

Effective Connectivity and Intersubject Variability: Using a Multisubject Network to Test Differences and Commonalities

Andrea Mechelli,^{*1} Will D. Penny,^{*} Cathy J. Price,^{*} Darren R. Gitelman,[†] and Karl J. Friston^{*}

^{*} Wellcome Department of Imaging Neuroscience, Institute of Neurology, 12 Queen Square, London, WC1N 3BG, United Kingdom; and [†] Cognitive Neurology and Alzheimer's Disease Center, Department of Neurology and Department of Radiology, Northwestern University, Chicago, Illinois 60611

Received January 29, 2002

This article is about intersubject variability in the functional integration of activity in different brain regions. Previous studies of functional and effective connectivity have dealt with intersubject variability by analyzing data from different subjects separately or pretending the data came from the same subject. These approaches do not allow one to test for differences among subjects. The aim of this work was to illustrate how differences in connectivity among subjects can be addressed explicitly using structural equation modeling. This is enabled by constructing a multisubject network that comprises m regions of interest for each of the n subjects studied, resulting in a total of $m \times n$ nodes. Constructing a network of regions from different subjects may seem counterintuitive but embodies two key advantages. First, it allows one to test directly for differences among subjects by comparing models that do and do not allow a particular connectivity parameter to vary over subjects. Second, a multisubject network provides additional degrees of freedom to estimate the model's free parameters. Any neurobiological hypothesis normally addressed by single-subject or group analyses can still be tested, but with greater sensitivity. The common influence of experimental variables is modeled by connecting a virtual node, whose time course reflects stimulus onsets, to the sensory or "input" region in all subjects. Further experimental changes in task or cognitive set enter through modulation of the connections. This approach allows one to model both endogenous (or intrinsic) variance and exogenous effects induced by experimental design. We present a functional magnetic resonance imaging study that uses a multisubject network to investigate intersubject variability in functional integration in the context of single word and pseudoword reading. We tested whether the effect of word type on the reading-related coupling differed significantly among subjects. Our results

showed that a number of forward and backward connections were stronger for reading pseudowords than words, and, in one case, connectivity showed significant intersubject variability. The discussion focuses on the implications of our findings and on further applications of the multisubject network analysis. © 2002 Elsevier Science (USA)

INTRODUCTION

This article is about intersubject variability in the functional integration of activity in different brain regions. Differences in functional integration may reflect varied cognitive strategies adopted to perform a task or degenerative solutions the brain can adopt to complete the same strategy. First, we review briefly how previous neuroimaging studies of functional and effective connectivity have dealt with differences and commonalities among subjects. Second, we describe how the issue of intersubject variability can be addressed directly and illustrate the approach with empirical data.

Functional integration refers to the interaction among functionally specialized brain regions and can be characterized in terms of functional or effective connectivity (e.g., McIntosh *et al.*, 1994; Horwitz and Sporns, 1994; Friston *et al.*, 1993a,b). Functional connectivity is defined as the temporal correlations among neurophysiological events in different neural systems, whereas effective connectivity is defined as the influence that one neural system exerts over another. Over the past decade, a number of positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) studies have tried to characterize functional and effective connectivity. While some PET studies have estimated connectivity by exploiting subject-to-subject variability (e.g., see Horwitz *et al.*, 1998), most studies have discounted intersubject variability by treating data from different subjects as if they came from the same subject (Iidaka *et al.*, 2001; Bokde *et al.*, 2001; Bullmore *et al.*, 2000; Maguire *et al.*, 2000; Coull

¹ To whom correspondence should be addressed. Fax: 0044-020-78131420. E-mail: andream@fil.ion.ucl.ac.uk.

et al., 1999). Here estimates of the coupling among regions of interest are based on item-to-item or block-to-block variability over subjects. This approach assumes that the pattern of connectivity estimated over subjects is a good approximation to the underlying connectivity in all the subjects studied. Treating data from different subjects in this way allows one to draw inferences about the group of subjects studied under the assumption that variations in connectivity from subject to subject are random, well-behaved, and uninteresting. It also increases sensitivity by augmenting the number of data points per voxel. However, this approach may be problematic when intersubject variability is pronounced. For example, when there are subjects with opposite patterns of connectivity, the estimated connections may be far from the true values in any of the subjects studied. This was illustrated by Gonçalves *et al.* (2001), who found that the results of the group analysis did not reflect any of the features seen in individual analyses. In addition, estimating connectivity over subjects does not allow one to relate clinical or demographic measures (e.g., task performance or age) to indices of functional integration (e.g., correlations among regions) in a subject-specific fashion.

An alternative approach involves analyzing the data by performing subject-specific analyses (Gonçalves *et al.*, 2001; Mechelli *et al.*, 2001; Büchel and Friston, 1997, 1998). This method allows one to estimate connectivity independently for each subject. In addition, it enables one to correlate behavioral measures with indices of functional integration over subjects. However, this approach indicates apparently inconsistent results among subjects (Gonçalves *et al.*, 2001; Mechelli *et al.*, 2001). This raises the question of whether inconsistent effects correspond to significant differences in connectivity.

In summary, unless the model of the data can accommodate subject-specific variations in connectivity, differences among subjects cannot be evaluated. The aim of the present work is to show how individual differences in connectivity can be addressed directly using structural equation modeling (SEM) (Bollen, 1989). SEM of functional imaging data identifies connection strengths that best predict the variance–covariance structure of the empirical data (i.e., fMRI time series), under the constraints of a prespecified anatomical model. Generally, this is achieved by using iterative methods such as maximum likelihood or generalized least squares. Iteration begins by assigning a set of start values to the free parameters of the prespecified anatomical model, from which an implied variance–covariance structure is computed and compared with the variance–covariance structure of the empirical data. Iteration continues until the difference between the implied and the observed variance–covariance structures cannot be minimized any further. This pro-

duces a single number (i.e., value of the fitting function) that is a function of the discrepancy between the two variance–covariance structures. Such discrepancy is asymptotically distributed as χ^2 . χ^2 values are derived by multiplying the value of the fitting function by the sample size minus 1. χ^2 values are therefore an index of how “badly” the prespecified anatomical model fits the data and can be used to compare alternative models.

The model used in this article is based on a multi-subject network that comprises m regions of interest for each of the n subjects studied, resulting in a total of $m \times n$ nodes. Regions pertaining to the same subject are (reciprocally) connected to allow for coupling between different areas of the same brain. In contrast, regions from different subjects are not connected. Differences in functional integration across subjects are tested by comparing a model that allows for intersubject variability with one that does not. This is implemented by setting up a model in which one or more effects of interest (e.g., the modulation of coupling by stimulus type) may vary over subjects versus one in which these effects are held to be equivalent in all subjects. The best model is the one that explains most of the observed variance–covariance structure for the least cost, in terms of the number of model parameters. This is analogous to testing for subject \times condition interactions in classic ANOVA.

Constructing a network that comprises regions from different subjects has two key advantages. First, it allows one to test whether changes in connectivity differ significantly among subjects. Note that significant differences among subjects may occur not only when subject-specific analyses of effective connectivity give inconsistent inferences, but also when subjects show the same effect to a different extent. Second, constructing a network that comprises m regions from each of the n subjects studied provides one with enough non-redundant elements in the covariance matrix $e = \frac{1}{2}(m \times n)(m \times n + 1)$ to estimate all potential connections among regions pertaining to the same brain $p = n \times (m \times (m - 1))$ and all residual variances $m \times n$. In essence, by pretending that all the subjects are scanned at the same time, a multisubject network can be formed that ensures a high degree of sparsity. Note that $e - p \geq 1$ is a necessary but not sufficient condition for a model to be globally identifiable (this is known as the “ t rule”; see Bollen, 1989, p.93). A model is globally identifiable if it is possible to obtain unique values for each free parameter (e.g., connection strength). Global estimability is not always simple to evaluate in practice, especially when the model contains many free parameters, but can be easily detected using computer programs for SEM such as LISREL (Jöreskog and Sörbom, 1990) and EQS (Bentler, 1992).

In contradistinction, when SEM is used to estimate connectivity averaging over subjects or independently

for each subject, the number of parameters that can be estimated $e = \frac{1}{2} m \times (m + 1)$ is fewer than the maximum number of connections that can be specified $p = m \times (m - 1)$ and the residual variances (m) that need to be computed. Usually, the t rule is enforced either by constraining the number of specified connections or by assigning a priori arbitrary values to some of them. However, assuming that one region does not exert an influence over another may be problematic because primate anatomical data show that each region is reciprocally interconnected to a large number of other regions. In addition, assuming a particular a priori value for the coupling between two regions may be difficult when data from previous studies are not available.

The multisubject network approach we used here also differs from previous applications of SEM to functional imaging data in terms of the type of variance that is modeled. Most SEM studies in functional neuroimaging do not explicitly model the influence of exogenous or experimental variables. Rather, they assume that the estimated path coefficients reflect endogenous or intrinsic variance only. However, identical regions in different subjects are likely to be correlated because of the common influence of exogenous variables. If not accommodated, these induced covariances would compromise the goodness of fit of the model. The common influence of the exogenous variables is modeled by connecting a virtual node, whose time course reflects stimulus onsets, with the first sensory or "input" region in all subjects. Further experimental changes in task or cognitive set enter through modulation of the connections. Note that the modeling of both endogenous and exogenous variance is not specific to the use of a multisubject network but is also possible in the context of individual subject and group analysis. It is an important device because it allows explanatory variables or design variables to explain the interregional covariance induced. This approach is particularly essential in multisubject networks if the subjects are all exposed to the same paradigm.

Below a study is presented that uses a multisubject network to investigate intersubject variability of functional integration during single word reading. Specifically, we tested whether the modulatory effect of word type on the reading-induced coupling was significantly different among subjects. Neuropsychological data indicate that there are at least two different cognitive strategies that can be used to read: one that relies on access to lexical semantic information and one that involves direct links between spelling (orthography) and sounds (phonology) (see Patterson and Shewell, 1987 and Seidenberg and McClelland, 1989). Familiar words with unusual spellings (e.g., CHOIR) rely on lexical semantic access whereas unfamiliar "pseudo-words" that have no meaning (e.g., CHOIN) rely on sublexical spelling-to-sound relationships. Neverthe-

less, functional imaging studies have not shown a clear double dissociation at a neuroanatomical level for word and pseudoword reading. Pseudoword reading usually enhances inferior frontal activation relative to word reading (Xu *et al.*, 2001; Fiez *et al.*, 1999; Brunswick *et al.*, 1999; Hagoort *et al.*, 1999; Herbster *et al.*, 1997) but there are no brain areas that consistently show enhanced activation for word reading (Mechelli *et al.*, in press). Possible explanations for the functional imaging results are that lexical and sublexical reading "routes" are activated by both word and pseudoword reading; and reduced activation for word relative to pseudoword reading reflects familiarity effects (Hagoort *et al.*, 1999). Studies of effective connectivity may provide an alternative means for dissociating the different reading routes (Horwitz *et al.*, 1998). This approach relies on a priori anatomical models. In the study below, we use the anatomical model of reading developed in Price (2000) (see Fig. 1).

This model distinguishes two routes for reading: one that is shared by objects and involves the left posterior inferior temporal cortex (the semantic route) and one that is not activated by object naming and involves the left posterior superior temporal cortex (Wernicke's area). Although both routes may be activated irrespective of whether words or pseudowords are being read, the effective connectivity between the inferior and the superior posterior temporal areas may differ for word or pseudoword reading. Furthermore, connectivity changes may depend on individual reading strategies. For example, reading words with regular spelling may depend on sublexical processing for some subjects but lexical processing for other subjects. In the last part of the article we discuss our findings and further applications of the multisubject network.

METHODS

The study was approved by the National Hospital for Neurology and Institute of Neurology Medical Ethics Committee.

Subjects

Informed consent was obtained from 13 right-handed volunteers (7 males), aged 20 to 34 (mean age of 24), with English as their first language.

Design

Each subject was presented with blocks of words or pseudowords alternating with blocks of fixation (a cross in the middle of the screen). They were instructed to read each word/pseudoword silently as soon as it appeared on the screen. The variables were (i) stimulus type (words and pseudowords) and (ii) stimulus duration (e.g., the time that stimuli remained on the screen: 200, 600, and 1000 ms), with each of these experimen-

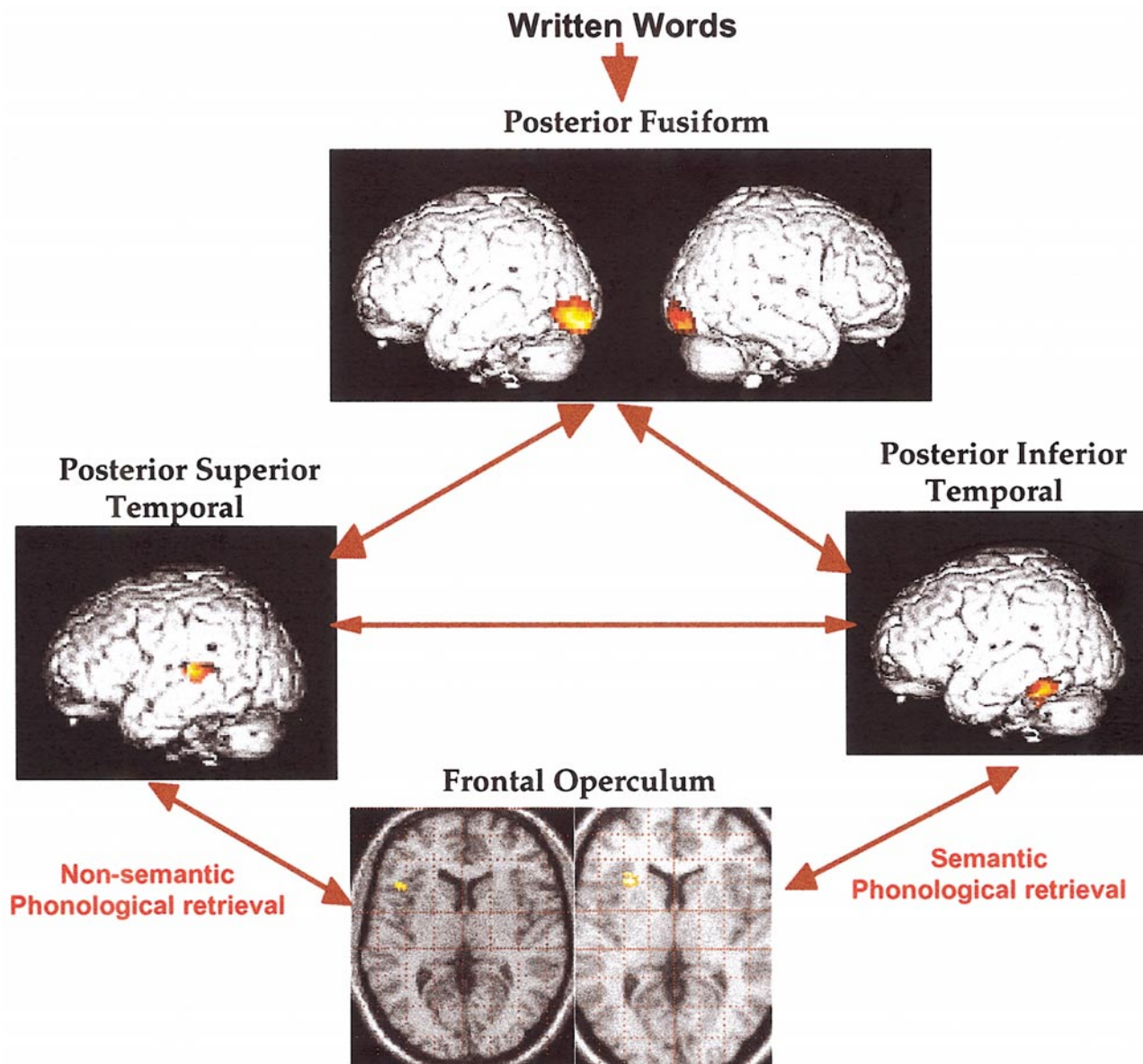


FIG. 1. Anatomical model of reading developed by Price (2000).

tal conditions repeated five times in a counterbalanced order across subjects. Pseudowords (e.g., lenner) refer to letter strings that are not real words, do not have semantic representations, but can be pronounced on the basis of sublexical spelling-to-sound relationships. Words were matched for frequency (Kucera and Francis, 1967), length, and number of syllables between blocks and were as regular as possible in terms of spelling-to-sound rule conversion. Presentation rate was 40 stimuli per minute for a stimulus onset asynchrony of 1500 ms. Each block of 14 stimuli lasted 21 s and was followed by 16 s fixation. An eye tracker was used to ensure that the subjects kept their eyes open and attended to the stimuli.

Data Acquisition

A 2-T Siemens VISION system (Siemens, Erlangen, Germany) was used to acquire both T_1 anatomical volume images ($1 \times 1 \times 1.5$ mm voxels) and T_2^* -weighted echo planar images ($64 \times 64 \times 3$ -mm pixels, $TE = 40$ ms) with BOLD contrast. Each echo planar image comprised thirty-five 1.8-mm axial slices with a 1.2-mm slice interval, giving a resolution of 3 mm. A total of 366 volume images were obtained continuously with an effective repetition time (TR) of 3.15 s per volume, the first six volumes in each session being discarded to allow for T_1 equilibration effects. Stimulus presentation was arranged so that every 90 ms of peristimulus

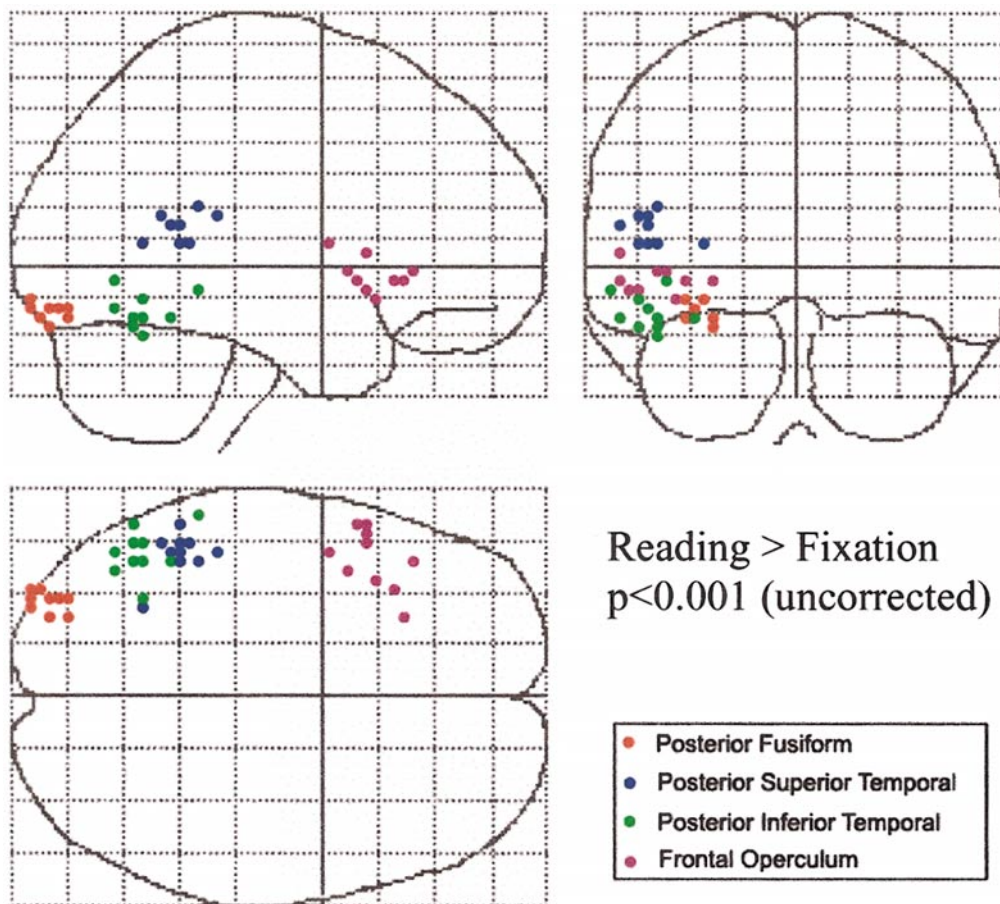


FIG. 2. Subject-specific regions that entered into the SEM. These comprised four left-lateralized regions: posterior fusiform (average coordinates $x = -32$, $y = -87$, $z = -13$); inferior temporal ($x = -48$, $y = -57$, $z = -13$); superior temporal gyrus ($x = -47$, $y = -51$, $z = 13$); and frontal operculum ($x = -45$, $y = 19$, $z = -2$).

time was sampled equally over the session (see Price *et al.*, 1999).

Data Analysis

The regions entering into SEM were defined using a conventional statistical parametric mapping analysis (SPM99: Wellcome Department of Imaging Neuroscience, London, UK. <http://www.fil.ion.ucl.ac.uk>). All volumes from each subject were realigned using the first as a reference and resliced (using sinc interpolation) adjusting for residual motion-related signal changes. A mean volume was created using the realigned volumes and the anatomical magnetic resonance image was coregistered to it. This ensured that the functional and structural images were spatially aligned. The functional images were spatially normalized (Friston *et al.*, 1995a) with respect to the MNI-305 template using nonlinear-basis functions. Functional data were spatially smoothed with a 6-mm full width at half-maximum, isotropic Gaussian kernel. A boxcar waveform convolved with a synthetic hemodynamic response function (HRF) was used as the reference

waveform for each condition. The data were high-pass filtered using a set of discrete cosine basis functions with a cutoff period of 156 s. Each of the six experimental conditions was modeled independently. Parameter estimates and variance were derived in a subject-specific fixed-effect analysis. A statistical parametric map of the main effect of reading was generated independently for each subject. In addition, the contrast images from each subject were entered into one-sample t tests to permit inferences at the group level (i.e., a random effects analysis).

Specification of the Multisubject Network

On the basis of the anatomical model of reading developed by Price (2000), we selected four left-lateralized anatomical regions in the left posterior fusiform, posterior inferior temporal cortex, posterior superior temporal cortex, and frontal operculum. As predicted by the model, these areas were activated for reading relative to fixation in 10 of 13 subjects ($P < 0.001$ uncorrected) and at the group level ($P < 0.05$ corrected for multiple comparisons). The three subjects who

showed no activation of the inferior temporal gyrus for reading relative to fixation were excluded from further analysis. Each region (6-mm radius) was selected specifically for each subject, to correspond with the most significant voxel in the categorical comparison. Regional activity was defined as the first eigenvariate of any region as in Büchel and Friston (1997). All time series were adjusted for confounds (e.g. global mean, low frequency components) after applying the general linear model with condition-specific predictors (Friston *et al.*, 1995b). Figure 2 shows the positions of the selected regions in the remaining 10 subjects.

We constructed a multisubject structural equation model that comprised the four regions above and all potential forward and backward connections among them within each subject. The resulting multisubject network contained $10 \times 4 = 40$ nodes and $10 \times 12 = 120$ connections. In addition, we connected a virtual node, whose time course corresponded to stimulus onsets, with the posterior fusiform region in all subjects which resulted in a total of 41 nodes and 130 connections. This allowed us to model the common influence of exogenous design variables on the induced interregional covariance in all models tested (see Introduction).

SEM Analysis

Structural equation modeling was performed using the SEM Toolbox of SPM99 (Rowe *et al.*, 2002; Maguire *et al.*, 2000; Coull *et al.*, 1999), which employs an iterative maximum likelihood algorithm (Higham, 1993) to estimate covariances that best predict the observed variance-covariance structure of the empirical data. The maximum likelihood algorithm assumes multinormality in the data, whereas the multisubject network contained variables such as word type that were not normally distributed. However, the maximum likelihood algorithm is robust to normality deviations (see Bollen, 1989, p.126) and the final results should not differ markedly if distribution-free methods were to be used. The residual influences were dealt with by fixing them to unity as in Büchel and Friston (1997) in all regions/subjects.

Evaluation of the goodness of fit of the multisubject network. First, we evaluated the goodness of fit of the multisubject network by performing a χ^2 test. Here we set reciprocal connections (e.g., posterior fusiform \rightarrow inferior temporal and inferior temporal \rightarrow posterior fusiform) to have the same value within each subject to ensure that the model could be globally identified. In contrast, no connections were constrained to be the same over different subjects. This allowed us to see whether the multisubject network, which allowed for intersubject variability, accommodated the variance-covariance structure of the empirical data reasonably (i.e., was not an implausible model). Since the meaning

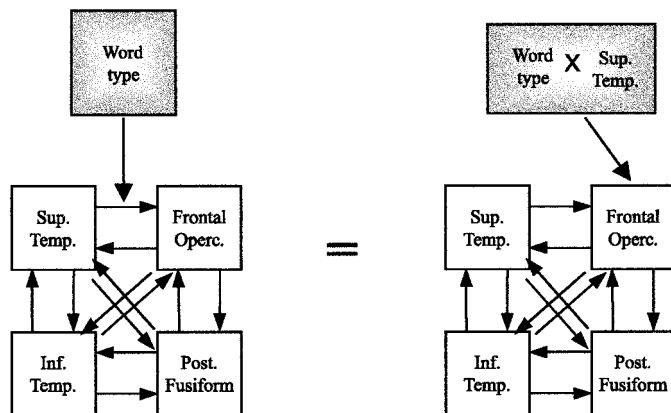


FIG. 3. Subject-specific model for the modulatory effect of word type on connectivity (left) and its implementation in the context of SEM (right). Activity in the “target” region (i.e., frontal operculum) was characterized in terms of a linear combination of (i) activity in the “source” region (i.e., superior temporal); (ii) the design variable (in our analysis a vector that encoded the presentation of pseudowords vs words); and (iii) an interaction term constructed by multiplying the mean-centered time series of the “source” region by the design variable. The influence of the interaction term on the “target” region corresponds to the effect of word type on the coupling between the “source” and “target” regions.

of significant path coefficients in a poorly fitting model is unclear (MacCallum, 1995), we required the multisubject network to provide a good fit according to the χ^2 index (i.e., P value > 0.05).

SEM of word type effects. Second, we tested for the impact of word type on each connection independently (e.g., the impact of words versus pseudowords on the connection posterior fusiform \rightarrow inferior temporal). Here we allowed the connection of interest to differ from its reciprocal connection. However, we set the remaining reciprocal connections to have the same value within each subject. This allowed the impact of word type on each specific connection to be tested without compromising the global estimability of the model.

The impact of word type on connectivity was modeled using a set of virtual nodes whose time course represented the interaction between source activity and the appropriate design variable, as described in Büchel and Friston (1997). These interaction terms were constructed by multiplying the mean-centered time series of the source region by the design variable (in our analysis a vector that encoded the presentation of pseudowords versus words). Activity in any “target” region was characterized in terms of a linear combination of (i) activity in the source region, (ii) the design variable, and (iii) the interaction term. The influence of the interaction term on the “target” area corresponds to the effect of word type on the coupling between the “source” and “target” regions, as depicted in Fig. 3. The interaction terms play exactly the same role as psychophysiological interactions (PPIs) in simple regression

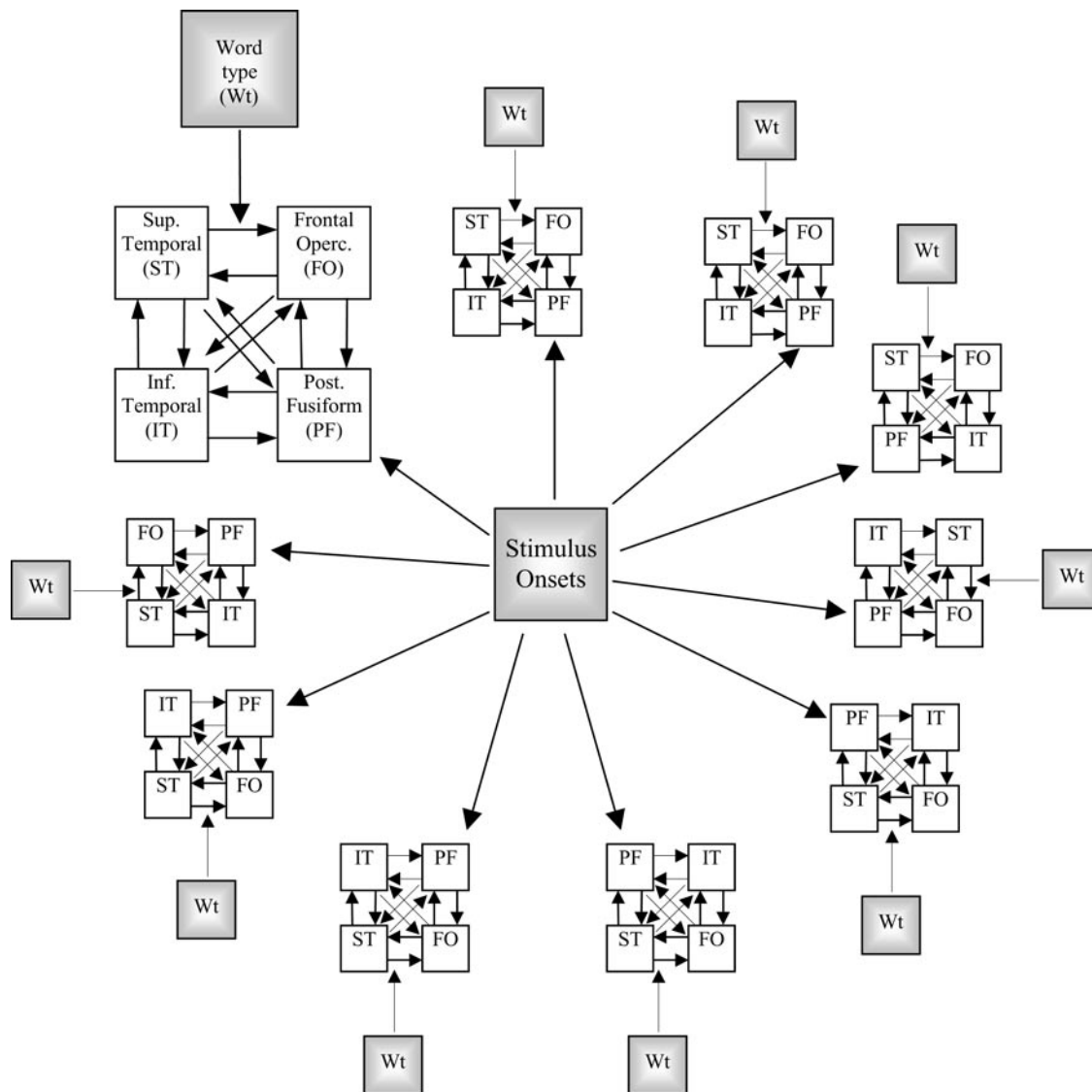


FIG. 4. Multisubject network used to look at the effects of word type on the reading-induced coupling between the superior temporal gyrus and the frontal operculum (similar networks were used to investigate the effects of word type on the remaining connections). The network was composed of the subject-specific regions in Fig. 2. A virtual node, encoding stimulus presentation, was connected to the posterior fusiform gyrus in all subjects to account for the common influence of the exogenous design variables. Word type-dependent changes in connectivity were modeled using a further set of virtual nodes whose time course represented the interaction between source activity and the appropriate design variable. Anatomical nodes are in white and virtual nodes are in gray.

analyses of interregional coupling (Friston *et al.*, 1997).

We assessed the significance of the modulation of connectivity by word type, independently for each connection, by comparing a model in which this stimulus-specific effect was set to be zero with one in which it was the same in all subjects. The best model was identified as the one that explained as much of the observed variance-covariance structure as possible but required the fewest number of model parameters. The significance of the difference between null and alternative models was expressed as the difference in χ^2 goodness of fit with 1 *df* ($P < 0.05$). A total of 12 comparisons

(i.e., one for each connection) were performed to investigate the impact of word type on connectivity, which involved 24 models (12 models in which the effect of word type on a connection of interest was set to be the same in all subjects and 12 models in which the same effect was set to be zero in all subjects). The multisubject network used to investigate the effect of word type on the forward connection between the superior temporal cortex and the frontal operculum is shown graphically in Fig. 4. Similar networks were used to investigate the effects of word type on the remaining connections. Regional nodes are in white and virtual nodes are rendered in gray.

It should be noted that, although reciprocal connections were constrained to be the same, context- or condition-sensitive changes in connectivity were free to be asymmetric. This is because the modulations by word type in each direction were modeled by different virtual nodes. This is important because, although anatomical connections are generally reciprocal, they show asymmetries in their functional expression, with backward connections being more modulatory.

Although the overall impact of word type on connectivity would be the same in a model based on covariances over subjects, such a model would not allow one to estimate subject-specific effects and, most importantly, to test whether they differ significantly across subjects. In contrast, the multisubject network used here allowed us to estimate the overall impact of word type on connectivity over subjects and to test whether the effects showed significant variability across subjects.

SEM of intersubject variability. Intersubject variability of this modulatory effect was assessed independently for each connection by comparing a free model (in which the effect of word type on connectivity was free to vary across subjects) with a restricted model (in which the effects of word type on connectivity were set to be the same in all subjects). Again, the best model was identified as the one that explained as much as possible of the empirical correlations with the fewest number of estimated parameters. In this case, the significance of the difference between models was expressed as the difference in χ^2 goodness of fit with 9 *df* (i.e., number of subjects minus one). A total of 12 comparisons (i.e., one for each connection) were performed to investigate whether the impact of word type on connectivity varied across subjects. This involved comparing the 12 models in which the effect of word type on a connection of interest was set to be the same in all subjects with 12 models in which the same effect was free to vary across subjects.

RESULTS

The goodness of fit of the multisubject network, in which reciprocal connections were set to be the same within subject but free to vary among subjects, was 795.58. This corresponded to a *P* value of 0.12 under an asymptotic χ^2 distribution with 750 effective *df* (computed after correction for autocorrelation in the fMRI time series; see Worsley and Friston, 1995). The model was therefore not refuted (i.e., it provided a sufficient account of the data).

With respect to the modulation of connectivity by word type, we found that a number of functional connections were stronger when reading pseudowords relative to words. These included both forward (posterior fusiform \rightarrow inferior temporal, posterior fusiform \rightarrow frontal operculum, and superior temporal \rightarrow frontal

operculum) and backward (frontal operculum \rightarrow superior temporal) connections. Significant decreases in connectivity when reading pseudowords relative to words were not found. These findings are summarized in Table 1. The values reported represent the modulation by word type on the coupling between a "source" region and a "target" region. A value significantly higher than zero indicates that the coupling between the two regions is stronger during reading pseudowords than words, while a value significantly lower than zero indicates that the coupling is weaker.

All forward effective connections were better characterized by a model in which changes in connectivity for pseudowords relative to words were set to be the same across subjects. In contrast, the backward connection between the frontal operculum and the superior temporal gyrus showed significant intersubject variability. In 7 of 10 subjects, this connection was stronger when reading pseudowords but the size of this effect was highly variable. The remaining subjects showed either a decrease in connectivity for reading pseudowords relative to words or no effect at all. *P* values of the difference between the free model (which allowed for intersubject variability) and the restricted model (in which effects were set to be the same across subjects) are reported in the last column of Table 1. Results are represented graphically in Fig. 5.

DISCUSSION

In the present study, we used a multisubject network to investigate intersubject variability in functional integration in the context of single word and pseudoword reading. Within each subject, each region was reciprocally connected to the remainder by both forward and backward connections. This was motivated by the idea that functional specialization is not an intrinsic property of any region, but depends on both forward and backward connections (Friston and Price, 2001; McIntosh, 2000; Hebb, 1949). To ensure that the model could be globally identified, we forced reciprocal connections to have the same value. In addition, we modeled exogenous variance to account for the common influence of experimental effects over subjects. These influences were encoded (i) by design variables that entered directly the "input" regions to model sensory evoked responses and (ii) indirectly through interactions with regional activity to model context-sensitive changes in connection strength.

First, we evaluated whether the multisubject network which allowed for intersubject variability provided a good explanation for the variance-covariance structure of the empirical data. A critical principle in model specification and evaluation is that, by definition, all of the models are wrong or incomplete to some degree. In other words, models at their best can provide only a close approximation to the observed data,

TABLE 1
Effects of Reading Pseudowords versus Words on Connectivity^a

	Free model (allowing for inter-subject variability)										Restricted model subj.1–10	Free vs restricted model, <i>P</i> value of difference
	subj.1	subj.2	subj.3	subj.4	subj.5	subj.6	subj.7	subj.8	subj.9	subj.10		
Forward connections												
Post. Fusiform → Inf. Temp.	0.22	0.11	0.34	-0.08	0.19	-0.13	0.08	-0.03	0.33	0.25	0.11*	0.15
Post. Fusiform → Sup. Temp.	0.24	0.05	0.00	-0.02	0.06	0.00	-0.18	-0.25	-0.15	-0.23	-0.03	0.24
Post. Fusiform → Frontal Operc.	0.13	0.05	-0.10	0.08	0.24	-0.05	0.14	0.17	0.15	0.11	0.10*	0.89
Inf. Temporal → Sup. Temp.	0.17	0.12	0.22	0.03	0.27	-0.11	-0.05	-0.06	-0.16	-0.24	0.01	0.19
Inf. Temporal → Frontal Operc.	0.08	0.06	-0.06	-0.02	-0.06	0.01	0.08	0.01	-0.09	0.05	0.00	0.99
Sup. Temporal → Frontal Operc.	0.41	0.09	0.00	0.19	-0.09	0.16	0.24	0.07	0.00	-0.13	0.10*	0.11
Backward connections												
Inf. Temp. → Post. Fusiform	0.14	0.12	0.18	-0.14	0.06	-0.07	0.08	-0.02	0.28	0.00	0.05	0.61
Sup. Temp. → Post. Fusiform	0.15	0.07	0.01	0.01	0.00	-0.01	-0.13	-0.09	-0.08	-0.04	-0.01	0.92
Frontal Operc. → Post. Fusiform	0.13	-0.05	0.12	0.04	0.06	-0.08	-0.03	0.08	0.11	-0.11	0.02	0.83
Sup. Temp. → Inf. Temp.	0.07	0.06	0.13	0.08	0.19	-0.12	-0.05	-0.01	-0.01	-0.14	0.01	0.67
Front. Operc. → Inf. Temp.	0.06	0.03	0.01	0.00	-0.11	0.05	0.10	0.00	-0.18	0.18	0.01	0.68
Front. Operc. → Sup. Temp.	0.48	0.11	0.00	0.06	-0.06	0.17	0.23	0.05	0.04	-0.21	0.10*	0.01

^a The effect of word type on the coupling between a “source” and a “target” region in the free model (which allows for intersubject variability) and the restricted model (in which effects are constrained to be the same in all subjects). A value significantly higher than zero indicates that the coupling between the two regions is stronger during reading pseudowords than words, while a value significantly lower than zero indicates that the coupling is weaker. The asterisk (*) indicates significant effects of word type on the reading-related covariance, when changes in connectivity are set to be the same in all subjects (P value < 0.05). The last column refers to the P values of the difference in goodness of fit between the free and restricted models. It can be seen that, for the backward connection between the frontal operculum and the superior temporal gyrus, the free model accounts for the data better than the restricted model (i.e., $P < 0.05$).

rather than an exact fit. In the case of our study, the neuronal dynamics underlying the reading of words and pseudowords are likely to involve more regions (e.g., right hemisphere areas) than those included in the model. In addition, such dynamics are likely to be more complex than we can represent using a structural equation model that embodies only first-order interactions. The finding that our multisubject network fitted the observed data means only that our model provides a “plausible” representation of the real neuronal dynamics. This enabled us to use the multisubject network to investigate the effect of word type on connectivity.

Our results showed that a number of forward and backward connections were modulated by word type.

Contrary to our intention, we did not reveal a double dissociation in the pattern of reading activity for words and pseudowords because the change in connection strength was in one direction only—stronger for pseudowords. In most cases, only the forward connection was modulated by word type whereas the backward connection was less affected. However, in the case of the posterior superior temporal region and the frontal operculum, both the forward and backward connections were stronger during reading pseudowords than words. Our findings illustrate that differences between processing words and pseudowords do not simply lie in the degree of activation in one or more regions of the language system. Rather, such differences can be characterized in terms of context-sensi-

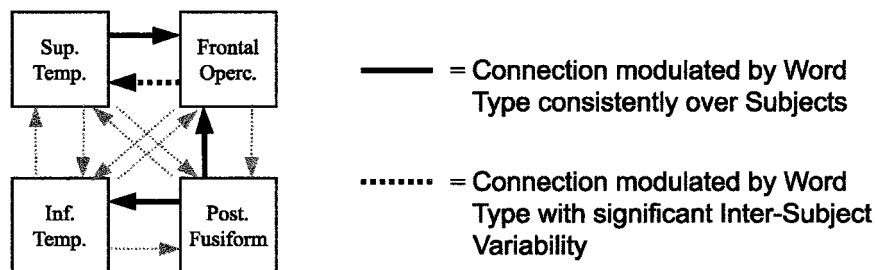


FIG. 5. Graphical representation of the SEM results. Solid black lines indicate connections that were modulated by word type but did not show significant intersubject variability. Broken black lines indicate connections that were significantly modulated by word type and showed significant intersubject variability. Gray broken lines indicate connections that showed neither a modulation by word type nor significant intersubject variability.

tive interactions among brain areas. We found that connectivity in both semantic and nonsemantic reading routes was stronger for pseudowords. One possibility for this result is that, if functional connectivity is a function of task difficulty (Furey *et al.*, 2000), pseudowords are associated with greater connectivity than words because they are more difficult to read.

The main aim of the present study was to characterize intersubject variability by testing for differences in connectivity among subjects. We found that the backward connection between the frontal operculum and the superior temporal gyrus was modulated by word type differentially across subjects. This illustrates that a pattern of connectivity, which is estimated over subjects, may not be a good approximation of the underlying patterns of connectivity in all of the subjects studied. Rather, in some cases a model that allows for intersubject variability would be preferred to a model in which one single pattern of connectivity is assumed. Our results suggest that the subjects may have adopted different strategies for reading either words or pseudowords, which influenced the backward connections between frontal operculum and posterior superior temporal region. For instance, pseudowords are fairly ambiguous stimuli that may engage a search for (missing) semantic and phonological representations in some subjects but not in others (Fiez and Petersen, 1998; Price *et al.*, 1996). However, our findings should be considered in the context of the experimental paradigm adopted in the present study. Silent reading of 40 single words per minute is a relatively undemanding task which is likely to leave additional attentional resources that may be directed at other cognitive processes. One possibility is that, if subjects performed a more engaging task, functional integration would have been more consistent across subjects. Finally, differences in functional integration may reflect degenerative solutions the brain can adopt to complete the same strategy.

We have described how the issue of intersubject variability in functional integration can be addressed directly and have illustrated the approach with analyses of experimental data. Another possible application involves comparing changes in connectivity among different groups of subjects. This may be particularly useful in the context of patient studies, in which one may want to compare functional integration between patients and normal subjects (e.g., dyslexic vs normal) or between different groups of patients (e.g., surface vs phonological dyslexia). In this case, a multisubject network can be treated as a multigroup network. If the model, which allows for group differences, accounts for the empirical data better than a restricted model in which the group are the same, then the two groups express differential connectivity. The crucial advantage of using a multigroup network over individual group analyses is that one can make inferences about

differences among groups in addition to inferences pertaining to single groups, in the same way classic analysis of variance allows one to test for group \times condition interactions. Finally, a multisubject network can be used to address intersubject variability in relation to behavioral measures. For instance, one can investigate whether differences in performance correspond to differences in functional integration among subjects or groups. This approach enables one to interpret changes in connectivity meaningfully by establishing a relation between changes in connectivity and behavior.

ACKNOWLEDGMENT

This work was funded by the Wellcome Trust.

REFERENCES

- Bentler, P. M. 1992. *EQS Structural Equations Program Manual*. BMDP Statistical Software, Los Angeles.
- Bokde, A. L., Tagamets, T. A., Friedman, R. B., and Horwitz, B. 2001. Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. *Neuron* **30**: 609–617.
- Bollen, K. A. 1989. *Structural Equations with Latent Variables*. Wiley, New York.
- Brunswick, N., McCrory, E., Price, C. J., Frith, C. D., and Frith, U. 1999. Explicit and implicit processing of words and pseudowords by adult developmental dyslexics: A search for Wernicke's Wortschatz. *Brain* **122**: 1901–1917.
- Büchel, C., and Friston, K. J. 1997. Modulation of connectivity in visual pathways by attention: Cortical interactions evaluated with structural equation modelling and fMRI. *Cereb. Cortex* **7**: 768–778.
- Büchel, C., and Friston, K. J. 1998. Dynamic changes in effective connectivity characterized by variable parameter regression and Kalman filtering. *Hum. Brain Mapp.* **6**: 403–408.
- Büchel, C., Coull, J. T., and Friston, K. J. 1999. The predictive value of changes in effective connectivity for human learning. *Science* **283**: 1538–1541.
- Bullmore, E., Horwitz, B., Honey, G., Brammer, M., Williams, S., and Sharma, T. 2000. How good is good enough in path analysis of fMRI data? *NeuroImage* **11**: 289–301.
- Coull, J. T., Büchel, K. J., Friston, K. J., and Frith, C. D. 1999. Noradrenergically mediated plasticity in a human attentional neuronal network. *NeuroImage* **10**: 705–715.
- Fiez, J. A., and Petersen, S. A. 1998. Neuroimaging studies of word reading. *Proc. Natl. Acad. Sci. USA* **95**: 914–921.
- Fiez, J. A., Balota, D. A., Raichle, M. E., and Petersen, S. E. 1999. Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron* **24**: 205–218.
- Friston, K. J., and Price, C. J. 2001. Dynamic representations and generative models of brain function. *Brain Res. Bull.* **54**: 275–285.
- Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. J. 1993a. Functional connectivity: The principal component analysis of large (PET) data sets. *J. Cereb. Blood Flow Metab.* **13**: 5–14.
- Friston, K. J., Frith, C. D., and Frackowiak, R. S. J. 1993b. Time-dependent changes in effective connectivity measured with PET. *Hum. Brain Mapp.* **1**: 69–80.
- Friston, K. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., and Frackowiak, R. S. J. 1995a. Spatial registration and normalization of images. *Hum. Brain Mapp.* **2**: 1–25.
- Friston, K. J., Holmes, A., Worsley, K. J., Poline, J.-B., Frith, C. D., and Frackowiak, R. S. J. 1995b. Statistical parametric maps in

- functional imaging: A general linear approach. *Hum. Brain Mapp.* **2**: 189–210.
- Friston, K. J., Büchel, C., Fink, G. R., Morris, J., Rolls, E., and Dolan, R. J. 1997. Psychophysiological and modulatory interaction in neuroimaging. *NeuroImage* **6**: 218–229.
- Furey, M. L., Pietrini, P., Alexander G. E., Schapiro, M. B., and Horwitz, B. 2000. Cholinergic enhancement improves performance on working memory by modulating the functional activity in distinct brain regions: A positron emission tomography regional blood flow study in healthy humans. *Brain Res. Bull.* **51**: 213–218.
- Gonçalves, M. S., Hall, D. A., Johnsrude, I. S., and Haggard, M. P. 2001. Can meaningful effective connectivities be obtained between auditory cortical regions? *NeuroImage* **14**: 1353–1360.
- Hagoort, P., Brown, C., Indefrey, P., Herzog, H., Steinmetz, H., and Seitz, R. 1999. The neural circuitry involved in the reading of German words and pseudowords: A PET study. *J. Cogn. Neurosci.* **11**: 383–398.
- Hebb, D. O. 1949. *The Organization of Behavior*. Wiley, New York.
- Herbster, A. N., Mintun, M. A., Nebes, R. D., and Becker, J. T. 1997. Regional cerebral blood flow during word and nonword reading. *Hum. Brain Mapp.* **5**: 84–92.
- Higham, N. J. 1993. Optimization by direct search in matrix computations. *SIAM J. Matrix Anal. Appl.* **14**: 317–333.
- Horwitz, B., and Sporns, O. 1994. Neural modeling and functional neuroimaging. *Hum. Brain Mapp.* **1**: 269–283.
- Horwitz, B., Rumsey, J. M., and Donohue, B. C. 1998. Functional connectivity of the angular gyrus in normal reading and dyslexia. *Proc. Natl. Acad. Sci. USA* **95**: 8939–8944.
- Iidaka, T., Otori, O., Murata, T., Kosaka, H., Yonekura, Y., Okada, T., and Sadato, N. 2001. Neural interaction of the amygdala with the prefrontal and temporal cortices in the processing of facial expressions as revealed by fMRI. *J. Cogn. Neurosci.* **13**: 1035–1047.
- Jöreskog, K.G., and Sörbom, D. 1990. Model search with TETRAD II and LISREL. *Sociol. Methods Res.* **19**: 93–106.
- Kucera, H., and Francis, W. H. 1967. *Computational Analysis of Present-Day American English*. Brown Univ. Press, Providence, RI.
- MacCallum, R. C. 1995. Model specification: Procedures, strategies and related issues. In *Structural Equation Modelling: Concepts, Issues and Applications* (R. H. Hoyle, Ed.), pp. 16–36. Sage, Thousand Oaks, CA.
- Maguire, E. A., Mummery, C. J., and Büchel, C. 2000. Patterns of Hippocampal–cortical interaction dissociate temporal lobe memory subsystems. *Hippocampus* **10**: 475–482.
- McIntosh, A. R. 2000. Towards a network theory of cognition. *Neural Netw.* **13**: 861–870.
- McIntosh, A. R., Grady, C. L., Ungerleider, L. G., Haxby, J. V., Rapoport, S. I., and Horwitz, B. 1994. Network analysis of cortical visual pathways mapped with PET. *J. Neurosci.* **14**: 655–666.
- Mechelli, A., Price, C. J., and Friston, K. J. 2001. Effective connectivity during reading: The effects of word type and stimulus rate. *NeuroImage* **13**: S1292.
- Mechelli, A., Gorno-Tempini, M. L., and Price, C. J. Neuroimaging studies of word and pseudoword reading: Consistencies, inconsistencies and limitations. *J. Cogn. Neurosci.*, in press.
- Patterson, K. E., and Shewell, C. 1987. Speak and spell: Dissociations and word class effects. In *The Cognitive Neuropsychology of Language* (M. Coltheart, G. Sartori, R. Job, K. E. Patterson, and J. C. Marshall, Eds.). Erlbaum, London.
- Price, C. J. 2000. The anatomy of language: Contributions from functional neuroimaging. *J. Anat.* **197**: 335–359.
- Price, C. J., Wise, R. J. S., and Frackowiak, R. S. J. 1996. Demonstrating the implicit processing of visually presented words and pseudowords. *Cereb. Cortex* **6**: 62–70.
- Price, C. J., Veltman, D., Ashburner, J., Josephs, O., and Friston, K. 1999. The critical relationship between the timing of stimulus presentation and data acquisition in fMRI. *NeuroImage* **10**: 36–45.
- Rowe, J., Stephan, K. E., Friston, K. J., Frackowiak, R., Lees A., and Passingham, R. 2002. Attention to action in Parkinson's disease: Impaired effective connectivity among frontal cortical regions. *Brain* **125**: 276–289.
- Seidenberg, M. S., and McClelland, J. L. 1989. A distributed developmental model of word recognition and naming. *Psychol. Rev.* **96**: 523–568.
- Worsley, K. J., and Friston, K. J. 1995. Analysis of fMRI time-series revisited—again. *NeuroImage* **2**: 45–53.
- Xu, B., Grafman, J., Gaillard, W. D., Ishii, K., Vega-Bermudez, F., Pietrini, P., Reeves-Tyler, P., DiCamillo, P., and Theodore, W. 2001. Conjoint and extended neural networks for the computation of speech codes: The neural basis of selective impairment in reading words and pseudowords. *Cereb. Cortex* **11**: 267–277.