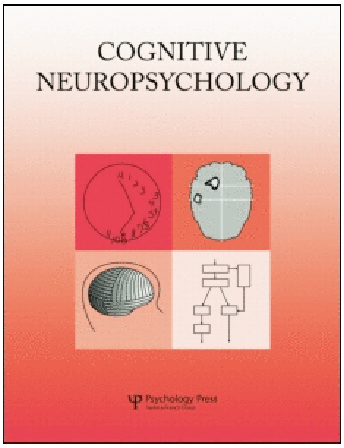
This article was downloaded by: *[Novick, Jared M.]*
On: *5 February 2010*
Access details: *Access Details: [subscription number 919103496]*
Publisher *Psychology Press*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Cognitive Neuropsychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713659042

A case for conflict across multiple domains: Memory and language impairments following damage to ventrolateral prefrontal cortex

Jared M. Novick [a][b]; Irene P. Kan [c][d]; John C. Trueswell [c]; Sharon L. Thompson-Schill [c]
[a] University of Pennsylvania, PA, USA [b] Center for Advanced Study of Language & Department of Psychology, University of Maryland, College Park, MD, USA [c] University of Pennsylvania, Philadelphia, PA, USA [d] Villanova University, Villanova, PA, USA

Online publication date: 05 February 2010

To cite this Article Novick, Jared M., Kan, Irene P., Trueswell, John C. and Thompson-Schill, Sharon L.(2009) 'A case for conflict across multiple domains: Memory and language impairments following damage to ventrolateral prefrontal cortex', Cognitive Neuropsychology, 26: 6, 527 — 567

To link to this Article: DOI: 10.1080/02643290903519367

URL: http://dx.doi.org/10.1080/02643290903519367

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

COGNITIVE NEUROPSYCHOLOGY, 2009, 26 (6), 527–567

Psychology Press
Taylor & Francis Group

A case for conflict across multiple domains: Memory and language impairments following damage to ventrolateral prefrontal cortex

Jared M. Novick
University of Pennsylvania, PA, USA, and Center for Advanced Study of Language & Department of Psychology, University of Maryland, College Park, MD, USA

Irene P. Kan
University of Pennsylvania, Philadelphia, PA, USA, and Villanova University, Villanova, PA, USA

John C. Trueswell and Sharon L. Thompson-Schill
University of Pennsylvania, Philadelphia, PA, USA

Patients with focal lesions to the left inferior frontal gyrus (LIFG; BA 44/45) exhibit difficulty with language production and comprehension tasks, although the nature of their impairments has been somewhat difficult to characterize. No reported cases suggest that these patients are Broca's aphasics in the classic agrammatic sense. Recent case studies, however, do reveal a consistent pattern of deficit regarding their general cognitive processes: They are reliably impaired on tasks in which conflicting representations must be resolved by implementing top-down cognitive control (e.g., Stroop; memory tasks involving proactive interference). In the present study, we ask whether the language production and comprehension impairments displayed by a patient with circumscribed LIFG damage can best be understood within a general conflict resolution deficit account. We focus on one patient in particular—patient I.G.—and discuss the implications for language processing abilities as a consequence of a general cognitive control disorder. We compared I.G. and other frontal patients to age-matched control participants across four experiments. Experiment 1 tested participants' general conflict resolution abilities within a modified working memory paradigm in an attempt to replicate prior case study findings. We then tested language production abilities on tasks of picture naming (Experiment 2) and verbal fluency (Experiment 3), tasks that generated conflict at the semantic and/or conceptual levels. Experiment 4 tested participants' sentence processing and comprehension abilities using both online (eye movement) and offline measures. In this task, participants carried

---

Correspondence should be addressed to Jared M. Novick, Center for Advanced Study of Language, University of Maryland, Box 25, College Park, MD 20742–0025, USA (E-mail: jnovick1@umd.edu)

This research was supported by National Institutes of Health (NIH) Grants R01-HD37507, awarded to J.C.T., and R01-MH67008, awarded to S.L.T.-S. We thank David January and J. Ben Hutchinson for assisting with stimulus creation and participant testing, and Rebecca Stone, Jennifer Greene, Zoe Clancy, and David January for help with data coding. Additionally, we are extremely grateful to H. Branch Coslett and David Kraemer for assistance with lesion identification and figures. A portion of this work was presented at the Seventeenth Annual City University of New York (CUNY) Conference on Human Sentence Processing in College Park, MD, and at the 45th Annual Meeting of the Psychonomics Society in Minneapolis, MN. We thank the audiences at these conferences for their invaluable feedback. Finally, we are grateful to all the patients who participated in this research and their families, especially for their endurance, their commitment to being involved in this work, and for the joy in getting to know them.

http://www.psypress.com/cogneuropsychology
DOI:10.1080/02643290903519367

out spoken instructions containing a syntactic ambiguity, in which early interpretation commitments had to be overridden in order to recover an alternative, intended analysis of sentence meaning. Comparisons of I.G.'s performance with frontal and healthy control participants supported the following claim: I.G. suffers from a general conflict resolution impairment, which affects his ability to produce and comprehend language under specific conditions—namely, when semantic, conceptual, and/or syntactic representations compete and must be resolved.

Keywords: Left inferior frontal gyrus (LIFG); Broca's area; Conflict resolution and cognitive control; Language processing; Ambiguity resolution.

As we go about our day, navigating and perceiving the world around us, we occasionally encounter moments when we have to rein in our initial cognitive or emotional reactions to something or someone in light of other knowledge that promotes different interpretations, goals, and/or plans of action. Some of these moments fall into the category of politeness or social norms: We prevent ourselves from laughing when something embarrassing happens to a friend or colleague. Other such moments are more subtle but provoke an arguably quite similar experience: We can prevent ourselves from calling someone by an incorrect name that inexplicably comes to mind; we can avoid coming to the wrong interpretation of an ambiguous sentence (consider reading the CNN headline: "Ohio bodies are missing New Hampshire children"). We often successfully achieve our immediate goals with ease by relying on a collection of automatic behaviours and mental computations, but as just illustrated, there are circumstances under which we must instead attend to the currently relevant rules or knowledge at hand, which impose the need to override familiar or "gut" responses. Cognitive psychologists and cognitive neuroscientists commonly refer to this process of discounting dominant or prepotent behaviours as involving *cognitive* (or *executive*) control—a top-down function that guides behaviour in accordance with current goals and task demands.

The need to exert cognitive control can arise at multiple levels, from competing representations of a stimulus to incompatible response options (for recent reviews see Kan & Thompson-Schill, 2004b; Novick, Trueswell, & Thompson-Schill, 2005). In the laboratory, cognitive control abilities have been studied using a variety of tasks, perhaps most notably the classic Stroop task (Stroop, 1935). In this task, participants must name the colour of the ink in which colour terms are printed instead of reading the words themselves. Competition—and thus the need for cognitive control—is created on trials in which the ink colour and the colour term do not match: for instance, the word "red" printed in blue ink.

The type of conflict created by the Stroop paradigm—namely, the need to override the prepotent reading response—is one example of a situation necessitating cognitive control. Conflict requiring *prepotent response override* can be distinguished from conflict that arises from competition among multiple, incompatible representations, none of which is more compelling than the others (Botvinick, Braver, Barch, Carter, & Cohen, 2001). For instance, neuropsychological assessments of language impairments often include a category fluency task, in which patients are asked to generate as many exemplars as possible given a particular category cue, such as "animals". This cue can give rise to *underdetermined response conflict* because all exemplars within this category are equally valid response candidates (e.g., "cow", "rabbit", and "chicken", among others).

In this study we examine the situations that create—and the processes that then resolve—both types of conflict. Specifically, we characterize conflict that arises in the course of language processing, both during language production (i.e., picture naming and verbal fluency) and during sentence comprehension. We address these issues with a comparison of normal language abilities to those of a neurological patient who was selected on the basis of his specific neuroanatomical profile—namely, circumscribed damage to the inferior frontal gyrus of the left hemisphere. As we review briefly below, damage to this region is hypothesized to produce a general deficit in conflict resolution and cognitive control abilities. This case illustrates the very specific linguistic deficits that are predicted to arise from a more general failure of cognitive control.

Conflict resolution and cognitive control: The role of left inferior frontal gyrus (LIFG)

For more than 50 years, the function of the prefrontal cortex (PFC) has been tied to flexible behaviour, or the attention-demanding ability to adapt to new rules, goals, and constantly changing task demands (see Miller & Cohen, 2001, for a review). Recently, researchers have begun to fractionate the PFC into more specific functional–anatomical associations, in order to more clearly define which regions within PFC support particular types of cognitive control functions.

The link between the left inferior frontal gyrus (LIFG; in particular BA 44/45) and processes involved in conflict resolution and cognitive control has been established by functional neuroimaging data obtained in a wide variety of experimental settings (e.g., Stroop task: Milham, Banich, & Barad, 2003; working memory: Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; Nelson, Reuter-Lorenz, Sylvester, Jonides, & Smith, 2003; verb generation: Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; picture naming: Kan & Thompson-Schill, 2004a). Across these tasks, one finds examples of the LIFG responding to conflict arising both from prepotent response conflict and from underdetermined response conflict; further, some neuroimaging data (e.g., Milham et al., 2003; Nelson et al., 2003) have indicated a specific role for LIFG in the resolution of conflict that is distinguishable from response selection (henceforth, *representational conflict*). These data all point to a general role for LIFG in the resolution of conflict, broadly defined, although questions about functional subdivisions within the LIFG remain (cf. Badre, Poldrack, Paré-Blagoev, Insler, & Wagner, 2005; Badre & Wagner, 2002; see also Donohue, Wendelken, Crone, & Bunge, 2005; Nagel, Schumacher, Goebel, & D'Esposito, 2008).

Compelling results from neuropsychological case studies corroborate and extend these findings (Hamilton & Martin, 2005; Robinson, Blair, & Cipolotti, 1998; Robinson, Shallice, & Cipolotti, 2005; Thompson-Schill et al., 2002; Thompson-Schill et al., 1998). In one study, a patient (M.L.) with a lesion in LIFG including the frontal and parietal operculum showed exaggerated difficulty completing conflict-related trials on the Stroop task, evidenced by reaction times and error rates well outside the normal range (e.g., Hamilton & Martin, 2005). M.L. also exhibited a selective impairment resolving proactive interference on a modified item recognition task within a working memory paradigm, in which familiar items from previous memory sets obstructed his ability to respond quickly and accurately to probes from the current and immediately relevant sets (for a similar finding, see Thompson-Schill et al., 2002).

Other neuropsychological evidence comes from Robinson and colleagues (Robinson, et al., 1998), who report a patient (A.N.G.) with a left frontal meningioma impinging on LIFG (BA 45). On verbal generation tasks in which stimuli activated several competing and therefore underdetermined candidate responses (for instance, a sentence completion task with an open-ended context, e.g., *The man went into his house and . . .*), A.N.G. showed a selective impairment generating completions relative to conditions in which the continuation was much more constrained and therefore predictable (e.g., *The man went into the movie theatre and . . .*, where the most likely response is "watched a movie"). The authors reasoned that the unconstrained conditions—akin to an underdetermined response conflict task—offered too many competing alternative continuations for the patient to resolve, and thus A.N.G.'s damage to BA 45 resulted in "an inability to select a verbal response in situations where the stimulus activate[s] many competing response options" (Robinson et al., 1998, p. 82). Taken together, the findings from both brain-imaging and patient data indicate that LIFG is involved in, and necessary for, the successful resolution of representation-based competition (for a review, see Novick et al., 2005).

The primary goal of the current investigation was to explore the consequences of a putative cognitive control deficit for language abilities. In other words, can the specificity of the linguistic impairments that result from focal damage to LIFG be gainfully understood as a more general failure of cognitive control? Although most of the cognitive control tasks that pose difficulty for LIFG patients include general verbal components (e.g., naming colours or remembering lists of letters), which span across a variety of different experimental paradigms, they do not involve deploying complex linguistic processes that serve the purpose of communicating information through combinatory syntactic, semantic, or other grammatical operations, as is the case in everyday language production and comprehension. Interestingly, patients with lesions confined to LIFG, which includes Broca's area (BA 44/45) (see Amunts et al., 1999), appear to have only transient and limited language production and comprehension difficulties (Hamilton & Martin, 2005; Robinson et al., 1998; Thompson-Schill et al., 2002). Crucially, they do not necessarily suffer from Broca's aphasia (Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004; Lindenberg, Fangerau, & Seitz, 2007; for a review, see Novick et al., 2005), which is diagnosed on the basis of a constellation of *behavioural* impairments, and not on lesion site per se. Robinson and colleagues have in fact argued that LIFG patients' impairments are reduced to cases in which input stimuli activate many competing verbal responses or when context does not sufficiently (or easily) constrain and guide production (Robinson et al., 1998; Robinson et al., 2005). However, it has yet to be demonstrated that the transient language difficulties seen in such patients are attributable to a general failure of conflict resolution and cognitive control.

Cognitive control and language use: Preliminaries and predictions

We address here the extent to which general conflict resolution and cognitive control mechanisms, and thereby the injury to cortical areas that support such mechanisms, contribute to specific language production and comprehension skills. In light of mounting evidence from both neuroimaging and neuropsychological studies suggesting that LIFG is consistently involved in (and necessary for) resolving conflict among competing representations (Hamilton & Martin, 2005; Jonides et al., 1998; Kan & Thompson-Schill, 2004b; Nelson et al., 2003; Thompson-Schill et al., 2002; for a review see Novick et al., 2005), we expect that LIFG ought to be involved in only special—that is, limited—cases of language use within the spheres of both production and comprehension. In particular, when conceptual, lexical, syntactic, and/or semantic representations compete for a response, thereby creating high conflict resolution demands, patients with circumscribed lesions to LIFG should have particular difficulty resolving the conflict.

Consider, for example, the impact of stimulus indeterminacy on language production. On a confrontation picture-naming task, patients are asked to name common objects as they are presented. Pictures that are associated with multiple possible (i.e., underdetermined) names, such as a drawing of an item of furniture that might be called a couch, a sofa, or a loveseat, are hypothesized to introduce a serious challenge because all the name options, generated by their conceptual representations, should compete, and none of the options is more compelling than the next. By contrast, objects with high name agreement, such as an apple, should present no such difficulty, as the conflict demands are relatively limited. We present the details of these experimental investigations of conflict resolutions during word production in Experiments 2 and 3.

Similarly, in the domain of processing and comprehending written or spoken language, temporary ambiguity abounds; thus different representations of sentence meaning compete. For example, evidence from language-processing studies suggests that readers and listeners process sentences in real time, as the words and phrases are encountered; when reading or interpreting speech, we do not wait until the end of a sentence or even the end of a single word before we start assigning an interpretation to what's being produced by the writer or speaker (e.g., Altmann & Kamide, 1999; Marslen-Wilson, 1987; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; inter alia). As we discuss in Experiment 4 below, incremental processing of sentences can place high demands on the parsing and executive systems especially when the initial interpretation of a temporary ambiguity must be overridden in light of later arriving linguistic material (i.e., the so-called garden-path sentence). Under such circumstances, a patient with a circumscribed lesion to LIFG is predicted ultimately to be unable to reprioritize highly supported—but ultimately incorrect—initial parsing commitments: He will have difficulty overriding first interpretations associated with the most reliable probabilistic evidence, perhaps in some cases never arriving at the appropriate target meaning of the sentence at all. This view is motivated in part by previous results reported in the functional magnetic resonance imaging (fMRI) and patient literatures demonstrating a critical function for LIFG in ambiguity resolution across multiple levels of representation (see Novick et al., 2005)—for example, resolving word sense ambiguity (Bedny, Hulbert, & Thompson-Schill, 2007; Bedny, McGill, & Thompson-Schill, 2008; Mason & Just, 2007; Rodd, Davis, & Johnsrude, 2005; Zempleni, Renken, Hoeks, Hoogduin, & Stowe, 2007), ambiguous phonetic categories (Blumstein, Myers, & Rissman, 2005), and syntactic ambiguity (e.g., January, Trueswell, & Thompson-Schill, 2009; Mason, Just, Keller, & Carpenter, 2003; Ye & Zhou, 2009)—including lexical (noun/verb) ambiguities embedded in sentences, which results in a form of syntactic ambiguity (see Snijders et al., 2009).

Profile of the single-case patient and experimental prospectus

We report a case study of a patient with a restricted lesion to LIFG (see Figure 1A) and compare his performance on a variety of tasks to groups of age-matched non-LIFG neuropsychological patients and healthy control participants (lesion sites of patients in the neuropsychological group are also illustrated in Figure 1 B–D). All participants in this report were native speakers of American English. None had any history of psychiatric impairment or substance abuse, nor did they present with any visual field cuts or eye movement disorders.

Our patient of interest, patient IG363 (henceforth I.G.),[1] has damage to the left frontal operculum, the same region as that previously highlighted in at least three separate case studies, which have shown patients' exaggerated deficits in the ability to resolve competing verbal representations in memory (Hamilton & Martin, 2005; Robinson et al., 1998; Thompson-Schill et al., 2002). Thus, based on this anatomical characterization alone, one would predict that I.G. has impaired cognitive control. After confirming this prediction (Experiment 1), the remainder of the paper is devoted to a careful scrutiny of I.G.'s language abilities and disabilities, using production and comprehension tasks, each with conditions that place varying demands on cognitive control.

I.G. is a right-handed, college-educated male, who suffered an occlusion to the precentral branch of the left middle cerebral artery two years prior to testing, which resulted in focal damage to LIFG (left frontal operculum and pars triangularis; BA 44/45). At the time of testing he was 66 years old. I.G.'s language production was intelligible, clear, and free of any articulation errors; however, occasional word-finding problems and rare pauses made his production mildly dysfluent, though he was not agrammatic. His formal lexical comprehension score on the Philadelphia

---

[1] Throughout the paper, we refer to patients using their anonymous identification code in the University of Pennsylvania patient database system. For ease of exposition, we refer to our patient of interest as I.G. in the text, which is a shortening of his database ID number (IG363).

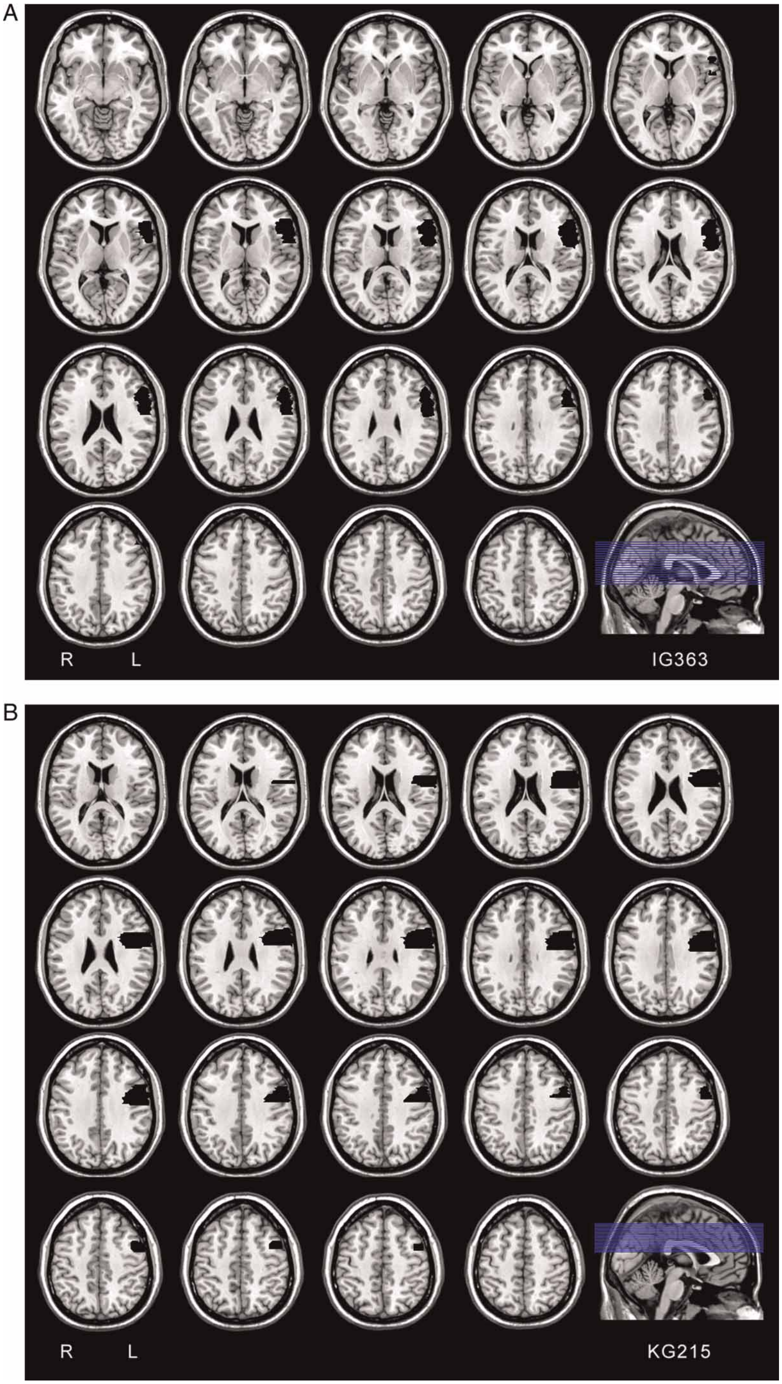


Figure 1. *Radiological scans showing the full extent of the lesions for (A) patient IG363, whose lesion involved circumscribed damage to the left frontal operculum and pars triangularis; (B) patient KG215; (C) patient GU412; and (D) patient NN454. The sagittal slice for each patient in the figure provides a sense of scale. A board-certified neurologist verified that I.G. had substantially more involvement of BA 44/45 than did the other three patients. To view a colour version of this figure, please see the online issue of the Journal.*

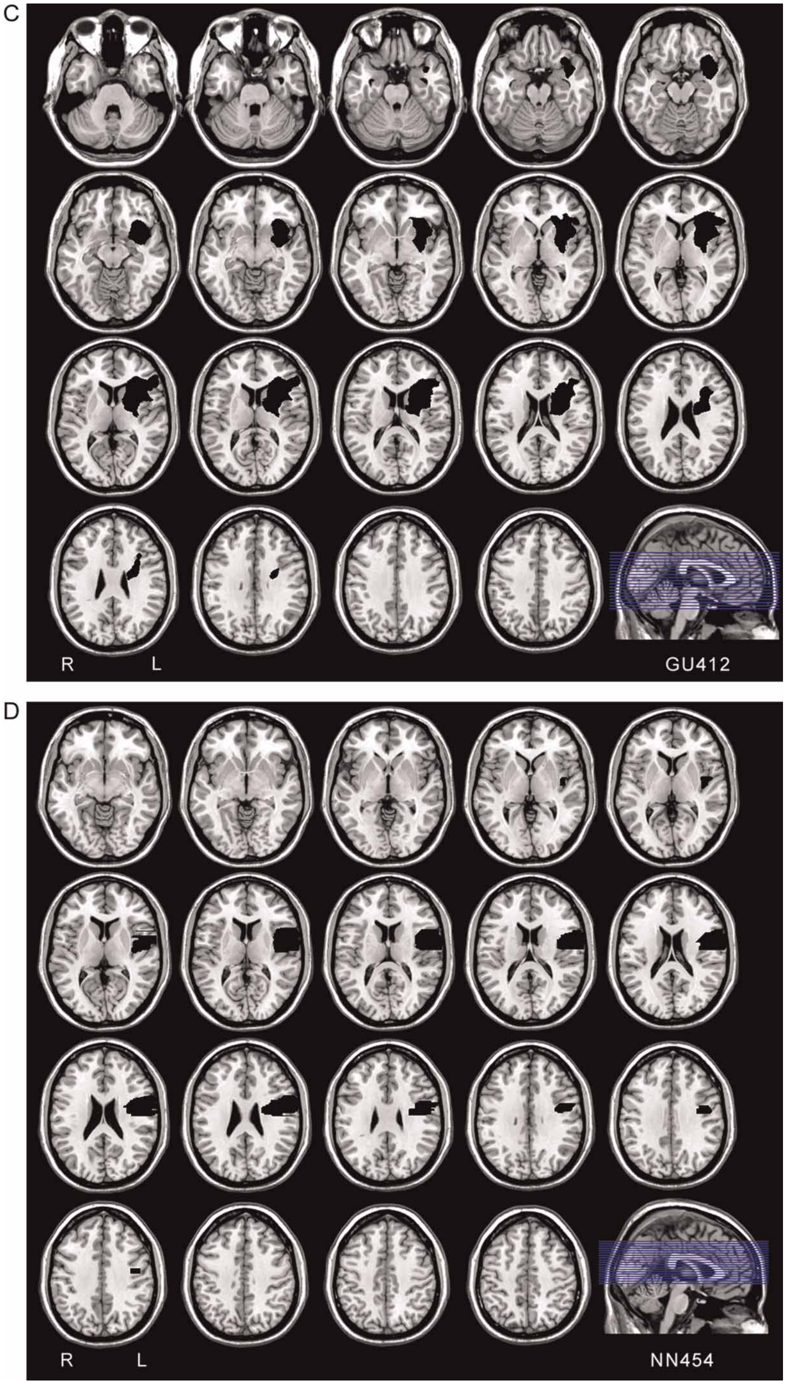


Figure 1. *Continued*

Comprehension Battery (PCB; Saffran, Schwartz, Linebarger, Martin, & Bochetto, 1988) was 98%. On the lexical comprehension task, he was required to listen to a word and to choose the corresponding picture given four choices. The distractor pictures were either in the same or in a different category from the target. For the across-category section, one picture was either a phonemic or a visual distractor. I.G. made only one error on this subtest (98% correct), which was within category, and he later commented that he realized he had chosen the wrong picture.[2] He scored 90% and 77% on grammaticality judgement and synonymy triplets, respectively. His overall score was 93% on the PCB and 87% on the Western Aphasia Battery (Kertesz, 1982).

Finally, I.G. was administered the American version of the National Adult Reading Test (AMNART; Grober, Sliwinski, Schwartz, & Saffran, 1989) to estimate his premorbid verbal intelligence (see Crawford, Deary, Starr, & Whalley, 2001; Grober & Sliwinski, 1991). He scored a 115, which was well within the normal range (mean = 100; *SD* = 15) and indistinguishable from the general intellectual functioning of the other patients examined in this investigation (see below).

In sum, I.G. presented with some language difficulties that were more evident in production than comprehension. Here we ask: does I.G. have an impairment of his language faculties per se, or are his deficits more usefully characterized in nonlinguistic terms?

In what follows, Experiment 1 demonstrates that I.G.—akin to patients from prior studies with a similar neuroanatomical profile—loses normal cognitive functioning when interfering and incompatible characterizations of memory traces arise and need to be resolved. Experiments 2 through 4 expand upon this further in the language production and comprehension domains and test whether the specificity of his production and comprehension deficits can also be linked to increased levels of representational conflict. In particular, Experiments 2 and 3 focus on whether the competition that arises from relatively unconstrained semantic or conceptual representations leads to deficits in language production, and Experiment 4 addresses whether any comprehension impairment can be explained by competition in the form of temporary syntactic ambiguity during spoken language comprehension.

EXPERIMENT 1: RESOLVING PROACTIVE INTERFERENCE WITHIN A WORKING MEMORY PARADIGM

In order to assess I.G.'s general conflict resolution abilities within working memory, and to replicate the findings of prior case studies studying similar patients, we administered the same proactive interference task that previous neuropsychological and brain-imaging studies have employed to show LIFG's role in resolving among competing mnemonic representations (Hamilton & Martin, 2005; Jonides et al., 1998; Nelson et al., 2003; Thompson-Schill et al., 2002). In this task, participants have to indicate whether a probe (e.g., "K") is part of an immediately preceding set of items (e.g., "m k d p"). For the majority of trials, this can be achieved relatively easily by relying on stimulus familiarity. That is, if the probe representation is quickly recognizable, it is usually safe to say that it recently appeared, and therefore the response should be "yes". However, on a small proportion of trials, conflict is generated by presenting a recognition probe (e.g., "R") that is not a member of the current memory set (e.g., "t b w f") but was a member of the memory set on the previous trial (e.g., "r m b t"; see Method below). Reaction times and error rates increase on these so-called "recent–no" trials compared to trials in which the probe item did not recently appear,

---

[2] A lack of sensitivity to phonological distractors on picture–word matching is revealing for I.G.'s comprehension profile. We note, however, that although one might assume that such a lack of phonological confusion is also true for production, this is not necessarily the case. Unfortunately, we do not have enough information from the assessments to permit such a phonological-distractor analysis for I.G.'s production.

because participants have to come up with a response that reconciles the conflict between the generally reliable—but currently misleading—familiarity-based information and the foremost activation of the letters from the current set. It was on these trials, compared to trials with nonrecently appearing probes, that patient M.L. showed an exaggerated interference effect for accuracy (see also patient R.C. in Thompson-Schill et al., 2002). Such trials also routinely activate LIFG in neuroimaging studies of healthy participants (e.g., Jonides et al., 1998; Nelson et al., 2003). We have argued elsewhere (see Novick et al., 2005) that conflict-related "recent–no" trials in this task and incongruent trials in the Stroop task have in common the need to engage control mechanisms that are necessary to bias representations toward what is immediately relevant for the current task instructions. Hence, similar patterns of boosted activation in LIFG during functional neuroimaging are found across both task types (e.g., Jonides et al., 1998; Milham et al., 2003; Nelson et al., 2003), and inflated interference effects in behavioural performance are observed in patients with restricted damage to LIFG (see Novick et al., 2005, for an expanded view of this idea and discussion of other related tasks). The primary purpose of Experiment 1 was to establish whether the current patient of interest, like other patients with LIFG damage, would exhibit a comparable deficit resolving conflict induced by proactive interference.

Method

Participants

In addition to patient I.G., two left frontal patients participated in this task, whose lesions, crucially, spared LIFG (see text below; Figure 1). In addition, these three patients were compared to the same healthy control group ($n = 6$) as that reported in Thompson-Schill et al. (2002). This "elderly" control group's mean age (62.3 years) was matched to I.G. and showed a reliable interference effect for both reaction time and errors on this task. All patients in the current study were recruited from the Department of Neurology and Center for Cognitive Neuroscience at the University of Pennsylvania, provided informed consent, and were compensated $15 per hour. Each of the two control groups—that is, the two other frontal (non-LIFG) patients and the healthy group from Thompson-Schill et al. (2002)—are described separately below.

Left frontal patients ($n = 2$). We identified two frontal patients who suffered diffuse damage to left PFC as a result of a cerebral vascular accident to the left hemisphere; crucially though they had unilateral damage to the left lateral prefrontal cortex that spared the critical posterior LIFG site of interest. Patient GU412 had damage to more medial portions of inferior frontal regions with negligible involvement of the frontal operculum and BA 44/45 more broadly (Figure 1C), and more white matter damage than other patients in the current study. She was 40 years old at testing (no AMNART score available), and on the Western Aphasia Battery (WAB) she made few errors. She was fluent, and her comprehension was above average, but her main difficulty on this battery was repetition, not confrontation naming; her behavioural diagnosis was conduction aphasia. Patient NN454 suffered damage primarily to the left perisylvian region (Figure 1D). She was 62 years old at testing (AMNART score = 115).

Healthy control group (from Thompson-Schill et al., 2002). The healthy control group comprised 6 right-handed individuals (3 female) who were matched to I.G. for age ($M = 62.3$ years, range = 54–81). None had any history of substance abuse or neurological or psychiatric disorders.

Procedure

This experiment employed the same modified item recognition paradigm as that described in Thompson-Schill et al. (2002; see also Jonides et al., 1998; McElree & Dosher, 1989; Monsell, 1978; Nelson et al., 2003). Participants completed 160 trials as follows: Each trial began with a "Get Ready" prompt followed by a fixation cross, which was displayed for 500 ms. Four lower-case letters were then presented around the cross for

1,500 ms, and participants were instructed to remember this memory set. After a 3-s delay, a single upper-case letter was presented in the centre of the screen. A participant's task was to determine whether the probe was a member of the previously studied set. The recognition probe appeared in upper case so that participants could not merely perform a visual judgement to correctly identify whether the probe was a member of the presented set of items. Participants responded by pressing a YES or NO button on a button box with their dominant hand (there were no cases of hemiplegia of the dominant hand).

Participants completed 16 practice trials and 160 test trials in one session. A brief break was permitted after Trials 40, 80, and 120, in which a "Take a Break" screen appeared that remained until the participant was ready to begin again. The entire task lasted approximately 45 minutes.

Experimental design

Half of all trials required a "yes" response, and half required a "no" response. However, half of all "yes" and half of all "no" trials contained probe items that were also members of the previous trial's memory set; these trials are called "recent" trials. The other half of "yes" and "no" trials are called "nonrecent" trials, in which the probe items were not members of either of the previous two memory sets. The resulting four trial types ("nonrecent–yes", "nonrecent–no", "recent–yes", and "recent–no") were pseudorandomly arranged throughout the experiment. For all trials, two of the four letters in the memory set from the previous trial were repeated in the subsequent trial so that item recurrence in the memory set was not confounded with the type of trial.

At the end of the experiment, participants completed a one-block control condition of 40 trials that were identical to the target trials except that no target or probe items were repeated between trials—that is, the four letter stimuli that were used on each trial, including targets and probes, did not appear on either of the two prior trials. This control design tested baseline working memory abilities more generally. There were both "yes" and "no" control trials. Table 1 provides examples of each trial type.

Results

For each of the six trial types, accuracy and median response time (RT) to correct items were calculated for each participant. The RT data for each condition, including interference effects (RT for recent–no trials minus RT for nonrecent–no trials), appear in Table 2. As can be seen for the neurologically intact age-matched control group (see also Thompson-Schill et al., 2002), participant reaction times were on average 90 ms longer for recent–no trials than for nonrecent–no trials.

Given that the paradigm employed in Thompson-Schill et al. (2002) is identical to the one reported here, we use precisely the same method of patient/control group comparison as that established previously to straightforwardly connect to prior work examining LIFG patients' performance on this task (Thompson-Schill

Table 1. *Sample trials from six conditions in the modified item recognition task with proactive interference, Experiment 1*

		Trial n − 1		*Trial n*	
Condition (Trial n)		*Target set*	*Probe*	*Target set*	*Probe*
Control	Yes	g k p n	D	s f h m	F
	No	w b t q	X	c z r v	J
Nonrecent	Yes	m d s k	M	s f k t	T
	No	v n b f	C	r b j n	D
Recent	Yes	k r z b	N	b t k s	B
	No	h l w p	L	k p w n	H

Table 2. *Reaction times and interference effects from item recognition task with proactive interference, Experiment 1*

	Trial type						
	Control		*Nonrecent*		*Recent*		
Participants	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*Interference effect*[b]
Controls[a] ($n = 6$)	1,347 (185)	1,212 (155)	1,374 (190)	1,331 (299)	1,464 (215)	1,260 (279)	90 (149)
Patient NN454	1,978	1,194	2,229	1,160	2,393	925	164
Patient GU412	1,119	1,016	1,130	1,084	1,146	1,050	16
Patient IG363	1,460	1,466	1,611	1,460	1,698	1,441	87

Note: Reaction times in ms. Standard deviations in parentheses.
[a]The healthy control group is the same group as that tested in Thompson-Schill et al. (2002). [b]The interference effect is the difference between the recent–no and the nonrecent–no trials.

Table 3. *Error rate and interference effects from item recognition task with proactive interference, Experiment 1*

	Trial type						
	Control		*Nonrecent*		*Recent*		
Participants	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*Interference effect*[b]
Controls[a] ($n = 6$)	5.8 (9.7)	6.2 (10.7)	3.8 (6.6)	10.5 (15.3)	11.3 (10.2)	9.6 (8.6)	7.5 (4.3)
Patient NN454	65.0	20.0	62.9	15.4	69.2	12.8	6.4
Patient GU412	15.0	0	0	12.8	12.8	15.4	12.8
Patient IG363	15.0	25.0	5.7	30.8	20.5	13.1	14.8[c]

Note: Error rates in percentages. Standard deviations in parentheses.
[a]The control group is the same group as that tested in Thompson-Schill et al. (2002). [b]The interference effect is the difference between the recent–no and the nonrecent–no trials. [c]interference effect beyond 1.64 standard deviations (95th percentile) of control group's mean.

et al., 2002; see also Hamilton & Martin, 2005). Following earlier single-case approaches, then, each patient's interference effect was tested against that of the healthy control group. Impaired performance was defined as an interference effect that fell 1.64 standard deviations above the control mean (approximately the 95th percentile; see Thompson-Schill et al., 2002). Based on this definition, the cut-off for normal reaction time interference was 339 ms. As can be seen in Table 2, all 3 patients, including I.G., performed within the normal range.[3]

Table 3 presents mean proportion of errors for the healthy controls and the patients. The mean

---

[3] Across the four experiments in the current study, we compare the scores of all patients, including I.G., to the mean performance of the healthy control group. Some researchers (e.g., Haarmann & Kolk, 1991) have pointed out that this practice has potential to be flawed due to the possibility of considerable variation within the control group: Some control participants, despite the overall mean, could in principle be performing exactly the same as the "extreme" patients. However, inspection of our data reveals that this was never the case, and that the range for individual healthy control participants was rather dissimilar from I.G.'s range of performance on all tasks. Thus, we collapsed the control participants into a single group mean for comparison with the individual patients and are confident that our data are not vulnerable to this concern.

interference magnitude for healthy participants' error rate was 7.5% ($SD = 4.3$). Again, any patient who scored more than 1.64 standard deviations beyond the normal mean (corresponding to 14.5% or higher) was considered impaired. As can be seen in Table 3, patient I.G. was the only participant to demonstrate an interference magnitude that exceeded this cut-off. He made on average 20.5% errors on recent–no trials but only 5.7% errors on nonrecent–no trials, yielding an interference effect of 14.8%. This magnitude of interference is similar to the effects previously observed for two other patients with I.G.'s neuroanatomical profile who also performed this task (Hamilton & Martin, 2005; Thompson-Schill et al., 2002; though Hamilton and Martin's patient, M.L., exhibited an even larger effect). By contrast, the error rate interference magnitudes for patients GU412 and NN454 were 12.8% and 6.4%, respectively, both of which fell within the normal range of errors.[4] However, we note that although GU412's interference effect for errors was not outside the normal range by this measure (i.e., by previously established measures; see Hamilton & Martin, 2005; Thompson-Schill et al., 2002), this patient's effect was only slightly smaller than I.G.'s (12.8% vs. 14.8%). It is therefore unlikely that the interference effect for these two patients is reliably different.

We also employed an additional statistical method that is specifically devised to account for the small size of a control group, against which a single case is to be compared (e.g., Crawford & Garthwaite, 2006; Crawford & Howell, 1998). In particular, this test is an adapted *t* test, which uses the following formula:

$$t = (x^* - \bar{x})/(s\sqrt{(n + 1/n)}$$

Here, x^* is the patient's score; $\bar{x}$ and s are the mean and standard deviation, respectively, of the control group's scores; and n is the size of the control sample. "If the t value obtained from this test exceeds the one-tailed 5% critical value for t on (n − 1) df then it can be concluded that the patient's score is sufficiently low to reject the null hypothesis that it is an observation from the scores of the control population, and the patient is considered to exhibit an impairment on the task in question" (Crawford & Garthwaite, 2006, p. 878). According to this test, I.G.'s interference effect for errors is marginally significant, $t(5) = 1.57$; $p < .08$. The effect sizes for the other two frontal patients do not reach statistical reliability, even marginally: G.U., $t(5) = 1.14$; $p > .15$; NN, $t(5) = -0.24$; $p > .4$. Again however, it may be unlikely that G.U.'s effect size is different from I.G.'s. Taken alongside the 1.64 measure, I.G. shows the greatest interference effect for errors, and this pattern puts him in the same class of deficit as patient R.C. (Thompson-Schill et al., 2002).

Discussion of Experiment 1

Consistent with previous reports, I.G. demonstrated a significantly elevated interference effect for errors (>1.64 *SD*s beyond the control group's mean) that patterns with earlier effects for patients with similar neuroanatomical profiles (patients M.L.—though M.L. showed an even larger effect than I.G.—and R.C.; Hamilton & Martin, 2005; Thompson-Schill et al., 2002). Crucially, other patients in the current study, whose lesions spared BA 44/45, did not show this effect, nor did any of the healthy control participants they were compared to (Thompson-Schill et al., 2002). Nevertheless, GU412's interference effect for errors did trend toward I.G.'s and thus may also be considered impaired, even though her

---

[4] It should be noted that patient NN454 committed a large proportion of errors in both the recent–no and the nonrecent–no conditions (Table 3). The critical assessment of conflict resolution performance, however, is made on the basis of a comparison between these two conditions (recent–no minus nonrecent–no), and not by examining error rates from each condition alone (see Hamilton & Martin, 2005; Thompson-Schill et al., 2002; inter alia). As such, patient NN454's interference effect is within the normal range, despite her lower accuracy rate overall. Her performance could signal a general verbal working memory impairment that is not restricted to conflict resolution and cognitive control. Nevertheless, the lack of interference effect for this particular patient should be viewed with caution given the higher error rates across conditions relative to other participants.

performance was not significantly different from that of healthy controls across both statistical measures employed. As can be seen in Figure 1C, however (and as verified by a board certified neurologist), GU412's lesion includes substantially less involvement of left frontal operculum and pars triangularis. Thus, while non-LIFG patients may also express inflated effects on occasion as a result of their damage (and in some cases wider damage), they are decidedly not expected to show the entire constellation of impairments across the remaining tasks reported below. Only I.G., whose damage is restricted to LIFG, is expected to exhibit the full conflict resolution deficit pattern.

Demonstrating the proactive interference replication in I.G. is important because this interference task has been widely used in both previous imaging and patient work and has implicated an important and necessary role for LIFG in resolving competition that is brought about by incompatible representations in working memory. That is, the lingering familiarity of a probe letter that appeared in the previous memory set, but not the current memory set, temporarily impedes on how that probe should be characterized in memory: The familiarity suggests a "yes" response, but its absence in the current set of items imposes a "no" response, requiring the need to override the reliable familiarity cue. I.G.'s significantly increased error rate for these interference conditions, relative to noninterference conditions, shows an inflated failure to resolve familiarity-based competing characterizations of the probe representation.

By contrast, an exaggerated interference effect for reaction time was not observed for patient I.G.. We believe, however, that despite response times within normal limits, I.G.'s inability to override familiar representations remains evident in his error data. That is, the presence of a significant interference effect in the accuracy data clearly suggests that I.G.'s ability to override familiar representations is impaired under high- versus low-conflict conditions.

In what follows, we assess the role of LIFG in representational conflict resolution across a variety of language production and comprehension tasks. As outlined in the introduction to this paper, we hypothesize that one fruitful way to explain the linguistic impairments exhibited by patients with restricted LIFG damage may be to characterize their deficits in domain-general terms, specifically within a conflict resolution framework.

EXPERIMENT 2: RESOLVING CONFLICT IN LANGUAGE PRODUCTION—PICTURE NAMING

In a previous neuroimaging study using a picture-naming task (Kan & Thompson-Schill, 2004a), increased LIFG activity was found in healthy participants when naming line drawings of objects with multiple possible and therefore competing name alternatives (e.g., couch, sofa, loveseat), compared to naming of drawings of objects that have a single, predominant name (e.g., apple). The assumption underlying the source of competition in this confrontation naming task hinges upon the notion of "name agreement": The degree of competition related to picture naming varies depending on how many names are likely to come to mind for any given object. That is, competition is assumed to be stronger when multiple names apply to a single object (low name agreement; e.g., couch/sofa/loveseat) than when a unique, reliable name unfailingly suggests itself for another object (high name agreement; e.g., apple). Here, and in previous studies (Kan, Kable, Van Scoyoc, Chatterjee, & Thompson-Schill, 2006; Kan & Thompson-Schill, 2004a), we used the variability in name agreement across subjects as a proxy measure of the degree of lexical competition within individuals.

Parallel neuroimaging and patient results have also been reported using a similar picture-naming task involving a conflict resolution manipulation that does not centre on the name agreement variable per se, but rather on a manipulation of semantic interference (Schnur et al., 2009). Thus, LIFG-mediated conflict resolution mechanisms get deployed during picture-naming tasks under a variety of conditions involving competition, and the putative role that the LIFG plays in completing such tasks involves conflict resolution at the broadest level.

Such boosted LIFG recruitment under conditions of increased competition during picture naming (e.g., low name agreement; Kan & Thompson-Schill, 2004a) may suggest that this region mediates the resolution of competing representational alternatives and, further, may provide insight into understanding more precisely the nature of naming deficits observed in some patients with restricted LIFG damage. However, functional imaging results alone cannot certify that specific regions of cortex are necessary for a particular cognitive process; establishing such a dependent functional–anatomical relationship needs substantiation from lesion-deficit analyses of patients who have lost the ability to perform certain cognitive functions. Thus, specific predictions can be made for a patient with a circumscribed lesion to LIFG if this area is necessary for resolving among competing representations—in this case, of conceptual or semantic representations. In particular, such a patient ought to be expected to show an impairment in naming pictures associated with high competition items (e.g., couch) compared to pictures associated with low competition items (e.g., apple; for a similar experiment with related predictions, see Schnur et al., 2009). Thus, I.G. participated in a picture-naming task modelled after Kan & Thompson-Schill (2004a), which manipulated representational competition on the basis of name agreement.

Method

Participants

In addition to patient I.G., two groups of individuals also participated in this task: a left frontal group without LIFG damage and a healthy control group. I.G. participated in this task during a separate testing session from Experiment 1.

Left frontal patients (n = 2). The left frontal group consisted of 2 right-handed patients with left frontal damage, and MRI scans revealed, crucially, that their lesions spared LIFG. Patient GU412 was 40 years old at the time of testing with 12 years of formal education. This patient also participated in Experiment 1 but, like I.G., did so during a separate testing session. Patient KG215 was 51 years old at testing with 14 years of formal education and an AMNART score of 112 (see Figure 1B for lesion location).

Healthy control group. The healthy control group comprised 7 right-handed individuals who were matched to I.G. for age ($M = 62.3$, $SD = 13.9$) and years of formal education ($M = 16.9$, $SD = 1.6$). These participants had no history of neurological disease or psychiatric disorder, nor did they have any history of substance abuse.

All participants provided informed consent and were compensated $15 per hour.

Materials

Line drawings used in this experiment were identical to those used in Kan and Thompson-Schill, Experiment 1 (2004a). A total of 192 black-and-white line drawings were divided into two sets—high name agreement and low name agreement—on the basis of name agreement norms reported by Snodgrass and Vanderwart (1980). Mean name agreement for high-agreement (low-competition) items was 98% (range: 95–100%), and mean name agreement for low-agreement (high-competition) items was 71% (range: 33–86%). An unpaired t test indicated a significant difference in name agreement between the two sets of items, $t(190) = 20.37$, $p < .001$. Furthermore, items in the two sets were matched for mean familiarity (M for high agreement = 3.41; M for low agreement = 3.15); $t(190) = 1.88$; $p > .05$, and on median Kucera–Francis (KF) written frequency for picture names (high agreement $M = 10.50$; low agreement $M = 10.00$); $t(173) = 0.39$, $p > .05$. KF written frequencies were not available for 17 of the low-agreement items.

Procedure

Drawings were presented using PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993).

Participants viewed the series of drawings one at a time with self-paced presentation and named each object aloud as they were seen. Pictures from the two agreement conditions were randomly intermixed. Each picture remained on the screen until a response was made.

Results

Participants' responses were compared to names collected in a separate norming study ($n = 90$) and were considered correct if they matched one of the previously collected responses. Naming accuracy in each condition was calculated for each participant, and these data are summarized in Table 4. Healthy control participants were significantly more accurate at naming high-name-agreement items than low-name-agreement items (Wilcoxon Signed Ranks Test, $Z = 2.4$, $p < .02$). All 3 patients (i.e., I.G. plus the left frontal control group) performed within the normal range on the high-agreement trials (i.e., within 1.64 standard deviations of the healthy control group's mean). For low-agreement trials, the 2 non-LIFG frontal patients performed within the normal range whereas patient I.G. was significantly impaired at generating a response (in fact, >3 *SD*s below normal performance).[5] The effect of name agreement was also evaluated as a function of difference score in naming performance between high-agreement and low-agreement conditions. Whereas the two frontal patients performed within normal range ($Z = 0.76$ and $Z = 0.88$ for G.U. and K.G., respectively), I.G.'s performance was well outside the normal range ($Z = 3.46$).[6]

Table 4. *Picture-naming accuracy, Experiment 2*

Participants	*High name agreement (% accuracy)*	*z* score	*Low name agreement (% accuracy)*	*z* score
Controls	94.9 (1.2)	n/a	91.7 (3.0)	n/a
Patient IG363	93	1.5	82	3.2*
Patient GU412	93	1.5	89	0.9
Patient KG215	94	0.7	92	−0.1

Note: Standard deviations in parentheses.
$^*p < .05$.

Discussion of Experiment 2

In the picture-naming task, patient I.G.—the one patient with a restricted lesion to LIFG—was the only patient to demonstrate a significant inability to resolve competition due to multiple alternative response options. Under high-name-agreement conditions, when competition among alternatives is arguably lower, I.G.'s performance on picture naming was comparable to that of healthy controls. High-name-agreement items (e.g., apple) were assumed to limit the degree of representational competition because only one name is reliably evoked. I.G. performed at 93% accuracy for high-name-agreement items, which was well within the normal range of performance. On the other hand, when name agreement was low for a

---

[5] The types of incorrect response that I.G. made under low-name-agreement conditions were mixed. Most errors involved semantic substitutions. On occasion, he failed to make any response at all before deciding to move on to the next trial (or he said "I don't know"). On other error trials, I.G. generated a response that was not among one of the names given by healthy controls and was therefore labelled incorrect. Among these incorrect responses were trials on which he used a category name (e.g., furniture) rather than the specific category member that was presented (e.g., couch). An error analysis would be interesting in order to tease apart which stimuli in the high-competition condition gave rise to a certain kind of response. However, while his error rate was outside the normal range, I.G.'s overall number of errors was not high enough to permit this kind of analysis (he was over 80% correct in the low-agreement condition).

[6] Given that the comparison in familiarity ratings between high-agreement items and low-agreement items approaches significance ($p = .06$), one might wonder whether the observed effect in fact reflects a consequence of the name agreement manipulation. To better match familiarity, we removed two items from the high-agreement condition ("sun" and "key") and two items from the low-agreement condition ("spinning wheel" and "ostrich"). With this new set of items, familiarity ratings are now better matched ($p = .13$); KF frequency remains matched ($p = .72$), and name agreement remains significantly different ($p < .001$). Reanalysis of the data revealed the exact same pattern, suggesting that the finding was not due to a marginal significant difference in familiarity ratings.

picture (e.g., couch)—that is, when several competing alternative names typically come to mind—I.G. was accurate only 82% of the time, significantly below the normal mean. Other left frontal patients, by contrast, whose damage spared LIFG, showed no evidence of impairment under conditions of high competition (low name agreement), performing well within the normal range of healthy participants. Taken together, this pattern of findings implicates LIFG specifically as a region of cortex that is necessary for the resolution of competing representations in semantic/conceptual memory.

In the current experiment, we derived our name agreement measure based on across-subjects variability in picture naming and used that as a proxy measure of the degree of lexical competition within subjects. One might question whether the two types of variability reflect the same underlying representations. We reason that the two types of variability are comparable. According to Levelt et al.'s (1991) two-stage model of picture naming, an input concept is automatically generated upon perceiving an object, and after perceptual analyses of the visual input, a set of semantically driven candidate lemmas become activated. A selection process then follows, from which only one of these semantically related alternatives will survive, and this is the only item that is encoded phonologically and ultimately articulated as a response (see Fraisse, 1969, for a similar account). As such, although an individual may consistently choose to use one name to identify an object (e.g., couch), so long as the other names are represented in association with that same object (e.g., sofa, loveseat), conflict will arise.

The current findings corroborate the results of Kan and Thompson-Schill (2004a), which observed increased LIFG activation in an fMRI study under these same conditions of competition. Moreover, the current experiment extends that finding: Restricted damage to LIFG produces a loss of function related to the ability to resolve representational competition during a confrontation naming task; thus, not only is LIFG involved in this process, it is necessary for it to be carried out successfully.

The results from this picture-naming experiment begin to reveal a productive way of how researchers and clinicians might characterize the nature of the nonfluent language impairments demonstrated by patients with focal LIFG damage. Patient I.G. is not aphasic, again, in the classic Broca's sense despite the restricted quality of his lesion site to LIFG and Broca's area (see Dronkers et al., 2004). That is, instead of having language production impairments under circumstances related primarily to grammar or syntax, I.G. appears to have such deficits under only limited conditions—namely, those that increase the level of competition among the linguistic representations, in this case of nameable objects, that he is attempting to produce (see Robinson et al., 1998, 2005). In other words, describing I.G.'s production impairments within a conflict resolution framework has thus far proven to be useful: During confrontation naming of pictured objects, he fails only when there are high conflict demands; otherwise, his production is relatively normal.

This finding raises the question of whether I.G.'s deficit is specific to resolving competition during only confrontation naming, or whether his impairment extends more generally to other language production tasks of fluency as well. We therefore tested in Experiment 3 the generality of his production impairment within the scope of our conflict resolution hypothesis. This experiment examined his performance on a task that manipulates the effects of competition on generative production and verbal fluency.

EXPERIMENT 3: RESOLVING CONFLICT IN LANGUAGE PRODUCTION—VERBAL FLUENCY

Verbal fluency tasks instruct participants to generate, for instance, as many exemplars of a category (e.g., animals) as possible within a restricted time frame and have been used to evaluate word retrieval processes in neurologically impaired populations. In addition to assessing a patient's lexical selection abilities though, verbal fluency tasks can also be employed to measure cognitive control processes; more precisely, they can be used to measure the effects of cognitive control abilities on language production. One reason is that participants do not typically produce exemplars arbitrarily; they tend to produce clusters of semantically related items (e.g., farm animals: horse, cow, pig) and then "switch" to new clusters of items (e.g., pets: cat, dog, hamster; e.g., Troyer, Moscovitch, & Winocur, 1997). Hirshorn and Thompson-Schill (2006) have argued that switching between semantic subcategories (e.g., farm animals and pets) requires cognitive control to navigate from one semantic space to another: Lexical selection must be guided toward weakly activated (i.e., not-yet-named) representations and away from active (but already named and therefore incompatible) clusters. Notably, the authors showed that under such switching conditions, LIFG activity increases (Hirshorn & Thompson-Schill, 2006).

In addition, competition effects during semantic fluency might be expected to vary as a function of the size of the semantic category being retrieved from memory; smaller categories, for instance, should be more constrained by nature and thus have fewer competitors. Although it is difficult to empirically measure category size per se, one way to avoid the problem is to develop pairs of stimuli in which one item from each pair is a subordinate category (e.g., farm animals) of the other item (e.g., animals) in that pair (e.g., Randolph, Braun, Goldberg, & Chase, 1993). Thus, although the absolute category size—and consequent degree of competition—may vary from category to category, the relative category size will be inherently smaller for the subordinate categories than for the superordinate ones.

This paradigm is adopted in Experiment 3 and is modelled after a study in which category fluency was measured in patients with Alzheimer's, Parkinson's, or Huntington's disease and in one nonaphasic patient with bilateral frontal lobe lesions (Randolph et al., 1993). Critically, the patient with the frontal lesions (along with patients with Parkinson's disease or Huntington's disease) was impaired at generating exemplars of the superordinate category "animals", but was normal at generating exemplars of the subordinate category "farm animals". In the current experiment, we test whether the specificity of this impairment depends on focal damage to LIFG. Specifically, we predict that controls will benefit from the larger category size and produce more exemplars in the superordinate categories than in the subordinate categories. In contrast, I.G. is not expected to experience such a benefit because the increased number of available exemplars will actually increase competition. As such, the increased competition is expected to lead to an impairment in exemplar generation in I.G.

Method

Participants

In addition to I.G., the same patients as those who completed the picture-naming experiment (Experiment 2) also participated in this experiment. The two experiments took place on separate days. The healthy control group in this experiment was different from the picture-naming study and comprised 8 right-handed individuals who were matched to I.G. for age ($M = 61.3$, $SD = 13.2$) and years of formal education ($M = 16.8$, $SD = 1.5$).

Materials

A total of 14 categories and their corresponding subcategories were used: (a) U.S. states/U.S. states in the Eastern Time Zone; (b) movies/action movies; (c) fruits/citrus fruits; (d) cars/foreign cars; (e) hobbies/arts and crafts; (f) clothing/winter clothing; (g) animals/farm animals; (h) world and U.S. cities/U.S. cities; (i) furniture/bedroom furniture; (j) famous people/actors and actresses; (k) U.S. presidents/20th-century U.S. presidents; (l) appliances/kitchen appliances; (m) magazines/women's magazines; and (n) vegetables/leafy vegetables.

Procedure

All participants were tested in two sessions spaced at least 6 weeks apart. On each trial, participants were presented with the category (or subcategory) name and were given 45 seconds to generate as many responses as possible. All participants were presented with 14 trials (seven categories and seven subcategories) in each session, and a category and its corresponding subcategory were never presented in the same session.

Participants' responses were recorded and tallied. To ensure that each subcategory was truly a subset of the broader category, each healthy participant's responses to each subcategory were compared to those that were generated for the corresponding superordinate category. Category and subcategory pairs were removed from the analysis if either of the following conditions were true: (a) if the number of items generated in the subcategory (e.g., farm animals) was greater than the number of items generated in the broader category (e.g., animals); and (b) if there was no overlap between items generated in the subcategory and those generated in the superordinate category. Both of these criteria were adopted to ensure the appropriateness of our category and subcategory designations. By definition, a subcategory should contain both fewer items than its superordinate category and items that overlap with members of its corresponding superordinate category pair. This procedure eliminated seven pairs, but the following pairs remained: U.S. states/U.S. states in the Eastern Time Zone; fruits/citrus fruits; cars/foreign cars; clothing/winter clothing; animals/farm animals; U.S. presidents/20th-century U.S. presidents; and vegetables/leafy vegetables.

Results

On average, healthy control participants generated significantly more items in response to the superordinate categories (e.g., animals) than in response to the subcategories (Wilcoxon Signed Ranks Test, $Z = 2.5$, $p < .02$). As described earlier, by definition, more exemplars ought to be generated for the superordinate categories than the subcategories. In other words, this finding confirms that our manipulation was successful.

To assess the effect of category size on verbal fluency, a difference score between the number of items generated in the superordinate category and the number of items produced in the subcategory was calculated for each superordinate and subordinate category pair for each participant. Mean difference scores across all pairs were calculated. As summarized in Table 5, healthy control participants generated on average 6.8 more items in response to the superordinate categories (e.g., animals) than in response to the subcategories. Compared to controls, all patients' difference scores fell outside the normal range (i.e., >1.64 *SD*s above the normal mean). However, I.G., at 2.5 standard deviations away from the normal mean, showed the greatest impairment on this task.

Table 5. *Verbal fluency performance, Experiment 3*

Participants	*Category*	*Subcategory*	*Difference score*	*z* score
Controls	16.2 (3.6)	9.4 (2.6)	6.8 (1.5)	n/a
Patient IG363	10.1	7.1	3.0	2.5*
Patient GU412	8.7	4.6	4.1	1.8
Patient KG215	11.4	7.4	4.0	1.9

Note: Performance is number of items. Difference score is number of items produced for category minus number of items produced for subcategory. Standard deviations in parentheses.

$^*p < .05$.

Discussion of Experiment 3

Age-matched healthy control participants generated significantly more exemplars of superordinate categories (e.g., animals) than of subordinate categories (e.g., farm animals; see Table 5), which reflects their ability to effectively use the larger category set size to produce a greater number of exemplars. Both non-LIFG patients also showed this pattern. However, among all the participants, the difference in the number of exemplars that I.G. produced between the two conditions was the smallest, suggesting that I.G. was least effiective at using the larger category set size to produce more responses. Whereas healthy control participants generated more exemplars when there was a larger pool of possible responses (i.e., in the superordinate categories) than when there was a smaller pool of possible responses (i.e., in the subordinate categories), I.G.'s performance was by contrast not facilitated by larger set size, presumably due to the increased competition in this condition. Indeed, he actually produced more "farm animals" (10) than "animals" (9). In fact, of the 11 individuals who took part in this task, I.G. was the only individual who showed this pattern on any trial. These data support the claim that LIFG is necessary to guide semantic retrieval when stimulus cues do not sufficiently constrain the response.

Two left frontal patients showed a similar pattern of not being facilitated by larger set size, though less so than I.G. These same control patients performed like healthy adults on Experiments 1 and 2 (proactive interference and picture naming); thus I.G. is the only patient to show the entire constellation of impairments across the three experiments thus far.

In conjunction with the findings from Experiment 2 (picture-naming task), we believe that it is appropriate to interpret I.G.'s relatively nonfluent production skills as being tied to a general conflict resolution impairment, as evidenced by an inability to resolve proactive interference in a working memory paradigm (Experiment 1). Generating a response in the superordinate category condition (e.g., animals) is akin to an underdetermined response conflict task because there are more response options available for production, none of which is more compelling than another. Thus, "chicken", "rabbit", and "cow" are all equally good candidates. Although generating a response in the subcategory condition (e.g., farm animals) also requires resolving competition among a number of equally good candidates, the number of available responses is inherently smaller than that in the superordinate category condition; thus competition demands are lower. A general conflict resolution impairment has been reported across at least two different case studies of patients with similarly confined lesions to LIFG (Hamilton & Martin, 2005; Thompson-Schill et al., 2002); the current data, however, extend previous findings and test the role of LIFG in language production (for similar results, see Randolph et al., 1993; Robinson et al., 1998, 2005).

In the introduction to this paper, we reviewed how LIFG patients have rather transient and limited language disorders and that, as a result, the nature of their deficits has been difficult to characterize concerning precisely under what circumstances they arise. Luria (1973) used the term "dynamic aphasia" to refer to a pattern of language impairments suffered by patients with frontal damage whose symptoms increased under certain conditions but diminished under others (see also Robinson et al., 1998). In our view, such a fluctuating pattern may be best characterized in broad cognitive terms—namely, that the deficits have to do with the occurrence and resolution of conflict during language-related tasks. Indeed, the results from Experiments 2 and 3 suggest that these patients' language skills, as far as production is concerned, may become temporarily impaired because of generally inadequate representational conflict resolution abilities. The two production tasks examined thus far clearly fall under the general rubric of underdetermined response conflict (Botvinick et al., 2001), because in both tasks there are cases of when multiple response representations are simultaneously generated, none of which is more relevant or decisive than another. Under such conditions, I.G. has difficulty settling on a single alternative, which ultimately has clear and important consequences for producing language, such as an inability to name an object or an inability to produce a normal number of exemplars given a particular category cue. By contrast, when the conflict is comparatively minimal, I.G.'s fluency is well within normal limits.

In the experiment that follows, we predict I.G.'s sentence processing and comprehension abilities within the same theoretical space: that any departure from a healthy, normal performance ought to be modulated expressly by comprehension environments in which representational conflict is at its relative highest. In this task, participants are asked to carry out spoken instructions to move toy objects around a table in front of them.

EXPERIMENT 4: RESOLVING CONFLICT IN SPOKEN LANGUAGE COMPREHENSION—SYNTACTIC AMBIGUITY RESOLUTION

In this experiment, we turn our attention to the real-time nature of sentence processing and comprehension in patients, using a technique that allows us to get a moment-by-moment glimpse into a listener's online interpretation of linguistic input as it reaches the ear. Evidence from sentence-processing studies suggests that readers and listeners achieve comprehension in real time as the words and phrases are encountered, rather than waiting until the end of an utterance to start assigning an interpretation to what they have just encountered (e.g., Altmann & Kamide, 1999; Ferreira & Clifton, 1986; Rayner, Carlson, & Frazier, 1983; Trueswell, Tanenhaus, & Garnsey, 1994; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). One important consequence of real-time processing is that it comes at the cost of having to deal with temporary ambiguity. Consider for instance Sentence 1a:

1a. Put the apple on the napkin into the box. (Temporarily ambiguous)
1b. Put the apple that's on the napkin into the box. (Unambiguous)

As the information in Sentence 1a is unfolding, the listener begins to develop a characterization of the input that is consistent with his or her prior lexical knowledge of the verb *put*—namely, that it is very likely to appear with a goal, probably in the form of a prepositional phrase (e.g., *Put the apple* ***on the napkin***.). However, such probabilistic evidence is insufficient to be an exact guide to how the rest of the sentence will unfold, as illustrated in Sentence 1a, where a second prepositional phrase, *into the box*, is encountered and unambiguously signals the intended goal of the verb *put*. Upon encountering *into the box*, listeners must override their initial interpretation of *on the napkin* as a goal in favour of a new interpretation that allows this phrase to be a modifier—that is, specifying which apple to move (the one on the napkin). It is precisely this kind of syntactic ambiguity resolution scenario, which is hypothesized to be akin to other nonsyntactic response override tasks (e.g., Stroop), that we exploit here to assess the specificity of I.G.'s comprehension abilities (see Novick et al., 2005, for theoretical development). This is further motivated by a recent brain imaging finding, which reveals that LIFG activation colocalizes within healthy participants who complete both an ambiguity resolution task and a Stroop task while undergoing fMRI (January et al., 2009).

We predict that only when I.G. has to override a well-supported developing interpretation will he fail to arrive at the intended analysis of a sentence. In the example above, I.G. would have to reject the highly supported unfolding interpretation (i.e., the goal interpretation of *on the napkin*) and promote an initially disfavoured alternative (i.e., the modifier interpretation of *on the napkin*). However, his impaired conflict resolution ability is predicted to hinder revision of an early syntactic commitment. By contrast, under syntactically unambiguous conditions, in which there is little or no competition among representations of sentence meaning—as in 1b, which is otherwise identical to 1a—I.G.'s comprehension abilities should be unaffected and will not differ from the normal pattern of interpretation. Indeed, there are fMRI studies of healthy adults that support and motivate this notion, and which have found increased activation in LIFG during trials containing various sorts of syntactic ambiguities (Fiebach, Vos, & Friederici, 2004; January et al., 2009; Mason et al., 2003; Ye & Zhou, 2009; see Novick et al., 2005, for a review). However, to date, only a few real-time sentence parsing studies have involved patients (e.g., Dickey, Choy, & Thompson, 2007; Dickey & Thompson, 2009; Hagoort, Wassenaar, & Brown, 2003; Yee, Blumstein, & Sedivy, 2008), and these studies have focused expressly on patients with (agrammatic) aphasia. But patient samples defined by behavioural profiles tend to have variable and extensive cortical and subcortical damage (Dronkers et al., 2004). In contrast, the current experiment aims to predict and characterize the type of language comprehension impairments that a nonagrammatic patient with restricted LIFG damage will demonstrate. It is also worth noting that the kind of sentence comprehension problem predicted here differs from the kinds of syntax-oriented sentence comprehension problems that are noted in cases of agrammatic aphasia, such as patients' difficulty with function words or inflected morphology.

Real-time sentence processing and the visual-world paradigm

Much of what is known about healthy listeners' ability to resolve temporary ambiguity during spoken language has come from studies using the so-called "visual-world paradigm", which records listeners' eye movements as they carry out spoken instructions to move toy objects around a table (e.g., Tanenhaus et al., 1995; Trueswell, Sekerina, Hill, & Logrip, 1999). Eye movements are closely time locked to the speech stream (Cooper, 1974), landing on a spoken referent within only a few hundred milliseconds of hearing enough phonemic information to distinguish it from all other candidate referents in the environment (see Allopenna, Magnuson, & Tanenhaus, 1998; inter alia). Thus, eye fixations are construed as indicating a listener's current focus of interpretation, and researchers can therefore get a fine-grained, moment-by-moment glimpse into a listener's comprehension process as it is unfolding in real time. Furthermore, it is also possible to obtain a video record of how the listener eventually carries out an instruction offline, which reveals the ultimate interpretation that an individual has assigned to the utterance.

Several prior eye movement and ambiguity resolution studies have used the sentence materials found in Sentence 1a and Sentence 1b (e.g., Novick, Thompson-Schill, & Trueswell, 2008; Spivey, Tanenhaus, Eberhard, & Sedivy, 2002; Tanenhaus et al., 1995; Trueswell et al., 1999). For such sentences, the listener's immediate visual environment has featured the following objects: an apple sitting on a napkin (the target referent), a banana in a bowl (the competitor referent), an empty napkin (an incorrect goal), and a box (the correct goal). Upon hearing the word *napkin* in *Put the apple on the napkin . . .*, healthy listeners show an increased tendency to look over at the incorrect goal (a potential goal for the *putting* action) as compared to syntactically unambiguous sentences: *Put the apple* ***that's*** *on the napkin. . . .* This suggests that listeners are briefly considering a goal interpretation of *on the napkin*. Upon hearing *into the box*, listeners tend to show some general confusion, looking around the visual scene more when hearing ambiguous sentences than when hearing unambiguous sentences. Yet they ultimately arrive at the correct interpretation, moving the apple into the box in most cases.

However, even under ambiguous conditions, healthy adult listeners' consideration of the incorrect goal (e.g., the empty napkin) can be attenuated by context (Novick et al., 2008; Spivey et al., 2002; Tanenhaus et al., 1995; Trueswell et al., 1999). When two apples are present, one on a napkin and one in a bowl (replacing the banana), consideration of the incorrect goal and listeners' general confusion are greatly reduced. This is because reference to an apple in a two-apple scene requires a speaker to offer further linguistic or nonlinguistic specification—for example, in the form of noun phrase modification, describing the apple in question as, for instance, *the apple on the napkin*. Healthy adult listeners appear to be sensitive to this fact and use it to reduce consideration of the goal interpretation and promote consideration of the modifier interpretation; listeners rarely look at the incorrect goal in these two-apple scenes. Nevertheless, Novick et al. (2008) have shown that temporary consideration of the goal interpretation is not completely blocked by these referentially supportive contexts. For example, a small but reliable increase in errors occurs for healthy adults in two-apple contexts when responding to temporarily ambiguous instructions (8% errors) as compared to unambiguous ones (0%). Five-year-old children, by contrast, appear not to be as sensitive as adults to such supportive contexts but instead rely consistently, as they do in one-apple scenes, on the lexical bias of the verb *put*; they commit errors involving the incorrect goal on over 60% of ambiguous two-apple trials (Trueswell et al., 1999; Weighall, 2008). In sum, healthy adult listeners temporarily consider the goal interpretation because of strong syntactic and semantic evidence associated with the verb *put*, but rapidly override this erroneous interpretation based on contextual evidence and later arriving linguistic information. Young children do not use these sources of evidence and fail to override the incorrect analysis of *on the napkin* as goal regardless of referential context. We believe that this shows that children, who have poor cognitive control, rely on the strongest probabilistic evidence—that *put* requires a goal—despite what the weaker contextual evidence supports. As such, we predict that I.G. should demonstrate a similar pattern (see Novick et al., 2005).

We adopt the visual-world eye-tracking paradigm to test the specificity of the conflict resolution hypothesis to I.G.'s comprehension abilities. This approach permits real-time insight into his developing interpretations in the face of competing syntactic representations (e.g., well-supported goal vs. less well-supported modifier), and it also allows an analysis of his offline hand actions, which indicate his final interpretation of the sentence. We hypothesize that the ambiguity in (1a), akin to a nonparsing response override task (Novick et al., 2005; Ye & Zhou, 2008), increases the representational conflict among alternative interpretations among sentence meaning, which should therefore result in I.G.'s selective inability to arrive at the intended analysis of the sentence. In particular, this should be exhibited by a high proportion of fixations on the incorrect goal (e.g., the empty napkin) in both one- and two-apple scenes, because in either condition the strong lexical evidence associated with the verb *put*—that it needs a goal—will have to be overridden in light of later disambiguating evidence (e.g., *into the box*) and/or weaker contextual evidence such as the presence of two apples (Novick et al., 2008; Novick et al., 2005). In addition, we also predict that I.G. will in the end be unable to disengage from pursuing this highly reliable goal response, which might be demonstrated by actions involving the incorrect goal when carrying out the instruction (e.g., the empty napkin; see Trueswell et al., 1999; Weighall, 2008). On the other hand, I.G.'s comprehension of syntactically unambiguous sentences (e.g., 1b) ought to be comparable to non-LIFG frontal patients' and healthy adults' comprehension patterns because these sentences effectively remove the representational competition between two alternatives of sentence meaning.

Method

Participants

In addition to patient I.G., two control groups also participated in this task: a left frontal group and a healthy control group. The left frontal patients were the same as those who participated in Experiment 1. The healthy control group comprised 3 right-handed individuals (2 female) who were matched to I.G. for age ($M = 63$ years, $SD = 2.1$) and education (16.1 years, $SD = 1.3$). None had any history of substance abuse or neurological or psychiatric disorders.

Procedure

The study procedure was similar to that of Novick et al. (2008; see also Spivey et al., 2002; Tanenhaus et al., 1995; Trueswell et al., 1999). Participants were told that they would hear and follow prerecorded instructions to manipulate stuffed animals and other toy objects around a table. Each participant sat in front of an inclined platform. At the centre of the platform was an aperture behind which a digital video camera was positioned to focus on the participant's face. In each quadrant of the platform there was a shelf on which one of several toy props could be placed (see Figure 2). A second camera, placed behind the participant, recorded hand actions and the locations of the props (i.e., where the props were initially set up and where the listener placed them after carrying out the instruction).

At the beginning of a trial, one experimenter laid out the props and introduced each of them (e.g., *This is a frog, a plate ...*). Prerecorded sound files were then played from a laptop connected to external speakers and to the video camera. On each trial, the participant was first told to look at a fixation point at the centre of the display. Then she or he was given two or three single sentence commands involving the props. The participant heard the first command, performed an action in response to it, and then heard the second command. If the participant wished to hear the sentence played a second time, the experimenters did so, but the eye movements were coded for only the first presentation of the sentence.

A total of 12 target toy stimuli were constructed and modelled after Trueswell et al. (1999) and were a subset of the stimuli used in Novick et al. (2008). The first sentence of each critical trial contained the verb *put* and appeared in one of two sentence types as shown in Sentence 1, shown here again (and adapted slightly) in Sentence 2:

2a. Put the frog on the napkin into the box. (Temporarily ambiguous)
2b. Put the frog that's on the napkin into the box. (Unambiguous)

In Sentence 2a, the prepositional phrase *on the napkin* is temporarily ambiguous between indicating either a destination/goal (i.e., where the frog should be put) or a restrictive modifier (i.e., indicating that the frog to be put somewhere is currently on a napkin). The inclusion of *that's* in Sentence 2b removes this temporary ambiguity and strictly imposes the modifier analysis of *on the napkin*. A female research assistant prerecorded all instructions. The prosody in these utterances was intended to be unbiased with respect to the two parsing alternatives: large prosodic breaks and pauses were avoided so as not to favour either the modifier or the goal interpretation.

The referential scene was also manipulated by changing critical features of the display configuration between target trials. One-referent Scenes (Figure 2a) contained, for example: a target animal (e.g., a toy frog sitting on a napkin); an incorrect goal (a second, unoccupied napkin); a correct goal (e.g., a box); and a competitor animal (e.g., a toy horse sitting in a bowl). These scenes should further encourage the goal interpretation of the phrase *on the napkin* in ambiguous sentences because modification of *the frog* would be referentially redundant due to the presence of only one frog. Two-referent scenes, by contrast, contained two toy referents of the same kind of animal—for example, two frogs (Figure 2b). These scenes were configured just like one-referent scenes

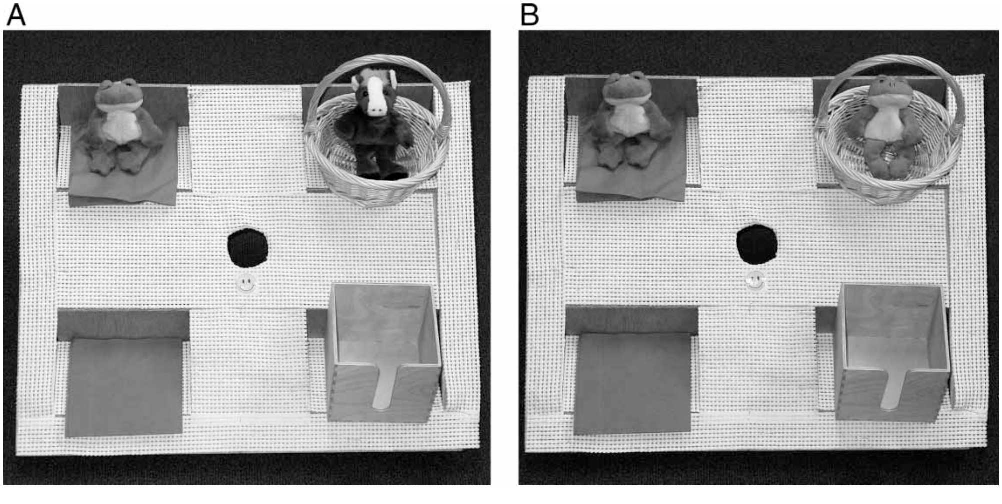


Figure 2. *Example scene types from Experiment 4: (A) One-referent scene (supports goal interpretation of* on the napkin*); (B) two-referent scene (supports modifier interpretation of* on the napkin*).*

except that a second toy frog replaced the competitor toy horse. (For any given participant, one or the other of these paired displays was used, but not both; i.e., no critical trials were repeated.) Two-referent displays should lend support to a noun phrase (NP) modifier interpretation of *on the napkin* because the definite NP *the frog* does not uniquely specify which frog is the intended referent. Thus, *on the napkin* is necessary to pick out which frog is meant.

Also like Novick et al. (2008), but unlike other previous studies using these visual and linguistic stimuli (e.g., Trueswell et al., 1999), the target and competitor animals in the current study were designed to differ spatially in terms of whether one was *on* something and the other was *in* something (e.g., a frog on a napkin vs. a horse/frog in a bowl). Half of all target trials had the target animal on some flat-surfaced platform and the competitor animal inside some container, whereas the opposite was true for the other half of the trials (e.g., *Put the horse in the bowl onto the plate*). This contrast in the spatial set-up permitted us to examine how quickly participants arrive at a modifier interpretation of *on the napkin* instead of the goal analysis in two-referent scenes. In particular, if the presence of two frogs is enough to allow participants to quickly override the goal analysis in favour of a modifier interpretation of *on the napkin* (i.e., with only little consideration of the incorrect goal), then the preposition itself (*on*) should allow listeners to distinguish between the two referents—the frog that's on a napkin and not the one that's in a basket (Figure 2b; see Chambers, Tanenhaus, & Magnuson, 2004; Novick et al., 2008). As mentioned above, Novick and colleagues (2008) found that healthy adults' consideration of the competitor in two-referent scenes (the other frog) was slightly yet reliably greater for ambiguous than for unambiguous items, suggesting brief difficulty overriding the goal interpretation and arriving at the modifier interpretation. Thus any patient with difficulty overriding this interpretation should show exaggerated consideration of the competitor and may even choose it as the object to act upon. Taken together, increased eye movements to, and hand actions involving, the frog in the *bowl*—alongside increased looks to the incorrect goal—would reflect difficulty arriving at the modifier analysis of *on the napkin*. That is, this would illustrate a failure to use the weaker contextual evidence to guide comprehension, but instead an explicit focus on the stronger lexical evidence encoded in *put* (that it requires a goal). Given a two-referent design more sensitive to distinguishing goal/modifier considerations, I.G. is predicted to exhibit a pattern showing very little commitment to the modifier, indexing an overall preference to go with the strongest evidence consistent with the early input.

Furthermore, we expect I.G. to show an inflated ambiguity effect in two-referent versus one-referent cases, reflected in higher error rates like young children (Trueswell et al., 1999; Weighall, 2008). For those without good cognitive control (five-year-olds and I.G.), listeners may completely adopt a goal interpretation, which they cannot revise regardless of context (one or two referents present). However, one-referent scenes might slightly mask the difficulty they are having arriving at the modifier interpretation because the only frog present in the scene is already on a napkin. That is, persisting with a goal interpretation of the first prepositional phrase (*on the napkin*) need not require moving the frog to the other napkin because this action appears to have already been carried out (it is already on a napkin), resulting in participants carrying out the second goal phrase (*into the box*). In the two-referent scenes, the presence of a second frog not on a napkin provides greater opportunity for the goal-parse to be manifested in the actions themselves, which is indeed what has been observed in young children. We thus anticipate a similar pattern from I.G.—that is, an asymmetry in performance across referential scene type, where one-referent scenes result in fewer errors than two-referent scenes.

Experimental design

Both ambiguity (temporarily ambiguous vs. unambiguous) and scene type (one-referent vs.

two-referent) factors were manipulated within participants. A presentation list was designed containing 12 critical trials pseudorandomly intermingled with 16 filler trials. Of the target items, eight were ambiguous (four within one-referent scenes and four within two-referent scenes), and four were unambiguous (two within one-referent scenes and two within two-referent scenes). Lists were run in both forward and reverse orders, alternating between participants. Unambiguous filler trials were also included; these displays looked similar to those of critical item displays (which were associated with ambiguous or unambiguous sentences) so that participants could not predict what sort of instruction they would hear. The second and third sentences of every trial (ambiguous, unambiguous, and filler trials) also served as built-in filler items as they were syntactically unambiguous and distracted from the experimental manipulations. These instructions, for instance, asked participants to *Now spin the duck around* or *Now make the other animal stand on its head*. Filler items also included unambiguous forms of the verb *put* and indicated its goal in various ways (e.g., *Put the X in the Y*,...*next to the Z*, etc.). Listeners also completed 5 practice trials before the start of the task in order to get familiar with the procedure; practice trials were similar to filler trials. After correctly completing practice items participants were asked if they were comfortable with the task and ready to proceed to the actual experiment. Critical trials (either ambiguous or unambiguous) never appeared as the first item in the experiment—only fillers did, so in effect participants had roughly 6 or 7 trials on the task before ever seeing a critical item. Thus, all participants were familiar and comfortable with the task before encountering experimental trials.

Coding

Following Snedeker and Trueswell (2004), gaze direction was coded from the video of the listener's face on a frame-accurate digital VCR (SONY DSR-30) with audio-lock so that direction of eye gaze could be determined with respect to the speech stream on a scale of every 33 ms. The trained coder documented the beginning of each target trial by logging the onset of the verb *put* and the listener's eye position at that moment. From that point on, frame-by-frame changes in gaze direction were recorded with the VCR's timestamp indicating when a new eye movement occurred and the coder noting the new direction of fixation. Fixation direction was coded as one of the following: on one of the platform's quadrants (e.g., upper left, lower right, etc.), in the centre of the platform display, or elsewhere (i.e., away from the experimental scene). "Track loss" was coded if the participant's eyes were closed or occluded—for example, by a reaching arm in front of the camera's lens. These trials were dropped from analyses if they accounted for more than 33% of the frames. A trial's offset was coded when a participant released an object after performing an action. The actual object on which a listener was fixating was later confirmed by the scene videotapes, which captured where in the scene each object was placed, so that *upper left* could be matched with *empty napkin* (i.e., the incorrect goal), for example. This procedure allowed the trained coder to be blind to the experimental conditions when logging each direction of gaze.

Participants' hand actions on critical trials were coded based on an inspection of the scene videotapes in order to evaluate the final interpretation that a patient assigned to a spoken instruction. An action was coded as correct if the target animal was moved directly to the correct goal without any involvement of the incorrect goal (or any other intervening steps). Actions were also coded as correct if the target animal was moved directly to the correct goal along with the modifying object that shared the target's quadrant (e.g., the napkin on which the frog was sitting; see Trueswell et al., 1999). Any actions involving the incorrect goal were coded as errors and can be categorized most frequently as "hopping" errors, in which the target animal, or competitor animal in two-referent scenes, was moved first to the incorrect goal (e.g., the empty napkin) and then to the correct goal (e.g., the box; see Trueswell et al., 1999).

Results

Eye movements and actions are reported separately below. In particular, we provide data for the patients' and control group's eye movements to the incorrect goal in both referential contexts and to the competitor animal in two-referent contexts. Recall from above that both of these measures tap listeners' early parsing commitments to interpretations that are guided by probabilistically reliable though currently unsupported syntactic and/or referential evidence (see Novick et al., 2008; Spivey et al., 2002; Tanenhaus et al., 1995; Trueswell et al., 1999).

The magnitude of individual ambiguity effects was computed as the difference in looking proportions to these objects between ambiguous and unambiguous conditions (ambiguous minus unambiguous). For offline hand actions in response to the sentences, we report whether any participant ultimately failed to override initial parsing commitments by using the incorrect goal and/or competitor animal (in two-referent scenes) to carry out an instruction.

Eye movements

An informal assessment of the videotapes containing listeners' eye movements revealed that the patients in this study were capable of examining their visual environments incrementally with respect to the temporally unfolding speech. In other words, objects in the display were often fixated soon after they were mentioned in the instructions, and, if an utterance could have temporarily referred to more than one object, patients' fixations would oscillate between the competing referents until a uniqueness point was specified in the speech stream (see Tanenhaus et al., 1995; Trueswell et al., 1999). Additionally, patients looked at objects just before reaching for them, which is akin to behaviours demonstrated by both young children and healthy adults (Ballard, Hayhoe, & Pelz, 1993; Tanenhaus et al., 1995; Trueswell et al., 1999). Thus, the way in which patients interacted with the "visual-world" set-up mirrored that of healthy populations. This brief inspection of the videotapes confirmed that no eye movement abnormalities were apparent with respect to the task, and, therefore, patients were unfettered by the eye-tracking paradigm employed here.

Looks to the incorrect goal. For each trial, the proportion of looks to the incorrect goal was analysed beginning from the onset of the word *napkin* until an action was completed. Participant means per condition were calculated excluding trials that contained more than 33% track loss (this accounted for approximately 13% of the data, evenly distributed across conditions). The magnitude of the ambiguity effect was computed from these conditional means, which was the difference in looking proportions to the incorrect goal between ambiguous and unambiguous conditions.

Table 6 reports each patient's mean looking proportions, including ambiguity effect sizes for

Table 6. *Mean proportion of looks to incorrect goal in ambiguous versus unambiguous conditions from the onset of* napkin, *also split by one- and two-referent scene types, Experiment 4*

	Both scene types			*One-referent*			*Two-referent*		
Participants	*Ambig*	*Unambig*	*Ambiguity effect*	*Ambig*	*Unambig*	*Ambiguity effect*	*Ambig*	*Unambig*	*Ambiguity effect*
Controls ($n = 3$)	.05 (.038)	.007 (.012)	.043 (.026)	.071 (.028)	0 (0)	.071 (.048)	.03 (.023)	.023 (.023)	.007 (.008)
Patient NN454	.068	.076	−.008	.016	0	.016	.052	.076	−.024
Patient GU412	.14	.007	.133*	.111	0	.111	.029	.007	.022
Patient IG363	.32	.017	.303**	.123	.017	.106	.197	0	.197***

Note: Ambig = ambiguous. Unambig = unambiguous. Ambiguity effect = ambiguous minus unambiguous. Standard deviations in parentheses.

*$p \leq .05$. **$p \leq .01$. ***$p \leq .005$.

each condition (the age-matched healthy controls were collapsed into a single group for comparison). Each patient's ambiguity effect size (ambiguous minus unambiguous) was compared to that of the healthy control group, first collapsing across referential conditions—that is, regardless of whether one or two frogs were present. A patient's early commitment to the goal interpretation was considered impaired if his or her looking proportion to the incorrect goal (ambiguous minus unambiguous) was more than 1.64 standard deviations above the healthy control group's mean (about the 95th percentile). As can be seen in Table 6, patients GU412 and I.G. both fit this criterion. The amount of time I.G. spent considering the incorrect goal upon hearing *napkin* was in fact 10 standard deviations above the healthy control group's mean fixation time to the same object (ambiguous minus unambiguous). In other words, by this measure, I.G. considered the incorrect goal (e.g., the empty napkin) for a significantly longer period of time upon hearing *Put the frog on the napkin* . . . under temporarily ambiguous conditions than under unambiguous conditions, suggesting that he was considering the goal analysis more than any other participant. Patient GU412 showed an exaggerated pattern too, though less so than I.G.; she scored approximately 3 standard deviations beyond the normal mean.

Furthermore, we employed Crawford's statistical method to account for the small size of the healthy control group (e.g., Crawford & Garthwaite, 2006; Crawford & Howell, 1998). According to this test, I.G.'s ambiguity effect (ambiguous minus unambiguous) for looks to the incorrect goal is reliably outside the normal range: $t(2) = 8.67, p < .01$. Using the same analysis, G.U.'s pattern is also inflated, $t(2) = 3.00, p = .05$, albeit less so than I.G.'s; N.N.'s pattern, by contrast, is not significant, $t(2) = -1.7, p > .1$.

We applied these same analyses to the ambiguity effect for each patient for both one- and two-referent scenes separately. In one-referent scenes, the healthy control group spent on average 7% more time looking at the incorrect goal under ambiguous conditions than under unambiguous conditions (Table 6). Such a large effect of ambiguity is consistent with findings from previous studies, as one-referent scenes further support the goal analysis of *on the napkin* (e.g., Tanenhaus et al., 1995; inter alia). The patients in the current study also displayed such ambiguity effects, but no patient's commitment to the incorrect goal, as evidenced by the proportion of looks to this object, was considered to be abnormal—every patient did initially consider the goal analysis upon hearing *napkin*; that is, they looked to the incorrect goal (e.g., the empty napkin)—but not more so than healthy controls (Table 6).

Under two-referent conditions, however, the healthy control group's ambiguity effect was much smaller on average (as expected) because they were presumably using contextual evidence (e.g., the presence of two frogs) to mitigate an early preference to consider the goal interpretation and instead arrive at the modifier interpretation of *on the napkin* (see Novick et al., 2008; Spivey et al., 2002; Tanenhaus et al., 1995; Trueswell et al., 1999). By contrast, an inspection of Table 6 reveals that only one patient, patient I.G., had difficulty using the contextual support to avoid large consideration of the incorrect goal, spending an abnormal amount of time looking at this object (e.g., the empty napkin). We note that the proportion of time that patient GU412 spent considering the incorrect goal in two-referent ambiguous conditions compared to unambiguous conditions was 1.87 standard deviations beyond the healthy control group's mean. However, employing Crawford's more sensitive test reveals no such impairment for G.U., $t(2) = 1.62, p > .1$. I.G., by contrast, clearly spent the most time considering the goal analysis of *on the napkin* in these cases according to both statistical tests; the time he spent looking at this object was actually more than 23 standard deviations outside the range of the healthy control group, which was highly significant: Crawford's test, $t(2) = 20.59; p < .005$. This suggests that even when two frogs were present, I.G. was still considering the goal analysis of *Put the frog on the napkin* . . . for a considerably longer period (within the same designated time window) than healthy controls and other left frontal patients. This pattern is highly suggestive of I.G.'s impaired ability to override the goal analysis due to the strong lexical evidence of *put*, despite other sources of information (e.g., the presence of two frogs), which makes a modifier analysis increase in likelihood. This predicted pattern of results is perfectly consistent with young children's two-referent parsing profile, which we address in greater detail later in the discussion.

Looks to the competitor animal. For each trial in two-referent conditions, the proportion of looks to the competitor animal (e.g., the frog in the basket in Figure 2b) was analysed beginning at the onset of the preposition *on* until an action was made. As outlined earlier, if patients, like healthy adults, can quickly override the goal analysis of *on the napkin* and arrive at the modifier analysis of this phrase in light of two-referent scenes, then the preposition *on* should help distinguish between the referents due to the spatial contrast in the setup: The target animal is *on* something, whereas the competitor animal is *in* something (see Novick et al., 2008). Thus, looks to the competitor animal at the preposition's onset are indicative of not interpreting *on the napkin* as a modifying phrase and that the listener is instead attempting to (incorrectly) satisfy the syntactic constraints of *put*: The competitor animal is currently in a basket—not on a napkin—and therefore must need to be put somewhere else in order to (incorrectly) satisfy the goal interpretation. This measure for such referential scene types are highly sensitive to analyses involving individuals' use of relatively weaker contextual evidence to override strong lexical constraints (*put*; *in* vs. *on*; see Novick et al., 2008).

Table 7 reports each patient's proportion of looks to the Competitor Animal in two-referent scenes.[7] As with all analyses reported throughout this paper (see also Thompson-Schill et al., 2002), a patient's effect size was considered impaired if found to be more than 1.64 standard deviations greater than the control group's mean. Crawford's test was also employed to account for a small control sample. Together, these analyses revealed only one such patient: patient I.G. Patient I.G. exhibited the greatest proportion of looks to the competitor animal (e.g., the frog in the basket) upon hearing *Put the frog on the napkin* ... in the presence of a visual scene containing two frogs: His looking time was more than 2.2 standard deviations above that of the healthy control group; Crawford's test revealed marginal significance, $t(2) = 1.94$, $p < .1$. This suggests (see Novick et al., 2008) that I.G. was not considering the modifier analysis from the moment he heard the temporarily ambiguous phrase *on the napkin*. This pattern of eye movements, coupled with his impaired ambiguity effect for looks to the incorrect goal in two-referent scenes, suggests that I.G. was rapidly pursuing an interpretation consistent with only the most reliable information, the lexical preference of *put*, despite (weaker) extrasentential referential evidence like visual context. He was therefore exhibiting a selective impairment both in rejecting the goal analysis of *on the napkin* and arriving at the modifier analysis of this phrase.

Table 7. *Mean proportion of looks to competitor animal[a] from the onset of the preposition[b] in two-referent ambiguous versus unambiguous conditions, Experiment 4*

	Two-referent		
Participants	*Ambig*	*Unambig*	*Ambiguity effect*
Controls	.056 (.007)	.04 (.07)	.015 (.071)
Patient NN454	.187	.362	−.175
Patient GU412	.054	.045	.09
Patient IG363	.174	0	.174[†]

Note: Standard deviations in parentheses.
[a] e.g., frog in basket. [b] *on the napkin* . . .
[†] $p < .1$.

Thus far, the multiple eye movement analyses reported suggest that during the real-time processing of a sentence containing a temporary ambiguity, compared to an unambiguous control sentence,

---

[7] Akin to the analysis for looks to the incorrect goal, participant means were calculated for each condition excluding trials that contained more than 33% track loss. This accounted for approximately 15% of the data.

patient I.G. clearly shows the most robust online commitment to an interpretation consistent with only the strongest constraints, in this case the lexico-syntactic biases of the verb *put*. Other patients and healthy controls show this pattern too, but their eye movements suggest that they begin to pursue an alternative interpretation—that is, they revise their syntactic commitments—upon gaining additional information as the sentence continues to unfold, like disambiguating evidence (e.g., *into the box*), or in light of contextual evidence (e.g., the presence of two frogs). In other words, healthy controls and non-LIFG patients spend less time considering the wrong interpretation in view of an ambiguous phrase, whereas I.G. appears to have trouble overriding the first interpretation he commits to.

In what follows, we assess this explicitly—that is, the extent to which I.G. can ultimately override these initial syntactic commitments—by examining his final interpretation of the sentences, evidenced by his offline hand actions in carrying out the instructions.

Action responses

A record was kept of how participants carried out the target instructions. Errors were coded as action responses that involved the incorrect goal—for example, if in response to hearing *Put the frog on the napkin into the box*, a participant picked up a frog and then put it on the empty napkin (the incorrect goal), either leaving it there and thus ignoring *into the box*, or placing it on the empty napkin first and then sequentially moving it to the box afterward. Both types of action reveal an interpretation consistent with *on the napkin* being incorrectly construed as the goal (see Trueswell et al., 1999; Weighall, 2008).

The data show that the healthy controls, like college-age adults from previous studies, were largely able to revise initial parsing commitments to a goal interpretation (evidenced by eye movement patterns during the ambiguous phrase *on the napkin*) and recover a dispreferred but currently relevant characterization of the input. Thus, their resulting hand actions were consistent with the modifier analysis of the ambiguous phrase *on the napkin* across all conditions. Only 1 healthy control participant made an error involving the incorrect goal, which was observed in a one-referent ambiguous condition (see Figure 3). Such an error rate for this group (8.3% overall) is comparable to that of young adults reported in the Trueswell et al. (1999) study for this condition. Also as expected, the healthy controls in the current study performed perfectly on all unambiguous trials (e.g., *Put the frog that's on the* . . .).

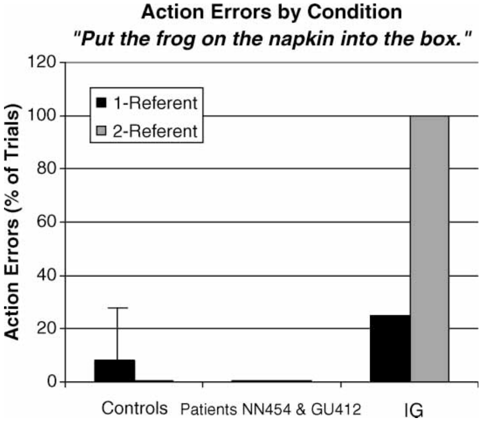


Figure 3. *Proportion of action errors involving the incorrect goal (e.g., empty napkin) under ambiguous conditions. The healthy control group was collapsed together, as were the two left frontal control patients.*

Similarly, the non-LIFG frontal patients, patients NN454 and GU412, did not make any errors on ambiguous trials, also showing their ability to arrive at the intended meaning of *on the napkin* despite syntactic ambiguity and hence representational conflict between two alternatives of sentence meaning (goal vs. modifier; see Figure 3). Taken together, healthy controls and frontal control patients with damage sparing LIFG were able to quickly revise initial parsing commitments to recover from temporarily ambiguous sentences in which a preferred interpretation, derived on the basis of well-supported and accumulating syntactic and semantic evidence as the sentence unfolds, must be overridden in favour of another analysis.

By contrast, I.G.'s ability to revise early parsing commitments was strikingly different from that of the frontal group and the healthy control group. He performed errors involving the incorrect goal on over 62% (5 out of 8) of all ambiguous trials. Broken down by referential context, I.G. made one error out of four ambiguous trials in one-frog scenes (25%), and four errors out of four ambiguous trials in two-frog scenes (100%). Thus, his resulting actions—particularly in two-referent contexts—were reliably consistent with the goal analysis of the ambiguous phrase *on the napkin* and rarely, if ever, with the modifier analysis. This pattern indicates that he failed to recover the intended interpretation of *on the napkin* when it conflicted with highly reliable lexical evidence, even though that evidence was ultimately misleading (because of later information in the speech stream, like *into the box*, which ultimately signalled the correct goal).

Furthermore, in two-referent scenes, I.G. always used the *competitor* animal (e.g., the frog in the basket) to carry out an instruction, suggesting that he never considered *on the napkin* to be a modifying phrase intended to provide more information about one of the two frogs in his visual environment. In other words, in every two-referent ambiguous case, I.G. would pick up the competitor (e.g., the frog that was sitting in the basket) and then move it to the incorrect goal (e.g., the empty napkin). Consistent with his online eye movement patterns to the competitor animal, this suggests that I.G. committed to the goal analysis early (certainly not the correct modifier analysis) and followed through with this interpretation, never arriving at all at a modifier interpretation for *on the napkin*. This phrase, to him, was the intended goal, and because there was a frog that was not already on a napkin but instead in a basket, he interpreted the sentence to mean that this frog should be put on a napkin, again, to satisfy incorrectly the syntactic constraint of *put*. No other patient exhibited this pattern.

No errors were observed for I.G. at all for the unambiguous trials, despite similar length and syntactic complexity to those of their ambiguous counterparts. This is to be expected under the conflict resolution deficit account, because the syntactic conflict was removed on these trials between goal and modifier analyses of *on the napkin*, thereby eliminating the need to override an initial goal response.

Discussion of Experiment 4

Considering both I.G.'s online eye movement patterns and his resulting hand actions to the instructions, the following picture emerges: I.G. consistently becomes entrenched in a syntactic representation of the sentence, and, when later disambiguating information or extrasentential evidence, such as visual–referential context, begins to support an alternative analysis that is incompatible with his early representation (i.e., the modifier analysis now competes with the goal analysis), he cannot disengage from his initial parse and override the way he initially characterized the input. We believe that I.G.'s general conflict resolution impairment (Experiment 1) has extensive implications for language processing as evidenced by his deficient performance in both production tasks (Experiments 2–3) and a comprehension task (Experiment 4), but only when representational conflict demands are at their relative peaks.

Interestingly, it is worth noting that I.G.'s inability to resolve syntactic ambiguity in the current comprehension experiment is markedly similar to that of five-year-old children, who also show a failure to revise initial parsing commitments on roughly 60% of ambiguous trials, but show the normal comprehension pattern in response to unambiguous sentences (Trueswell et al., 1999; Weighall, 2008). By age eight, children perform indistinguishably from healthy adults on the *Put the frog* . . . task. This late developmental progression of parsing skills, specifically ambiguity resolution, makes sense in the context of children's generally delayed cognitive control abilities on tasks such as the Wisconsin Card Sorting Task, the Simon task, Go/No-Go tasks, and others (e.g., Cepeda & Munakata, 2007; Diamond & Taylor, 1997; Yerys & Munakata, 2006). The commonality among all these tasks, including ambiguity resolution, we argue, is the need to implement top-down attention control by overriding a highly familiar or routine response in the face of either newly implemented rules or newly developed evidence or information that is in conflict with the standard way of carrying out the target objective (e.g., understanding a sentence; naming a word's ink colour rather than reading the word itself). We have recently suggested that this maturational lag pertaining to general cognitive control abilities as well as to sentence-processing and comprehension abilities may be due in part to late-developing prefrontal brain systems, including LIFG (Novick et al., 2005).

But why is I.G.'s ambiguity effect more pronounced in two-referent cases? This pattern at first blush may appear counter to expectations given that healthy adults rapidly use contextual evidence to override the goal analysis. Yet we believe that I.G.'s parsing pattern is just as predicted under the cognitive control account, especially when closely examined alongside young children's patterns. In fact, I.G.'s ambiguity resolution profile across both referential conditions mirrors that of five-year-olds in Trueswell et al. (1999) and Weighall (2008). Children in Trueswell et al. (1999) made slightly more errors on ambiguous two-referent trials than one-referent trials: 61% errors versus 55% errors, respectively. Although this effect is small (and most likely unreliable in a pairwise comparison), Weighall (2008) replicated it with a much larger sample: The size of the effect was bigger, in fact, and statistically significant: 66% errors in two-referent ambiguous compared to 50% errors in one-referent ambiguous (see their Table 1). Weighall (2008) also ran older children (8-year-olds and 11-year-olds) and found a marginal interaction between referential scene (one- vs. two-referent) and age for ambiguous items: "An interaction between number of referents and age group was found to approach significance, $F(2, 57) = 2.971$, $p = .059$, partial $g2 = .094$. Inspection of the means indicated that this interaction arose because 5-year-olds made many more correct responses when there was one referent than when there were two referents, whereas this difference was less pronounced for the older children" (Weighall, 2008, p. 85).

In this light, we believe that the interpretation of I.G.'s pattern is reasonably straightforward under the cognitive control account. For those without good cognitive control (5-year-olds and I.G.), listeners almost completely adopt a goal interpretation, which they cannot revise, regardless of context (one- or two-referent). But one-referent scenes slightly conceal the difficulty they are having arriving at the modifier interpretation. The only frog present is already on a napkin, so they (young children, I.G.) rationalize that they should move it to the box since they also heard *into the box*—they are just unable to use this new information to block the prepotent goal-parse of *on the napkin*. Indeed, these participants may even treat the napkin under the frog as the "first goal", like a hopping error, where one would not be able to observe the first "hop" as part of the incorrect action. That is, as Trueswell et al. (1999) argued, participants in the one-referent condition are sometimes acting correctly but for the wrong reasons. In two-referent scenes, participants also go for the goal interpretation and so prefer to move the frog that is not already on a napkin over to a napkin and then sometimes into the box. In the current study, I.G. picks the competitor frog in two-referent scenes all the time, which certain children did as well. And interestingly, whenever children chose to act on the competitor, they almost always moved it to the incorrect goal (the empty napkin), consistent with the reasoning they show in one-referent scenes. Weighall (2008) also found this: "Like Trueswell and colleagues (1999), we observed that the object selected first (in the two referent conditions) was closely linked with whether the correct destination was selected or not. It was found that the correct animal was moved to the correct destination on 68% of trials. In contrast, the incorrect animal was moved directly to the correct destination on only 3% of trials" (pp. 88–89).

To summarize I.G.'s inflated two-referent ambiguity effect, it is not that these scene types hurt him per se, but rather the one-referent scene helps people with poor cognitive control perform better, sometimes getting the action "right" despite having the wrong parse of the sentence. It may be the case that the smarter the listener, the better he will do on one-referent scenes and the worse he will do on two-referent scenes. This may be because the logical thing to do, if one has a goal-parse of *on the napkin* in two-referent scenes, is to move the frog that is not already on a napkin over to the napkin and then into the box. In one-referent scenes, the logical choice is not entirely clear. If the only frog is already on a napkin, should it be lifted up and put back down on the napkin it was sitting on (which some of Weighall's, 2008, kids did)? Or should it be moved over to the empty napkin and then into the box? Certainly I.G. is more knowledgeable than 5-year-olds, and hence he shows accentuated differences between two- and one-referent scenes. He made the logical choices given that he had a goal-parse of *on the napkin*.

In terms of I.G.'s syntactic abilities, it is unlikely that a syntactic deficit is what gives rise to his processing and comprehension impairment under temporarily ambiguous conditions. If that were the case, one might expect that unambiguous sentences—which are equally as complex as ambiguous ones—should have also resulted in difficulty. But I.G. made no action errors on any unambiguous trial, and his online ambiguity effect, again, was calculated as an increase in looking proportions to the incorrect goal compared to unambiguous trials. For these same reasons, it is doubtful that I.G. possesses a deficit in verbal working memory (i.e., capacity of a verbal store). If he did, he presumably would have trouble integrating the words online because there would be too many to keep track of in his working memory buffer.

Moreover, a separate anecdotal comment can be made that is consistent with this argument: I.G. was observed repeating the instructions flawlessly on every critical trial after hearing it, without being prompted to do so, prior to executing a response. He recalled the entire sequence of words in an utterance, *Put the frog (that's) on the napkin into the box*, and would repeat it sometimes two (or even three) times while carrying out an action, and he still made errors on ambiguous trials despite this. It seems reasonable to argue that if his verbal working memory capacity was impaired, then he would not have been able to repeat the instructions with such precision.

GENERAL DISCUSSION

The principal objective of the current research is to help characterize the nature of the cognitive and language impairments suffered by a patient with focal damage to LIFG. Our hypotheses are grounded in domain-general terms—in particular, that LIFG is a region that broadly supports top-down biasing of representations in the face of competition, and that a circumscribed lesion here ought to result in a discernable impairment in which patients are incapable of resolving conflict at multiple levels, especially when such resolution involves reversing probabilistically supported or routine processes (response override), or biasing response candidates for selection when multiple possible and equally conceivable ones compete (underdetermined response conflict). This notion is rooted in previous functional neuroimaging work in which LIFG activation is boosted under both types of conflict conditions (e.g., Stroop, name agreement) and other case studies of patients with LIFG damage who fail on these same cognitive tasks. However, none of these prior studies has focused on language production or comprehension tasks that rely on conflict resolution and cognitive control processes.

Our focus therefore sought to extend these findings by narrowing in on the language production and comprehension abilities of a patient with restricted damage to LIFG. This seemed a productive avenue to pursue for several reasons. First, a general cognitive control deficit account has potentially large-scale ramifications for language use. As we motivated in our introduction, and reported in our data, damage to a brain region that supports conflict resolution abilities—namely, LIFG—makes clear predictions of when production and comprehension skills will break down. In the former case, we argued that language production should become halted and dysfluent especially when speaking requires the selection of a lexical item that competes semantically and/or conceptually with other candidate lexical items. In the latter case, we argued that comprehension should fail in the face of syntactic ambiguity, specifically when a patient's highly supported developing interpretation conflicts with newly encountered input and, consequently, when such initial parsing commitments need to be reanalysed in order to be correctly understood.

Second, the current work helps pin down the qualitative character of such patients' language impairments in broader terms, beyond traditional accounts that suggest a priority for failed syntactic representations to be the source of their linguistic crash. The reason this distinction is important is that, as addressed earlier in the introduction, part of LIFG includes Broca's area (BA 44/45), the very area that has conventionally been equated with Broca's aphasia when damaged. Broca's aphasia, again, is characterized behaviourally by syntactic impairments in both production and comprehension, but the work by Dronkers and colleagues (2004) tells us that Broca's aphasics do not have restricted damage to BA 44/45, and that damage to BA 44/45 does not always yield Broca's aphasia (see also Alexander, 2006). Thus, our goal has been to test, within a conflict resolution framework, what sort of language impairments patients with focal damage here do demonstrate. It has been clear that such patients exhibit some form of language impairment (see, e.g., Robinson et al., 1998, 2005), albeit transient ones; we just have not known exactly what kind and under what conditions they arise. What the case studies by Hamilton and Martin (2005) and by Thompson-Schill and colleagues (2002) have revealed is that these patients have cognitive deficits anchored in ineffective conflict resolution abilities broadly construed. Here we asked whether this general conflict resolution deficit could also account for such patients' language difficulties.

We first showed that our patient, I.G., had a general conflict resolution impairment on a working memory task that manipulated proactive interference (Experiment 1) in the verbal domain. The reason for including this task was to replicate previous findings showing that other patients with a similar neuroanatomical profile (M.L. and R.C.) exhibited an exaggerated interference effect. The idea was to connect I.G. with earlier patients who have the same lesion contour and, as a result, the same conflict resolution impairment. We have predicted elsewhere (Novick et al., 2005) that patients like R.C. and M.L., given their conflict resolution deficits, should have language impairments on tasks that give rise to heightened conflict demands, like those reported here in Experiments 2–4. Hence, we tested I.G. against those claims. In addition, the proactive interference task employed here does not require reliance on any sophisticated combinatory linguistic (e.g., syntactic) mechanisms, and so I.G.'s impairment is thus not restricted to syntactic tasks per se (the single-word naming tasks in Experiments 2 and 3 also demonstrate this). On this proactive interference task, I.G. was the only patient with an inflated error rate beyond the normal range (the 95th percentile) for interference compared to noninterference trials.

In Experiments 2 and 3, we honed in on I.G.'s language production abilities in both a confrontation naming task and a task of verbal fluency. In the naming task (Experiment 2), when the to-be-named images had low name agreement (underdetermined response conflict; e.g., couch/sofa/loveseat) and therefore high conflict demands due to multiple possible response candidates, I.G. made a proportion of commission and omission errors more than 3 standard deviations above the performance of a healthy control group. By contrast, when he had to name objects with high name agreement and therefore fewer alternative response options (e.g., apple), I.G.'s performance was within normal limits. No other patient showed this pattern.

In the verbal fluency task (Experiment 3), I.G.'s impairment revealed itself in a similar fashion: When category set size was large and relatively unconstrained (e.g., animals), compared to smaller categories with a comparatively more restrictive set of exemplars (e.g., farm animals), I.G. was unable to use the larger categories to produce a greater number of category members, unlike healthy controls. Under such conditions, he hesitated, offered an incorrect response, or could not answer altogether. Two other left frontal patients also showed similar patterns, but I.G.'s performance, at 2.5 standard deviations below the mean for large categories, was clearly the lowest. Moreover, I.G. is the only patient to show the entire pattern of impairment consistently across the three experiments discussed thus far. Taken together, the data from Experiments 2 and 3 support the notion that LIFG is necessary to guide semantic retrieval when stimulus cues do not sufficiently constrain the response. In other words, I.G.'s language production can be broadly defined within a general conflict resolution deficit account: Regardless of the type of production task, I.G. has problems generating spoken words when the candidate options mutually compete on a lexical–semantic level.

In Experiment 4 we turned our attention to patients' online spoken language comprehension abilities. We focused centrally on syntactic ambiguity resolution because, under such conditions, listeners must reject highly supported developing linguistic analyses in light of other sources of evidence. Such a process of recovering alternative interpretations has previously been compared to nonsyntactic response override cognitive control tasks like Stroop (see Novick et al., 2005, for a discussion; see also January et al., 2009). In our task, participants had to carry out spoken instructions that required them to override an early preference to interpret a temporarily ambiguous phrase (e.g., *on the napkin*) as a goal of where to put an object (e.g., a frog) in light of later arriving evidence, such as *into the box*, signalling that this new phrase is actually where the object should be put. We observed that, via two online measures in particular—proportion of looks to the incorrect goal (e.g., empty napkin) and proportion of looks to the competitor animal (e.g., frog in basket) in two-referent conditions—I.G. showed an exaggerated preference in ambiguous conditions, compared to unambiguous conditions, to consider an analysis consistent with early parsing commitments. In particular, I.G. spent reliably more time fixating the incorrect goal (the empty napkin) under ambiguous conditions, suggesting that this was the location he considered to be the correct goal for the putting action to take place.

When separating by referential scene type, we discovered that this inflated proportion of looks to the incorrect goal was especially pronounced in two-referent contexts—that is, when two frogs were visually present. This pattern is important for the following reasons. First, this scene type supports the modifier interpretation of *on the napkin* because the presence of two frogs requires the speaker to specify which one should be acted upon (e.g., Altmann & Steedman, 1988). This prepositional phrase (*on the napkin*) denotes exactly that. Second, in spite of this strong contextual support for a modifier analysis, I.G. still considered the empty napkin as the likely goal upon hearing the word *napkin*, much more so than the healthy control group and the non-LIFG left frontal group. That is, he appears to pursue more than others an early commitment to *on the napkin* as the goal referent, which is strongly supported by the syntactic constraints of the verb *put* (it must take a goal); he does this regardless of two frogs being present, much like young children (Trueswell et al., 1999; Weighall, 2008). Healthy adults and non-LIFG patients routinely and effectively use these two sources of evidence—*into the box* and the presence of two frogs—to recover the correct interpretation of the input (i.e., that *on the napkin* is simply providing more information about which frog is meant, and *into the box* signals the intended goal; see also Novick et al., 2008; Spivey et al., 2002; Tanenhaus et al., 1995; Trueswell et al., 1999).

The conclusions drawn from the incorrect goal and competitor animal eye movement analyses in two-referent cases are further supported by I.G.'s offline hand actions. He makes roughly 60% errors under ambiguous conditions (collapsing across referential scene type) compared to the healthy control group's 8%, a rate comparable to that in previous studies using these materials and healthy young participants (see Novick et al., 2008, for a discussion). No other patient showed this pattern; in fact, no other patient demonstrated an error. Moreover, examining two-referent scenes separately, I.G. committed an error involving the false goal 100% of the time, and, further, he selected the unintended referent (the competitor animal) in every one of these cases to execute an action. This bolsters the argument that I.G. never arrived at the modifier interpretation and always pursued the goal interpretation despite conflicting sources of evidence: When he hears *Put the frog on the napkin into the box*, he picks up the frog that's in a bowl and moves it to an empty napkin. Such a pattern is especially expected under the cognitive control account in light of the strikingly similar parsing patterns demonstrated by young children in two-referent ambiguous conditions (Trueswell et al., 1999; Weighall, 2008).

It is important to reiterate that in unambiguous cases—*Put the frog that's on the napkin into the box*—I.G. never made an error involving the incorrect goal, and he never chose the competitor animal to act upon; he always picked up the frog on the napkin and moved it to the box. Thus, his comprehension failure here is not a direct result of a linguistic inability to resolve definite reference; in other words, he understands simple noun phases like *the frog*, and it is not a direct result of being unable to understand long complicated sentences as unambiguous forms are longer than ambiguous ones. I.G.'s language deficit arises in only select cases, ingrained in the ambiguity and the need to override a probabilistically supported interpretation. Healthy participants and other frontal patients, whose damage spares LIFG, are able to resolve this conflict. A subset of the frontal patients who also participated in this study showed some of the patterns exhibited by I.G., but these patterns were much smaller in magnitude by comparison, and, moreover, I.G. is the only patient who showed a significant conflict resolution deficit trend across all four experiments.

So, does I.G. have a language impairment? We would argue that the answer is both yes and no in light of the results of our study. It is clear that I.G.'s impairment is not specific to grammar per se. For evidence, we again point the reader to the proactive interference task in Experiment 1, which does not rely on any grammatical machinery to complete, nor do the name agreement and fluency tasks in Experiments 2 and 3. Moreover, I.G. comprehends syntactically complex yet unambiguous sentences with healthy precision in Experiment 4. What's common to all of the tasks reported here is a broad need to resolve among competing and/or conflicting representations in the verbal domain (though not in a syntactic domain). A general conflict resolution deficit has critical implications for language. In production (Experiments 2 and 3), when the semantic and/or conceptual features of a to-be-produced word are shared by, and therefore compete with, many potential alternative candidate words, fluent production becomes sparse and difficult. By comparison, language understanding involves the continuous updating of a sentence's meaning as new input arrives to the eye or ear, and when this new information conflicts with early developed representations (Experiment 4), comprehension fails.

These findings certainly cohere with those of Robinson and colleagues (e.g., Robinson et al., 2005), who report the results of several verbal and nonverbal conflict resolution tasks completed by a patient (C.H.) with restricted damage to posterior LIFG (BA 44/45). Notably, C.H. demonstrated a severe deficit for word, phrase, and sentence generation tasks when numerous potential responses were activated by a stimulus and therefore competed for selection (i.e., in cases of underdetermined response conflict). On the other hand, C.H. could produce language without difficulty when a dominant response was unambiguously activated by a stimulus. In remarkable contrast, C.H. demonstrated normal patterns when asked to produce a number of nonverbal responses (e.g., gestures) across several different measures despite conflict demands. This strongly suggests that LIFG is necessary for resolving conflict in the verbal but not the nonverbal domain (see also Hamilton & Martin, 2005, who demonstrate a similar noteworthy dissociation in patient M.L.). Although we did not test I.G. on any nonverbal tasks, the verbal/nonverbal evidence from patients C.H. and M.L. is rather compelling.

Moreover, another appeal of the general conflict and cognitive control framework is that it fits well with traditional views of the executive system. For example, in Norman and Shallice's (1986) model, a supervisory attentional system (SAS) oversees routine (i.e., "contention-scheduling") processes and actions, overriding them when necessary. As Botvinick and colleagues mention: "Although the theory does not explicitly indicate what particular events within contention scheduling serve to trigger SAS intervention, it is emphasized that contention scheduling serves primarily to prevent conflict among potentially relevant schemas (Norman & Shallice, 1986). Thus, the theory seems to imply that control is recruited when conflicts occur that contention-scheduling processes are not able to resolve efficiently" (Botvinick et al., 2001, p. 626). In classic terms, then, conflict resolution and cognitive control are inherent components to this model. One might argue particularly that the SAS would have to intervene upon encountering the prepotent response conflict discussed in Experiment 4—that is, when coming across new linguistic evidence (*into the box*) that conflicts with one's highly supported developing interpretation (e.g., *on the napkin* as the goal-parse of *put*). Similar parallels can be drawn from the types of underdetermined response conflict discussed throughout the paper as well.

An alternative interpretation of our findings is that damage to LIFG may affect memory retrieval.[8] Some researchers claim that this region supports controlled (as opposed to automatic) retrieval from semantic memory (Badre et al., 2005; Badre & Wagner, 2002). Such an explanation could in principle account for a large interference effect in the first experiment (recent−no minus nonrecent−no), because I.G. may have had difficulty recovering the contextual information indicating that the familiar but incorrect item (i.e., the recent−no item) is not in the current list. Similarly, the effect of name agreement might result because the existence of alternative names implies that none is as strongly associated with the concept as are concepts with only a single common name. Likewise, variations in picture familiarity could influence the degree of difficulty in retrieving the names of objects (though this is unlikely to account for the effect given that familiarity was well matched across our high- and low-conflict conditions; see Footnote 6). In the sentence comprehension task (Experiment 4), I.G. might have had difficulty retrieving an infrequent syntactic structure (i.e., the structure with the modifying prepositional phrase following the object noun may occur less frequently than a structure without the modifier). Taken together, these possibilities are worthy of serious consideration; however, it is not entirely apparent how the memory retrieval story can straightforwardly capture the category fluency data. As such, we believe that the conflict resolution and cognitive control account may be better equipped to wholly describe the data patterns across all four experiments, particularly given their designs. Nevertheless, it is crucial that future work attempt to disentangle these two positions.

Lastly, could an "impulse control" deficit alternatively explain I.G.'s pattern of performance? Perhaps, though two caveats should be considered before adopting such an interpretation. First, it is not altogether clear whether there are any specific tests that are designed to identify a general impulse control impairment, or what existing measures one might use to depict the problem. Thus, such a deficit may be hard to diagnose and capture. As such, we remain agnostic about whether I.G. suffers from a general lack of impulse control.[9] More to the point, however, a story that ascribes I.G.'s difficulty to a problem controlling impulses cannot account for his consistent pattern of performance across the four experiments. That is, the specificity of I.G.'s effects are hard to explain in impulse control terms: Why should impulsivity

---

[8] We thank Randi Martin for suggesting this, including many of the following important and plausible considerations.

[9] Anecdotally, however, we doubt this is the case, due to his perfectly appropriate social interactions.

result in an *acute failure* to (a) resolve proactive interference in working memory; (b) select a single name for an object that has multiple possible labels; (c) provide a greater number of exemplars from large versus small categories; and (d) resolve syntactic ambiguity (i.e., recover from misinterpretation)? Because our experiments were designed to manipulate both prepotent response conflict and underdetermined response conflict (not just the former), we believe that a lack of impulse control is unlikely to account for I.G.'s uniform collection of impairments. In other words, it is unclear why an "impulse control deficit" would be so specific to the patterns of performance observed here.

In view of our discussion on conflict and control, it is worth mentioning that none of the patients studied here suffered lesions to anterior cingulate cortex (ACC). This is notable because ACC appears to be dedicated to detecting response-based conflict, which we believe is different from the representational conflict described throughout this article. For instance, prior work has dissociated the roles of the ACC and LIFG in conflict resolution for evaluative processes versus representation. In one study (Carter et al., 2000), event-related fMRI of the Stroop task revealed early bilateral LIFG activity, which was interpreted to reflect strategic processes to reduce cognitive conflicts (i.e., competition between representations). This pattern of activity was distinguished from activity in ACC, which was thought to reflect monitoring and detecting cognitive states. Taken together, the authors suggested "the ACC performs an evaluative function, reflecting the degree of response conflict elicited by the task" whereas "inferior frontal cortex [might be] engaged according to the selection demands of the task" (Carter et al., 2000, p. 1947).

In many cases, conflict at the representational level and response conflict may be hard to separate; indeed, both LIFG and ACC are routinely activated in the Stroop task (e.g., Kerns et al., 2004; MacDonald et al., 2000; Milham et al., 2001). Other studies have also attempted to disentangle the relative contributions of these two regions to response-based and representational conflict resolution (e.g., Milham et al., 2003; Milham et al., 2001; Nelson et al., 2003), most notably using the proactive interference experiment reported in our Experiment 1. For instance, Nelson et al. proposed that recent–no trials in this task predominantly generate internal conflict rather than response-based conflict (i.e., how to characterize the probe letter). They tested this by comparing recent–no trials with trials that should induce strong response-based conflict. On these "response-conflict" trials, the current probe letter was not a member of the current set (a "no" response) but was both a member of the set on the previous trial and the probe item on that previous trial (a "yes" response). Here, conflict was expected at the response level—because participants just responded "yes" to the same probe—as well as internally, because the probe is familiar given its recent presentation. Indeed, these trials generated both ACC and LIFG activity relative to nonconflict trials, whereas the other recent–no trials (i.e., the ones used in our study) generated the greatest activation in LIFG. All in all, there clearly exist different proposals regarding the types of cognitive control mechanism that different areas within prefrontal cortex may subserve. However, the goal of this study was not to test among those.

Before closing, we wish to note that LIFG is a relatively large cortical area and includes at least two cytoarchitecturally separate regions (BAs 44 and 45; Amunts et al., 1999); furthermore, it is possible that even these regions can be further subdivided (Amunts et al., 1999). From the widest view, LIFG may be responsible for conflict resolution on tasks that give rise to underdetermined response conflict and/or prepotent response conflict (in the verbal domain), but that it is organized into dedicated circuitries depending upon the information types involved. Namely, resolution processes for conflicting representations may be implemented in LIFG broadly, but may be structured into compartmentalized subregions on the basis of what type of conflict is being adjudicated. However, a general conflict resolution system that deals with information from multiple modalities and various information types may reason against information specialized circuitry (imagine the combinatory upsurge of circuitry for each case of conflict). Consequently, we are agnostic about whether subregions within LIFG engage control in response to specific types of linguistic information. Instead, we make the broader claim, for now, that LIFG engages conflict resolution processes more generally, which include, of course, both underdetermined response conflict and prepotent response conflict. Exploring the possibility of specialized resolution functions, though, is an important empirical question and should be the focus of future work (see also Nagel et al., 2008, and references therein).

Additional research aims might be to further define the scope of how general such a conflict resolution impairment is in patients with restricted LIFG damage and whether they arise in the production domain, beyond the single word level. The conflict resolution deficit account should also be able to capture and explain other language impairments that LIFG patients have beyond the ones studied here—namely, tasks involving object naming, word production, and sentence comprehension. One prospective area to pursue with patients such as I.G., for instance, would be discourse-level situations in which a developing situational context first supports an initial set of inferences about a particular scenario, but whose completion reveals that background assumptions need to be updated in light of new events in the story (e.g., Hamm & Hasher, 1992). Whether LIFG patients can alter inferences initially drawn from a story's (or even a conversation's) propositional content could further reveal the extent and nature of their language impairments.

Manuscript received 12 January 2009
Revised manuscript received 17 November 2009
Revised manuscript accepted 17 November 2009

REFERENCES

Alexander, M. P. (2006). Aphasia I: Clinical and anatomical issues. In M. J. Farah & T. E. Feinberg (Eds.), *Patient-based approaches to cognitive neuroscience* (2nd ed.). Cambridge, MA: MIT Press.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419–439.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*, 247–264.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*(3), 191–238.

Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H. B. M., & Zilles, K. (1999). Broca's region revisited: Cytoarchitecture and intersubject variability. *Journal of Comparative Neurology*, *412*, 319–341.

Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, *47*, 907–918.

Badre, D., & Wagner, A. D. (2002). Semantic retrieval, mnemonic control, and prefrontal cortex. *Behavioral and Cognitive Neuroscience Reviews*, *1*(3), 206–218.

Ballard, D., Hayhoe, M., & Pelz, J. (1993). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, *7*, 66–80.

Bedny, M., Hulbert, J. C., & Thompson-Schill, S. L. (2007). Understanding words in context: The role of Broca's area in word comprehension. *Brain Research. Special Issue: Mysteries of Meaning*, *1146*, 101–114.

Bedny, M., McGill, M., & Thompson-Schill, S. L. (2008). Semantic adaptation and competition during word comprehension. *Cerebral Cortex*, *18*(11), 2574–2585.

Blumstein, S. E., Myers, E. B., & Rissman, J. (2005). The perception of voice onset time: An fMRI investigation of phonetic category structure. *Journal of Cognitive Neuroscience*, *17*(9), 1353–1366.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.

Carter, C. S., MacDonald, A. M., Botvinick, M., Ross, L. L., Stenger, V. A., Noll, D., et al. (2000). Parsing executive processes: Strategic versus evaluative functions of the anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, *97*, 1944–1948.

Cepeda, N. J., & Munakata, Y. (2007). Why do children perseverate when they seem to know better: Graded working memory, or directed inhibition? *Psychonomic Bulletin & Review*, *14*, 1058–1065.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *30*, 687–696.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments & Computers*, *25*(2), 257–271.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*(1), 84–107.

Crawford, J. R., Deary, I. J., Starr, J., & Whalley, L. J. (2001). The NART as an index of prior intellectual functioning: A retrospective validity study covering a 66-year interval. *Psychological Medicine*, *31*, 451–458.

Crawford, J. R., & Garthwaite, P. H. (2006). Methods of testing for a deficit in single-case studies: Evaluation of statistical power by Monte Carlo simulation. *Cognitive Neuropsychology*, *23*(6), 877–904.

Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist*, *12*, 482–486.

Diamond, A., & Taylor, C. (1997). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do." *Developmental Psychobiology*, *29*, 315–334.

Dickey, M. W., Choy, J., & Thompson, C. K. (2007). Real-time comprehension of wh-movement in aphasia: Evidence from eyetracking while listening. *Brain and Language*, *100*, 1–22.

Dickey, M. W., & Thompson, C. K. (2009). Automatic processing of wh- and NP-movement in agrammatic aphasia: Evidence from eyetracking. *Journal of Neurolinguistics*, *22*, 563–583.

Donohue, S. E., Wendelken, C., Crone, E. A., & Bunge, S. A. (2005). Retrieving rules for behavior from long-term memory. *NeuroImage*, *26*, 1140–1149.

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Jr., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, *92*, 145–177.

Ferreira, F., & Clifton, C. E. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348–368.

Fiebach, C. J., Vos, S. H., & Friederici, A. D. (2004). Neural correlates of syntactic ambiguity in sentence comprehension for low and high span readers. *Journal of Cognitive Neuroscience*, *16*, 1562–1575.

Fraisse, P. (1969). Why is naming longer than reading? *Acta Psychologica*, *30*, 96–103.

Grober, E., & Sliwinski, M. (1991). Development and validation of a model for estimating premorbid verbal intelligence in the elderly. *Journal of Clinical and Experimental Neuropsychology*, *13*, 933–949.

Grober, E., Sliwinski, M., Schwartz, M., & Saffran, E. (1989). *The American version of the NART for predicting premorbid intelligence*. Unpublished manuscript.

Haarmann, H. J., & Kolk, H. H. J. (1991). Syntactic priming in Broca's aphasics: Evidence for slow activation. *Aphasiology*, *5*, 247–263.

Hagoort, P., Wassenaar, M., & Brown, C. (2003). Real-time semantic compensation in patients with agrammatic comprehension: Electrophysiological evidence for multiple-route plasticity. *Proceedings of the National Academy of Sciences*, *100*, 4340–4345.

Hamilton, A. C., & Martin, R. C. (2005). Dissociations among tasks involving inhibition: A single-case study. *Cognitive, Affective, & Behavioral Neuroscience*, *5*, 1–13.

Hamm, V. P., & Hasher, L. (1992). Age and the availability of inferences. *Psychology and Aging*, *7*, 56–64.

Hirshorn, E. A., & Thompson-Schill, S. L. (2006). Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency. *Neuropsychologia*, *44*, 2547–2557.

January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of Stroop and syntactic ambiguity resolution in Broca's area: Implications for the neural basis of sentence processing. *Journal of Cognitive Neuroscience*, *21*, 2434–2444.

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proceedings of the National Academy of Sciences*, *95*, 8410–8413.

Kan, I. P., Kable, J. W., Van Scoyoc, A., Chatterjee, A., & Thompson-Schill, S. L. (2006). Fractionating the left frontal response to tools: Dissociable effects of motor experience and lexical competition. *Journal of Cognitive Neuroscience*, *18*, 267–277.

Kan, I. P., & Thompson-Schill, S. L. (2004a). Effect of name agreement on prefrontal activity during overt and covert picture naming. *Cognitive, Affective, & Behavioral Neuroscience*, *4*, 43–57.

Kan, I. P., & Thompson-Schill, S. L. (2004b). Selection from perceptual and conceptual representations. *Cognitive, Affective, & Behavioral Neuroscience*, *4*, 466–482.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., III, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate, conflict monitoring, and adjustments in control. *Science, 303*, 1023–1026.

Kertesz, A. (1982). *Western Aphasia Battery*. Orlando, FL: Grune & Stratton.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, *98*, 122–142.

Lindenberg, R., Fangerau, H., & Seitz, R. J. (2007). "Broca's area" as a collective term? *Brain and Language*, *102*, 22–29.

Luria, A. R. (1973). *The working brain: An introduction to neuropsychology*. New York: Basic Books.

MacDonald, A. W., III, Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288*, 1835–1838.

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

Mason, R. A., & Just, M. A. (2007). Lexical ambiguity in sentence comprehension. *Brain Research. Special Issue: Mysteries of Meaning*, *1146*, 115–127.

Mason, R. A., Just, M. A., Keller, T. A., & Carpenter, P. A. (2003). Ambiguity in the brain: What brain imaging reveals about the processing of syntactically ambiguous sentences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 1319–1338.

McElree, B. M., & Dosher, B. A. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General*, *118*, 346–373.

Milham, M. P., Banich, M. T., & Barad, V. (2003). Competition for priority in processing increases prefrontal cortex's involvement in top-down control: An event-related fMRI study of the Stroop task. *Cognitive Brain Research*, *17*, 212–222.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., & Kramer, A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research, 12*, 467–473,

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.

Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology*, *10*, 465–501.

Nagel, I. E., Schumacher, E. H., Goebel, R., & D'Esposito, M. (2008). Functional MRI investigation of verbal selection mechanisms in lateral prefrontal cortex. *NeuroImage*, *43*, 801–807.

Nelson, J. K., Reuter-Lorenz, P. A., Sylvester, C. C., Jonides, J., & Smith, E. E. (2003). Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. *Proceedings of the National Academy of Sciences*, *100*, 11171–11175.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Vol. 4. Advances in research and theory* (pp. 1–18). New York: Plenum Press.

Novick, J. M., Thompson-Schill, S. L., & Trueswell, J. C. (2008). Putting lexical constraints in context into the visual-world paradigm. *Cognition*, *107*, 850–903.

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. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(3), 263–281.

Randolph, C., Braun, A. R., Goldberg, T. E., & Chase, T. N. (1993). Semantic fluency in Alzheimer's, Parkinson's, and Huntington's disease: Dissociation of storage and retrieval failures. *Neuropsychology*, *7*(1), 82–88.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing. *Journal of Verbal Learning and Verbal Behavior*, *22*, 358–374.

Robinson, G., Blair, J., & Cipolotti, L. (1998). Dynamic aphasia: An inability to select between competing verbal responses? *Brain*, *121*, 77–89.

Robinson, G., Shallice, T., & Cipolotti, L. (2005). A failure of high level verbal response selection in progressive dynamic aphasia. *Cognitive Neuropsychology*, *22*(6), 661–694.

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*, *15*, 1261–1269.

Saffran, E. M., Schwartz, M. F., Linebarger, M. C., Martin, N., & Bochetto, P. (1988). *The Philadelphia Comprehension Battery*, Unpublished test battery.

Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences*, *106*(1), 322–327.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*, 238–299.

Snijders, T. M., Vosse, T., Kempen, G., vanBerkum, J. J. A., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: An fMRI study using word category ambiguity. *Cerebral Cortex*, *19*, 1493–1503.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, *6*, 174–215.

Spivey, M., Tanenhaus, M., Eberhard, K., & Sedivy, J. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology, 45*, 447–481.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

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. *Proceedings of the National Academy of Sciences*, *94*, 14792–14797.

Thompson-Schill, S. L., Jonides, J., Marshuetz, C., Smith, E. E., D'Esposito, M., Kan, I. P., et al. (2002). Effects of frontal lobe damage on interference effects in working memory. *Cognitive, Affective, & Behavioral Neuroscience*, *2*, 109–120.

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. *Proceedings of the National Academy of Sciences*, *95*, 15855–15860.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology*, *11*(1), 138–146.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*, 89–134.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory & Language*, *33*, 285–318.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology*: *Learning, Memory and Cognition*, *31*(3), 443–466.

Weighall, A. R. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Child Psychology*, *99*, 75–95.

Ye, Z., & Zhou, X. (2008). Involvement of cognitive control in sentence comprehension: Evidence from ERPs. *Brain Research*, *1203*, 103–115.

Ye, Z., & Zhou, X. (2009). Conflict control during sentence comprehension: fMRI evidence. *NeuroImage*, *48*, 280–290.

Yee, E., Blumstein, S., & Sedivy, J. C. (2008). Lexical-semantic activation in Broca's and Wernicke's aphasia: Evidence from eye movements. *Journal of Cognitive Neuroscience*, *20*(4), 1–21.

Yerys, B. E., & Munakata, Y. (2006). When labels hurt but novelty helps: Children's perseveration and flexibility in a card-sorting task. *Child Development*, *77*, 1589–1607.

Zempleni, M., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *NeuroImage*, *34*, 1270–1279.