NeuroImage 84 (2014) 354–366

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Competitive mechanisms in sentence processing: Common and distinct production and reading comprehension networks linked to the prefrontal cortex

Gina F. Humphreys [a], Silvia P. Gennari [b,*]

[a] Neuroscience and Aphasia Research Unit, School of Psychological Sciences, University of Manchester, Zochonis Building 3F, Brunswick St., Manchester M13 9PL, UK
[b] Department of Psychology, University of York, Heslington, York YO10 5DD, UK

ARTICLE INFO

Article history:
Accepted 21 August 2013
Available online 3 September 2013

Keywords:
Language production and comprehension
Competition
Interactive networks
Production planning
Language processing
LIFG

ABSTRACT

Despite much interest in language production and comprehension mechanisms, little is known about the relationship between the two. Previous research suggests that linguistic knowledge is shared across these tasks and that the left inferior frontal gyrus (LIFG) may be commonly recruited. However, it remains unclear the extent to which production and comprehension share competition mechanisms. Here we investigate this issue and specifically examine competition in determining the event roles in a sentence (agent or affected participant). We used both behavioral and fMRI methods and compared the reading and production of high- and low-competition sentences, specifically targeting LIFG. We found that activity in *pars opercularis* (PO), independently identified by a competition-driven localizer, was modulated by competition in both tasks. Psychophysiological interaction analyses seeded in PO revealed task-specific networks: In comprehension, PO only interacted with the posterior temporal lobe, whereas in production, it interacted with a large network including hippocampal, posterior temporal, medial frontal and subcortical structures. Production and comprehension therefore recruit partially distinct functional networks but share competitive processes within fronto-temporal regions. We argue that these common regions store long-term linguistic associations and compute their higher-order contingencies, but competition in production ignites a larger neural network implementing planning, as required by task demands.

Introduction

Language production and comprehension are essential communicative behaviors involving message planning and understanding. Although a great deal of work in cognitive neuroscience has been dedicated to understanding language production and comprehension processes, very little is known about their relationship, particularly with respect to full sentence composition, a process absent in word-based studies. The present work aims to fill this gap and investigate the extent to which sentence production and reading comprehension share neural processes with particular reference to previously studied competition resolution mechanisms. Specifically, we examine competition in determining event roles in the sentence such as agents and affected participants, i.e., who is doing what to whom in the event.

Previous behavioral and imaging research has focused on the recruitment of similar lexical and grammatical information across production and comprehension (Bock et al., 2007; Chang et al., 2006; Gennari and MacDonald, 2009; MacDonald, 1999; Seidenberg and MacDonald, 2001). Gennari and MacDonald (2009), for example, have shown that structural configurations that readers find difficult are those that producers do not naturally produce, suggesting that the mapping of lexical concepts into sentence structure is shared across tasks. Similarly, a neuroimaging priming study comparing sentence production and auditory comprehension has shown that both tasks recruit a common fronto-temporal network including the posterior temporal lobe and the left inferior frontal gyrus (henceforth LIFG) (Menenti et al., 2011). However, the recruitment of common brain regions does not necessarily imply that similar processes take place in both production and comprehension — they could reflect the recruitment of common semantic and grammatical knowledge, common processes such as meaning access and sentence assembly or a mixture of both. It remains unclear therefore whether and how subcomponent processes in production and comprehension are related, and in particular, whether competitive processes are shared across the two tasks, despite the expected differences in input and output networks.

To elucidate the extent to which production and comprehension share cognitive and neural processes, rather than simply linguistic and/or grammatical knowledge, here we chose to focus on processes that arguably involve competition between alternative representations.

---

[*] Corresponding author at: Department of Psychology, University of York, York, YO10 5DD, UK.
E-mail address: silvia.gennari@york.ac.uk (S.P. Gennari).

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.08.059

Such competition may occur in comprehension when multiple alternative interpretations are activated and resolution between them takes place. In studies of lexical ambiguity, for example, equally frequent meanings of an ambiguous word like *bowl* are initially activated, and competition resolution is required to arrive at the intended meaning, often with the help of contextual cues (Duffy et al., 1988; Gennari et al., 2007; Rayner and Duffy, 1986; Rayner et al., 1983; Rodd et al., 2005, 2010; Seidenberg et al., 1982; Swinney, 1979; Tanenhaus et al., 1979). At the sentence level, similar semantic competition processes take place when alternative conflicting interpretations are activated as the sentence unfolds (Novick et al., 2005; Rodd et al., 2010; Tanenhaus and Trueswell, 1995). Competition resolution may also occur in production when alternative utterances, e.g., alternative sentence structures or words, are associated with the intended meaning (Gennari et al., 2012), or in the case of cued production, when alternative utterances are strongly associated with the cue (Barch et al., 2000; Spalek and Thompson-Schill, 2008). For example, in elicitation studies in which participants must produce a verb/action in response to a cuing noun, strong associations like *scissors-cut* are easier to utter than weaker associations, e.g., *wheel-turn*, because these weaker associations compete with other available alternatives (Barch et al., 2000; Thompson-Schill et al., 1997). Cue-based production, which is used in the present study, is thus similar to ambiguity resolution in comprehension in that the appropriate response to the cue is essentially ambiguous, as more than one alternative utterance is activated (see for example numerous picture–word interference studies (Costa et al., 2005)). Therefore both production and comprehension can involve competition between available alternative interpretations or utterances. Finding evidence of brain regions and behaviors that are similarly sensitive to conflicting demands in both production and comprehension would suggest that processes of a competitive nature are shared across tasks.

Although some neurocognitive studies suggest that competition occurs in both production and comprehension, particularly within LIFG, psycholinguistic approaches have typically argued that these tasks involve distinct component processes that act upon a shared linguistic knowledge base (Bock and Levelt, 1994; MacDonald et al., 1994). Indeed, production requires word retrieval and planning of a sentence structure, whereas comprehension involves word recognition and the understanding of syntactic–semantic relations between words. These processes have long been argued to give rise to marked asymmetries between these tasks. Sentence production for example often involves false starts, disfluencies and errors, which are thought to arise from production planning and articulation processes not shared with comprehension (Bock and Levelt, 1994; Garrett, 1980; Levelt et al., 1999). Furthermore, with specific reference to competition, sentence comprehension involves activating multiple alternative interpretations of polysemous stimuli (MacDonald et al, 1994), but in production polysemy is sometimes unnoticed by speakers, who have a single intended meaning in mind (Ferreira et al., 2005). This suggests the possibility that generally assumed competition mechanisms in fact differ across production and comprehension tasks. Differences in the types of tasks faced by speakers and comprehenders might lead to differences in the underlying processing systems. These observations are consistent with many production studies recruiting partially distinct neural networks from those typically found in comprehension studies, including supplementary motor regions, the cingulate cortex, and subcortical structures that are associated with motor responses and decisions (Barch et al., 2000; Sörös et al., 2006; Tremblay and Small, 2011; Whitney et al., 2008). Nonetheless, patients with pre-frontal or LIFG lesions show both production and comprehension impairments in cognitively demanding tasks (Jefferies and Lambon Ralph, 2006; Novick et al., 2009) and both readers and speakers show similar brain correlates of assumed competition processes in LIFG and nearby regions, albeit across different studies (Bedny et al., 2008b; de Zubicaray et al., 2001; Heim et al., 2009; Indefrey, 2011; Spalek and Thompson-Schill, 2008; Tyler and Marslen-Wilson, 2008). This thus indicates that both tasks may recruit the LIFG under competition demands (Thompson-Schill et al., 2005).

Together these findings suggest that the LIFG may play a role in both production and comprehension when resolving competition, but that there may also be differences in the networks implementing competition in each task and cooperating with LIFG. To evaluate this possibility, we used behavioral measures (Experiment 1) and functional magnetic resonance imaging (Experiment 2) to examine the networks interacting with LIFG in the production and reading comprehension of sentences taken to vary in competition demands (high- vs. low-competition). We used reading rather than auditory comprehension because reading is more appropriate for a noisy environment like the MRI scanner, this task circumvents debated design decisions such as appropriate baseline (Adank, 2012) and because many studies comparing reading and listening with non-competitive stimuli report little or no differences in supramodal language regions such as LIFG and the posterior middle temporal gyrus that are likely to be engaged in processing the present stimuli (Adank, 2012; Buchweitz et al., 2009; Jobard et al., 2007; Lindenberg and Scheef, 2007; Michael et al., 2001). Example stimuli can be found in Table 1.

In the reading comprehension task, we chose subordinate clauses such as *the director that the movie impressed*… and *the movie that the director watched*… for our stimuli because these structures have been amply demonstrated to vary in difficulty in previous studies (Mak et al., 2002, 2006; Traxler et al., 2002). Specifically, high-competition clauses take longer to read and elicit more activity in LIFG than low-competition ones (Caplan et al., 2008; Chen et al., 2006; Gennari and MacDonald, 2008). This difficulty plausibly arises because conflicting lexical and structural cues lead to incorrect predictions regarding the event roles in the sentence: animate and inanimate nouns like *director* and *movie* cue an agent and an affected-participant event role respectively, however structural markers and the subordinate verb cue the opposite roles. The first high-competition structure in Table 1, for example, is incompatible with an interpretation in which the director writes or watches the movie — the situations most typically associated with these nouns at the point at which the subordinate verb *impressed* is encountered. At this point, readers anticipate potential relations between *director* and *movie*, e.g., the director writing or watching the movie, or the movie being about the director, none of which is the interpretation implied by the verb *impress* in the subordinate clause, generating conflict between currently available and incoming interpretations. As in cases of ambiguity, alternative interpretations are entertained at this point, as the reader must inhibit the initially constructed interpretation in favor of the one cued by incoming words. In the low-competition condition, in contrast, the nouns and the structure lead to the ultimately correct interpretation in terms of event roles (the movie being watched by the director). These interpretation preferences have been

Table 1
Example stimuli for the production and the comprehension task.

Task	Condition	Example stimulus items
Comprehension	High-competition	1. The director that the movie impressed had received a prize 2. The banker that the loan worried was short of money
	Low-competition	1. The movie that the director watched had received a prize 2. The loan that the banker refused was a lot of money
Completion	High-competition	1. The director that the movie … 2. The banker that the loan ….
	Low-competition	1. The movie that the director … 2. The loan that the banker ….

demonstrated by eliciting readers' interpretations at each point in the subordinate clause. Critically, the number of alternative interpretations available at the point in which the subordinate verb is encountered predicts reading times at subsequent word positions, thus supporting competition claims (Gennari and MacDonald, 2008).

In the production task, we used a cue-based paradigm in which participants completed sentence fragments from the comprehension stimuli into full sentences (Table 1). This production task showed parallel competition results to those obtained in comprehension. Specifically, here we found that high-competition fragments take longer to complete than low-competition fragments and that the number of alternative continuations or plans produced for these fragments (by different speakers) correlates with production initiation times. Taken together, the present and previous findings with our stimuli strongly suggest the presence of competition between alternatives in interpreting the sentences or providing a continuation for prompt fragments (production). Otherwise, it would be unexplained why activating more alternatives leads to longer processing times, and why the number of alternatives available correlates with performance in both tasks, given that only one of the alternatives is eventually produced or understood. Indeed, ambiguity resolution processes and cue-based multiple-alternative productions have long been conceived as competitive processes in psycholinguistics and cognitive neuroscience, both at the lexical and the sentential level (Costa et al., 2005; de Zubicaray et al., 2001; MacDonald et al., 1994; Novick et al., 2005; Simpson, 1994; Spalek and Thompson-Schill, 2008), given that such cases appear distinctive from other difficult processes, for example, the difficulty experienced in retrieving the meaning of a low frequency word. Our behavioral production and comprehension tasks therefore provide appropriate support to examine the brain correlates of competition mechanisms in each task.

In Experiment 2 using fMRI, we conducted the same production and reading comprehension tasks as above, except that production was covert to minimize movement. We also conducted a localizer Stroop task, which was primarily used to identify a relevant region of interest associated with competition processes within LIFG — a functionally and anatomically heterogeneous region (Amunts et al., 2010; Chein et al., 2002; Dapretto and Bookheimer, 1999; Fedorenko et al., 2012; Petrides, 2005). We focused on LIFG rather than other regions typically activated by the Stroop task (Milham et al., 2001) because, as suggested above, portions of the LIFG, unlike other regions of the Stroop network, are very likely candidates to be recruited by both production and comprehension, as they have already been argued to play a regulatory, control or unifying role in many models of language processing (Hagoort, 2005; Indefrey, 2011; Tyler and Marslen-Wilson, 2008). The Stroop task was designed in such a way that allowed the identification of voxels sensitive to the representational conflict in this task, rather than the response conflict (Hindy et al., 2012; January et al., 2009). Within the LIFG, we identified these voxels for each individual participant and evaluated the activity elicited by the conditions in each production and comprehension. Individual-based analyses guarantee a high degree of anatomical specificity that is not influenced by normalization procedures or anatomical differences across participants (Fedorenko et al., 2010; Glezer and Riesenhuber, 2013; Juch et al., 2005; Nieto-Castañón and Fedorenko, 2012; Saxe et al., 2006; Swallow et al., 2003). Additionally, we conducted psychophysiological interaction analyses using the identified LIFG cluster as seed region to determine the networks cooperating with it as a function of competition condition.

Based on the view outlined above, if the LIFG is involved in resolving competition in both production and comprehension, we predicted that the levels of competition in both production and comprehension should modulate activity within a region of the LIFG also sensitive to semantic competition in the Stroop task (localizer). Critically, if each task implements competition processes differently and contains specific component processes, as suggested by psycholinguistics research, distinct task-specific neural networks should interact with the identified portion of LIFG as a function of competition resolution demands, i.e., the competition networks revealed by the psychophysiological interaction analysis should differ according to task. These predictions assume that our stimuli indeed capture competitive processes, as suggested by our behavioral results and previous literature, and that the Stroop task, and particularly, the specific contrast utilized, also captures some aspects of the conflict resolution between representations within LIFG (Hindy et al., 2012; January et al., 2009; Milham et al., 2001).

Methods

Experiment 1: behavioral study

Participants

39 native English speaking students at the University of York participated in the experiment and received course credit for their participation (mean = 22.5 years, SD = 3.08, range = 18–32, 22 females).

Materials

Items in the comprehension task had two versions varying in the configuration of the noun animacy and the subordinate verb within the subordinate clause, yielding a high-competition and a low-competition version (see Table 1). As indicated above, the syntactic configuration of the animate and inanimate nouns in the high-competition condition causes competition between alternative interpretations as the sentence unfolds: the event roles strongly associated with the animacy of the nouns (animate and inanimate nouns are typically agents and affected-participants respectively) must be inhibited in favor of the infrequent event role interpretation provided by the subordinate verb (Gennari and MacDonald, 2008; Mak et al., 2002; Traxler et al., 2002, 2005). Previous work has carefully documented the presence of competition between alternative interpretations involved during reading and is consistent with similar views proposed in the cognitive neuroscience literature (Novick et al., 2005).

There were 84 stimulus sentences in total. 42 experimental items with subordinate clauses such as those in Table 1 were used, along with 42 filler sentences. Participants were asked to read these sentences and fully understand their meaning (see below for details). The nouns and the verbs of the stimulus sentences were matched for length and frequency across conditions (see Table 2). The majority of the filler sentences (80%) were subject–verb–object main clause structures. The remaining filler sentences included some form of subordinate or conjoined structure.

Items in the production task were similar to those in the comprehension task but only included the initial fragment of the sentences (Table 1). The task instructions required the participant to complete the fragment into a full sentence (see below for details). This thus required speakers to assign event roles to the nouns of the fragment and to select an appropriate verb phrase to continue it that matched the meaning of the nouns. Our previous and present results suggest that more alternative continuations are available in the high-competition fragments than the low-competition ones (Gennari and MacDonald,

Table 2
The mean log word frequencies and length for the nouns and verbs from the high-competition and low-competition conditions.

Condition	Example	Log word frequency	Length
High-competition noun	*Director*	4.18	6.88
Low-competition noun	*Movie*	4.27	6.02
High-competition verb	*Impressed*	4.05	7.64
Low-competition verb	*Watched*	4.08	7.57

2008, henceforth G&M), and thereby, more competing utterance plans, within the semantic and syntactic constraints of the fragment.

Task design

Two tasks, sentence comprehension and sentence completion, were conducted within the same experimental session arranged in two separate blocks with separate instructions. For each task, the items were split across two lists with each list containing 42 experimental items (21 high-competition and 21 low-competition sentences) and 42 filler sentences. Each participant saw only one list for each task. The lists were constructed in such a way that the high-competition and low-competition versions of the same item would be on a separate list for a given task (Latin-square design), but the two versions would be seen once in the production task and once in the comprehension task. The order of the items was randomized and randomly intermixed with filler sentences. The order of the production and the comprehension task was counterbalanced across participants.

Procedure

In the comprehension block, participants were told to read the words one at a time in a self-paced moving-window paradigm (G&M) and fully understand the meaning of the sentence. After each sentence, a comprehension question was asked, requiring a YES/NO answer (e.g., *did the movie impress the director?*). 53% of these questions referred to the content of the subordinate clause. Reading times and comprehension accuracy were recorded. In the production block, participants read the fragments one word at a time as in comprehension, and were instructed to fully understand their meanings before initiating their completions out loud. We recorded their spoken responses to compute response accuracy, and the time it took them to initiate or plan their utterances, as measured from the offset of the last word processed in the initial fragment. We also checked that participants read each word of the fragment carefully, as instructed, rather than waiting until the end of the phrase to comprehend the fragment, by comparing the reading times of the fragments across tasks. Reading times for the fragments in the production task were indeed much slower than in the comprehension task (for contrasts at all word positions, p < .001). In particular, the reading times of the last word of the fragment (the last noun) was on average 2205 ms longer in the completion task than in reading comprehension, thus suggesting that participants read carefully in anticipation of, or in preparation for, their upcoming completion task.

Data coding and analysis

To analyze the production responses, we identified completion errors by simple inspection of the transcribed utterances. Continuations that did not make sense were considered errors. For example, for a fragment such as *the scientist that the book*, the completion *had written* was considered an error, as books typically do not write scientists. Two independent researchers coded the data and they agreed on more than 95% of the cases coded. Cases of disagreement were discussed among researchers and a final coding was jointly decided. Additionally, we computed a measure indexing competition: the proportion of different verbs produced across speakers for each item in the completion task (out of the total number of completions for that item). This measure represents how many different verb continuations there are in general for each item and thus, reflects the number of alternative continuations potentially entertained during planning, even though a given participant would have ultimately produced only one of them. Two continuations were considered to contain the same verb if the same lexical root of the verb was used, independently of tense or auxiliaries (e.g., *hit*, *had hit*). For example 1 in Table 1, the continuations were *displeased*, *pleased*, *impressed*, *upset*, *used*, *was about*, *watched*, *was directed by*, and *had watched*, which amounted to 8 different completions (0.88 proportion). In our analyses, we used this measure to predict planning difficulty.

Data analyses of processing times (planning and reading times) were all carried on correct responses only (correct production trials and sentences for which participants answered the comprehension question correctly). For the analysis of reading times, residual reading times were used as the dependent measure, as standard in psycholinguistics, although raw reading times display the same effects, as shown in Fig. 1. Residuals are calculated from regressing out word length from raw reading times (G&M). For the analysis of production planning times, we used log transformed times to approximate normally distributed data points, although analyses of raw planning times yield the same results. Responses that deviated by more than 3.5 SDs from the mean initiation times were removed from the analysis.

Results

Comprehension task

Results indicated that as expected, high- vs. low-competition sentences differ in comprehension difficulty: participants were less accurate and reading times were slower for high-competition sentences than low-competition sentences, and there was no interaction between task order and competition conditions. Fig. 1 plots the reading times for each word position as a function of condition. The overall mean accuracy across participants for the comprehension questions was 89% (SD 2.71). Response accuracy for the high-competition items was significantly lower than the low-competition items suggesting increased comprehension difficulty in the high competition case (by-item analysis: high-competition M = 84%, SD = 16.15; low-competition M = 94%, SD = 11.22; t(39) = −3.23, p = .003; by-participant analysis: high-competition M = 84.33%, SD = 10.41; low-competition M = 93.86%, SD = 6.60; t(38) = 5−.01, p < .0001). A repeated measures ANOVA with residual reading time as the dependent variable and condition (high- vs. low-competition) and word-position (subordinate noun, e.g. *movie*, subordinate verb, e.g. *pleased*, and the main verb region, e.g. *had received a*) as factors revealed a main effect of condition (by-item analysis: F(1, 39) = 7.55, p = .009, $\eta p^2 = .16$; by-participant analysis: F(1, 38) = 10.34, p = .003, $\eta p^2 = .21$). No main effect of position was found and there was no condition by position interaction. Planned comparisons showed significantly slower residual reading times for the high-competition vs. low-competition condition at the subordinate verb (by-item analysis: t(39) = 2.31, p = .026; by-participant analysis: t(38) = 2.54, p = .015), and also at the main verb region (by-item analysis: t(39) = 2.80, p = .008; by-participant analysis: t(38) = 3.85, p < .0001). Finally, a mixed ANOVA with residual reading times as dependent variable, competition conditions as repeated factors and task order as between subjects factor revealed no interaction of competition with task order at any of the critical verb positions. Taken together, these results replicate previous findings in the literature (Gennari and MacDonald, 2008; Mak et al., 2002; Traxler et al., 2002) confirming the suitability of our stimuli, and suggest that as the sentence proceeds, readers encountered more difficulty at the subordinate verb position, which indicates who is doing what in the subordinate structure. This difficulty continues into the main verb phrase of the sentence.

Production task

Results indicated that high- vs. low-competition sentences differ in production difficulty: response accuracy was lower and initiation times longer for the high-competition condition than the low-competition condition and there was no interaction between competition conditions and task order. Across participants and items, there was a significantly lower proportion of correct responses for the high-competition condition than the low-competition condition (by-item analysis: high-competition M = .81, SD = .16; low-competition M = .97, SD = .06; t(39) = −6.47, p < .001; by-participant analysis: high-competition M = .81, SD = .21; low-competition M = .97, SD = .09; t(36) = −5.92, p < .001).

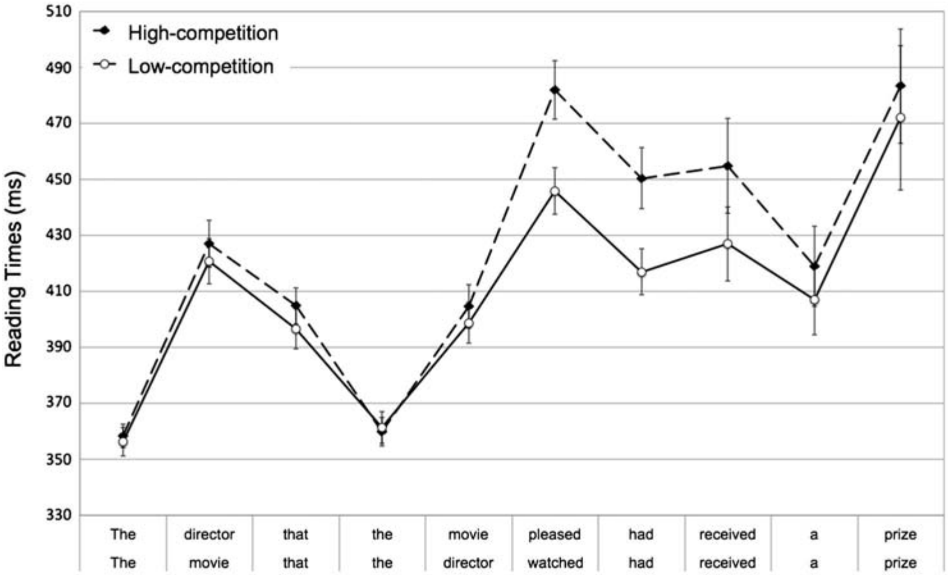

Fig. 1. Reading times as a function of condition and word position in Experiment 1. Error bars indicate standard error.

Interestingly, the majority of errors in these responses (60.4%) were event role errors, i.e., errors in which the produced verb implies the wrong sentential roles for the nouns of the fragment, as in *The director that the movie… "watched was good"*. These errors clearly suggest that participants found it difficult to inhibit prepotent verb responses associated with the nouns, e.g. the verb *watch* is strongly associated with *director* and *movie*. Moreover, log transformed planning times were longer for the high-competition compared to low-competition items (item analysis: t(39) = 5.13, p < .001; subject analysis: t(38) = 6.51, p < .001). Raw planning times are shown in Fig. 2. These results remained significant after controlling for utterance length in a hierarchical regression analyses across items. Length alone explained 13% of the variance in initiation times, a significant predictor (F(1, 78) = 11.73, p < .001), but adding condition (high-competition vs. low-competition) into the model significantly increased the variance explained to 22% (F change (1, 77) = 8.85, p < .005). This confirms that high-competition fragments are more difficult to complete, irrespective of utterance length. Finally, a mixed ANOVA with planning time as the dependent variable, competition conditions as repeated factors, and task order as between subjects factor revealed no interaction of competition with task order (F(1,37) = 1.37, p = n.s.).

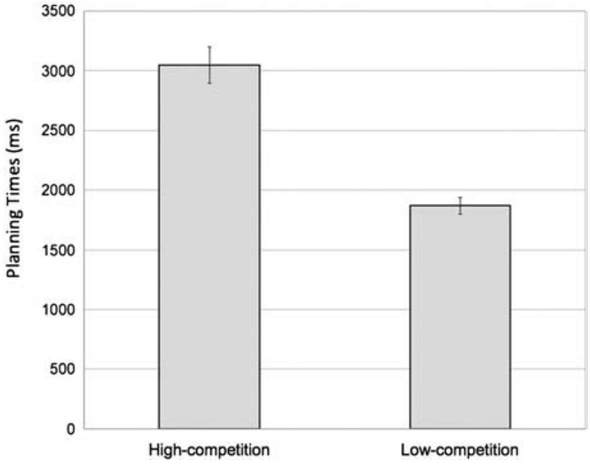

Fig. 2. Production planning times as a function of condition in Experiment 1. Error bars indicate standard error.

Correlation

To shed light on whether the differences in difficulty that we have revealed in production planning suggest underlying competitive processes, as previously reported for the comprehension of our items, we correlated the number of alternative continuations with initiation times. We used the proportion of different verbs produced across speakers for each item in the completion task (see Data coding and analyses) and regressed it onto the average planning times for each item, including both high- and low-competition cases. We therefore expected that if competition between alternative continuations took place during production planning, this correlation should be significant. Indeed, the proportion of alternative verbs accounted for 40% of variance in planning times (p < .0001). This result parallels previously reported results in the comprehension of these structures in that the number of interpretations entertained up to the relative clause verb predicts reading difficulty at subsequent word positions (G&M).

Taken together, these results suggest that both production and comprehension show processing difficulty in the high-competition condition and both tasks are sensitive to the number of alternative interpretations or plans entertained, thus strongly suggesting competitive processes.

Experiment 2 — fMRI study

Participants

17 right-handed native English speaking students at the University of York, who did not take part in Experiment 1, participated in this study. All participants had no history of psychiatric illness or neurological damage (10 males, 7 females, mean age = 23.2 years, SD = 3.72, range = 19–33 years).

Materials

The same 42 items and fragments used in Experiment 1 (see Table 1), were used in this experiment (84 experimental items in total).

Task design

Three fMRI scans were conducted: comprehension task, production task, and a color Stroop task. The Stroop task was used as a localizer to identify a region of interest involved in semantic competition resolution within LIFG. The order of the production and comprehension scans was counterbalanced across participants, as in Experiment 1, and the Stroop task was always completed last. The production and comprehension scans were presented in an event related design. The most efficient ordering of events was determined using Optseq (http://www.freesurfer.net/optseq). Null time was intermixed between trials and randomly varied between 0 and 18 s, with an averaged inter-trial time of 4.67 s (Josephs and Henson, 1999). Every event type followed each other with equal probability (e.g., for three event types, every type followed every other across the experiment 33% of the time). Due to practical time constraints, we did not include fillers in this experiment.

Reading comprehension task. In each event a sentence was visually presented (white font on a black background) split into two consecutive halves of 3 s each. The first half contained the subordinate clause (e.g. *The movie that the director watched*) and the second half contained the remaining of the sentence. Whole sentence presentation has been successfully used with relative clause stimuli both in behavioral studies (Traxler et al., 2002, 2005) and fMRI studies (Chen et al., 2006; Caplan et al., 2008). This is a more appropriate paradigm to use with fMRI, compared to self-paced reading, given that activity to individual words in the sentence could not be estimated in any case due to the sluggishness of the hemodynamic response. Furthermore, we did not follow the self-paced method of Exp. 1 because we did not want to contaminate reading comprehension activity with motor activity involved in button presses, which could possibly make comprehension more similar to production. In order to ensure that the participants were processing the meaning of the sentences, 16 catch-trials were randomly intermixed with experimental trials. Catch trials took the form of YES/NO comprehension questions as in Experiment 1. The participants were given 4 s in which to answer the question. Between trials a series of "+"s was presented which was the same length as the average sentence length. This acted as a visual baseline. In total the scan lasted 906 s.

For each subject, the high-competition and low-competition version of an item were assigned to different halves of the experiment so that if one version appeared in the first half of the experiment, the remaining version appeared in the second half. The two halves of the experiments were counterbalanced across subjects so that each subject saw the two versions of an item in a different order. Across all items, half were presented first in the high-competition condition and half in the low-competition condition, so that in each half of the comprehension task, there were equal numbers of sentences from each condition. This careful counterbalancing eliminates potential influences of item order on both group statistics (each participant saw a different order) and the average activity for a condition within a participant (across items, each half was presented in a different order). This design followed Chen et al. (2006) and Caplan et al. (2008) who have found positive results using a similar paradigm.

Production task. Each trial presented a sentence fragment for 2 s after which the symbols "???" appeared for 6 s. Participants were instructed to covertly complete the sentence fully and meaningfully when they saw these symbols and to press a button when they were finished. Note that this task was not aimed at removing articulatory aspects of production, since overt and covert production sometimes similarly activate common regions, e.g., supplementary motor cortex and insula, but not LIFG (Huang et al., 2001). On the contrary, whatever motor planning and articulatory processes take place in production ought to be reflected in the production network, but not in comprehension. To act as a reading baseline, 84 read-only items were intermixed with the production items. Here, the participant was instructed to simply read the fragment for meaning and press a button when they were finished reading. These items and trials were identical to the completion fragments except that they were presented in red font and followed by "XXX". After practice, the font color thus clearly indicated from the beginning of the trial whether participants would complete or read the stimulus fragments. By contrasting the activity from the completion trials and the read-only trials, we aimed to remove the reading component of the trial, revealing only production-related activity. Note that regions found to be active only in the completion task, as compared to the comprehension task above, which constitute the majority of our results, cannot be attributed to any influence of the reading component of the completion task. If for example, the reading component of the completion task was not successfully removed (via the contrast between the completion trials and the read-only trials), we would expect the production results to be similar to the comprehension results from the main comprehension task above, which was clearly not the case.

The production scan was divided into two runs, one lasting 842 s and the other lasting 830 s. As in the comprehension task, the order for the runs was counterbalanced across participants. The runs were structured such that if a high-competition production item was presented in run 1, then the read-only version of this item would be presented in run 2 and vice versa. Also, the items from the high- and low-conditions were ordered such that if the high-competition version of an item was presented in the first run, then low-competition version would not be presented until the second run and vice versa. Thus the order in which the conditions occurred was also counterbalanced across runs. Across all items, half were presented first in the high-competition condition and half in the low-competition condition, so that in each half of the comprehension task, there were equal numbers of sentences from each condition. Additionally, the order of event types was determined by Optseq and was such that each event type followed each other equally often. This careful counterbalancing of conditions across runs and participants makes it unlikely that content repetition between some items (e.g., read-only and completion trials) would have an influence on the group results. Similarly, any potential influence of task changes in the production session (completion vs. reading-only) would not unduly influence any specific condition, thus allowing safe comparisons between high- and low-competition trials.

Stroop task (localizer). This task was taken from previous studies (Milham et al., 2001), and has successfully been used in similar language studies in the literature (Hindy et al., 2012; January et al., 2009). The responses were restricted to three font colors — yellow, green and blue. The task was to indicate the font color by pressing a button on a button box, whilst ignoring the word meaning. There were three conditions: incongruent-eligible, incongruent-ineligible, and neutral. In the incongruent-eligible trials, the word and ink color were incongruent and the word denoted a color that was a potential response (yellow, green, or blue). In the incongruent-ineligible trials, the word denoted a color that was not a potential response (orange, brown, or red). The distinction between eligible and ineligible trials is that eligible trials will involve both conflict at the motor response level (since the word indicates a possible response), as well as conflict at the representational level (since both the word and the ink involve color representations). However, ineligible trials will only involve conflict at the representational level (January et al., 2009). Neutral trials consisted of non-color words that were length and frequency matched to the color items (e.g. *plenty*, *horse*, *deal*). The trial structure and design were identical to that in January et al. (2009) and Hindy et al. (2012).

Acquisition parameters. Imaging was carried out using a 3 T Signa Excite MRI scanner at the York Neuroimaging Centre (YNiC). High-resolution whole brain T1-weighted structural images were obtained for all participants (1 mm × 1 mm × 1 mm). Functional images were obtained using a gradient-echo EPI sequence with 38 contiguous axial slices per volume (TR = 2000 ms, TE = 20 ms, flip angle 90°, matrix 64 × 64, FOV 24 cm, slice thickness = 3.5 mm). A T1-weighted FLAIR image was also obtained to aid co-registration.

Data analysis. Data analyses were carried out using FSL tools (the software library of the Oxford Centre for Functional MRI of the Brain (FMRIB); www.fmrib.ox.ac.uk/fsl). First- and higher-level analyses were carried out using FEAT (FMRI Expert Analysis Tool). The two production runs were combined by concatenating their demeaned data. Pre-processing of the data included motion correction (Jenkinson et al., 2002), slice-timing correction, brain extraction, spatial smoothing using a Gaussian kernel of full-width-half-maximum (FWHM) 8 mm, and high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, sigma = 25.0 s). Time series analyses were conducted using FILM (FMRIB's Improved Linear Model) with local autocorrelation correction (Woolrich et al., 2001).

The data from each task were modeled separately with event-based explanatory variables corresponding to the condition and trial structure, which were convolved with a hemodynamic response function (gamma function). For all tasks, hemodynamic response functions (HRFs) were aligned to the beginning of the event and lasted for the duration of the event except for the production trials, which were modeled as lasting 4.5 s. This aimed to capture the reading portion of the trial (2 s) plus the planning and initial encoding of the utterance, rather than the articulation. This timing was determined using data from Experiment 1 (the mean initiation times were around 2.5 s, see Fig. 2). Thus, if there were length differences across participants' utterances, there were not taken into account in the modeling, which only captured the initial planning.

Several contrasts were computed for each data set. For the comprehension task, we contrasted all the sentences relative to the visual baseline (language-baseline), and each condition separately relative to baseline (high-competition-baseline; low-competition-baseline). We performed similar contrasts on the production data but this time contrasting the experimental trials with the read-only control trials. There were three production contrasts: all production > control, high-competition production > control, low-competition production > control. Higher-level group analyses comparing activity across conditions or across tasks were conducted by entering participants' lower-level contrast parameter estimates into high-level statistics (cluster corrected, $z > 2.3$, $p = .05$). In turn, the Stroop data was modeled with covariates for each condition. The main contrast of interest was the representational competition contrast in which incongruent ineligible items were contrasted with neutral items.

Region of interest analysis. The Stroop data was used to identify a portion of LIFG most strongly responding to semantic conflict in each participant, using the contrast between ineligible vs. neutral trials to capture the representational conflict (see description of the Stroop task above). This portion, which turned out to be within the *pars opercularis* (BA44) was significantly (and more strongly) active in a group level analysis (cluster corrected $z > 3$, $p = .05$), and was also consistently activated in all participants, as revealed by individual results. This region is considered to play a central role in competition resolution in language processing and other cognitive tasks (Miller and Cohen, 2001; Thompson-Schill et al., 2005). For our region of interest analysis, we identified this region in each individual brain by visual inspection of the high-resolution contrast images for the ineligible vs. neutral Stroop trials. We initially drew a mask on the active cluster falling within the *pars opercularis*, which was delimited by the inferior precentral sulcus posteriorly, the inferior frontal sulcus dorsally, the lateral fissure ventrally and the vertical ramus of the lateral fissure anteriorly. From these clusters, we then extracted the 35 most active voxels by thresholding the clusters at the appropriate level for each participant and transformed each resulting ROI into each individual's functional space. Within each ROI, we computed average measures of language activity for each participant. For the comprehension results, we extracted the mean contrast parameter estimates for the high-competition vs. baseline and low-competition vs. baseline contrasts. For the production results, we extracted the corresponding parameter estimates, i.e., high-competition vs. read-only control and low-competition vs. the read-only controls. These mean contrast parameter estimates were then converted into percent signal change and used as dependent variables in random-effect group analyses reported for each task to test whether there was an effect of condition across participants.

Psychophysiological interaction analyses. In order to examine the networks that are engaged during competition resolution in the production and the comprehension tasks, we conducted functional connectivity analyses using the psychophysiological interaction method. Specifically, we aimed to determine whether the functional coupling between the *pars opercularis* and other brain regions interacts with task demands (levels of competition) in different ways for production and comprehension. This allows an examination of differences and similarities across production and comprehension in the underlying connectivity of the *pars opercularis* during competition resolution. The PPI analyses were carried out in FSL using the methods described in the FSL documentation. A mask of the *pars opercularis* cluster obtained from the Stroop group results was used as the seed region. First level analyses were carried out in FEAT for each of the tasks. The average time course of the seed region was used as physiological variable in the analysis of both the production and comprehension data. In comprehension, the contrast high-competition > low-competition was used as the psychological variable (using the contrast 1, −1) to examine interactions of our physiological regressor with levels of competition. In the production task, the psychological variables was the contrast of the high-competition condition with both the low-competition condition and the reading control condition (2, −1, −1). The psychological regressor was convolved with the gamma function and temporal derivatives with temporal filtering added. The PPI regressor was the interaction term between the zero-centered psychological regressor and the de-meaned physiological regressor. (Note that FSL differs from SPM in convolving the psychological regressor, which is treated as nuisance variable in the interaction, rather than deconvolving the physiological regressor. Although this has been suggested to lead to loss of statistical power for event-related designs, the results that do survive are nevertheless valid (Gitelman et al., 2003)). To evaluate group statistics, including comparisons across tasks, higher-level analyses were carried out on participants' lower-level contrast parameter estimates using a threshold of $p < .001$ (uncorrected) and of $p = .05$ (cluster corrected, $z = 2.3$).

Results

Behavioral performance during scanning sessions

During the comprehension task, participants responded correctly to the catch trials 90% of the time on average, indicating that they maintained attention on the task. In the production task, participants took 5.35 s on average to complete the fragments and took about 2.60 s to read the fragments in the read-only control trials. Both these measures are consistent with our behavioral results in Exp. 1. Interestingly, there was a significant difference between the completion times of the high- and low-competition fragments, as measured from the beginning of the trial ($M = 5.58$ vs. 5.11 s, $t(16) = 6.06$, $p < .0001$). In the Stroop localizer, we observed the typical competition effects in which incongruent trials, both eligible and ineligible, elicited longer response times than the neutral trials (M incongruent (eligible and ineligible) = 689 ms, M neutral = 627 ms, $t(16) = 2.17$, $p < .05$). Overall, the behavioral measures suggest that participants remained focused on the tasks at hand and followed the instructions.

Whole brain results

To examine the extent to which the comprehension and production networks overlap, we conducted whole brain analyses contrasting each task vs. baseline or controls (cluster-corrected, $Z > 2.3$, $p < .05$). These broad contrasts guarantee that potential commonalities across tasks are not missed in more specific or direct contrasts, where common activity may be canceled out. These analyses revealed that both sentence comprehension and production recruited a wide network of

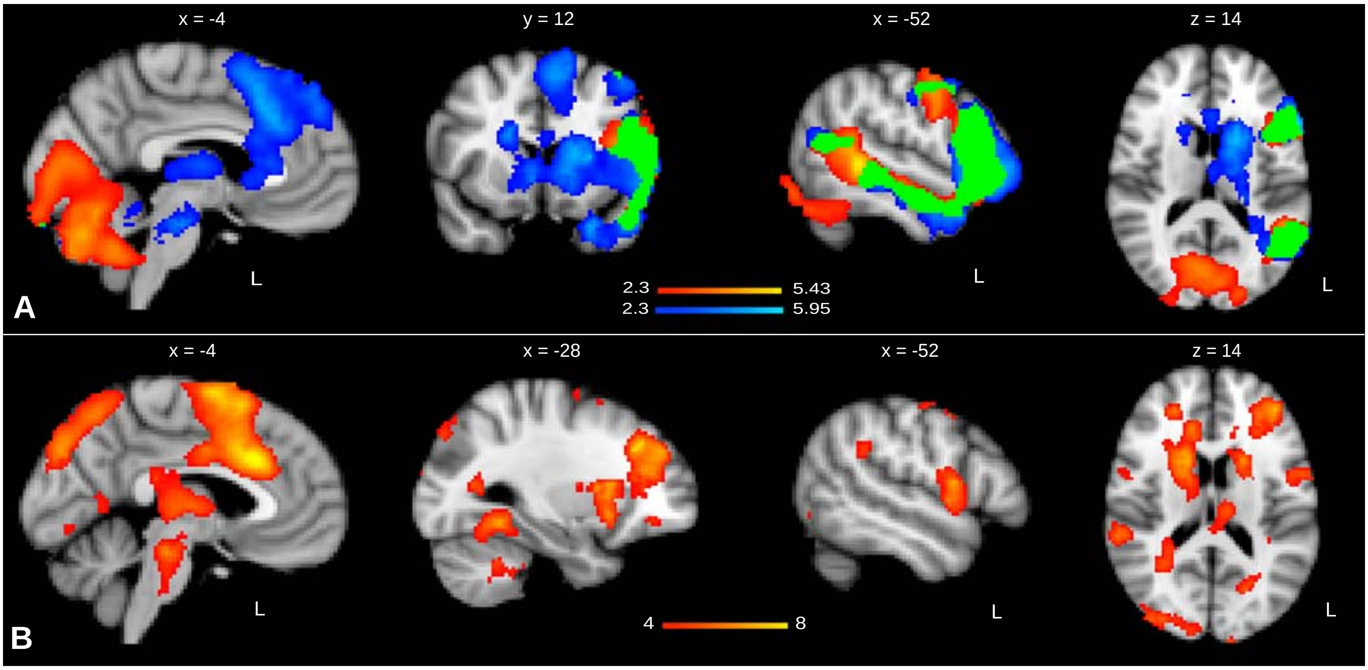

Fig. 3. Whole brain analysis from Experiment 2 showing in panel A the production (blue) and comprehension (red) networks (cluster corrected, $z > 2.3$, $p < .05$). The overlap (conjunction) between the two networks is shown in green. The production results show the contrast between all sentences and the reading controls. The comprehension results show the contrast between all sentences vs. the visual baseline. Panel B shows in red the direct contrast production > comprehension.

left-lateralised fronto-temporal areas. Fig. 3, panel A, shows the comprehension results in red, the production results in blue and the overlap between the two in green, which corresponds to a conjunction analysis as suggested by Nichols and colleagues (Nichols et al., 2005) (see also Table 3). Both tasks shared the involvement of several regions — the left inferior frontal gyrus (BA44, BA45, BA47), the precentral gyrus (BA6) and the posterior and anterior divisions of the middle temporal gyrus — although more extensive pMTG activity was found for comprehension. This is consistent with multiple studies reporting this core fronto-temporal network in language processing, which is thought to implement interactions between lexico-semantic retrieval and prefrontal control or competition processes (Gennari et al., 2007; Tyler and Marslen-Wilson, 2008; Whitney et al., 2011). However, the production task also recruited an additional network of medial structures, including the supplementary motor area (SMA, BA6), the anterior cingulate cortex (ACC), the caudate and the thalamus, which were not active during comprehension, as can be seen in Fig. 3. These regions have long been associated with production (Eickhoff et al., 2009; Ketteler et al., 2008; Murphy et al., 1997; Sörös et al., 2006; Turkeltaub et al., 2002; Wise et al., 1999), and the recruitment of the SMA and ACC in particular, is well-known to play a critical role in response conflict or competition resolution in action and speech planning (Barch et al., 2000; Botvinick et al., 2004; Cohen et al., 2000; de Zubicaray et al., 2006; Mostofsky and Simmonds, 2008; Mukamel et al., 2010; Murphy et al., 1997; Simmonds et al., 2008; Tremblay and Small, 2011; Wilson et al., 2009).

Additionally, to determine the regions that were more active in one or another task, a higher-level analysis directly compared production vs. comprehension (cluster-corrected, $Z > 2.3$, $p < .05$). These contrasts revealed that the production task elicited more activity than comprehension in a wide network of regions, including the precentral gyrus, middle frontal gyrus and most medial frontal and subcortical structures mentioned above (see Fig. 3 panel B, Table 3). The contrast comparing comprehension > production only showed more activation for comprehension in a portion of the temporal lobe and occipito-temporal structures associated with visual processing (see Table 3). Overall, the pattern of results indicates that although a fronto-temporal network is common to production and comprehension (with some differences in activation levels across tasks in a portion of the posterior temporal lobe), a more extensive network including motor control and planning regions is recruited in sentence production.

ROI results

To establish whether the high- and low-competition conditions modulated activity within the Stroop-defined *pars opercularis* region in both production and comprehension, we compared percent signal change for each condition across participants, which was obtained for each individual participant from contrasting each condition relative to its baseline or control within the 35 most active voxels in the Stroop task (see Methods, *region of interest analysis*). We found a similar pattern of results across the production and comprehension tasks, as shown in Fig. 4. In comprehension, the activity elicited by the high-competition items relative to baseline was higher than that of the low-competition items relative to baseline ($M = .37$, $SE = .08$; $M = .28$, $SE = .05$; $t(16) = 2.31$, $p = .03$). Similarly, production revealed higher activation for the high-competition condition than the low-competition condition relative to their read-only controls ($M = .084$, $SE = .017$; $M = .069$, $SE = .014$; $t(16) = 2.90$, $p = .01$). Note that the overall mean activity in the production task is much smaller than that of comprehension due to removal of common reading activity (production minus reading-control contrast). Taken together, these results clearly suggest common mechanisms across tasks operating at the level of each individual brain in the *pars opercularis*.

PPI results

To establish how the *pars opercularis* interacts with high vs. low task demands in production and comprehension, we conducted PPI analyses using this region as seed region. The results of the comprehension task revealed a stronger coupling of activity between *pars opercularis* and the pMTG as a function of competition (uncorrected, $p = .001$), consistent with numerous studies indicating the involvement

Table 3
Results from whole-brain analysis (cluster corrected, p < .05).

			Voxel (MNI)		
Contrasts	Brain area	*Z*	x	y	z
Comprehension > rest	L IFG (BA44/BA45)	5.01	−50	14	24
	L IFG (BA 47)	3.94	−54	26	−2
	L precentral G (BA6)	5.03	−46	−2	32
	L PMTG (BA 32)	5.81	−52	−44	2
	L MTG (BA 21)	4.71	−56	−16	−10
	L ATL (BA 21)	4.18	−50	14	−22
	L hippocampus	3.80	−30	−24	−10
	L IPL (BA 39)	3.07	−36	−60	44
Production > control	L IFG (BA44/BA45)	5.37	−58	20	10
	L IFG (BA 47)	4.21	−54	34	−12
	L precentral G (BA 6)	2.65	−54	−4	48
	L MFG (BA 6)	3.51	−46	6	52
	L SFG (BA 9)	4.28	−10	56	38
	L SFG (BA 6)	4.72	−12	16	60
	L PSTG (BA39)	3.47	−44	−62	16
	L MTG (BA 21)	4.46	−56	−30	−10
	L ATL (BA 21)	4.52	−50	12	−30
	L ITG (BA 20)	3.96	−42	−8	−42
	L caudate	4.43	−16	12	12
	R caudate	3.46	6	14	4
	L ACC (BA 32)	4.49	−10	26	32
	R ACC (BA 39)	3.09	14	36	20
	Thalamus	3.31	−2	−12	10
	R cerebellum	4.23	36	−58	−40
Production > comprehension	L ACC (BA 32)	8.82	−8	20	38
	R ACC (BA 32)	8.98	4	20	30
	R SFG (BA6)	8.79	12	6	62
	L SFG (BA6)	7.66	−8	2	68
	Precuneous	7.11	−10	−80	42
	R/L fusiform G	7.79	+/−32	−48	−12
	L MFG (BA 9)	7.11	−30	34	36
	R/L insular cortex	6.74	+/−30	14	−4
	L precentral G/insula	6.17	−46	6	2
	L precentral G (BA6)	7.69	−60	4	36
	R/L caudate	7.40	+/−18	12	14
	R/L putamen	6.48	−22	10	10
	R/L cerebellum	6.22	+/−36	−54	−42
Comprehension > production	L PMTG	7.40	−60	−36	0

Note: L = left hemisphere, R = right hemisphere, IFG = inferior frontal gyrus, PMTG = posterior middle temporal gyrus, PSTG = posterior superior temporal gyrus, MTG = middle temporal gyrus, ITG = inferior temporal gyrus, ATL = anterior temporal lobe, IPL = inferior parietal lobe, MFG = middle frontal gyrus, SFG = superior frontal gyrus, ACC = anterior cingulate cortex.

of this fronto-temporal network in comprehension (see references above). The production results in contrast, indicated that a much more distributed network of regions was coupled with the *pars opercularis* as a function of competition (uncorrected, p = .001). These results were stronger than those in comprehension, surviving more stringent correction thresholds (cluster corrected z = 2.3, p = .05) (see Fig. 5 and Table 4). This production network included bilateral medial superior frontal gyrus, middle and superior temporal gyrus, left parietal cortex (postcentral gyrus and precuneus), insular cortex and subcortical areas including the left basal ganglia, thalamus and bilateral hippocampus. Some of the regions in this network, particularly parietal and medial structures, survived correction thresholds when directly contrasting production and comprehension (production > comprehension), using the lower-level contrast parameter estimates from the PPI analysis in a higher level group analysis (uncorrected, p = .001; see Table 4). All these production regions have been typically shown to be involved in speech planning, particularly when competitive processes are involved, e.g., lexical selection (de Zubicaray et al., 2001, 2006; Eickhoff et al., 2009; Heim et al., 2009; Murphy et al., 1997; Tremblay and Small, 2011; Whitney et al., 2008; Wilson et al., 2009; Wise et al., 1999). Note that although the posterior temporal clusters show slightly different centers for each task, there was considerable overlap, particularly if the thresholds are lowered. This was despite the fact that reading was removed from the completion task and thus potential commonly involved regions could have also been removed. This analysis therefore, suggests that a similar portion of the posterior temporal lobe is functionally connected to *pars opercularis* in both tasks (see Tables 3 and 4). Overall, these results indicate that whereas *pars opercularis* and portions of the posterior temporal cortex are recruited by both sentence production and comprehension, the network of regions effectively connected to *pars opercularis* is quite different in each task.

Discussion

The present studies aimed to investigate whether sentence production and comprehension share mechanisms of a competitive nature within LIFG but differ in LIFG's supporting networks. Experiment 1 indicated that both during reading comprehension and production planning, readers and speakers take longer to process the high-competition condition. Importantly, production planning times were predicted by the number of alternative plans entertained in the current context, just as reading times in comprehension correlate with the number of interpretations available in the unfolding context (Gennari and MacDonald, 2008). These results suggest that the processes taking place in our tasks involve competition between alternative interpretations or plans, thus providing an appropriate testing ground to examine the brain correlates of competition in each task.

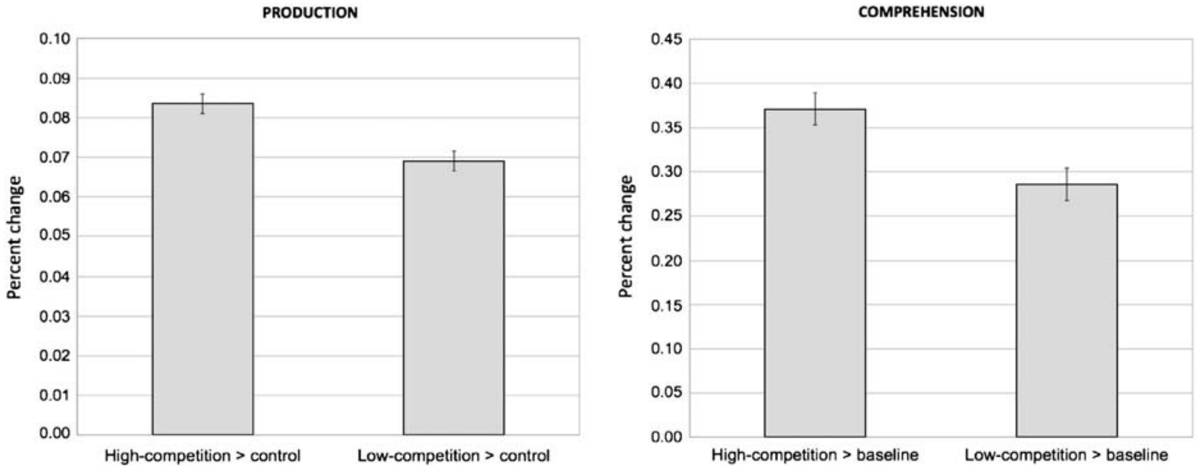

Fig. 4. Results of the region of interest analysis within the Stroop-defined *pars opercularis*. Percent signal changes calculated from contrast parameter estimates in each task (Experiment 2). Relevant contrasts are indicated in the x-axis. Error bars indicate standard errors.

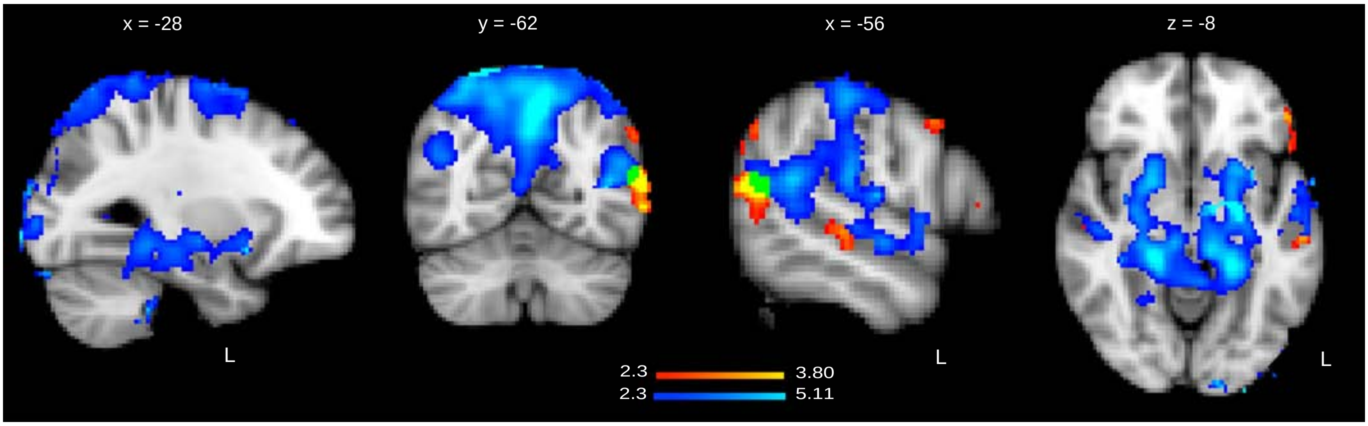

Fig. 5. PPI results from Experiment 2. Production results are shown in blue and comprehension results are shown in red (threshold at p = .001 (uncorrected) for illustration purposes). The green cluster indicates the overlap between production and comprehension.

Our fMRI experiment revealed several findings. Comparisons of the production and comprehension networks relative to rest or controls indicated that these tasks shared a fronto-temporal network, including LIFG and middle and superior temporal areas (Fig. 3). Within this network, the posterior temporal lobe is thought to store lexical knowledge, particularly that associated with verbs and their event structures (Bedny et al., 2008a; Dronkers et al., 2004; Humphreys et al., 2013; Tyler and Marslen-Wilson, 2008), whereas the *pars opercularis* is thought to perform general regulatory functions, including memory maintenance, controlled retrieval and encoding, integration and selection/inhibition (Barde and Thompson-Schill, 2002; D'Esposito et al., 1999; Fiebach et al., 2006; Fuster, 2001; Miller and Cohen, 2001; Thompson-Schill et al., 2005; Wagner et al., 2001). However, only production additionally engaged brain regions known to be critically involved in motor action control and/or speech planning, including the medial superior frontal regions (SMA), the ACC and subcortical regions such as caudate and putamen, all of which also emerged from direct contrasts across tasks (Alario et al., 2006; Barch et al., 2000; Botvinick et al., 2004; Chen et al., 2009; Ketteler et al., 2008; Mukamel et al., 2010; Murphy et al., 1997; Nachev et al., 2007; Simmonds et al., 2008; Sörös et al., 2006; Wahl et al., 2008; Wise et al., 1999).

Table 4
Results from whole-brain PPI analysis. Regions showing an interaction between level of competition (high vs. low) and activity within BA44.

			Voxel (MNI)		
Task (contrast)	Brain area	Z	x	y	z
Comprehension (high > low)	L PMTG	3.81	−60	−62	12
Production (high > low, control)	R SFG	3.90	24	12	60
	L SFG	3.50	−22	−2	66
	L PMTG	3.87	−44	−60	10
	L STG	3.93	−52	−24	0
	L caudate	3.10	−8	14	−2
	L putamen	3.43	−22	4	−10
	L hippocampus	3.80	−20	−38	−6
	R hippocampus	3.67	18	−34	−8
	Thalamus	4.01	−4	−8	4
	L postcentral G	3.78	−42	−36	68
	L precuneus	4.01	−4	−66	50
	L insular cortex	3.56	−32	12	−10
Production (high > low, control) > comprehension (high > low)	L IPL	3.44	−62	−24	26
	L precentral G	2.56	−62	4	12
	R/L MFG	2.64	−32	38	26
	L SFG	2.59	−14	0	60
	R ACC	2.6	6	36	16
	R SFG	2.6	18	14	64
	R MFG	3.2	30	4	56

Note: L = left hemisphere, R = right hemisphere, PMTG = posterior middle temporal gyrus, STG = superior temporal gyrus, SFG = superior frontal gyrus, IPL = inferior parietal lobe, MFG = middle frontal gyrus, ACC = anterior cingulate cortex.

Critically, competition conditions modulated activity for both tasks in the Stroop-defined *pars opercularis*, which targeted representational conflict, suggesting common competitive mechanisms. The strength and the extent of the LIFG's functional coupling, however, differed across tasks. Indeed, PPI analyses revealed that the *pars opercularis* established distinct patterns of connectivity in production compared to comprehension as a function of condition (Fig. 5). Incomprehension, high competition compared to low competition resulted in increased connectivity with the posterior temporal lobe. In contrast, high competition in production resulted in increased connectivity with additional regions relative to comprehension, such as the medial superior frontal gyrus (SMA), parietal lobe, basal ganglia structures such as the caudate, and the hippocampus, although these subcortical structures did not survive correction in the direct contrast between production and comprehension for these analyses. The SMA and these subcortical structures, particularly the caudate, are involved in motor/speech planning, as indicated above, as is the parietal lobe (Brownsett and Wise, 2010; Geranmayeh et al., 2012), whereas the hippocampus is well known to serve memory retrieval (Nadel and Moscovitch, 2001; Paller and Wagner, 2002; Whitney et al., 2008). These results therefore suggest that high levels of competition in production cause reverberations across a much wider network of regions compared to comprehension, specifically recruiting regions associated with motor response conflict and possibly, memory retrieval.

Production and comprehension thus appear as clearly distinct tasks in respect of the functional networks recruited for task performance in connection with *pars opercularis*, in line with psycholinguistic approaches, but they appear to share mechanisms of a competitive nature within *pars opercularis* and the posterior temporal lobe. This finding is consistent with numerous comprehension and production studies reporting activity in *pars opercularis* and the posterior temporal lobe (de Zubicaray et al., 2006; Gennari et al., 2007; Menenti et al., 2011; Tremblay and Small, 2011), as well as neuropsychological and TMS studies reporting either production or comprehension impairments when these regions are damaged or stimulated (Jefferies and Lambon Ralph, 2006; Spalek and Thompson-Schill, 2008; Thompson-Schill et al., 2002; Whitney et al., 2011). These common fronto-temporal regions thus operate together in both tasks, particularly when conflicting information is activated, and are involved in storing long-term associations between words, structures and their meanings (linguistic knowledge) and computing their higher order contingencies in sentence processing. Yet, production and comprehension handle linguistic knowledge differently by igniting associations in different regions. Production in particular may involve additional processes such as the resolution of competition between alternative plans in motor-related regions.

These results, together with behavioral studies, illuminate the nature of the processes that might be common or distinct in production and comprehension. Common processes may include the cue-based predictions and competition mechanisms suggested by Experiment 1, which have also been extensively investigated in the psycholinguistic literature (Altmann and Kamide, 1999; DeLong et al., 2005; Pickering and Garrod, 2007). Many reading studies for example indicate that comprehension difficulty is predicted by distributional probabilities in English corpora, and in particular, by the degree of fit between upcoming material and its probability of occurring in the preceding context, with poor fit leading to comprehension difficulty (Gennari and MacDonald, 2008, 2009; Hale, 2006; Levi, 2008; Lewis and Vasishth, 2005; MacDonald et al., 1994; Smith and Levy, 2013; Tanenhaus and Trueswell, 1995; Trueswell et al., 1994). This suggests that alternative continuations are entertained as the reading proceeds, leading to competition and inhibition processes. Similarly, our production study indicates that the number of alternative sentence continuations, as measured by overall preferences across speakers, predicts production planning times, suggesting that a greater number of available choices leads to greater difficulty and competition. These observations are consistent with the general executive functions often attributed to portions of LIFG and prefrontal cortex in general in that they mediate contingencies between cues and their associations, regardless of whether these cues are internally or externally generated (Fuster, 2001; Koechlin and Jubault, 2006; Passingham et al., 2000). The *pars opercularis* in particular is indeed involved in establishing linguistic contingencies during language learning (Opitz and Friederici, 2003, 2004) and it is necessary for producing word sequences and comprehending word order, grammatical cues and grammatical relations (Caplan et al., 1996; Thothathiri et al., 2010). Therefore, the *pars opercularis* and its interactive networks may manage linguistic contingencies, and thus, the anticipation or activation of information associated with the current task context.

The production-specific areas of activity reported here likely reflect production-specific processes such as word or structure retrieval, and linguistic and articulatory planning (or sub-vocalization), which nevertheless appear to involve competition as a result of the semantic properties of the context. Indeed, despite the lack of temporal resolution in fMRI, the fact that the level of competition modulates the interaction of *pars opercularis* with motor-related regions suggests that semantic competition percolates to putatively lower levels of phonological or articulatory planning. Thus, competition resolution does not appear restricted to word selection or semantic-role assignment, but instead, reverberates across a large network involving linguistic representations at different levels, syntactic, phonological, and articulatory. For example, it is possible that the alternative utterances entertained by the speakers do not only take place at a semantic level but also at the level of the phonological forms or motor plans associated with the different alternatives. Thus the finding that the influence of semantic competition pervades the production network challenges serial/encapsulated models of sentence production that argue for distinctive stages of lexical retrieval, planning and word sequencing (Bock and Levelt, 1994) and instead, suggests parallel processing and interactivity across an extensive network.

Our results have implications for models of language processing discussed in the literature such as dual-stream models. These models argue that there are dorsal and ventral streams in language processing, with different networks being engaged as a function of task, although they differ in the specific functional roles assigned to the individual regions within the streams (Friederici, 2009; Hickok and Poeppel, 2004, 2007; Petrides and Pandya, 2009; Saur et al., 2008; Wilson et al., 2011). The ventral stream including the fronto-temporal network is thought to be responsible for mapping sounds (and letters) to meaning, whereas the dorsal stream including the temporo-parietal junction and motor structures is thought to be involved in syntactic processing and in mapping sounds to action. This broad characterization is supported by what is known about the white matter tracts connecting sensory processing regions in posterior cortex to frontal cortex, with ventral and dorsal tracts converging in pre-frontal cortex (Anwander et al., 2007; Frey et al., 2008). Our results are broadly consistent with these processing routes but additionally suggest a potential common role of the fronto-temporal network in both production and comprehension.

More generally, together with existing literature, these results highlight the importance of understanding functional networks in the context of tasks demands, rather than focusing on a region's common role across tasks. Clearly, competition resolution must be achieved by the critical contribution of the interactive network at hand, and no single region is solely responsible for competition resolution. Indeed, many studies and reviews have suggested that *pars opercularis* is involved in a great variety of tasks, including perceptual decisions, response conflict resolution (e.g., Stroop task), working memory, memory retrieval and language processing (Crittenden and Duncan, 2012; Duncan, 2010; Duncan and Owen, 2000; Wagner et al., 2001). Across these tasks, verbal stimuli additionally engage temporal regions whereas goal and response-oriented tasks additionally engage parietal regions (Cisek, 2007; Duncan, 2010; Duncan and Owen, 2000; Tyler and Marslen-Wilson, 2008). These findings suggest that the networks engaged in each case is determined by stimulus characteristics and task demands and challenges attempts to attribute a given region such as the *pars opercularis* a unique specific function across all tasks. In the present study, the tightly controlled nature of our stimuli and behavioral results suggest that at least some aspects of evaluating alternatives within the common fronto-temporal network is likely to be shared by both production and comprehension, but this does not preclude that these fronto-temporal regions, and *pars opercularis* in particular, play a different role in the context of other task networks. In fact, the mere fact that *pars opercularis* in our results, interacts with more regions in one task than another suggests that at a neuronal level there must be some task differences with respect to the information being handled in this region and perhaps, the timing of activation across different regions. Establishing an adequate level of functional description with fMRI, which lacks temporal resolution, remains one of the main challenges in cognitive neuroscience.

Finally, note that our study specifically targeted LIFG, and found that the *pars opercularis*, which most strongly responded to the representational conflict in the Stroop task, was also sensitive to generally assumed competition processes both in production and comprehension. However, these findings do not preclude that other regions activated by both tasks (see Fig. 3) are also sensitive to competition processes, since we have not probed each one in turn. This is particularly the case for the *pars triangularis* and possibly *pars orbitalis*, which were also commonly activated by the two tasks. Moreover, our localizer only targeted word-level competition, and it is possible that sentence level competition requires additional portions of the pre-frontal cortex. Thus, more research is required to address this issue and further delineate the role of regions common to both tasks.

In sum, sentence production and reading comprehension appear to share competition mechanisms in a fronto-temporal network typically engaged in language processing. This network appears to manage the myriad of activations elicited by various types of linguistic cues, including the association between noun meanings and their event roles in a sentence. However, production recruits several additional regions to retrieve information and generate linguistically-based plans. Therefore, competition in production and comprehension reverberates across distinct functional networks, as determined by task demands.

Acknowledgments

This work was supported by research funds of the Department of Psychology, University of York. We thank Lauren Henshall and staff at the York Neuroimaging Centre for help with data collection and technical support. We also thank Nick Hindy and colleagues for sharing the Stroop task used in this work.

References

Adank, P., 2012. Design choices in imaging speech comprehension: an Activation Likelihood Estimation (ALE) meta-analysis. Neuroimage 63, 1601–1613.

Alario, F.-X., Chainay, H., Lehericyc, S., Cohen, L., 2006. The role of the supplementary motor area (SMA) in word production. Brain Res. 1076, 129–143.

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

Amunts, K., Lenzen, M., Friederici, A.D., Schleicher, A., Morosan, P., Palomero-Gallagher, N., Zilles, K., 2010. Broca's region: novel organizational principles and multiple receptor mapping. PLoS Biol. 8, e1000489.

Anwander, A., Tittgemeyer, M., von Cramon, D.Y., Friederici, A.D., Konsche, T.R., 2007. Connectivity-based parcellation of Broca's area. Cereb. Crotex 17.

Barch, D.M., Braver, T.S., Sabb, F.W., Noll, D.C., 2000. Anterior cingulate and the monitoring of response conflict: evidence from an fMRI study of overt verb generation. J. Cogn. Neurosci. 12, 298–309.

Barde, L.H., Thompson-Schill, S.L., 2002. Models of functional organization of the lateral prefrontal cortex in verbal working memory: evidence in favor of the process model. J. Cogn. Neurosci. 14, 1054–1063.

Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., Saxe, R., 2008a. Concepts are more than percepts: the case of action verbs. J. Neurosci. 28, 11347–11353.

Bedny, M., McGill, M., Thompson-Schill, S.L., Bedny, M., McGill, M., Thompson-Schill, S.L., 2008b. Semantic adaptation and competition during word comprehension. Cereb. Cortex 18, 2574–2585.

Bock, K., Levelt, W., 1994. Language production: grammatical encoding. In: Gernsbacher, M.A. (Ed.), Handbook of Psycholinguistics. Academic Press, San Diego, CA.

Bock, K., Dell, G.S., Chang, F., Onishi, K.H., 2007. Persistent structural priming from language comprehension to language production. Cognition 104, 437–458.

Botvinick, M.M., Cohen, J.D., Carter, C.S., 2004. Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn. Sci. 8, 539–546.

Brownsett, S.L., Wise, R.J., 2010. The contribution of the parietal lobes to speaking and writing. Cereb. Cortex 20, 517–523.

Buchweitz, A., Mason, R.A., Tomitch, L.M.B., Just, M.A., 2009. Brain activation for reading and listening comprehension: an fMRI study of modality effects and individual differences in language comprehension. Psychol. Neurosci. 2, 111–123.

Caplan, D., Hildebrandt, N., Makris, N., 1996. Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. Brain 119, 933–949.

Caplan, D., Stanczak, L., Waters, G., 2008. Syntactic and thematic constraint effects on blood oxygenation level dependent signal correlates of comprehension of relative clauses. J. Cogn. Neurosci. 20, 643–656.

Chang, F., Dell, G.S., Bock, K., 2006. Becoming syntactic. Psychol. Rev. 113, 234–272.

Chein, J.M., Fissellc, K., Jacobsd, S., Fiez, J.A., 2002. Functional heterogeneity within Broca's area during verbal working memory. Physiol. Behav. 77, 635–639.

Chen, E., West, C., Waters, G., Caplan, D., 2006. Determinants of BOLD signal correlates of processing object-extracted relative clauses. Cortex 42, 591–604.

Chen, C.Y., Muggleton, N.G., Tzeng, O.J., Hung, D.L., Juan, C.H., 2009. Control of prepotent responses by the superior medial frontal cortex. Neuroimage 44, 537–545.

Cisek, P., 2007. Cortical mechanisms of action selection: the affordance competition hypothesis. Philos. Trans. R. Soc. Lond. B Biol. Sci. 362, 1585–1599.

Cohen, J.D., Botvinick, M., Carter, C.S., 2000. Anterior cingulate and prefrontal cortex: who's in control? Nat. Neurosci. 3, 421–423.

Costa, A., Alario, F.X., Caramazza, A., 2005. On the categorical nature of the semantic interference effect in the picture–word interference paradigm. Psychon. Bull. Rev. 12, 125–131.

Crittenden, B.M., Duncan, J., 2012. Task difficulty manipulation reveals multiple demand activity but no frontal lobe hierarchy. Cereb. Cortex. http://dx.doi.org/10.1093/cercor/bhs333.

Dapretto, M., Bookheimer, S.Y., 1999. Form and content: dissociating syntax and semantics in sentence comprehension. Neuron 24, 427–432.

de Zubicaray, G.I., Wilson, S.J., McMahon, K.L., Muthiah, S., 2001. The semantic interference effect in the picture–word paradigm: an event-related fMRI study employing overt responses. Hum. Brain Mapp. 14, 218–227.

de Zubicaray, G.I., McMahon, K.L., Eastburn, M.M., Pringle, A., 2006. Top-down influences on lexical selection during spoken word production: a 4T fMRI investigation of refractory effects in picture naming. Hum. Brain Mapp. 27, 864–873.

DeLong, K.A., Urbach, T.P., Kutas, M., 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nat. Neurosci. 8, 1117–1121.

D'Esposito, M., Postle, B.R., Jonides, J., Smith, E.E., 1999. The neural substrate and temporal dynamics of interference effects in working memory as revealed by event-related functional MRI. Proc. Natl. Acad. Sci. U. S. A. 96, 7514–7519.

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

Duffy, S.A., Morris, R.K., Rayner, K., 1988. Lexical ambiguity and fixation times in reading. J. Mem. Lang. 27, 429–446.

Duncan, J., 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends Cogn. Sci. 14, 172–179.

Duncan, J., Owen, A.M., 2000. Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends Neurosci. 23, 475–482.

Eickhoff, S.B., Heim, S., Zilles, K., Amunts, K., 2009. A systems perspective on the effective connectivity of overt speech production. Philos. Trans. R. Soc. Lond. A Math. Phys. Eng. Sci. 367, 2399–2421.

Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., Kanwisher, N., 2010. New method for fMRI investigations of language: defining ROIs functionally in individual subjects. J. Neurophysiol. 104, 1177–1194.

Fedorenko, E., Duncan, J., Kanwisher, N., 2012. Language-selective and domain-general regions lie side by side within Broca's area. Curr. Biol. 22, 2059–2062.

Ferreira, F.S., Slevc, L.R., Rogers, E.S., 2005. How do speakers avoid ambiguous linguistic expressions? Cognition 96, 263–284.

Fiebach, C.J., Rissman, J., D'Esposito, M., 2006. Modulation of inferotemporal cortex activation during verbal working memory maintenance. Neuron 51, 251–261.

Frey, S., Campbell, J.S.W., Pike, G.B., Petrides, M., 2008. Dissociating the human language pathways with high angular resolution diffusion fiber tractography. J. Neurosci. 28, 11435–11444.

Friederici, A.D., 2009. Pathways to language: fiber tracts in the human brain. Trends Cogn. Neurosci. 13, 175–181.

Fuster, J.M., 2001. The prefrontal cortex—an update: time is of the essence. Neuron 30, 319–333.

Garrett, M.F., 1980. Levels of processing in sentence production. In: Butterworth, B. (Ed.), Language Production. Academic Press, London, pp. 177–220.

Gennari, S.P., MacDonald, M.C., 2008. Semantic indeterminacy in object relative clauses. J. Mem. Lang. 58, 161–187.

Gennari, S.P., MacDonald, M.C., 2009. Linking production and comprehension processes: the case of relative clauses. Cognition 111, 1–23.

Gennari, S.P., MacDonald, M.C., Postle, B.R., Seidenberg, M.S., 2007. Context-dependent interpretation of words: evidence for interactive neural processes. Neuroimage 35, 1278–1286.

Gennari, S.P., Mirkovic, J., MacDonald, M.C., 2012. Animacy and competition in relative clause production: a cross-linguistic investigation. Cogn. Psychol. 65, 141–176.

Geranmayeh, F., Brownsett, S.L., Leech, R., Beckmann, C.F., Woodhead, Z., Wise, R.J., 2012. The contribution of the inferior parietal cortex to spoken language production. Brain Lang. 121, 47–57.

Gitelman, D.R., Penny, W.D., Ashburner, J., Friston, K.J., 2003. Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. NeuroImage 19, 200–207.

Glezer, L.S., Riesenhuber, M., 2013. Individual variability in location impacts orthographic selectivity in the "visual word form area". J. Neurosci. 33, 11221–11226.

Hagoort, P., 2005. On Broca, brain, and binding: a new framework. Trends Cogn. Sci. 9, 416–423.

Hale, J., 2006. Uncertainty about the rest of the sentence. Cogn. Sci. Multidiscip. J. 30, 643–672.

Heim, S., Eickhoff, S.B., Friederici, A.D., Amunts, K., 2009. Left cytoarchitectonic area 44 supports selection in the mental lexicon during language production. Brain Struct. Funct. 213, 441–456.

Hickok, G., Poeppel, D., 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92, 67–99.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402.

Hindy, N.C., Altmann, G.T.M., Kalenika, E., Thompson-Schill, S.L., 2012. The effect of object state-changes on event processing: do objects compete with themselves? J. Neurosci. 32, 5795–5803.

Huang, J., Carr, T.H., Cao, Y., 2001. Comparing cortical activations for silent and overt speech using event-related fMRI. Hum. Brain Mapp. 15, 39–53.

Humphreys, G.F., Newling, K., Jennings, C., Gennari, S.P., 2013. Motions and actions in language: semantic representations in occipito-temporal cortex. Brain Lang. 125, 94–105.

Indefrey, P., 2011. The spatial and temporal signatures of word production components: a critical update. Front. Psychol. 255.

January, D., Trueswell, J.C., Thompson-Schill, S.L., 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. J. Cogn. Neurosci. 21, 2434–2444.

Jefferies, E., Lambon Ralph, M.A., 2006. Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. Brain 129, 2132–2147.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimisation for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17, 825–841.

Jobard, G., Vigneau, M., Mazoyer, B., Tzourio-Mazoyer, N., 2007. Impact of modality and linguistic complexity during reading and listening tasks. Neuroimage 34, 784–800.

Josephs, O., Henson, R.N., 1999. Event-related functional magnetic resonance imaging: modelling, inference and optimization. Philos. Trans. R. Soc. Lond. B Biol. Sci. 354, 1215–1228.

Juch, H., Zimine, I., Seghier, M., Lazeyras, F., Fasel, J.H.D., 2005. Anatomical variability of the lateral frontal lobe surface: implication for intersubject variability in language neuroimage. Neuroimage 24, 504–514.

Ketteler, D., Kastrau, F., Vohn, R., Huberd, W., 2008. The subcortical role of language processing. High level linguistic features such as ambiguity-resolution and the human brain; an fMRI study. Neuroimage 39, 2002–2009.

Koechlin, E., Jubault, T., 2006. Broca's area and the hierarchical organization of human behavior. Neuron 50, 963–974.

Levelt, W.J.M., Roelofs, A., Meyer, A.S., 1999. A theory of lexical access in speech production. Behav. Brain Sci. 22, 1–75.

Levi, R., 2008. Expectation-based syntactic comprehension. Cognition 106, 1126–1177.

Lewis, R.L., Vasishth, S., 2005. An activation-based model of sentence processing as skilled memory retrieval. Cognit. Sci. 29, 375–419.

Lindenberg, R., Scheef, L., 2007. Supramodal language comprehension: role of the left temporal lobe for listening and reading. Neuropsychologia 45, 2407–2415.

MacDonald, M.C., 1999. Distributional information in language comprehension, production, and acquisition: three puzzles and a moral. In: MacWhinney, B. (Ed.), The Emergence of Language. Erlbaum, Mahweh, NJ.

MacDonald, M.C., Pearlmutter, N., Seidenberg, M., 1994. Lexical nature of syntactic ambiguity resolution. Psychol. Rev. 101, 676–703.

Mak, W.M., Vonk, W., Schriefers, H., 2002. The influence of animacy on relative clause processing. J. Mem. Lang. 47, 50–68.

Mak, W.M., Vonk, W., Schriefers, H., 2006. Animacy in processing relative clauses: the hikers that rocks crush. J. Mem. Lang. 54, 466–490.

Menenti, L., Gierhan, S.M.E., Segaert, K., Hagoort, P., 2011. Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. Psychol. Sci. 22, 1173–1182.

Michael, E.B., Keller, T.A., Carpenter, P.A., Just, M.A., 2001. fMRI investigation of sentence comprehension by eye and by ear: modality fingerprints on cognitive processes. Hum. Brain Mapp. 13, 239–252.

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. Cogn. Brain Res. 12, 467–473.

Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. Annu. Rev. Neurosci. 24, 167–202.

Mostofsky, S.H., Simmonds, D.J., 2008. Response inhibition and response selection: two sides of the same coin. J. Cogn. Neurosci. 20, 751–761.

Mukamel, R., Ekstrom, A.D., Kaplan, J., Iacoboni, M., Fried, I., 2010. Single-neuron responses in humans during execution and observation of actions. Curr. Biol. 20, 750–756.

Murphy, K., Corfield, D.R., Guz, A., Fink, G.R., Wise, R.J., Harrison, J., Adams, L., 1997. Cerebral areas associated with motor control of speech in humans. J. Appl. Physiol. 83, 1438–1447.

Nachev, P., Wydell, H., O'Neill, K., Husain, M., Kennard, C., 2007. The role of the pre-supplementary motor area in the control of action. Neuroimage 36, T155–T163.

Nadel, L., Moscovitch, M., 2001. The hippocampal complex and long-term memory revisited. Trends Cogn. Neurosci. 228–230.

Nichols, M., Brett, M., Andersson, J., Wager, T., Poline, J.-B., 2005. Valid conjunction inference with the minimum statistic. Neuroimage 25, 653–660.

Nieto-Castañón, A., Fedorenko, E., 2012. Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. Neuroimage 63, 1646–1669.

Novick, J.M., Trueswell, J.C., Thompson-Schill, S.L., 2005. Cognitive control and parsing: reexamining the role of Broca's area in sentence comprehension. Cogn. Affect. Behav. Neurosci. 5, 263–281.

Novick, J.M., Kan, I.P., Trueswell, J.C., Thompson-Schill, S.L., Novick, J.M., Kan, I.P., Trueswell, J.C., Thompson-Schill, S.L., 2009. A case for conflict across multiple domains: memory and language impairments following damage to ventrolateral prefrontal cortex. Cogn. Neuropsychol. 26, 527–567.

Opitz, B., Friederici, A.D., 2003. Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. Neuroimage 19, 1730–1737.

Opitz, B., Friederici, A.D., 2004. Brain correlates of language learning: the neuronal dissociation of rule-based versus similarity-based learning. J. Neurosci. 24, 8436–8440.

Paller, K.A., Wagner, A.D., 2002. Observing the transformation of experience into memory. Trends Cogn. Sci. 6, 93–102.

Passingham, R.E., Toni, I., Rushworth, M.F., 2000. Specialisation within the prefrontal cortex: the ventral prefrontal cortex and associative learning. Exp. Brain Res. 133, 103–113.

Petrides, M., 2005. Lateral prefrontal cortex: architectonic and functional organization. Phil. Trans. R. Soc. B 360, 781–795.

Petrides, M., Pandya, D.N., 2009. Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. PLoS Biol. 7, e1000170.

Pickering, M.J., Garrod, S., 2007. Do people use language production to make predictions during comprehension? Trends Cogn. Sci. 11, 105–110.

Rayner, K., Duffy, S., 1986. Lexical complexity and fixation times in reading: effects of word frequency, verb complexity, and lexical ambiguity. Mem. Cognit. 14, 191–201.

Rayner, K., Carlson, M., Frazier, L., 1983. The interaction of syntax and semantics during sentence processing: eye movements in the analysis of semantically biased sentences. J. Verbal Learn. Verbal Behav. 22, 358–374.

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

Rodd, J.M., Longeb, O.A., Randall, B., Tyler, L.K., 2010. The functional organisation of the fronto-temporal language system: evidence from syntactic and semantic ambiguity. Neuropsychologia 48, 1324–1335.

Saur, D., Kreher, B.W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M.-S., Umarova, R., Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., Weiller, C., 2008. Ventral and dorsal pathways for language. Proc. Natl. Acad. Sci. U. S. A. 105, 18035–18040.

Saxe, R., Brett, M., Kanwisher, N., 2006. Divide and conquer: a defense of functional localizers. Neuroimage 30, 1088–1096.

Seidenberg, M.S., MacDonald, M.C., 2001. Constraint satisfaction in language acquisition and processing. In: Christiansen, M.H., Charter, N. (Eds.), Connectionist Psycholinguistics. Ablex, Westport, CT, pp. 281–318.

Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.M., Bienkowski, M., 1982. Automatic access of the meanings of ambiguous words in context: some limitations of knowledge-based processing. Cogn. Psychol. 14, 489–537.

Simmonds, D.J., Pekar, J.J., Mostofsky, S.H., 2008. Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. Neuropsychologia 46.

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

Smith, N.J., Levy, R., 2013. The effect of word predictability on reading time is logarithmic. Cognition 128, 302–319.

Sörös, P., Guttman Sokoloff, L., Bose, A., McIntosh, A.R., Graham, S.G., Stuss, D.T., 2006. Clustered functional MRI of overt speech production. Neuroimage 32, 376–387.

Spalek, K., Thompson-Schill, S.L., 2008. Task-dependent semantic interference in language production: an fMRI study. Brain Lang. 107, 220–228.

Swallow, M.K., Braver, T.S., Snyder, A.Z., Speer, N.K., Zacks, J.M., 2003. Reliability of functional localization using fMRI. Neuroimage 20, 1561–1577.

Swinney, D.A., 1979. Lexical access during sentence comprehension: (Re) consideration of context effects. J. Verbal Learn. Verbal Behav. 18, 523–534.

Tanenhaus, M.K., Trueswell, J.C., 1995. Sentence comprehension. In: Miller, J., Eimas, P. (Eds.), Speech, Language and Communication. Academic Press, San Diego, CA, pp. 217–262.

Tanenhaus, M.K., Leiman, J.M., Seidenberg, M.S., 1979. Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. J. Verbal Learn. Verbal Behav. 18, 427–440.

Thompson-Schill, S.L., D'Esposito, M., Aguirre, G.K., Farah, M.J., 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. Proc. Natl. Acad. Sci. U. S. A. 94, 14792–14797.

Thompson-Schill, S.L., Jonides, J., Marshuetz, C., Smith, E.E., D'Esposito, M., Kan, I.P., Knight, R.T., Swick, D., 2002. Effects of frontal lobe damage on interference effects in working memory. Cogn. Affect. Behav. Neurosci. 2, 109–120.

Thompson-Schill, S.L., Bedny, M., Goldberg, R.F., 2005. The frontal lobes and the regulation of mental activity. Curr. Opin. Neurobiol. 15, 219–224.

Thothathiri, M., Schwartz, M.F., Thompson-Schill, S.L., 2010. Selection for position: the role of left ventrolateral prefrontal cortex in sequencing language. Brain Lang. 113, 28–38.

Traxler, M.J., Morris, R.K., Seely, R.E., 2002. Processing subject and object relative clauses: evidence from eye movements. J. Mem. Lang. 47, 69–90.

Traxler, M.J., Williams, R.S., Blozis, S.A., Morris, R.K., 2005. Working memory, animacy, and verb class in the processing of relative clauses. Journal of Memory and Language 53, 204–224.

Tremblay, P., Small, S.L., 2011. Motor response selection in overt sentence production: a functional MRI study. Frontiers in Psychology 2, 253.

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 and Language 33, 285–318.

Turkeltaub, P.E., Guinevere, F.E., Jones, K.M., Zeffiro, T.A., 2002. Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. Neuroimage 16, 765–780.

Tyler, L.K., Marslen-Wilson, W., 2008. Fronto-temporal brain systems supporting spoken language comprehension. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363, 1037–1054.

Wagner, A.D., Paré-Blagoev, E.J., Clark, J., Poldrack, R.A., 2001. Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. Neuron 31, 329–338.

Wahl, M., Marzinzik, F., Friederici, A.D., Hahne, A., Kupsch, A., Schneider, G.-H., Saddy, D., Curio, G., Klostermann, F., 2008. The human thalamus processes syntactic and semantic language violations. Neuron 59, 695–707.

Whitney, C., Weis, S., Krings, T., Huber, W., Grossman, M., Kircher, T., 2008. Task-dependent modulations of prefrontal and hippocampal activity during intrinsic word production. J. Cogn. Neurosci. 21, 697–712.

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M.A., Jefferies, E., 2011. The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. Cereb. Cortex 21, 1066–1075.

Wilson, S.M., Isenberg, A.L., Hickok, G., 2009. Neural correlates of word production stages delineated by parametric modulation of psycholinguistic variables. Hum. Brain Mapp. 30, 3596–3608.

Wilson, S.D., Galantucci, S., Tartaglia, M.C., Rising, K., Patterson, D.K., Henry, M.L., Organ, J.M., DeLeon, J., Miller, B.L., Gorno-Tempini, M.L., 2011. Syntactic processing depends on dorsal language tracks. Neuron 72, 397–403.

Wise, R.J., Greene, J., Buchel, C., Scott, S.K., 1999. Brain regions involved in articulation. Lancet 353, 1057–1061.

Woolrich, M.W., Ripley, B.D., Brady, J.M., Smith, S.M.T., 2001. Temporal autocorrelation in univariate linear modelling of FMRI data. Neuroimage 14, 1370–1386.