Different roles of cytoarchitectonic BA 44 and BA 45 in phonological and semantic verbal fluency as revealed by dynamic causal modelling

Stefan Heim *, Simon B. Eickhoff, Katrin Amunts

Department of Psychiatry and Psychotherapy, RWTH Aachen University, Aachen, Germany
Research Centre Jülich, Institute for Neurosciences and Medicine (INM-1 and INM-2), Jülich, Germany
JARA Jülich-Aachen Research Alliance, Jülich, Germany

ARTICLE INFO

ABSTRACT

The interactions of left cytoarchitectonic BA 44 and BA 45 during semantic and phonological verbal fluency tasks were investigated using dynamic causal modelling (DCM). Three different models were tested, all of which featured BA 44 and BA 45 as top-down driven interconnected nodes projecting to the motor cortex as the final output region. Model #1 represents the hypothesis that BA 45 is involved in lexical retrieval including both semantic and phonological processes, while BA 44 supports other phonological processes. Model #2 reflects the notion of a clear-cut segregation of computational processes sustained by BA 44 (phonological processing) and BA 45 (semantic processing). Model #3 was based on the hypothesis that both BA 44 and BA 45 support semantic and phonological processing. When these models were compared against each other by Bayesian model selection, evidence emerged in favour of the first model, implying that BA 45 supports word retrieval processes whereas BA 44 is involved in processing phonological information during word generation. In a subsequent analysis of the derived model parameters for model #1, all connection strengths were significantly positive except for the inhibitory coupling between BA 44 and BA 45. This inhibition may reflect how the phonological analysis in BA 44 during word generation constrains lexical word retrieval in BA 45. To conclude, DCM provided additional insights into the roles of BA 44 and BA 45 during verbal fluency revealing the involvement of BA 45 in lexical retrieval and the relevance of BA 44 for phonological processing during word generation.

Introduction

Verbal fluency tasks test the ability to generate words according to a given criterion. For example, in a semantic fluency task, subjects are asked to produce words of a semantic category (e.g. animal names), whereas in phonological fluency they have to find words starting with a given letter or phoneme. Functional neuroimaging studies have repeatedly demonstrated that these verbal fluency tasks recruit Broca's region in the left inferior frontal cortex, which comprises at least Brodmann's areas (BA) 44 and 45 (Amunts et al., 2004). In a recent review Costafreda et al. (2006) suggested that the functional specificity of BA 44 and BA 45 for phonological and semantic processing, respectively, which has been found for language comprehension (e.g. Bookheimer, 2002; Friederici, 2002; Hagoort, 2005), also holds for verbal fluency. This position was recently challenged by a functional magnetic resonance imaging (fMRI) study that directly compared semantic vs. phonological verbal fluency (Heim et al., 2008). Instead of the proposed functional segregation of BA 44 and BA 45 for phonological vs. semantic fluency, both areas support either

type of fluency. BA 44, however, showed an additional increase in activation in the phonological fluency condition. These results hence suggested that BA 45 unspecifically supports the retrieval of words from the mental lexicon independent of the retrieval criterion (be it semantic or phonological), whereas BA 44 may subserve phonological processing.

In the present study, we used dynamic causal modelling (DCM; Friston et al., 2003) in order to gain further insight in the functions of BA 44 and BA 45 in verbal fluency. DCM is used in order to find a causal model of the underlying neuronal interactions, which may explain the pattern of activation observed in the previous general linear model analysis. That is, standard GLM approaches merely describe the pattern of condition specific activation in the brain, without any formal explanatory power of how these emerged. On the other hand, a DCM analysis provides a causal model of the effective connectivity underlying the activation pattern seen in this particular dataset by conventional statistical approaches. More precisely, DCM as a method for analysing effective connectivity between brain regions aims at modelling the influence these regions exert over one another under given experimental conditions. One key feature of effective connectivity as assessed by DCM is its theoretically motivated approach, which requires the a priori, hypothesis-driven definition of competing network models. These models, which reflect different assumptions of

* Corresponding author. RWTH Aachen University, Department of Psychiatry and Psychotherapy, Pauwelsstraße 30, 52074 Aachen, Germany. Fax: +49 2461 61 2820.
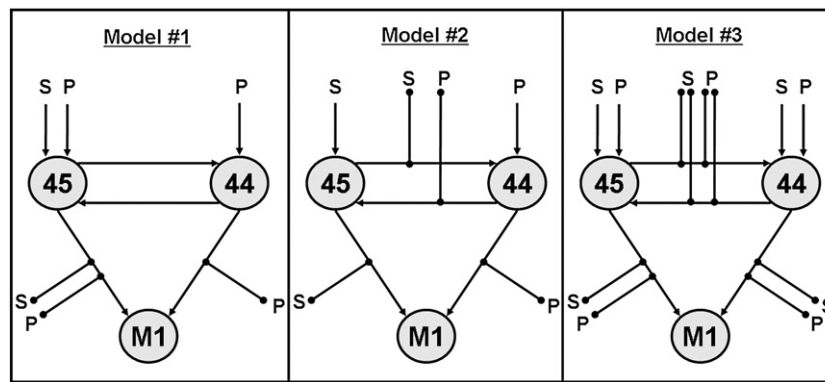   *E-mail address:* sheim@ukaachen.de (S. Heim).

Fig. 1. Outline of the three DCM models tested in the present study (see text for details). Arrows represent driving inputs into regions, or intrinsic connections between regions. Lines with black dots at their ends indicate modulations of the intrinsic connections by the tasks. Abbreviations: S, semantic fluency; P, phonological fluency, M1, motor cortex.

how a particular task may be sustained by inter-regional interactions, are then compared to each other based on the evidence they receive from the current data using a Bayesian model selection procedure (Penny et al., 2004). The posterior estimates of the parameters for that model which is most supported by the experimental data are then further analysed to enable inference on their magnitude and valence (positive or negative) across the group of analysed subjects.

We aimed at comparing three alternative models for the contribution of left BA 44 and 45 during verbal fluency (Fig. 1). In all of these left BA 44 and left BA 45 were designated as input regions that were driven intrinsically by task-related effects (cf. Bitan et al., 2007). The motor cortex, on the other hand, always represented the output of the cortical language production network. The three compared models varied, however, with respect to the hypothesised roles of BA 44 and BA 45 during verbal fluency. Model #1 is based on the idea that BA 45 is involved in lexical retrieval independent of the criterion (semantic or phonological) whereas BA 44 supports additional phonological processing. Thus, the model is characterised by both specialised and shared processes within BA 44 and BA 45. The other two models represent only the one or only the other aspect. Model #2 features the hypothesis of a clear functional segregation between BA 44 and BA 45. In line with the view by Costafreda et al. (2006), BA 44 is supposed to support phonological retrieval whereas BA 45 is involved in semantic retrieval. In contrast, model #3 represents a shared functionality of BA 44 and BA 45 during phonological and semantic retrieval where both regions support either process. The models and the neurofunctional hypotheses they represent are outlined in detail below.

Materials and methods

Participants

Data from a subset of participants from our previous study (Heim et al., 2008) who showed activation in BA 44, BA 45, and motor cortex in the conjunction of the phonological and the semantic fluency task were included in the present analysis (19 subjects included in Heim et al., (2008) plus one additional subject). Subjects who failed to have their activation maximum in all three regions within the pre-defined search radius (see below) at a threshold of $p<0.05$ (uncorrected) were excluded from the original sample. The remaining 20 healthy right-handed participants (mean age 28.3 years; 14 women) were all native German speakers and had normal or corrected-to-normal vision. Participants had no history of neurological or psychiatric disorders. Informed consent was obtained from all participants. The experimental standards were approved by the local ethics committee of RWTH Aachen University. A new second-level random effects analysis was run with these 20 participants in order to ensure that the

originally reported pattern of activation could be replicated in the smaller sample (see below).

Six out of the excluded subjects (mean age 25.2 years) also had activation overlapping with all three regions in both fluency tasks, although they had not met the strong criteria for time course extraction applied above. In order to generalise the findings of the first 20 subjects who met the more stringent inclusion criterion, these other six subjects were analysed with the same methods as the 20 included subjects. Their data are reported separately.

Tasks

A total of four overt German word generation tasks were applied in a block design: semantic, syntactic, phonological, and free (no explicit criterion). Each generation task was performed in six blocks (see below). For the present analysis only blocks of semantic and phonological fluency will be considered as we explicitly aimed at testing competing hypotheses for the contribution of BA 44 and BA 45 to these two tasks. In the semantic fluency task subjects had to overtly generate examples for six categories: birds, mammals, food, weapons, tools, and toys. In the phonological fluency task the participants generated nouns starting with the phonemes /b/, /f/, /k/, /m/, /sh/, and /t/.

Stimulus presentation

Visual stimuli were presented as written strings in Helvetica font at 48 pts via goggles (VisuaStim™, Resonance Technology, CA, USA). Stimulus presentation was controlled by a computer placed in the control room using Presentation software (Neurobehavioral Systems, Albany, CA, USA).

In the experimental session, we tested six blocks for each condition and 24 resting blocks separating the task blocks. A written instruction was presented for 6 s before each task block. The blocks started immediately after that instruction and lasted for 20 s. All conditions were presented in a pseudo-randomised order, with different randomisations for each participant. The total duration of the experiment was 19 min. Each block lasted 20 s, including ten trials of 2 s each. The fMRI data were acquired in the first 1.04 s of each trial using a bunched-early sequence (see below). After this time, a fixation cross appeared for the remaining 0.96 s of silence, indicating the subject that he or she could now utter the next word during a silent period (Heim et al., 2008; de Zubicaray et al., 2001). This setup prevents motion-induced susceptibility artefacts, since subjects only speak when no fMRI data are recorded. Moreover, this paced word generation reduces (imaging relevant) head motion during speaking as compared to unpaced word generation (Basho et al., 2007).

Data acquisition and analysis

Functional imaging data

The experiment was performed on a 3 T Siemens Trio scanner. A standard birdcage head coil was used with foam paddings reducing head motion. Functional data were recorded from 17 sagittal slices in the left hemisphere using a gradient-echo EPI sequence with echo time = 30 ms, flip angle = 90°, and repetition time (TR) = 2 s. The sagittal orientation of the slices was chosen in order to correct head motion in-plane, which is highest in the $y–z$ plane. Acquisition of the slices within the TR was arranged so that all slices were acquired in the first 1040 ms, followed by a 960 ms period of no acquisition to complete the TR during which the subjects spoke. The field of view was 200 mm, with an in-plane resolution of 3.1 mm × 3.1 mm. The slice thickness was 3 mm with an inter-slice gap of 1 mm.

Data processing and analysis was performed using MATLAB 6.5 (The Mathworks Inc., Natick, USA), and SPM5 (Wellcome Department of Cognitive Neurology, UK). Two dummy scans acquired before the beginning of the experiment to allow for magnetic saturation were discarded. Data pre-processing included the standard procedures of realignment, normalisation to the MNI single subject template, and spatial smoothing (FWHM = 8 mm). For the statistical analysis at the single subject level, the block functions for each word generation condition were convolved with a canonical haemodynamic response function (HRF). For each participant, the contrasts Semantic > Rest and Phonological > Rest were calculated. For the general linear model (GLM) group analysis, the individual contrast images were entered into a repeated-measures ANOVA (including non-sphericity correction) as a second-level random effects analysis to identify the location of common and differential activations.

Anatomical localisation

For the anatomical localisation of the maxima in the group analysis we used cytoarchitectonic probability maps of BA 44 (maximum at MNI coordinates $[x,y,z]$ −46,10,4) and BA 45 (maximum at −44,28,22) in Broca's region (Amunts et al., 1999, 2004) and BA 4p (maximum at −44, −14,36) in the motor cortex (Geyer et al., 1996). These maps are based on an observer-independent analysis of the cytoarchitecture in a sample of ten post-mortem brains (Schleicher et al., 1999; Zilles et al., 2002). They provide information about the location and variability of cortical regions in standard MNI space (http://www.fz-juelich.de/ime/SPM_Anatomy_Toolbox).

Extraction of time courses

Left BA 44, BA 45, and motor cortex were selected as volumes of interest (VOI). For each subject and each VOI, the individual local maximum ($p < 0.05$ uncorrected; cf. Mechelli et al., 2005; Eickhoff et al., 2009) was automatically identified that was closest to the group maximum in the conjunction analysis across all conditions at the single subject level. Local maxima were only accepted if their distance to the group maximum was less than 16 mm. Moreover, an iterative approach to the definition of local maxima was used to ensure that only such maxima were selected that were closer to the other maxima representing the same regions (own cluster) as opposed to those representing any other region (neighbouring clusters), hereby ensuring homogeneity of definitions across subjects. For each VOI, a time series was extracted as the first principal component of all voxel time series within a sphere (radius 4 mm) centred on the individual local maximum. The average MNI coordinates (x,y,z) at which the time series were extracted were −49.1,9.3,4.6 (BA 44), −45.1,26.6,23.5 (BA 45), and −44.8, −10.8,38.9 (BA 4p). Further inspection ensured that the spherical extraction volume around the individual local maxima overlapped with the corresponding cytoarchitectonic area.

Since all individual datasets had previously been normalised into standard MNI coordinate space, the probabilistic brain atlas applied equally to all participants.

For the additional six subjects, time courses were extracted from a local maximum closest to the group maximum, which was allowed to be outside the cytoarchitectonic region as long as the activation cluster still overlapped with this region.

Definition of the DCM models

The DCM analysis was aimed at evaluating three competing hypotheses about the organisation of the inferior frontal network sustaining verbal fluency (Fig. 1). Importantly, all three may well account for the results observed in the GLM analysis, i.e. activation in BA 44 and BA 45 in both fluency tasks plus higher activation for phonological than semantic fluency only in BA 44.

Model #1

Model #1 is based on the assumption that BA 45 is involved in the retrieval of words independent of the retrieval criterion (semantic or phonological), whereas BA 44 supports phonological processing (cf. Heim et al., 2008). This is modelled such that BA 45 receives driving inputs from both tasks, which modulate the forward connectivity to motor cortex as well. Based on the intrinsic connectivity between BA 44 and 45 assumed in the model, activation also propagates into BA 44 rendering this area co-activated whenever BA 45 is activated in any fluency task. This co-activation is automatic and thus not specifically modulated by any of the fluency tasks. In addition to this automatic activation, BA 44 receives driving input by phonological processing as a second, direct influence on the activation in BA 44. The activation observed in the GLM analysis would hence result from the following flow of activation: BA 45 is activated by both tasks and co-activates BA 44. BA 44 receives additional task-driven top-down activation by the phonological fluency task which increases the level of activation in BA 44 over that in BA 45. Following the concept of (Gold et al.2005) of common and dissociable activation patterns associated with controlled semantic and phonological processing, model #1 features *domain-preferentiality* of BA 44, i.e. the general involvement in several tasks (here: semantic and phonological processing) with particular involvement in one process (here: phonology). Gold et al. (2005) take domain-preferentiality as a more appropriate description of the roles of BA 44 and BA 45 than *domain-specificity*, i.e. a clear functional segregation between regions.

Model #2

Model #2 assumes such domain-specificity, i.e. the functional segregation of BA 44 and BA 45 as suggested by Costafreda et al. (2006). As this hypothesis implicates BA 45 only in semantic and BA 44 only in phonological retrieval, the semantic fluency task was defined as the driving input into BA 45. Activation then propagates to motor cortex and to BA 44. This flow of activation is modulated solely by the semantic task. Complementarily, BA 44 is driven by the phonological fluency condition, which also is the only modulator of the forward connections into BA 45 and motor cortex. In this model, semantic activation in BA 45 results from the driving input whereas phonological activation in BA 45 comes via the intrinsic connection from BA 44. The reverse pattern holds true for BA 44. Higher activation in BA 44 for phonological than semantic fluency may be caused by the fact that the driving input into BA 44 is higher than that into BA 45, and that the levels of activation are modulated by the tasks via the modulation of the intrinsic connections BA 44 ↔ BA 45.

Model #3

Model #3 represents shared functionality of BA 44 and BA 45 during both phonological and semantic retrieval. Thus, differential activation results from relative participation and computational load

whereas there is no fundamental distinction between the roles of these two areas in the two fluency conditions. Consequently, model #3 represents the hypothesis that BA 44 and BA 45 of the left IFG function as a supramodal executive module selecting task-relevant information among competing alternatives (cf. Thompson-Schill, 2003). Accordingly, semantic and phonological fluency both serve as driving inputs into BA 44 and BA 45 in the corresponding DCM model. The intrinsic connections between both regions and the motor cortex are modulated by both tasks. Differences between phonological and semantic activations in BA 44 and BA 45 may result from differing amounts of driving input, differing strengths of the intrinsic connections, or a combination of both.

Model selection and parameter test

In order to compare the three different models, we used a Bayesian model selection (BMS) procedure (Penny et al., 2004; Heim et al., 2009). BMS rests on the so-called model evidence, which is given by the posterior probability $p(y|m)$ of observing the data y given a particular model m integrated over the model parameters. Given two models i and j and their evidences, the Bayes Factor (BF) B_{ij} is defined as the ratio $p(y|m=i)/p(y|m=j)$. The BF quantifies how much evidence in favour of model i (relative to model j) is provided by the

experimental data. In such a pair-wise model comparison procedure, a $BF=1$ indicates equivalence of models; a $BF>1$ means evidence in favour of the first model, while a $BF<1$ signifies evidence for the second model. This approach was implemented as follows. First, the BFs for the pair-wise comparisons of the three models for each subject were calculated using procedures implemented in SPM5. As the individual subjects represent independent observations, the average Bayes Factor (ABF) for each pair-wise comparison was calculated as the geometric mean of all subject-specific BFs (Smith et al., 2006; Stephan and Penny, 2007). After determining the optimal model with BMS, the posterior estimates of its model parameter (driving inputs, intrinsic connections, and modulations) from all subjects were entered into one-sample t-tests (SPSS 15.0, Bonferroni correction for multiple comparisons) to test whether they differed significantly from zero over our group of 20 participants. Inference on model parameters is hence conceptually equivalent to the idea of random effects analysis in classical statistics under the assumption of sphericity in the individual parameters.

Model simulation

The architectures of the to-be-tested models were motivated above. One common feature of these models is that regions receive
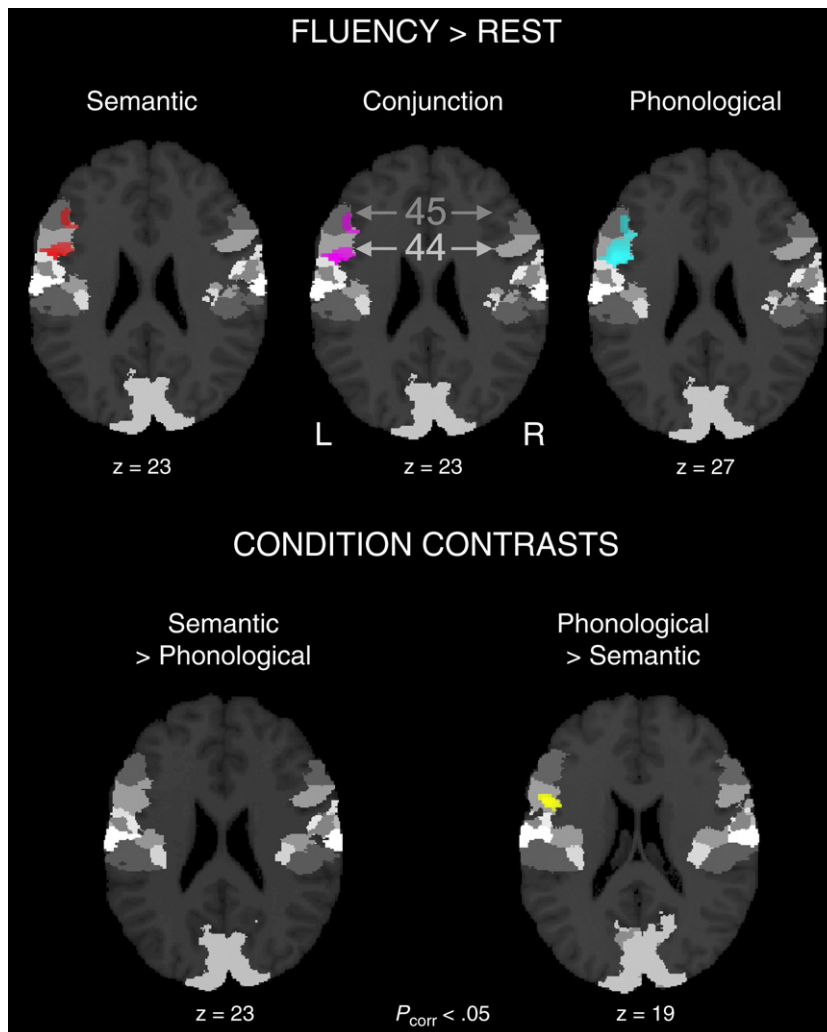

Fig. 2. Standard GLM analysis of the activation in cytoarchitectonic BA 44 (light grey) and BA 45 (dark grey) during semantic (red) and phonological (blue) fluency. The data reveal the involvement of both regions in both fluency tasks (see Conjunction analysis in purple), with equal activation levels in BA 45 but higher activation for phonological than for semantic fluency in BA 44 (yellow). All activation maps are thresholded at $p_{corr}<0.05$ FWE-corrected for the volume of the combined maximum probability maps of the left BA 44 and BA 45 (see Eickhoff et al., 2005, 2006), and masked inclusively with this search volume. Other, non-labelled grey shades refer to other cytoarchitectonically defined regions not relevant for the purpose of the present analysis.

input by one condition (e.g. semantic fluency), and at the same time the intrinsic connections in the model are modulated by the very same condition. Currently, there is a discussion that such model architectures may bias the resulting parameter estimates (cf. https://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind05&L=SPM&P=R568755). One possibility to test whether such bias is present is to generate synthetic datasets based on the model that received the highest evidence in the Bayesian model selection, set up a simulated model based on these synthetic data, and then compare the average parameters of the synthetically generated models with those of the model with the superior empirical evidence. Differences between the two sets of parameters would indicate the presence of a bias. Accordingly, we generated as many simulated DCMs containing the synthetic datasets as were subjects in the study ($N=20$) using the procedures *spm_dcm_create.m* and *spm_dcm_generate.m* contained in SPM5. These 20 synthetic DCMs were based on the individual parameters of each participant's model #1, i.e. the individual selected models, with a signal-to-noise ratio of 2. The synthetic DCMs were then estimated like the "real" models containing the empirical data, and parameter estimates were extracted as described above. These parameters were then compared to the parameters of the empirical model #1 using paired-sample *t*-tests.

Results

Performance data

The subjects produced on average 48.7 (range: 36–57; standard error of mean: ±1.29) words in the semantic condition and 46.7 (range: 34–58; standard error of mean: ±1.68) words in the phonological condition. This performance was statistically identical ($t_{19}=1.42$; $p=0.171$).

Table 1a
Individual Bayes Factors, Average Bayes Factors (ABF) and standard error of mean (SEM) for the comparison of models #1–3.

Subject	Original model selection			Model selection with replaced missings		
	#1 vs. #2	#1 vs. #3	#3 vs. #2	#1 vs. #2	#1 vs. #3	#3 vs. #2
01	–	11.36	5.52	1.00	11.36	5.52
02	392.78	200.49	–	392.78	200.49	1.00
03	–	7.06	2.87	1.00	7.06	2.87
04	4.04	115.76	28.64	4.04	115.76	28.64
05	9.83	111.31	11.33	9.83	111.31	11.33
06	–	376.18	151.08	1.00	376.18	151.08
07	7.48	86.75	11.59	7.48	86.75	11.59
08	16.46	84.25	5.12	16.46	84.25	5.12
09	13.80	345.04	25.01	13.80	345.04	25.01
10	23.21	370.83	15.98	23.21	370.83	15.98
11	5686.27	36.95	–	5686.27	36.95	1.00
12	7.67	90.15	11.75	7.67	90.15	11.75
13	89.52	8.73	–	89.52	8.73	1.00
14	655.71	711.64	–	655.71	711.64	1.00
15	2.79	379.38	136.10	2.79	379.38	136.10
16	29.86	231.90	7.16	29.86	231.90	7.16
17	–	30.12	13.86	1.00	30.12	13.86
18	12.68	379.77	29.95	12.68	379.77	29.95
19	–	46.90	26.75	1.00	46.90	26.75
20	–	11.36	5.52	1.00	11.36	5.52
ABF	32.79	88.18	15.65	11.51	88.18	9.03
SEM	402.44	42.27	11.27	338.66	42.27	10.46

Bayes Factors smaller than 1 would indicate that there is positive evidence for the second model in the comparison. According to the heuristics by Jeffreys (1961), Bayes Factors between 3 and 10 (or their inverse) can be regarded as substantial, those higher than 10 as strong. If no value is provided (–), the model comparison did not yield consistent evidence for either model. Consequently, no Bayes Factor was calculated by SPM5. In order to account for these missings, they were replaced by the value 1 in the right half of the table, by which the analysis becomes more conservative.

Table 1b
Model comparison for six subjects who had activation in all three VOIs but for which the local maximum was outside the defined search volume (cf. Materials and methods).

Subject	Original model selection			Model selection with replaced missings		
	#1 vs. #2	#1 vs. #3	#3 vs. #2	#1 vs. #2	#1 vs. #3	#3 vs. #2
21	836.82	12.87	65.06	836.82	12.87	65.06
22	283.45	38.36	7.39	283.45	38.36	7.39
23	348.68	54.79	6.37	348.68	54.79	6.37
24	212.54	2413.13	–	212.54	2413.13	1.00
25	319.49	1.00	320.31	319.49	1.00	320.31
26	33.31	5.07	6.57	33.31	5.07	6.57
ABF	239.15	26.29	23.01	239.15	26.29	13.64
SEM	109.67	398.54	60.85	109.67	398.54	51.45

General linear model analysis

In the general linear model analysis, we confirmed that the previously reported pattern of results (Heim et al., 2008) was replicated in the subset of volunteers tested in the present study. There was a common activation of BA 44 and BA 45 in both fluency tasks. In BA 45, the activation was equally high for semantic and phonological fluency, whereas the activation was higher for phonological than semantic processing in BA 44 (Fig. 2). The local maxima of the activations were located at MNI coordinates (x,y,z) −46,10,4 (BA 44), −44,28,22 (BA 45) and −44, −14,36 (motor cortex).

Bayesian model selection

The pair-wise comparison of the three models revealed superior evidence in favour of model #1. It was superior to model #2 (ABF = 32.79), and received more evidence than model #3 (ABF = 88.18). Moreover, there was positive evidence in favour of model #3 relative to model #2 (ABF = 15.65). The individual BFs for each pair-wise model comparison are presented in Table 1a.

The same pattern also emerged for the six subjects whose local maxima were in the VOIs but outside defined search volume (see Table 1b for details).

Parameters inference

The random effects analysis (one-sample *t*-tests) on the parameters of model #1 receiving the highest evidence in the BMS procedure yielded the following results (cf. Tables 2a, 2b and 2c). The values for the driving inputs into BA 44 and BA 45 were all significantly greater than zero ($p<0.05$ Bonferroni-corrected for multiple comparisons). This implies that activation in BA 44 and BA 45 is increased when one of the two verbal fluency tasks is performed. Similarly, the intrinsic, i.e., not context dependent, connections

Table 2a
Average parameter estimates of model #1, their standard errors, and their significances in one-sample *t*-tests.

	Mean	Standard error	t	df	p
Intrinsic connectivity BA 44 → BA 45	−0.12	0.04	−3.25	19	0.004
Intrinsic connectivity BA 44 → M1	0.26	0.04	6.15	19	<0.001
Intrinsic connectivity BA 45 → BA 44	0.38	0.08	5.10	19	<0.001
Intrinsic connectivity BA 45 → M1	0.43	0.06	7.14	19	<0.001
Input of SEM into BA 45	0.19	0.03	5.55	19	<0.001
Input of PHO into BA 44	0.06	0.01	5.36	19	<0.001
Input of PHO into BA 45	0.16	0.03	5.18	19	<0.001
Effect of SEM on connectivity BA 45 → M1	0.28	0.05	5.44	19	<0.001
Effect of PHO on connectivity BA 44 → M1	0.16	0.03	5.40	19	<0.001
Effect of PHO on connectivity BA 45 → M1	0.18	0.03	6.95	19	<0.001

Abbreviations: SEM, semantic fluency; PHO, phonological fluency; M1, motor cortex (BA 4p).

Table 2b
Average parameter estimates of model #1 for the remaining six subjects who also activated all three VOIs.

	Mean	Standard error	t	df	p
Intrinsic connectivity BA 44 → BA 45	−0.08	0.07	−1.18	5	0.29
Intrinsic connectivity BA 44 → M1	0.13	0.03	3.95	5	0.01
Intrinsic connectivity BA 45 → BA 44	0.31	0.08	3.94	5	0.01
Intrinsic connectivity BA 45 → M1	0.37	0.08	4.81	5	0.01
Input of SEM into BA 45	0.12	0.02	5.78	5	0.00
Input of PHO into BA 44	0.04	0.03	1.62	5	0.17
Input of PHO into BA 45	0.12	0.03	3.81	5	0.01
Effect of SEM on connectivity BA 45 → M1	0.22	0.05	4.49	5	0.01
Effect of PHO on connectivity BA 44 → M1	0.08	0.03	2.34	5	0.07
Effect of PHO on connectivity BA 45 → M1	0.12	0.04	2.79	5	0.04

For details cf. Table 2a; p values are uncorrected.

between BA 44, BA 45, and motor cortex all significantly differed from zero ($p < 0.05$ Bonferroni-corrected for multiple comparisons). Such positive intrinsic connections imply that activation in one of the regions propagates to the connected regions, thus increasing the activation level there. In the present model, this holds for the connections from BA 44 and BA 45 towards the motor cortex, indicating that task-induced activation in BA 44 and BA 45 results in activation increase in the motor cortex. Similarly, task-induced activation in BA 45 exerts facilitatory influence on the activation in BA 44 and thus adds to the activation induced into BA 44 by the phonological fluency task. Interestingly, there was a negative intrinsic influence from BA 44 to BA 45, implying that the activation of BA 44 inhibits and thus reduces the activation level in BA 45. Finally, the modulations of the intrinsic connections from BA 44 and BA 45 to the motor cortex by the semantic and phonological tasks were all significantly greater than zero, indicating an increase of the respective coupling strengths' when subjects had to engage in the verbal fluency task. A schematic of the inputs, intrinsic connections, and modulations of the model with the best evidence is shown in Fig. 3.

Model simulation

Finally, the parameters of each of the 20 simulated models based on the synthetic datasets were compared to the parameters of the empirical model #1. The parameters were identical to those of model #2, revealing that no bias was present in model #1.

Discussion

We here compared three dynamic causal models representing alternative hypotheses about the roles of cytoarchitectonic BA 44 and BA 45 during verbal fluency. The applied Bayesian model selection procedure revealed that a model featuring domain-preferentiality of BA 44 received higher empirical evidence than models reflecting a functional specialisation (i.e. BA 44 supporting phonological and BA
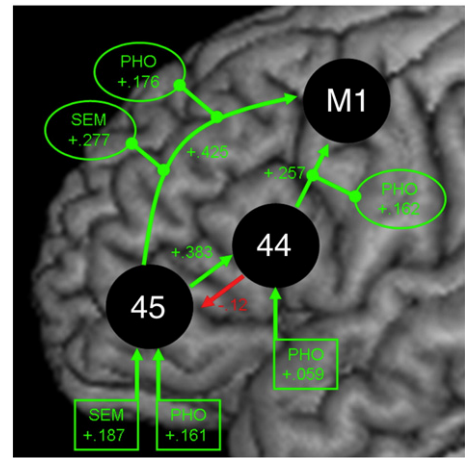
Table 2c
Average parameter estimates of model #1 for the original sample plus the remaining six subjects who also activated all three VOIs.

	Mean	Standard error	t	df	p
Intrinsic connectivity BA 44 → BA 45	−0.11	0.03	−3.47	25	0.002
Intrinsic connectivity BA 44 → M1	0.23	0.03	6.65	25	<0.001
Intrinsic connectivity BA 45 → BA 44	0.37	0.06	6.10	25	<0.001
Intrinsic connectivity BA 45 → M1	0.41	0.05	8.46	25	<0.001
Input of SEM into BA 45	0.17	0.03	6.42	25	<0.001
Input of PHO into BA 44	0.06	0.01	5.39	25	<0.001
Input of PHO into BA 45	0.15	0.02	6.04	25	<0.001
Effect of SEM on connectivity BA 45 → M1	0.27	0.04	6.52	25	<0.001
Effect of PHO on connectivity BA 44 → M1	0.14	0.03	5.68	25	<0.001
Effect of PHO on connectivity BA 45 → M1	0.16	0.02	7.41	25	<0.001

For details cf. Table 2a; p values are uncorrected.

Fig. 3. Significant parameter estimates (inputs, intrinsic connections, and modulations) of the DCM with the best evidence (model #1). Green colour indicates values greater than zero, red colour negative values. Arrows with numbers in rectangular boxes signify driving inputs. Arrows with numbers without boxes show intrinsic connections. Modulatory influences are indicated by oval boxes. Abbreviations: SEM, semantic fluency; PHO, phonological fluency; M1, motor cortex.

45 supporting semantic processing) or a shared functionality (i.e. BA 44 and BA 45 being equally involved in phonological and semantic processing). Domain-preferentiality in the model with the best evidence was characterised by the involvement of BA 44 (and BA 45) in both phonological and semantic fluency while BA 44 received additional task-driven input from phonological processing.

This result implies that a domain-specificity of BA 44 (for phonological processing) and BA 45 (for semantic processing) may not be sufficient to fully explain their contribution to verbal fluency. Such fundamental dichotomy of the roles of BA 44 and BA 45 had been suggested on the basis of a recent literature review of verbal fluency studies (Costafreda et al., 2006; Gough et al., 2005). Such view was not supported by the results of the present study. Rather, the model that was most supported by the experimental data is in line with the hypothesis that BA 45 is involved in processes of word retrieval in general (i.e. according to semantic and phonological criteria) whereas BA 44 is co-activated while additional phonological processing takes place (cf. Heim et al., 2008). Such position does not contradict earlier studies suggesting the involvement of BA 45 in the selection of words during semantic fluency (e.g. Amunts et al., 2004), but rather complements them by providing a combined account for semantic and phonological processing. At this point, it should be noted that the present findings and conclusions only apply to verbal fluency. Other studies (e.g. Gough et al., 2005) have convincingly demonstrated a double dissociation between the posterior and the anterior portion of the left IFG during word comprehension. When subjects performed semantic synonym judgements, transcranial magnetic stimulation interfered with the performance only at the anterior but not the posterior stimulation site. The reverse pattern was observed for phonological homophone judgements. In this study, however, no words were generated after retrieval from the mental lexicon; rather, the judgements were performed for given stimuli. Therefore, it is well possible that domain-specificity may, under certain task demands, characterise the roles of the posterior vs. anterior portion of the left IFG during word comprehension, whereas domain-preferentiality (as in the present study) is a better characterisation during word generation. An alternative explanation for the different effects in the study by Gough et al. (2005) and the present study could be that, overall, the semantic task in the Gough et al. (2005) study was more difficult than the phonological task, as indicated by the reduced accuracy during semantic judgements. If the anterior portion of the left IFG (presumably BA 45) is involved in procedures of selection and

retrieval, any interference should be more dramatic the more difficult the task becomes — exactly the pattern observed by Gough et al. (2005). On the other hand, a more pronounced interference effect for phonological processing in the posterior portion of the left IFG (presumably BA 44) is perfectly in line with the present data, which revealed additional phonological processing in BA 44. Thus, the seemingly different findings by Gough et al. (2005) can be explained in the framework of the present study. The least common determiner, however, is that domain-preferentiality is a good description of the roles of BA 44 and BA 45 during the retrieval of words during verbal fluency tasks.

The question arises as to how the functional interplay of BA 45 and BA 44 during verbal fluency can be further characterised. The co-activation of BA 44 when BA 45 is activated by semantic and phonological fluency, plus the additional driving input into BA 44 only by phonological fluency, is in line with the notion of domain-preferentiality of BA 44 that was proposed by Gold et al. (2005; see also Gold and Buckner, 2002). Domain-preferentiality implies that a brain region is preferentially involved in one type of cognitive process (here: phonological processing) but also supports other processes (here: word retrieval in general). In other words, the data suggest that BA 44 is implicated in (at least) two processes during verbal fluency, one process specific for phonological fluency and another process more generally involved in word retrieval. Evidence for the notion that different types of phonological processes are supported by BA 44 comes from a wealth of neuroimaging studies employing a number of different phonological tasks. Such tasks included phonemic decisions on syllable pairs (Burton et al., 2000), phonemic sequencing (Démonet et al., 1992), grapheme-to-phoneme conversion in words with high vs. low letter-to-sound correspondence (Fiez et al., 1999), tests of phonemic awareness (Katzir et al., 2005), or syllable counting (Poldrack et al., 1999). Assuming the occurrence of at least two different phonology-related processes in BA 44, a potential candidate process specific for phonological fluency is the processing of the phonemic cue for word generation in the phonological fluency task. Such processing would include the maintenance in phonological working memory and the comparison of the initial phoneme of a retrieved word with this cue. In line with this assumption, Zurowski et al. (2002) demonstrated the involvement of the opercular part of the left IFG (approximately BA 44) in both phonological working memory and phonological decisions. The second, more general process shared by phonological and semantic verbal fluency tasks could either be the retrieval of the phonological code (i.e. the sound form) of a to-be-produced word, or the subsequent syllabification (i.e. the grouping of phonemes into syllables; cf. the language production model by Levelt et al., 1999). The present study was not designed to distinguish between these two alternative processes, which are both related to phonological processing in language production. Yet, a tentative conclusion can be drawn on the basis of a review of neuroimaging studies involving overt speaking (Indefrey and Levelt, 2004). According to this review, syllabification crucially involves the posterior aspect of the left IFG (and thus BA 44) whereas the retrieval of the phonological code appears to recruit the posterior aspects of the left superior and middle temporal gyri. Thus, it seems likely that the second, unspecific process supported by BA 44 during verbal fluency involves the building of syllables out of phonemes as a preparation for subsequent articulation.

To summarise, we have so far identified three processes occurring in Broca's region during verbal fluency tasks: word retrieval from the mental lexicon (BA 45), syllabification (BA 44), and the processing of the phonemic cue according to which words must be selected during phonological fluency tasks (also BA 44). One test of this hypothetical scenario would be whether it is in accordance with the parameter values of the model, i.e. the pattern of intrinsic connections and their task-dependent modulations. The parameter tests revealed that all regions were connected to each other. Most of these connections were

positive, reflecting the enhancing flow of information from one region to the other. Interestingly, however, the BA 44 → BA 45 connection was negative. This pattern implies that the increase of activation in BA 44 results in a decrease of activation in BA 45, i.e., the processes in BA 44 inhibit processing in BA 45. Following the argumentation about the functions of BA 44 and BA 45 during verbal fluency, one would thus argue that the processing of the phonemic task cue or the syllabification (or both) inhibit the retrieval of words from the mental lexicon. In other words, the inhibitory influence of BA 44 onto BA 45 may have two functional aspects for the retrieval of words, one related to the processing of the phonemic task cue and the other to syllabification. First, when a phonemic cue is given, the search for an appropriate entry in the mental lexicon must be restricted to the cohort starting with that phoneme while all other entries must be inhibited. Second, when an entry has been retrieved from the mental lexicon and its phonological code is subsequently processed, no further lexical search is required until the initiation of the search for the next word. This implies that during phonological processing of a to-be-produced word the lexical search in BA 45 must be interrupted, which may require inhibition from BA 44 where the phonological processing takes place. These two options, i.e. limitation of the lexical search by the phonemic cue vs. interruption of the lexical search during syllabification of a retrieved entry from the mental lexicon, are not mutually exclusive. However, the present analysis does not provide sufficient information to distinguish between the two options, or to reveal whether both types of inhibitory processes occur in alternation. The investigation of these two potential processes would require an event-related design involving the independent manipulation of both processes and must thus be referred to future research. For the purpose of the present study it is important to stress that the pattern of parameters, and in particular the inhibitory connection from BA 44 to BA 45, is in line with the functions of BA 44 and BA 45 proposed above (processing of the phonological cue and syllabification in BA 44, word retrieval in BA 45).

Apart from such detailed analysis of processes, it is striking that BA 44 exerts inhibitory influence over BA 45. Only few studies of functional connectivity have so far distinguished between sub-regions of the left IFG (Bitan et al., 2007; Bokde et al., 2001; Heim et al., 2009; Mechelli et al., 2005) whereas others regarded the IFG as one functional unit (e.g. Allen et al., 2008; Bitan et al., 2005, 2006; Friston et al., 2003; Quaglino et al., 2008). Among those studies that did distinguish regions within the left IFG, only two (Bitan et al., 2007; Heim et al., 2009) analysed the effective connectivity between them. Both studies found positive mutual intrinsic connections between cytoarchitectonic BA 44 and BA 45 (Heim et al., 2009) and the ventral vs. dorsal aspect of the IFG, respectively (Bitan et al., 2007). Thus, at first sight, these studies seem to be in discordance with the finding of the present study that BA 44 has a negative intrinsic connection to BA 45. However, one must consider that the intrinsic connections depend on the experimental task that is used in a study. As Bitan et al. (2006) showed, the strengths of intrinsic connections between regions can vary with task demands, and even their polarities may change. Therefore, it is plausible that BA 44 and BA 45 may have different patterns of intrinsic connections during the reading of words vs. pseudo-words (Heim et al., 2009), rhyming judgements on visually presented words (Bitan et al., 2007), and the overt production of words in different types of verbal fluency tasks (the present study).

With respect to the generalisability of the findings, two aspects should be considered. (1) One might criticise that only about 4/7 of the data from the original GLM analysis could be included into the present DCM analysis, because only 20 out of 28 subjects had distinguishable activation in all three VOIs according to the sharp criteria we defined a priori. However, the additional analysis of six of the remaining subjects who also activated all three VOIs corroborates the findings of the primary DCM analysis. For these six subjects, model

#1 was also the model with highest positive evidence; moreover, the connectivity pattern was comparable, at least qualitatively. Although the input into BA 44 and the negative intrinsic connection from BA 44 to BA 45 failed to reach significance, this might well be explained by the weak power, considering the small sample size here. As Table 2c shows, collapsing the two samples provides again strong evidence for the pattern of results discussed above. Finally, the validity of the results of our original sample was additionally corroborated by the DCM model simulation using synthetic data, which revealed no evidence for any bias in the selected model #1 based on empirical data. (2) One potential limitation of the generalisability of the data arises from the fact that the three DCMs which were compared in the present study might be regarded as being unequally parsimonious, given the results of the previous GLM analysis. Considering the pattern of driving inputs, model #1, which received the highest amount of evidence in the Bayesian model selection procedure, resembles most closely the activation pattern in the GLM analysis, with equal input of semantic and phonological fluency into area 45 and additional phonological input into area 44. One could argue that the other two models require additional assumptions about the strengths of intrinsic connections in order to explain the GLM data equally well. This argument is partly correct, in particular because the driving inputs in a DCM reflect the activation strength in the GLM analysis. However, a DCM does not only contain driving inputs as a sole parameter set, but rather, three sets of parameters. Intrinsic connections and their modulations are as important as the driving inputs in order to model the influences the regions exert over one another. The notion of "additional assumptions" does not take this notion of three parameter sets with equal importance into account. In addition, there is a empirical evidence against the parsimony issue in the data. If model #1 was a priori the most parsimonious, one would expect high BFs for the comparisons of model #1 against the two other models, but low BFs for the comparison between the two other models. The empirical pattern is different. Whereas the comparison of models #1 and #3 indeed shows a high corrected BF (88.18), the BF in the comparison of models #1 and #2 (11.51) is only in the same range as that in the comparison of models #3 and #2 (9.03). In order to fully rule out the parsimony question, future research should test the three DCMs defined in the present study in a new sample of subjects undergoing a replication of the study, and corroborate the analysis presented here.

Conclusion

Here we used DCM in order to test alternative hypotheses about the roles of cytoarchitectonic BA 44 and BA 45 during verbal fluency. This DCM analysis provides empirical evidence for the differential involvement of these two distinct parts of Broca's region in the retrieval of words from the mental lexicon that could not be obtained from the conventional GLM analysis. Importantly, both the architecture and the parameters of the model with the best evidence contribute to the understanding of the processes taking place in Broca's region during verbal fluency. DCM suggested that BA 45 supports the retrieval of words from the mental lexicon in any of the two fluency tasks, whereas BA 44 is involved in phonological processing including the use of the phonemic task cue in phonological fluency and the syllabification of words in both fluency tasks. It appears that the functional data are best accounted for by a model combining shared and specialised processes sustained by BA 44 and BA 45, i.e. domain-preferentiality, rather than domain-specificity or the lack of any functional distinction.

Acknowledgments

References

Allen, P., Mechelli, A., Stephan, K.E., Day, F., Dalton, J., Williams, S., McGuire, P.K., 2008. Fronto-temporal interactions during overt verbal initiation and suppression. J. Cogn. Neurosci. 20 (9), 1656–1669.

Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H.B.M., Zilles, K., 1999. Broca's region revisited: cytoarchitecture and intersubject variability. J. Comp. Neurol. 412, 319–341.

Amunts, K., Weiss, P.H., Mohlberg, H., Pieperhoff, P., Eickhoff, S., Gurd, J.M., Marshall, J.C., Shah, N.J., Fink, G.R., Zilles, K., 2004. Analysis of neural mechanisms underlying verbal fluency in cytoarchitectonically defined stereotaxic space — the roles of Brodmann areas 44 and 45. NeuroImage 22, 42–56.

Basho, S., Palmer, E.D., Rubio, M.A., Wulfeck, B., Muller, R.A., 2007. Effects of generation mode in fMRI adaptations of semantic fluency: paced production and overt speech. Neuropsychologia 45, 1697–1706.

Bitan, T., Booth, J.R., Choy, J., Burman, D.D., Gitelman, D.R., Mesulam, M.M., 2005. Shifts of effective connectivity within a language network during rhyming and spelling. J. Neurosci. 25, 5397–5403.

Bitan, T., Burman, D.D., Lu, D., Cone, N.E., Gitelman, D.R., Mesulam, M.M., Booth, J.R., 2006. Weaker top-down modulation from the left inferior frontal gyrus in children. NeuroImage 33, 991–998.

Bitan, T., Burman, D.D., Chou, T.L., Lu, D., Cone, N.E., Cao, F., Bigio, J.D., Booth, J.R., 2007. The interaction between orthographic and phonological information in children: an fMRI study. Hum. Brain Mapp. 28, 880–891.

Bokde, A.L., Tagamets, M.A., Friedman, R.B., Horwitz, B., 2001. Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. Neuron 30, 609–617.

Bookheimer, S., 2002. Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. Annu. Rev. Neurosci. 25, 151–188.

Burton, M.W., Small, S.L., Blumstein, S.E., 2000. The role of segmentation in phonological processing: an fMRI investigation. J. Cogn. Neurosci. 12, 679–690.

Costafreda, S.G., Fu, C.H., Lee, L., Everitt, B., Brammer, M.J., David, A.S., 2006. A systematic review and quantitative appraisal of fMRI studies of verbal fluency: role of the left inferior frontal gyrus. Hum. Brain Mapp. 27, 799–810.

Démonet, J.F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.L., Wise, R., Rascol, A., Frackowiak, R., 1992. The anatomy of phonological and semantic processing in normal subjects. Brain 115, 1753–1768.

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.

Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., Zilles, K., 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. NeuroImage 25, 1325–1335.

Eickhoff, S.B., Heim, S., Zilles, K., Amunts, S., 2006. Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. NeuroImage 32, 570–582.

Eickhoff, S.B., Heim, S., Zilles, K., Amunts, S., 2009. A systems perspective on the effective connectivity of overt speech production. Phil. Trans. Royal Soc. A 367, 2399–2421.

Fiez, J.A., Balota, D.A., Raichle, M.E., Petersen, S.E., 1999. Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. Neuron 24, 205–218.

Friederici, A.D., 2002. Towards a neural basis of auditory sentence processing. Trends Cogn. Sci. 6, 78–84.

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. NeuroImage 19, 1273–1302.

Geyer, S., Ledberg, A., Schleicher, A., Kinomura, S., Schormann, T., Bürgel, U., Klingberg, T., Larsson, J., Zilles, K., Roland, P.E., 1996. Two different areas within the primary motor cortex of man. Nature 382, 805–807.

Gold, B.T., Buckner, R.L., 2002. Common prefrontal regions coactivate with dissociable posterior regions during controlled semantic and phonological tasks. Neuron 35, 803–812.

Gold, B.T., Balota, D.A., Kirchhoff, B.A., Buckner, R.L., 2005. Common and dissociable activation patterns associated with controlled semantic and phonological processing: evidence from FMRI adaptation. Cereb. Cortex 15, 1438–1450.

Gough, P.M., Nobre, A.C., Devlin, J.T., 2005. Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. J. Neurosci. 25, 8010–8016.

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

Heim, S., Eickhoff, S.B., Amunts, K., 2008. Specialisation in Broca's region for semantic, phonological, and syntactic fluency? NeuroImage 40, 1362–1368.

Heim, S., Eickhoff, S.B., Ischebeck, A.K., Friederici, A.D., Stephan, K.E., Amunts, K., 2009. Effective connectivity of the left BA 44, BA 45, and inferior temporal gyrus during lexical and phonological decisions identified with DCM. Hum. Brain Mapp. 30, 392–402.

Indefrey, P., Levelt, W.J.M., 2004. The spatial and temporal signatures of word production components. Cognition 92, 101–144.

Jeffreys, H., 1961. The Theory of Probability3rd ed. Clarendon Press, Oxford, UK.

Katzir, T., Misra, M., Poldrack, R.A., 2005. Imaging phonology without print: assessing the neural correlates of phonemic awareness using fMRI. NeuroImage 27, 106–115.

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

Mechelli, A., Crinion, J.T., Long, S., Friston, K.J., Lambon Ralph, M.A., Patterson, K., McClelland, J.L, Price, C.J., 2005. Dissociating reading processes on the basis of neuronal interactions. J. Cogn. Neurosci. 17, 1753–1765.

Penny, W.D., Stephan, K.E., Mechelli, A., Friston, K.J., 2004. Comparing dynamic causal models. NeuroImage 22, 1157–1172.

Poldrack, R.A., Wagner, A.D., Prull, M.W., Desmond, J.E., Glover, G.H., Gabrieli, J.D., 1999. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. NeuroImage 10, 15–35.

Quaglino, V., Bourdin, B., Czternasty, G., Vrignaud, P., Fall, A.S., Meyer, M.E., Berquin, P., Devauchelle, B., de Marco, G., 2008. Differences in effective connectivity between dyslexic children and normal readers during a pseudoword reading task: an fMRI study. Neurophysiol. Clin. 38 (2), 73–82.

Schleicher, A., Amunts, K., Geyer, S., Morosan, P., Zilles, K., 1999. Observer-independent method for microstructural parcellation of cerebral cortex: a quantitative approach to cytoarchitectonics. NeuroImage 9, 165–177.

Smith, A.P., Stephan, K.E., Rugg, M.D., Dolan, R.J., 2006. Task and content modulate amygdala–hippocampal connectivity in emotional retrieval. Neuron 49, 631–638.

Stephan, K.E., Penny, W.D., 2007. Dynamic causal models and Bayesian selection. In: Friston, KJ (Ed.), Statistical Parametric Mapping. Elsevier, Amsterdam, pp. 577–585.

Thompson-Schill, S.L., 2003. Neuroimaging studies of semantic memory: inferring "how" from "where.". Neuropsychologia 41, 280–292.

Zilles, K., Schleicher, A., Palomero-Gallagher, N., Amunts, K., 2002. Quantitative analysis of cyto- and receptor architecture of the human brain. In: Mazziotta, J., Toga, A. (Eds.), Brain Mapping, The Methods. Academic Press, San Diego, pp. 573–602.

Zurowski, B., Gostomzyk, J., Gron, G., Weller, R., Schirrmeister, H., Neumeier, B., Spitzer, M., Reske, S.N., Walter, H., 2002. Dissociating a common working memory network from different neural substrates of phonological and spatial stimulus processing. NeuroImage 15, 45–57.