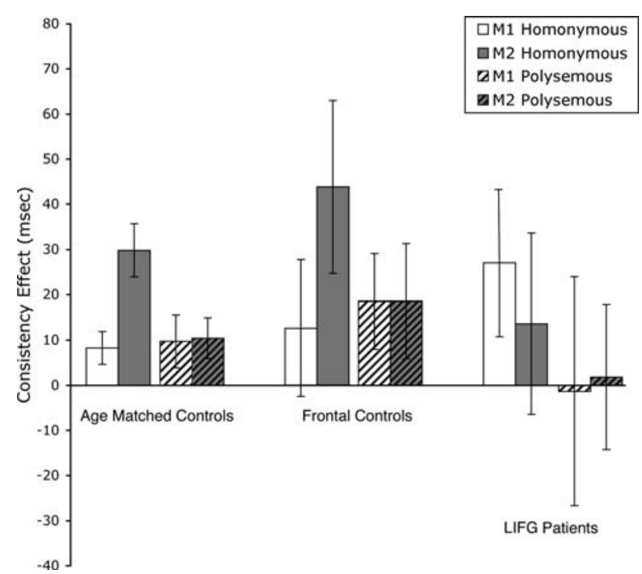

Fig. 2 – The interaction of meaning frequency and consistency (Experiment 1). The difference score between the inconsistent and consistent conditions is plotted on the *Y* axis in milliseconds. Each bar represents the consistency effect for homonymous and polysemous words as a function of meaning frequency. Error bars represented the standard error of the consistency effect.

$F_2(1,289) = 9.09$, $p < 0.01$] and ambiguity type [$F_1(1,38) = 24.97$, $p < 0.0001$; $F_2(1,289) = 7.26$, $p < 0.01$], but no consistency-by-ambiguity-type interaction [$F_1 < 1$; $F_2(1,289) = 2.80$, $p < 0.10$]. Non-LIFG patients were slightly faster to make lexical decisions than were LIFG patients. The effect of group was significant in the item-wise analyses [$F_2(1,289) = 67.40$, $p < 0.0001$], but did not reach significance in the subject-wise analyses ($F < 1$). The group-by-consistency interaction did not reach significance [$F_1(1,38) = 0.87$, $p = 0.36$; $F_2(1,95) = 0.55$, $p = 0.46$].

2.2. Discussion Experiment 1

2.2.1. Healthy control performance

Previous research has shown that healthy participants select the context-appropriate meaning of homonyms in a lexical decision task (Schvaneveldt and Meyer, 1976). Experiment 1 replicated this finding and extended it to polysemous words: healthy participants were faster to respond to target words in the consistent than inconsistent condition for both homonyms and polysemous words. However, the effect of consistency was greater for homonymous than for polysemous triplets, as evidenced by the marginal ambiguity-type-by-consistency interaction. These findings are consistent with the hypothesis that meaning selection is an integral part of word comprehension for both homonyms and polysemous words.

Notably, meaning selection was not necessary for the current task, since a lexical decision could be made without selecting a contextually appropriate meaning. We therefore conclude that meaning selection is engaged both for homonyms and polysemous words when context is available. In the triplet lexical decision task, participants select the context-appropriate meaning of the ambiguous word based on the preceding prime-word, even though selection is not required for a response to be made. The smaller size of the consistency effect for polysemous words is consistent with the idea that the senses of polysemous words are represented as partially overlapping, distributed patterns of connectivity. Consequently, in the inconsistent condition, only the non-overlapping part of the inconsistent meaning is context incongruous. The suppression of the incongruous part of the sense is perhaps partially counteracted by the priming of the congruous aspect of the sense.

While we found that meaning selection occurs for homonymous and polysemous words, the effect of meaning frequency on selection differs across these ambiguity types. Our finding of a meaning frequency by consistency interaction is in agreement with other research suggesting that the less frequent meaning of an ambiguous word is rapidly disregarded when the preceding context primes the more frequent meaning; in contrast, the more frequent meaning exerts an effect on processing even when context primes the less frequent meaning (Duffy et al., 1988). We hypothesize that the subordinate, but not the dominant meaning can be completely suppressed. In the inconsistent condition, the incomplete suppression of M1 results in partial priming of the M1 target, while the complete suppression of M2 results in no priming of the M2 target.

The effect of meaning frequency on meaning selection in homonyms but not polysemous words is an intriguing finding that is consistent with recent ERP results (Klepousniotou et al., 2006). In that study, the N400 component was sensitive to homonym meaning frequency, but not polysemous sense frequency during a word-pair priming paradigm. The current behavioral data, coupled with these ERP results, suggest that sense frequency does not affect lexical representations in the same way as meaning frequency. Together these results illustrate that homonym and polysemous word comprehension is similar in that context is used to modulate the semantic representation that is retrieved from memory. However, important distinctions exist between the way in which meanings and senses are represented in the neural networks that support language.

In combination with previous research, the data of Experiment 1 suggest that there are important similarities and differences between the processing of homonymous and polysemous words. Homonymous and polysemous words are similar in that for both word types meaning retrieval is affected by context. However, there are important representational differences between homonymous and polysemous words (for example, with respect to meaning frequency). As a result of the representational differences between homonymous and polysemous words, these word types behave differently in some tasks (Beretta et al., 2005; Frazier and Rayner, 1990; Klepousniotou, 2002). For example, when participants perform a lexical decision task in the absence of a biasing context, the meanings of homonyms compete, while the senses of polysemous words do not (Rodd et al., 2002). Additionally, whether polysemous senses compete for selection may depend on the type of polysemous words being studied. Sense competition may occur for polysemous words with saliently distinct, and frequent senses, but not for polysemous words with one primary senses and closely related infrequent senses (Traxler et al., 2002; Williams, 1992). In light of these data, we would argue that the goal of future research should not be to determine whether homonymous and polysemous words are generally similar or generally different, but rather to identify the specific similarities and differences between the comprehension of homonymous and polysemous words and to understand the processing differences that may exist between different types of polysemy.

2.2.2. Patient performance

Patients with frontal lobe damage that did not include the LIFG showed normal context-dependent meaning selection, as evidenced by a main effect of consistency. Non-LIFG patients were faster to respond in the consistent than inconsistent condition. Patients with lesions to the LIFG did not show a significant effect of consistency (although there was an effect of consistency in the item-wise analysis for homonyms). An additional indication that meaning selection is impaired in LIFG patients comes from the meaning frequency analysis. While no meaning frequency effects reached significance in the patient data, the pattern of results for non-LIFG patients was qualitatively similar to that of the aged matched controls. In contrast, the performance of LIFG patients was not similarly affected by meaning frequency (see Fig. 2).

These data indicate that damage to the LIFG reduces, but does not abolish context effects. Specifically, we hypothesize that LIFG damage interferes with the ability to suppress context-inappropriate meanings, but does not interfere with semantic priming more generally. According to this hypothesis, LIFG damage does not reduce priming in the consistent condition, but rather enhances priming in the inconsistent condition. Such a deficit would result in a smaller, non-significant, difference between the consistent and inconsistent conditions in the LIFG patient group.

In Experiment 2, we directly tested this interpretation. To accomplish this, we included a neutral condition and tested one of the patients who had damage to the entire LIFG region. In the neutral condition, two unrelated words preceded the target word. One might question whether this condition can be considered truly "neutral". However, for the present purposes, it is sufficient to assume that, on average, one meaning is not favored over another in the neutral condition. Including a neutral condition permitted us to examine whether damage to the LIFG reduces semantic priming in general, or impairs the selection of the context appropriate meaning and suppression of the inappropriate meaning in particular. That is, was the lack of a consistency affect in Experiment 1 because *both* conditions produced priming or because *neither* did? In Experiment 2, we used a different set of stimuli to test patient 363 repeatedly in a within subject design. We compare the performance of patient 363 to a set of 24 young control subjects, and one older control.

3. Experiment 2

3.1. Results

3.1.1. Young controls

A three-way ANOVA revealed a main effect of consistency [$F_1(3,152)=5.76$, $p<0.001$; $F_2(3,119)=7.80$, $p<0.0001$], no effect of ambiguity type [$F_1(3,152)=2.08$, $p=0.15$; $F_2(3,119)=0.3$, $p>0.5$], and no ambiguity by consistency interaction [$F_1(9,152)=1.79$, $p=0.15$; $F_2(3,119)=1.69$, $p>0.15$] (see Fig. 3). Because the different versions of Experiment 2 contained different target words that were not matched on frequency or length within each version, we modeled the effect of version in analyzing the young control subject data. There was indeed a significant main effect of version [$F_1(3,152)=6.11$, $p<0.001$; $F_2(3,119)=5.87$, $p<0.0001$], and a version by consistency interaction [$F_1(9,152)=7.40$, $p<0.0001$; $F_2(3,120)=1.73$ $p<0.1$]. A post hoc Tukey HSD test revealed a significant difference between the inconsistent and consistent conditions ($M_{con}=551$, SD=56, $M_{inc}=590$, SD=104; $p=0.05$). The consistent condition was also significantly faster than the neutral condition ($M_{neutral}=585$, SD=71; $p=0.05$), but the inconsistent condition was not different from the neutral condition. The effect of condition was independently significant for polysemous words [$M_{con}=541$, SD=51, $M_{inc}=599$, SD=117, $M_{neutral}=577$, SD=65, $M_{control\text{-}t}=553$, SD=58; $F_1(3,60)=4.96$, $p<0.01$; $F_2(3,54)=8.89$, $p<0.0001$], but not for homonyms [$M_{con}=561$, SD=50, $M_{inc}=581$, SD=69, $M_{neutral}=592$, SD=65, $M_{control\text{-}t}=579$, SD=80; $F_1(3,60)=2.16$, $p=0.1$; $F_2(3,54)=1.88$, $p=0.14$].

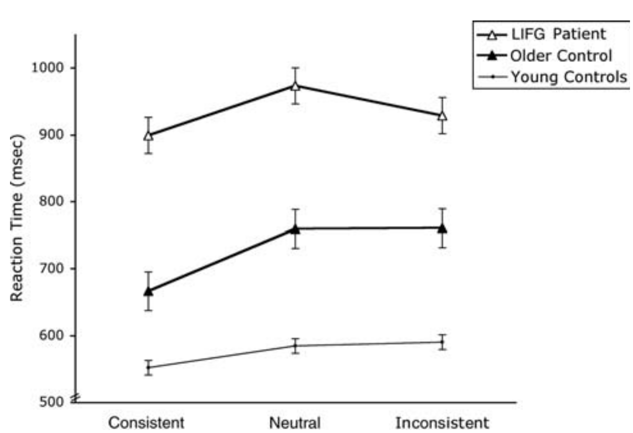

Fig. 3 – The effect of consistency on reaction time (Experiment 2). Reaction time as a function of condition for the young control group, the older participant, and the LIFG patient. Error bars represent the standard error of the consistency effect across participants for the young control group, and across items for the older control, and LIFG patient.

Participants responded correctly on 97.3% (SD=4.8) of the critical trials, and 94.4% (SD=5.0) of the filler trials. There was a main effect of ambiguity type on accuracy. Participants made significantly more errors on homonymous than on polysemous trials [$M_{hom}=95.6\%$, SD=5.6%, $M_{poly}=99.0\%$, SD=2.0%; $F_1(1,149)=12.63$, $p<0.001$; $F_2(1,119)=3.38$, $p=0.07$]. There was a marginal effect of consistency on accuracy [$F_1(3,149)=2.20$, $p=0.09$; $F_2(3,119)=1.37$, $p=0.25$], and a marginal ambiguity type by consistency interaction [$F_1(3,149)=2.32$, $p=0.08$; $F_2(3,119)=1.81$, $p=0.15$]. The accuracy data for homonyms exactly mirrored the reaction time data. That is, participants made more errors in those conditions where their responses were slower, though the effect of condition did not reach significance. For polysemous words, no accuracy effects were significant; however, participants made slightly (1.5%) fewer errors in the inconsistent than consistent condition. There was no effect of version, or a version by consistency interaction ($F<2$). In a post hoc Tukey HSD test, the differences between the consistent, inconsistent, neutral, and control terminal conditions were not significant. Accuracy data as a function of condition are summarized in Table 3.

3.1.2. Older control participant, and patient 363

The older participant responded accurately on 96.1% of the critical trials, and 82.2% of the filler trials. Patient 363 responded accurately on 96.9% of the critical trials, and 83.9% of the filler trials. Based on the results from the young control participants, we performed planned comparisons to test for a difference between the inconsistent and consistent conditions, and between the neutral and consistent conditions. Because there was no ambiguity-by-condition interaction, and to gain power in a within-subject analysis, we collapsed across homonymous and polysemous trials. The older participant was significantly faster to respond in the consistent condition ($M=667$ ms, SD=109) than in the inconsistent [$M=760$ ms, SD=183; $t(28)=2.53$, $p<0.05$], and neutral conditions [$M=758$ ms, SD=168; $t(28)=2.40$, $p<0.05$]. The inconsistent condition was not different from the neutral condition. There was no effect of condition on accuracy [$\chi^2(2, N=29)=1.18$, $p=0.56$].

Table 3 – Target response accuracy as a function of condition (Experiment 2)

	Critical trials				Fillers
	Consistent	Neutral	Inconsistent	Control terminal	
Patient 363	97.4	97.7	95.5	–	83.9
Older control	93.3	97.4	97.5	–	82.2
Young controls	98.8(3.2)	95.4(7.3)	97.3(4.5)	97.5(4.4)	94.1(5.0)

Mean percent correct for target words as a function of condition, for each group. Standard deviations of the mean appear in parenthesis.

Patient 363 was significantly faster to make a lexical decision in the consistent condition ($M=899$ ms, $SD=140$) than in the neutral condition [$M=973$ ms, $SD=161$; $t(30)=2.10$, $p=0.045$]. Replicating the results of Experiment 1, the inconsistent ($M=928$ ms, $SD=149$), and consistent conditions were not significantly different from each other [$t(30)=0.83$, $p=0.41$] (see Fig. 3). There were no effects of condition in the accuracy analyses [$\chi^2(2, N=31)<1$].

3.2. *Discussion Experiment 2*

The results of Experiment 2 replicated the findings of Experiment 1, both in the performance of the healthy control subjects, and the performance of patients with LIFG damage. As in Experiment 1, healthy participants were able to select the context-appropriate meaning of ambiguous words. However, in contrast to Experiment 1, the effect of consistency was more robust for polysemous words than for homonyms, though the interaction of consistency and ambiguity type was not significant. This result suggests that the magnitude of the consistency effect may depend on factors other than ambiguity type. One explanation for why the polysemous consistency effect was more robust in Experiment 2 may have to do with differences in how the stimuli in the two experiments were constructed: the stimuli in Experiment 2 were taken from prior studies (Klein and Murphy, 2001; Schvaneveldt and Meyer, 1976). In contrast, we designed the stimuli in Experiment 1 in such a way as to minimize the difference between the relationship of the first and third words in the triplets of the inconsistent and consistent conditions. It is possible that this procedure was more difficult for polysemous words than for homonyms, and consequently less effective prime words were chosen for polysemous triplets in Experiment 1. Taken together, the results of Experiments 1 and 2 support the conclusion that context-dependent meaning selection is a part of polysemous and homonymous word comprehension. However, further study is needed to determine whether the effect of consistency is reduced for polysemous words relative to homonyms.

The results of Experiment 2 replicate and extend the results of Experiment 1. As in Experiment 1, LIFG damage reduced the difference between the consistent and inconsistent conditions. Furthermore, the results of Experiment 2 refute the hypothesis that LIFG damage abolishes semantic priming in the triplet lexical decision paradigm. Like control subjects, patient 363 (who suffered a lesion affecting the entire left inferior frontal gyrus) showed a significant priming effect in the consistent relative to neutral condition. In contrast, he showed no difference between the inconsistent and consistent conditions. As hypothesized, the average reaction time in the inconsistent condition fell between that of the consistent and neutral conditions. This result suggests that LIFG damage reduces the ability to suppress the context inappropriate meaning, but does not interfere with the ability to prime the context appropriate meaning. This result suggests that LIFG damage interferes with resolving competition in particular, and not semantic priming more generally. However, it is worth noting that the conclusions drawn from Experiment 2 are limited by the fact that only one patient was tested. We chose to study the behavior of this patient in Experiment 2 based on the independent criterion of having a lesion encompassing the entire LIFG. In future research, it will be important to replicate this finding with a sample of LIFG patients, thus extending it to the population of patients with damage to the LIFG.

4. General discussion

Healthy participants and patients with frontal lobe damage that did not impinge on the LIFG were faster to make a lexical decision to target words in the consistent than inconsistent condition. This finding illustrates that participants select the context-appropriate meaning of homonymous and polysemous words. A number of previous studies have demonstrated context-dependent meaning selection for homonyms in the lexical decision task (e.g., Schvaneveldt and Meyer, 1976). The current study replicates this effect in homonyms and extends it to polysemous words. Several recent studies have reported consistency effects for polysemous words in a sensicality judgment task (Klein and Murphy, 2001, 2002). Our similar pattern of results in a lexical decision task suggests that sense selection is not peculiar to the sensicality judgment paradigm. Furthermore, while the lexical decision task is known to engage processes that are not thought of as bottom up aspects of lexical access (Neely et al., 1989), a lexical decision response does not require one meaning or sense to be selected. The present results therefore suggest that sense selection occurs even when selection does not seem to be necessary for the task. Combining the current results with previous work, we hypothesize that sense and meaning selection is an integral part of word comprehension.

These findings are consistent with a recent connectionist models of sense and meaning representations proposed by Rodd et al. (2004). The model's performance successfully simulated lexical decision times to polysemous, and homonymous words in single word presentation paradigm (Rodd et al., 2002, 2004). Based on this model, the authors predicted that polysemous words and homonyms would behave similarly with respect to contextual disambiguation. The present findings are consistent with this prediction.

While the current findings suggest that both homonyms and polysemous words are disambiguated based on context, an interesting difference between these word types emerged from the meaning frequency analysis. While meaning frequency interacted with consistency for homonyms, this was not so for polysemous words. This finding is in line with recent evidence that the N400 component is sensitive to meaning frequency but not sense frequency (Klepousniotou et al., 2006). An interesting avenue for future research is to examine how connectionist models of lexical-semantics can account for the presence of meaning but not sense frequency effects.

In contrast to patients with non-LIFG frontal lobe damage, patients with LIFG damage were not significantly faster to make lexical decisions to target words in the consistent than inconsistent condition. However, while the main effect of consistency was not significant in the overall analyses, there was an item-wise effect of consistency for homonyms in Experiment 1. This result suggested that context effects are reduced, but not eliminated by LIFG damage. Experiment 2 replicated the findings of Experiment 1, and illustrated that LIFG damage does not abolish semantic priming in the triplet lexical decision paradigm. A patient with a lesion encompassing the entire LIFG region was not faster in the consistent than inconsistent condition, but showed priming in the consistent relative to neutral condition. This result illustrates that damage to the LIFG does not abolish contextual priming, but rather increases priming in the inconsistent condition. These findings suggest that damage to the LIFG interferes with the ability to suppress the context inappropriate meaning of ambiguous words, but does not affect the activation of context appropriate information.

The present findings are consistent with the hypothesis that contextual meaning selection is delayed in Broca's aphasics during sentence comprehension (Swaab et al., 1998). When presented with sentences where the final word is a homonym, ERP data from Broca's aphasics show evidence of meaning selection by 1250 ms, but not 100 ms, after the sentence-final ambiguous word. In contrast to the paradigm used by Swaab et al. (1998), the triplet lexical decision task has a deadline, namely the lexical decision response. If the participant makes a response before the selection processes is completed, the reaction time does not reflect the effects of meaning selection. The hypothesis that LIFG patients are slower to select the context appropriate meaning predicts that they are more likely to show a consistency effect on trials where the response to the target word was delayed (perhaps by the length and frequency of the target word itself).

To examine whether there was evidence of delayed selection in the current task, we calculated a correlation coefficient for each participant (across items) between the mean reaction time for each target in the consistent and inconsistent conditions, and the consistency effect for that target word. We then evaluated whether the correlation coefficients, across participants, were reliably different from zero using the Signed-Rank test (the correlation coefficients were not normally distributed for one of the groups). For the LIFG patients, there was a marginally reliable, positive correlation between mean reaction time and the magnitude of the consistency effect [$M=0.07$, $SD=0.09$; $t(6)=11$, $p=0.08$]. This effect did not approach significance in the Frontal Controls [$M=0.13$, $SD=0.22$; $t(6)=7$, $p=0.30$] or Age Matched Controls [$M=0.02$, $SD=0.20$; $t(19)=16$, $p=0.56$]. These data are consistent with the idea that LIFG damage delays contextual selection. However, in the triplet lexical decision task, this delay results in a failure to show a consistency effect in the target word reaction times. In this respect, the triplet lexical decision task is similar to online comprehension. Selection must be completed in time for further processing to receive proper input. If selection does not occur in time, then integration and structure building processes may receive inappropriate input and will not reflect the results of the meaning selection process.

While the present data are consistent with the hypothesis that Broca's aphasics are impaired at integrating words into the sentence context, we suggest two caveats to this hypothesis. First, only Broca's aphasics with LIFG damage (but not the 15% of Broca's aphasics who do not have LIFG damage; Dronkers, 2000), will show impaired meaning selection. Conversely, patients with LIFG damage will show a diminished ability to select context appropriate semantic representations even if they do not have Broca's aphasia. This is an important point since only 50–60% of patients with damage to the LIFG have Broca's aphasia; and LIFG damage alone does not produce the syndrome (Dronkers, 2000).

Secondly, we contend that our data indicate a role for the LIFG in the resolution of competition between lexical-semantic representations, and not in other aspects of integrating meanings into the sentence context. Yet, elsewhere, it has been hypothesized that the LIFG mediates aspects of syntactic processing such as trace deletion (Grodzinsky, 2000). One could argue that the integration of word meanings into the sentence context breaks down because some aspect of an appropriate representation of the sentence cannot be computed. However, in the triplet lexical decision task processing the prime words does not require the formation of a complex sentence representation. Therefore, it is unlikely that a deficit in meaning selection results from the inability to parse the context. The hypothesis that LIFG damage reduces the ability to resolve competition more generally further suggests that the role of the LIFG in word comprehension need not be restricted to the integration stage. Rather, the LIFG may be involved in multiple stages of word comprehension, including lexical access, to the extent that competition arises during these stages. For example, recent evidence for the idea that the LIFG is involved in accessing phonological representations comes from fMRI research showing that the LIFG is involved in resolving competition between phonetic categories when participants are presented with ambiguous synthetic speech stimuli (varying in voice onset time) (Blumstein et al., 2005). Together with prior research, the present data support the hypothesis that the LIFG resolves competition at multiple levels of linguistic representations, and the involvement of the LIFG is not restricted to the integration stage of word recognition.

The present characterization of the LIFG may help reconcile the seemingly disparate functional neuroimaging and neuropsychological literatures on the role of the LIFG in word comprehension: increased LIFG activity is a typical finding in neuroimaging studies using a variety of word comprehension tasks. However, patients with damage to the LIFG do not display profound comprehension deficits. We hypothesize that the LIFG acts as a top down biasing mechanism that facilitates resolving competition between semantic representations. While the LIFG is engaged when competition arises among lexical semantic representations, its contribution may not always be necessary for word comprehension. We hypothesize that relative to other language tasks, word comprehension is less reliant on a top down biasing mechanism. Unlike in production tasks, context directly biases the competition between lexical semantic representations. Consequently, selection can occur even without LIFG contribution, albeit more slowly and less completely. Thus, rapid suppression of irrelevant information may not always be required for relatively good comprehension.

What, then, is the most parsimonious description of the role of the LIFG in word comprehension? We would argue that the answer to this question depends not only on evidence from word comprehension studies, but also on what we know about the function of LIFG more generally. The LIFG is involved in various levels of linguistic processing, including production, comprehension, word, and sentence level processing (e.g., Novick et al., 2005; Thompson-Schill et al., 1998). In a variety of paradigms, this region has been shown to respond to competition between semantic representations (Thompson-Schill et al., 2005). Based on the current study, taken together with prior literature, we conclude that the LIFG facilitates suppressing irrelevant lexical-semantic representations during word comprehension. An important goal of future research is to delineate the precise mechanism of the LIFG's influence on meaning selection.

5. Experimental procedures

5.1. *Experiment 1*

5.1.1. *Participants*

Fourteen patients with frontal lobe damage, and 20 age- and education-matched controls (eight females) participated in the study (Table 4). Participants gave informed consent in accordance with University of Pennsylvania Institutional Review Board, and were paid for their participation. All participants were native English speakers with no history of substance abuse or psychiatric disorders. None of the control participants had a history of neurological disease. Elderly control subjects were recruited from the Philadelphia area.

Table 4 – Participant information

	Demographics				Lesion			Inc–Con (ms)
Patient	Age	Education	Gender	Hand	Onset	Description	Etiology	
LIFG patients								
2	59	19	M	R	2001	LIFG	Stroke	−22
363	68	16	M	R	2001	LIFG	Stroke	21
3	63	11	M	R	1993	LIFG	Stroke	−6
412	42	13	F	R	2001	LIFG, LPu, LIns	Stroke	17
1	37	12	F	R	2002	LIFG	Stroke	14
517	57	12	F	R	2000	LIFG, LMFG	Stroke	34
567	40	18	F	R	1995	LIFG	Stroke	9
Mean (SD)	52.3(12.4)	14.4(3.2)						
Range	37–68	11–19						
Frontal controls								
215	55	14	M	R	1999	LMFG	Stroke	14
568	60	16	F	R	2002	LSTG, LMFG, LpreCG, LpostCG	Stroke	59
529	61	12	F	R	2002	OrbG, CingG	Stroke	4
384	66	12	M	R	2001	LSFG	Stroke	26
541	41	20	M	Ambi.	2003	LSFG	Glioma	5
440	55	16	F	R	1999	RSFG	Glioma	12
481	64	12	F	R	2002	RIFG	Stroke	48
Mean	57.4(8.3)	14.6(3.0)						
Range	55–66	12–20						
Age matched controls		($n=20$)						
Mean (SD)	59.1(16.5)	14.9(3.0)						
Range	29–83	12–21						

Demographic, lesion, and reaction time information for patients and age-matched controls in Experiment 1. The Inc–Con column refers to the reaction time difference between the inconsistent, and consistent conditions for Experiment 1. The anatomical abbreviations are as follows: LIFG and RIFG (left and right inferior frontal gyri respectively), LMFG (left middle frontal gyrus), LSFG and RSFG (left and right superior frontal gyru, respectively), LpreCG (left precentral gyrus), LpostCG (left postcentral gyrus), OrbG (orbitofrontal gyrus), CingG (cingulate gyrus).

Frontal lobe patients were selected based on the presence of a lesion to the frontal cortex as indicated by magnetic resonance imaging (MRI) or computerized tomography (CT) scans. Participants were divided into two groups based on whether the lesion did or did not impinge on the posterior aspect of the left inferior frontal gyrus (Brodmann area 44/45). For five of the seven patients in the LIFG group, the lesion was confined to the LIFG. For one of the patients in the LIFG group, the lesion extended to the left putamen, and insula. For another patient, the lesion extended into the left middle frontal gyrus (LMFG). Of the patients in the non-LIFG group: two patients suffered damage to the superior frontal gyrus (LSTG); one patient suffered from a lesion to the LSTG, LMFG, left precentral gyrus, and left postcentral gyrus; one patient suffered from a lesion to the LMFG, one patient suffered from a lesion to the orbital gyrus, and anterior aspect of the cingulate cortex; one patient had a lesion to the RSFG; and one patient had a lesion in the RIFG (Table 4). The Western Aphasia Battery (WAB) scores were available for five out of the seven LIFG patients. WAB scores ranged from 79.6 to 94 (M=87.2, SD=5.2).

5.1.2. *Materials*

One hundred critical word triplets were included in the experiment.[2] The first two words in each triplet served as the primes, and the last word as the target. For half of the critical triplets the middle word was polysemous, and for the other half the middle word was a homonym (Durkin and Manning, 1989; Gawlick-Grendell and Woltz, 1994; Twilley et al., 1994). Each homonym and polysemous word had two triplets associated with it: one in the consistent condition and a second in the inconsistent condition. The two triplets shared the middle and target words, and differed only in the first prime word (the target words in the two conditions were thus perfectly matched on frequency and length). In the consistent condition, the first and third words were related to the same meaning of the ambiguous word (e.g., "BACK, PACK, BAG"), while in the inconsistent condition, the first and third words were related to different meanings of the ambiguous word (e.g., "WOLF, PACK, BAG") (see Appendix A, Supplementary data).

For half of the triplets, the target word was related to the less frequent meaning of the ambiguous word (e.g., inconsistent PENCIL, PEN, CAGE; consistent FENCE, PEN, CAGE), for the other half the target word was related to the most frequent meaning of the ambiguous word (e.g., inconsistent ASPRIN, TABLET, STONE; consistent COMMANDMENTS, TABLET, STONE) (Durkin and Manning, 1989; Gawlick-Grendell and Woltz, 1994; Twilley et al., 1994). The current stimulus set contained biased and balanced ambiguous words. The average difference between the meaning frequency of the most frequent, and the second most frequent meanings was .38 (SD=0.23) for homonyms, and 0.40 (SD=0.27) for polysemous words. Homonyms and polysemous words did not differ on this measure of meaning bias [$t_2(95)=-0.27$, $p=0.78$].

[2] Two target words were subsequently dropped from analyses. In one case, one of the prime words was misspelled; in the second case, the target word was different in the inconsistent and consistent conditions.

In addition to the critical trials, there were 383 filler triplets that contained a pronounceable non-word in the first, second, and/or third positions. Some of the filler words repeated so that participants could not distinguish them from the critical trials on the bases of repetition. There were a total of 901 words and 723 non-words in the experiment.

5.1.3. *Procedure*

Each trial began with the words "START" (in pink) in the middle of the screen. Participants had to press a pink key with the middle finger of their left hand to start the trial. When the participant initiated a trial, three crosses appeared on the screen, for 1000 ms. One by one, a written string replaced each cross. Participants were instructed to respond to every string, by pressing a key marked "Yes" if the string was a real word, or the key marked "No," if the string was a nonsense word. Words and non-words were presented in white on a black screen. Because some patients experienced weakness in their right arm, all the participants were instructed to respond with their left hand, using their index finger to indicate a Yes response, their ring finger to indicate a No response, and their middle finger to start the trial. Participants were instructed to respond as quickly and accurately as possible. Each word remained on the screen for 1250 ms for control subjects, and for 1500 ms for all but three patients. One of the frontal control patients, and two of the LIFG patients required the maximum time to respond to be raised to 1750 ms in order to perform above 75% correct on the practice stimuli. Triplets were separated by a 1500 ms inter-trial interval.

Before the experimental trials, participants were given 20 practice trials with feedback. At the end of a triplet, the word "Incorrect" appeared on the screen if the participant responded incorrectly or took longer than the maximum allotted time on any one of the three words. If a participant made errors on more than a quarter of the trials, they repeated the practice session until their performance improved (a maximum of three times). Feedback was not given during the experimental trials.

The experimental trials were broken up into 10 blocks and participants were encouraged to take breaks between blocks. Each participant saw each target word twice throughout the course of the experiment: once in the consistent condition, and once in the inconsistent condition. Trials were presented in pseudorandom order, with the stipulation that a target word never appeared twice in a single block. Each participant saw 25 inconsistent, and 25 consistent trials in the first half of the experiment, and 25 inconsistent and 25 consistent trials in the second half of the experiment.

Stimuli were separated into two versions of presentation that counterbalanced the order of the different triplets. Triplets that were presented first in the inconsistent condition in version one were presented first in the consistent condition in version two. Half of the control participants saw version one, while the other half saw version two. We were unable to perfectly counterbalance versions one and two across patients (more patients saw version one than version two). However, when we included version as a predictor variable in an ANOVA looking at patient reaction times, and accuracy, there was neither a main effect of version nor a version by condition interaction.

Age-matched controls and patients were tested using E-prime software (http://www.pstnet.com/products/e-prime/) on a PC laptop. Control subjects were all tested in a dimly lit, quiet testing room. Some patients were tested in the same testing room as controls; however, most of the patients were tested in their homes. The experiment took a total of one and a half hours. Eleven of the patients were tested in one session. Three of the LIFG patients were tested in two sessions no more than two weeks apart.

5.1.4. *Analysis*

For the reaction time analyses, we excluded trials where subjects made an error in responding to either of the prime words or the target word. This procedure eliminated 5% of the critical trials for age-matched control subjects, 15% for the patient controls, and 13% for the LIFG patients. Prior to analyses, reaction times that were three or more standard deviations away from the mean for that subject, in that condition, were replaced with the mean for that subject in that condition. On average, this procedure replaced one critical trial per condition for each age-matched control, patient control, and LIFG patient. In analyzing accuracy to target words as a function of condition, we excluded trials in which subjects made an error in responding to one of the prime words. This procedure excluded 4% of the critical trials for elderly controls, 10% for LPFC controls, and 10% for LIFG patients.

5.2. *Experiment 2*

5.2.1. *Participants*

Twenty-four young controls (15 females; the mean was age was 25.5), one elderly control, and one patient with left inferior frontal gyrus damage participated in Experiment 2. Participants gave informed consent in accordance with University of Pennsylvania IRB, and were paid for their participation. All participants were right-handed, native English speakers who had no history of substance abuse, or psychiatric disorders. None of the control participants had a history of neurological disease. The elderly control subject was a 56-year-old female, recruited from the Philadelphia area. Patient 363 from Experiment 1 participated in Experiment 2. Demographic information about this patient is provided in Table 1. This patient suffered from a stroke affecting the entire left inferior frontal gyrus, including all of Brodmann areas 44/45.

5.2.2. *Materials*

Forty-eight critical word triplets were included in the experiment.[3] For half of the critical triplets, the middle word was polysemous, and for the other half, the middle word was a homonym. The homonymous triplets for this experiment were taken from the Schvaneveldt and Meyer, 1976 study: the polysemous triplets were taken from Klein and Murphy (2001). Each homonym and polysemous word had three triplets associated with it, one in the consistent condition, one in the inconsistent condition, one in the neutral condition. The consistent and inconsistent conditions were constructed in the same manner as those of Experiment 1. In the neutral condition, the target word was preceded by two unrelated words (e.g., consistent "COW, CALF, BABY"; inconsistent "KNEE, CALF, BABY"; neutral "MIRROR, CUT, BABY").

Additionally, there was a control terminal condition in the version of the experiment administered to the young control subjects. In this condition the first word was unrelated to the ambiguous word, but the ambiguous word was the same as in the consistent and inconsistent conditions. Because there were no predictions about the patient's performance on the control terminal condition, it was dropped from the version of the experiment administered to the elderly control and patient. There were 112 filler triplets, which contained a pronounceable non-word in the first, second and/or third positions.

5.2.3. *Procedure*

The procedure of Experiment 2 was similar to that of Experiment 1; below we highlight differences between the two. Before the experimental trials, subjects were given 15 practice trials. Feedback was given on practice and experimental trials. We created four versions of the experiment, such that every young control subject saw each target word once during the testing session. Each version of the experiment was divided into four blocks that contained forty trials each. The elderly control and patient were tested on all four versions of the experiment, after the control terminal condition was removed. Patient 363 was tested in four sessions, which were separated by at least three weeks. The elderly control subject was tested on all four versions of the experiment in one 150 minute session. The patient and elderly control subjects saw the four versions of the experiment in the same order.

5.2.4. *Analysis*

The analysis procedure was identical to that of Experiment 1. However, no outliers were removed because there were no reaction times three standard deviations away from the mean in the young, elderly control, or patient data. Trials on which participants made an error to any one of the prime words, or the target word were excluded from reaction time analyses. This procedure excluded 6% of the critical trials for young control subjects. To conduct a within-item comparison and control for frequency and length effects (in the analyses of the older control and patients' data), we excluded all trials associated with a given target word, if the subject made an error to that target word, in any one of the conditions. This eliminated 40% of the trials from each condition for the older subject and 35% for patient 363. While this procedure eliminated a large number of trials, it was necessary to control for frequency, and length effect in a within-item comparison. The procedure excluded a comparable number of trials for the older subject and the patient. We also inspected the reaction time data excluding only the error trials themselves. The relationship between the consistent, inconsistent, and neutral conditions was the same as in the above analyses for both the patient and the older participant.

[3] One polysemous triplet was subsequently dropped from analyses because the target word was different in the inconsistent and consistent conditions.

In analyzing accuracy to target words as a function of consistency, we excluded trials where subjects made an error in responding to one of the prime words. This procedure excluded 3% of the critical trials for the young control subjects, 13% for the elderly control subject, and 11% for the LIFG patient.

Acknowledgments

This research was supported by National Institutes of Health (R01-MH67008, R01-MH60414, NS045839) and the Searle Scholars Program. We would like to thank Marianna Stark for her invaluable help with patient testing, as well as Anjan Chatterjee and Branch Coslett for their help in identifying the neuroanatomical loci of the lesions discussed in this study. We would also like to thank Laura Barde for her help in recruiting several of the patients, the Moss Rehabilitation Center for the use of their facilities, and Jeffrey Ellenbogen for his comments on an earlier draft of this paper. Finally, we would also like to thank all of the participants and, in particular, the patients and their families who made this research possible.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.brainres.2006.10.012.

REFERENCES

Barde, L.H., Thompson-Schill, S.L., 2002. Models of functional organization of the lateral prefrontal cortex in verbal working memory: evidence in favor of the process model. J. Cogn. Neurosci. 14 (7), 1054–1063.
Beretta, A., Fiorentino, R., Poeppel, D., 2005. The effects of homonymy and polysemy on lexical access: an MEG study. Brain Res.: Cogn. Brain Res. 24 (1), 57–65.
Blumstein, S.E., Myers, E.B., Rissman, J., 2005. The perception of voice onset time: an fMRI investigation of phonetic category structure. J. Cogn. Neurosci. 17 (9), 1353–1366.
Broca, P., 1861. Remarques sur le siege de la faculte du langage articule; suives d'une observation D-aphemie. Bull. Soc. Anat. Paris 6, 330–357.
Burgess, C., Simpson, G.B., 1988. Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. Brain Lang. 33 (1), 86–103.
Caramazza, A., Zurif, E.B., 1976. Dissociation of algorithmic and heuristic processes in language comprehension: evidence from aphasia. Brain Lang. 3 (4), 572–582.
Cruse, D.A., 1986. Lexical Semantics. Cambridge Univ. Press, Cambridge, MA.
Dronkers, N.F., 2000. The gratuitous relationship between Broca's aphasia and Broca's area. Behav. Brain Sci. 23, 30–31.
Duffy, S.A., Morris, R.K., Rayner, K., 1988. Lexical ambiguity and fixation times in reading. J. Mem. Lang. 27, 429–446.
Duffy, S.A., Kambe, G., Rayner, K., 2001. The effect of prior disambiguating context on the comprehension of ambiguous words: evidence from eye movements. In: Gorfein, D.S. (Ed.), On The Consequences of Meaning Selection. American Psychological Association, Washington, DC, pp. 105–117.
Durkin, K., Manning, J., 1989. Polysemy and the subjective lexicon: semantic relatedness and the salience of intraword senses. J. Psycholinguist. Res. 18 (6), 577–612.
Frazier, L., Rayner, K., 1990. Taking on semantic commitments: processing multiple meanings vs. multiple senses. J. Mem. Lang. 29, 181–200.
Gawlick-Grendell, L.A., Woltz, D.J., 1994. Meaning dominance norms for 120 homographs. Behav. Res. Methods Instrum. Comput. 26, 5–25.
Gernsbacher, M.A., 1990. Language Comprehension as Structure Building. Erlaum Associates, Hillsdale, NJ.
Gernsbacher, M.A., St. John, M.F., 2001. Modeling suppression in lexical access. In: Gorfein, D.S. (Ed.), On the Consequences of Meaning Selection. American Psychological Association, Washington, DC, pp. 47–65.
Gorfein, D.S., 2001. An activation-selection view of homograph disambiguation: a matter of emphasis? In: Gorfein, D.S. (Ed.), On the Consequences of Meaning Selection. American Psychological Association, Washington, DC, pp. 105–117.
Grodzinsky, Y., 2000. The neurology of syntax: language use without Broca's area. Behav. Brain Sci. 23, 1–71.
Hagoort, P., 1989. Processing of lexical ambiguities: a comment on Milberg, Blumstein, and Dworetzky (1987). Brain Lang. 36 (2), 335–348.
Hagoort, P., 1993. Impairments of lexical-semantic processing in aphasia: evidence from the processing of lexical ambiguities. Brain Lang. 45 (2), 189–232.
Kan, I.P., Thompson-Schill, S.L., 2004. Effect of name agreement on prefrontal activity during overt and covert picture naming. Cogn. Affect. Behav. Neurosci. 4 (1), 43–57.
Katz, W.F., 1988. An investigation of lexical ambiguity in Broca's aphasics using an auditory lexical priming technique. Neuropsychologia 26 (5), 747–752.
Kawamoto, A.H., 1993. Nonlinear dynamics in the resolution of lexical ambiguity: a parallel distributed processing account. J. Mem. Lang. 32, 474–516.
Kawamoto, A.H., Farrar, W.T., Kello, C.T., 1994. When two meanings are better than one: modeling the ambiguity advantage using a recurrent distributed network. J. Exp. Psychol. Hum. Percept. Perform. 20 (6), 1233–1247.
Klein, D.E., Murphy, G.L., 2001. The representation of polysemous words. J. Mem. Lang. 45 (2), 259–282.
Klein, D.E., Murphy, G.L., 2002. Paper has been my ruin: conceptual relations of polysemous senses. J. Mem. Lang. 47, 548–570.
Klepousniotou, E., 2002. The processing of lexical ambiguity: homonymy and polysemy in the mental lexicon. Brain Lang. 81 (1–3), 205–223.
Klepousniotou, E., Pike, B., Steinhauer, K., Gracco, V., 2006. Meaning activation patterns of homonymy and polysemy: an EEG investigation. Paper presented at the Cognitive Neuroscience Society 2006 Annual Meeting, San Francisco, CA (April 8th–11th).
Martin, C., Vu, H., Kellas, G., Metcalf, K., 1999. Strength of discourse context as a determinant of the subordinate bias effect. Q. J. Exp. Psychol. A 52 (4), 813–839.
McNamara, D.S., McDaniel, M.A., 2004. Suppressing irrelevant information: knowledge activation or inhibition? J. Exp. Psychol.: Learn. Mem. Cogn. 30 (2), 465–482.
Metzler, C., 2001. Effects of left frontal lesions on the selection of context-appropriate meanings. Neuropsychology 15 (3), 315–328.
Milberg, W., Blumstein, S., 1989. Reaction time methodology and the aphasic patient: a reply to Hagoort (1988). Brain Lang. 36 (2), 349–353.
Milberg, W., Blumstein, S.E., Dworetzky, B., 1987. Processing of lexical ambiguities in aphasia. Brain Lang. 31 (1), 138–150.
Neely, J.H., Keefe, D.E., Ross, K.L., 1989. Semantic priming in the lexical decision task: roles of prospective prime-generated

expectancies and retrospective semantic matching. J. Exp. Psychol. Lear., Mem., Cogn. 15 (6), 1003–1019.
Novick, J.M., Trueswell, J.C., Thompson-Schill, S.L., 2005. Cognitive control and parsing: reexamining the role of Broca's area in sentence comprehension. Cogn. Affect. Behav. Neurosci. 5 (3), 263–281.
Prather, P.A., Zurif, E., Stern, C., Rosen, T.J., 1992. Slowed lexical access in nonfluent aphasia: a case study. Brain Lang. 43, 336–348.
Prather, P.A., Zurif, E., Love, T., Brownell, H., 1997. Speed of lexical activation in nonfluent Broca's aphasia and fluent Wernicke's aphasia. Brain Lang. 59, 391–411.
Pustejovsky, J., 1995. The Generative Lexicon. MIT Press, Cambridge, Mass.
Rodd, J.M., Gaskell, M.G., Marslen-Wilson, W.D., 2002. Making sense of semantic ambiguity: semantic competition in lexical access. J. Mem. Lang. 46, 245–266.
Rodd, J.M., Gaskell, M.G., Marslen-Wilson, W.D., 2004. Modeling the effects of semantic ambiguity in word recognition. Cogn. Sci. 28, 89–104.
Rodd, J.M., Davis, M.H., Johnsrude, I.S., 2005. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. Cereb. Cortex 15 (8), 1261–1269.
Schvaneveldt, R.W., Meyer, D.E., 1976. Lexical ambiguity, semantic context, and visual word recognition. J. Exp. Psychol. Hum. Percept. Perform. 2 (2), 243–256.
Sereno, S.C., Brewer, C.C., O'Donnell, P.J., 2003. Context effects in word recognition: evidence for early interactive processing. Psychol. Sci. 14 (4), 328–333.
Shivde, G., Anderson, M.C., 2001. The role of inhibition in meaning selection: insights from retrieval-induced forgetting. In: Gorfein, D. (Ed.), On the Consequences of Meaning Selection: Perspectives on Resolving Lexical Ambiguity. American Psychological Association, Washington, DC, pp. 175–190.
Simpson, G.B.A., A.C., 2001. Repeated homographs in word and sentence contexts: multiple processing of multiple meanings. In: Gorfein, D.S. (Ed.), On the Consequences of Meaning Selection. American Psychological Association, Washington, DC, pp. 105–117.
Swaab, T.Y., Brown, C., Hagoort, P., 1998. Understanding ambiguous words in sentence contexts: electrophysiological evidence for delayed contextual selection in Broca's aphasia. Neuropsychologia 36 (8), 737–761.
Swaab, T.Y., Brown, C., Hagoort, P., 2003. Understanding words in sentence contexts: the time course of ambiguity resolution. Brain Lang. 86 (2), 326–343.
Swinney, D.A., 1979. Lexical access during sentence comprehension: (Re)consideration of context effects. J. Verbal Learn. Verbal Behav. 18 (6), 645–659.
Tanenhaus, M.K., Leiman, J.M., Seidenberg, M.S., 1979. Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. J. Verbal Learn. Behav. 18, 427–440.
Thompson-Schill, S.L., D'Esposito, M., Aguirre, G.K., Farah, M.J., 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. Proc. Natl. Acad. Sci. U. S. A. 94 (26), 14792–14797.
Thompson-Schill, S.L., Swick, D., Farah, M.J., D'Esposito, M., Kan, I.P., Knight, R.T., 1998. Verb generation in patients with focal frontal lesions: a neuropsychological test of neuroimaging findings. Proc. Natl. Acad. Sci. U. S. A. 95 (26), 15855–15860.
Thompson-Schill, S.L., D'Esposito, M., Kan, I.P., 1999. Effects of repetition and competition on activity in left prefrontal cortex during word generation. Neuron 23 (3), 513–522.
Thompson-Schill, S.L., Bedny, M., Goldberg, R.F., 2005. The frontal lobes and the regulation of mental activity. Curr. Opin. Neurobiol. 15 (2), 219–224.
Traxler, M.J., Pickering, M.J., McElree, B., 2002. Coercion in sentence processing: evidence from eye-movements and self-paced reading. J. Mem. Lang. 47, 530–547.
Twilley, L.C., Dixon, P., Taylor, D., Clark, K., 1994. University of Alberta norms of relative meaning frequency for 566 homographs. Mem. Cogn. 22 (1), 111–126.
Weinreich, U., 1964. Webster's third: a critique of its semantics. Int. J. Am. Linguist. 30, 405–409.
Williams, J.N., 1992. Processing polysemous words in context: evidence from interrelated meanings. J. Psycholinguist. Res. 21 (3), 193–218.