ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 30 (2006) 885 – 890

Identification of degenerate neuronal systems based on intersubject variability

Uta Noppeney,[a,b,*] Will D. Penny,[a] Cathy J. Price,[a] Guillaume Flandin,[a] and Karl J. Friston[a]

[a]*Wellcome Department of Imaging Neuroscience, University College London, 12 Queen Square, WC1 N3BG London, UK*
[b]*Max-Planck-Institute for Biological Cybernetics, Spemannstr. 38, 72076 Tubingen, Germany*

Received 13 July 2005; revised 1 October 2005; accepted 7 October 2005
Available online 21 November 2005

Group studies implicitly assume that all subjects activate one common system to sustain a particular cognitive task. Intersubject variability is generally treated as well-behaved and uninteresting noise. However, intersubject variability might result from subjects engaging different degenerate neuronal systems that are each sufficient for task performance. This would produce a multimodal distribution of intersubject variability. We have explored this idea with the help of Gaussian Mixture Modeling and Bayesian model comparison procedures. We illustrate our approach using a crossmodal priming paradigm, in which subjects perform a semantic decision on environmental sounds or their spoken names that were preceded by a semantically congruent or incongruent picture or written name. All subjects consistently activated the superior temporal gyri bilaterally, the left fusiform gyrus and the inferior frontal sulcus. Comparing a One and Two Gaussian Mixture Model of the unexplained residuals provided very strong evidence for two groups with distinct activation patterns: 6 subjects exhibited additional activations in the superior temporal sulci bilaterally, the right superior frontal and central sulcus. 11 subjects showed increased activation in the striate and the right inferior parietal cortex. These results suggest that semantic decisions on auditory–visual compound stimuli might be accomplished by two overlapping degenerate neuronal systems.
© 2005 Elsevier Inc. All rights reserved.

Keywords: Degeneracy; Structure–function relationship; Gaussian mixture modelling; Bayesian model comparison

Introduction

Functional localization is a central aim in cognitive neuroscience. It entails associating a particular cognitive function with one brain structure, i.e., the identification of structure–function relationships. Structure–function relationships might be of different sorts: In the simplest case of a one-to-one mapping a single brain structure sustains one and only one function. Although many functional imaging studies implicitly assume a one-to-one mapping, their diverse findings cannot be accommodated within this simplistic framework: First, there is evidence for one-to-many structure–function relationships, in which the same structural configuration can support multiple functions (Friston and Price, 2003; Wilkinson and Halligan, 2004). For instance, the left prefrontal cortex has been implicated in multiple cognitive functions. Second, the remarkable ability of the human brain to maintain and recover cognitive functions after focal cortical damage suggests that multiple neuronal systems can sustain the same function. This many-to-one structure–function relationship has been referred to as *degeneracy* (Edelman and Gally, 2001; Price and Friston, 2002). Multiple degenerate sets of brain regions might sustain the same cognitive task either via similar mechanisms or by implementing different cognitive strategies. Thus, several cognitive models suggest that complex cognitive processes such as sentence comprehension, action retrieval or reading can be accomplished in multiple ways, i.e., by engaging different sub-processes. For instance, the sentence "the boy eats an apple" can be understood by simply combining the semantic knowledge of the single words with pragmatic knowledge of the world or by formally assigning a syntactic structure to the lexical items. Similarly, based on neuropsychological data, it has been suggested that the appropriate action for an object can be retrieved either directly from visual structural features or indirectly through accessing semantic (contextual, associative) knowledge (Phillips et al., 2002; Rumiati and Humphreys, 1998). These examples demonstrate that cognitive functions can be characterized at multiple levels of description (e.g., action retrieval as one cognitive process or as a series of subprocesses) in the same way as structural elements can constitute sets of multiple brain regions, single brain regions, neuronal populations/assemblies or single neurons. Hence, the characterization of structure–function relationships depends implicitly on the descriptive level at which the structural and functional elements are specified (Noppeney et al., 2004; Price and Friston, 2005).

How can we identify degenerate structure–function relationships using (1) lesion studies or (2) functional imaging? (1) *Single* lesion studies are of limited use because if only one system is

* Corresponding author. Max-Planck-Institute for Biological Cybernetics, Spemannstr. 38, 72076 Tubingen, Germany. Fax: +44 20 7813 1420.
E-mail address: uta.noppeney@tuebingen.mpg.de (U. Noppeney).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2005.10.010

lesioned, the cognitive function can be supported by the remaining degenerate systems. Only lesioning all other degenerate elements will impair task performance and thus reveal the functional contribution of a particular neuronal element (i.e., the decrement in performance after lesioning). Thus, in *multiple* lesion studies, degeneracy can be inferred if the functional contribution of a neuronal element depends on the state (i.e., lesioned or intact) of another degenerate neuronal element (Aharonov et al., 2003)—hence the interest in multifocal TMS. (2) Functional imaging allows us to define a set of regions that are sufficient for a particular function within a single subject—at least if sensitivity issues are ignored. However, it cannot determine whether all regions are necessary for task performance or whether they represent multiple – partly overlapping – degenerate neuronal systems that are co-activated and thus functioning redundantly (Barlow, 2001; Friston and Price, 2003; Shannon and Weaver, 1949). Furthermore, functional imaging cannot disclose degenerate neuronal systems that are inhibited by the pre-potent system or not activated.

Here, we explore how functional imaging can nevertheless help us to identify degenerate neuronal systems based on intersubject variability in functional activations. Using random effects analyses, functional imaging studies usually focus on activations that are consistent across a group of subjects. Activations that are observed only in a subset of subjects or within individual subjects are treated as uninteresting noise. However, intersubject variability might result from subjects engaging different degenerate neuronal systems (see Kherif et al., 2003). Establishing structure within the unexplained residuals at the between-subject level may therefore enable us to define candidate regions for degenerate neuronal systems.

Based on this rationale, our approach was as follows: first, we perform a random effects analysis to identify the set of regions that are consistently activated across subjects. Second, performing a singular value decomposition on the residuals, we reduce their dimensionality to a few dominant activation patterns (i.e., eigenimages). The expression of these activation patterns over subjects is reflected in their associated eigenvariates. Third, we investigate whether the differential expression of activation patterns (i.e., eigenvariates) provides evidence for subjects coming from a single or two distinct Gaussian distributions. This involves a cluster analysis of the eigenvariates using multivariate Gaussian mixture modeling and Bayesian model comparison procedures. Fourth, if we obtain strong evidence for two clusters, we compare their activations in a two-sample t test. This shows the set of regions that are differentially engaged by the two groups and thus disclose candidate regions for degenerate neuronal systems.

We illustrate our approach using a crossmodal priming paradigm, in which subjects performed a semantic decision on environmental sounds or their spoken names that are preceded by a semantically congruent or incongruent picture or written name.

Methods

Subjects

17 healthy right-handed English native speakers (5 females; median age: 25) gave informed consent to participate in the study. The study was approved of by the joint ethics committee of the Institute of Neurology and University College London Hospital, London, UK.

Experimental design

The paradigm was a two-choice forced categorization of auditory stimuli that were preceded by visual stimuli. The activation conditions conformed to a $3 \times 2 \times 2$ factorial design manipulating

(1) Priming: (i) congruent semantics and response, (ii) incongruent semantics and congruent response, (iii) incongruent semantics and incongruent response
(2) Prime modality: $\text{words}_{\text{written}}$, pictures
(3) Target modality: $\text{words}_{\text{spoken}}$, sounds

At the beginning of each trial, a visual prime (i.e., $\text{words}_{\text{written}}$ or pictures) was presented for 100 ms. After an additional ISI (interstimulus interval) of 100 ms, the auditory target (i.e., $\text{words}_{\text{spoken}}$ or sounds) was presented. The trial onset asynchrony was 3.25 s. Subjects performed a semantic decision on the auditory targets ('Is the target stimulus heavier than 4 kg?'). 50% of the trials were primed, i.e., prime and target were semantically congruent. In these semantically congruent trials, prime and target referred to the same object (e.g., a picture of a dog followed by the barking sound of a dog). 25% of the trials were response congruent but semantically incongruent. In those trials, prime and target referred to different objects but required the same response, i.e., both weighed more or less than 4 kg (e.g., a picture of an elephant followed by the sound of a car).

Altogether, there were 64 stimuli. Each stimulus was presented 8×2 times, four times in each modality for each subject (i.e., $64 \times 8 \times 2 = 512$ crossmodal trials).

Additional 24×2 intramodal visual trials (i.e., picture–picture, picture–written word, written word–picture, written word–written word) were included to maintain subjects' attention to the visual primes that were response irrelevant. 50% of the trials required a yes response. Yes/No responses to all conditions were indicated (as quickly and as accurately as possible) by a two-choice key press. The activation conditions were interleaved with 6-s fixation. The stimuli and order of conditions were randomized and the stimuli rotated across conditions within and between subjects.

fMRI scanning technique

A 3-T Siemens Allegra system was used to acquire both T1 anatomical volume images and T2*-weighted axial echoplanar images with blood oxygenation level-dependent (BOLD) contrast (GE-EPI, Cartesian k-space sampling, TE = 30 ms, TR = 2.47 s, 38 axial slices, acquired sequentially in descending direction, matrix 64×64, spatial resolution $3 \times 3 \times 3.4$ mm^3 voxels, interslice gap 1.4 mm, slice thickness 2.0 mm). To avoid Nyquist ghost artefacts, a generalized reconstruction algorithm was used for data processing (Josephs et al., 2000). There were two sessions with a total of 473 volume images per session. The first six volumes were discarded to allow for T1 equilibration effects.

Conventional SPM analysis

The data were analyzed with statistical parametric mapping (using SPM2 software from the Wellcome Department of Imaging Neuroscience, London; http//www.fil.ion.ucl.ac.uk/spm, Friston et al., 1995). Scans from each subject were realigned using the first as a reference, spatially normalized into standard space (Talairach and Tournoux, 1988), resampled to $3 \times 3 \times 3$ mm^3 voxels and spatially

smoothed with a Gaussian kernel of 8 mm FWHM. The timeseries in each voxel were highpass filtered to 1/128 Hz and globally normalized with proportional scaling. The fMRI experiment was modeled in an event-related fashion using regressors obtained by convolving each event-related stick function with a canonical hemodynamic response function and its first temporal derivative. We modeled the 12 conditions in our $2 \times 2 \times 3$ factorial design and the intramodal stimuli. Nuisance covariates included the realignment parameters (to account for motion artefacts). Condition-specific effects for each subject were estimated according to the general linear model and passed to a second-level analysis as contrasts. This involved creating contrast images (average of all crossmodal stimuli > fixation) for each subject and a second-level one-sample *t* test. Inferences were made at the second level to allow a random effects analysis and inferences at the population level (Friston et al., 1999). We only report activations that are significant ($P < 0.05$) corrected for the entire brain volume (extent threshold > 3 voxels).

Singular value decomposition, Gaussian mixture modeling and Bayesian model comparison

The residuals of the one-sample *t* test (subject-specific deviations from the group mean activation for crossmodal stimuli > fixation) were entered into a singular value decomposition (SVD) using the multivariate toolbox as implemented in SPM99. This SVD decomposed the original series of residual activations over the 17 subjects into two orthogonal sets of vectors: (1) eigenimages representing patterns of activation in space and (2) eigenvariates indicating the expression of the activation patterns in each subject. The importance of each eigenimage (i.e., the amount of variance it accounts for in the residuals) is indicated by the associated singular value.

Next, we investigated whether the differential expression of the 5 most dominant activation patterns (explaining 52% of the variance) across subjects provides evidence for subjects coming from a single or two distinct Gaussian distributions (i.e., unimodal or bimodal). For this, we fitted two models, one with a single Gaussian and the other with a mixture of two Gaussians. By comparing the likelihood of (or evidence for) the models, we were able to infer a unimodal or bimodal distribution. The parameters of the Gaussian mixture models were estimated using a Variational Bayesian algorithm (Attias, 2000) with shrinkage priors (Friston et al., 2002). This approach prevents overfitting by down-weighting the Gaussian components that are not supported by the data. The Variational Bayesian algorithm operates by iterating between two steps until convergence: in the responsibility update step, it assigns each data point to the most likely Gaussian cluster. In the parameter update step, it recomputes the means, variances and mixing proportions of the two Gaussian components such that they maximise the evidence of the model (i.e., the likelihood of the data given a model). As the Variational Bayesian algorithm is a non-linear optimization procedure, it may become trapped in local minima and potentially lead to sub-optimal solutions. Therefore, we repeated model fitting for 10,000 different initializations so as to obtain an optimal cluster assignment.

The Variational Bayesian algorithm also provides an approximation to the model evidence (Penny et al., 2003, 2004). Bayes Factors (the ratio of the model evidences, here of the one- and two-Gaussian component models, Kass and Raftery, 1995) were used for model comparison, i.e., to decide whether the one- or two-Gaussian component model was a better model of the data. If strong evidence (i.e., Bayes factors > 20) was provided for a two-cluster model, subjects were assigned to the two clusters based on their final responsibilities (Ueda et al., 2000). The activations of the two subject groups were then compared in a two-sample *t* test SPM to reveal the neuronal systems that were differentially engaged by the two groups using an uncorrected threshold of $P < 0.01$ (extent threshold > 50 voxels). This uncorrected threshold was chosen only for characterization of the two degenerate neuronal systems. The statistical inference is based alone on the model evidence and not the SPM.

Recently, tensor probabilistic independent component analysis (tensor PICA) has been proposed as a method to explore intersubject variability(Beckmann and Smith, 2005). Tensor PICA is a model-free approach that decomposes multisubject fMRI data into spatial, temporal and subject-dependent variations by iterating between (1) estimating a 2-dimensional PICA of the spatial modes and a mixing matrix M and (2) decomposing the mixing matrix M into temporal and subject factors using SVD. In contrast, the approach, proposed here, encompasses model-dependent and model-free elements: first, we model single subject data using experimentally designed stimulus functions in a conventional SPM analysis. Second, we perform a model-free SVD on the unexplained residuals of the random effects analysis. Combining model-dependent and model-free analysis steps enables us to investigate intersubject variability associated with particular cognitive processes and hence address the specific question whether different groups engage distinct semantic retrieval systems for crossmodal stimuli.

Results

Conventional SPM analysis

Semantic decision > fixation

Semantic decisions on crossmodal stimuli, relative to fixation, activated the superior temporal gyri extending into the superior temporal sulci bilaterally, the left superior precentral/inferior frontal sulcus and the fusiform gyri/cerebellum bilaterally consistently across subjects. Activations were also observed in the cingulate sulcus.

Semantic decision < fixation

Semantic decisions relative to fixation decreased activation in the superior frontal gyri/sulci, inferior/superior temporal gyri, calcarine and collateral sulci bilaterally. Deactivations were also found in the cingulate sulcus, right lingual gyrus, insula and several occipital extrastriate areas (Table 1; Fig. 1).

Gaussian mixture modeling and Bayesian model comparison

The optimal Gaussian mixture modeling using the 5 eigenvariates with the greatest eigenvalues sorted the 17 subjects into two groups of 6 and 11, with one female in the 6 and four females in the 11 group. Comparing the one- and two-cluster models provided very strong evidence for the two-cluster model (Bayes factors > 150).

Out of 10,000 cluster solutions, 9308 provided a higher model evidence for the two-relative to the one-cluster model.

The optimal Gaussian mixture modeling using only one, two, three or four eigenvariates consistently provided evidence for a two cluster model (strong evidence for 2 to 4 eigenvariates).

Table 1
Activations consistent across subjects

Region	Co-ordinates	*z* score
(a) Semantic decision > fixation		
L. superior temporal gyrus/	−63, −24, 9	6.8
L. superior temporal sulcus	−63, −12, 3	6.2
	−45, −27, 3	6.0
R. superior temporal gyrus	66, −27, 0	6.5
R. superior temporal sulcus	54, −12, −6	6.3
	45, −33, 3	5.8
L. superior precentral gyrus/L. inferior frontal sulcus	−39, 12, 24	6.0
L. fusiform g.	−42, −57, −24	5.5
R. cerebellum/fusiform g.	33, −63, −30	5.2
Cingulate sulcus	−9, 12, 48	5.2
(b) Fixation > semantic decision		
Cingulate sulcus	−6, 33, −12	6.6
	−15, −36, 48	6.0
	12, −48, 57	5.9
R. superior frontal g.	15, 39, 57	5.1
R. superior frontal sulcus	30, 36, 48	5.9
	21, 18, 57	5.2
L. superior frontal sulcus	−30, 30, 39	5.1
	−9, 54, 42	5.2
	−15, 57, 30	5.2
Superior frontal g.	0, −21, 63	5.7
R. insula	36, 15, −18	5.2
R. intraparietal sulcus	24, −84, 27	5.5
R. superior temporal sulcus	45, −69, 30	5.5
	42, −78, 12	5.0
R. temporal pole	33, 15, −36	5.5
	45, 0, −33	4.9
R. collateral sulcus	30, −48, −9	5.4
L. collateral sulcus	−21, −48, −6	5.0
R. lingual g.	18, −75, −6	5.1
L. superior occipital g.	−15, −93, 33	5.2
Calcarine sulcus	−15, −63, 12	5.7
	15, −57, 12	5.5

Comparison of fMRI activations between the two groups

Six relative to 11 subjects showed enhanced activation in the superior temporal sulci bilaterally and the right central and superior frontal sulci. 11 relative to 6 subjects exhibited increased activation in the striate cortex and the right inferior parietal cortex (Table 2).

Comparison of behavioural measures between the two groups

There were no significant (>0.05) differences in mean reaction times or accuracy for any of the 12 conditions between the two groups. This is important if we want to link our bimodal BOLD-response distribution to degeneracy. In this instance, we are inferring there is a two-to-one mapping between structure (areas activated in the two groups) and function. Function is defined operationally in terms of task performance (i.e., accuracy and reaction time in a speeded weight judgement).

Discussion

The current functional imaging culture is dominated by group studies focussing on activations that are observed consistently

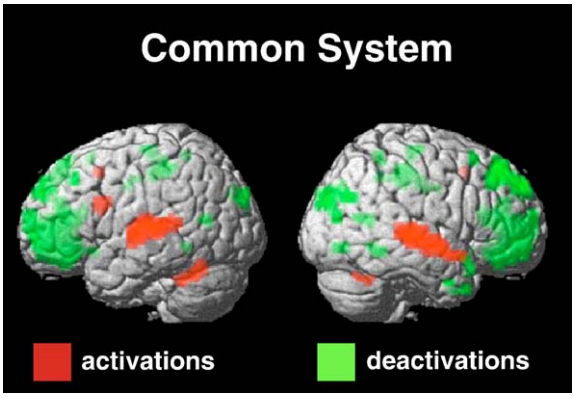

Fig. 1. Activations (red) and deactivations (green) for semantic decisions on crossmodal compound trials (for all normal subjects) relative to fixation are rendered on an averaged normalized brain. Height threshold: $P < 0.05$ corrected. Extent threshold > 3 voxels.

across subjects. Activations that are detected only variably in a subset of subjects are often discarded as uninteresting or spurious. Group studies implicitly assume that all subjects activate one common system to sustain a particular cognitive task, and that all deviations from this are well-behaved random effects. In the Introduction, we have introduced the notion of degeneracy, where multiple neuronal systems are each sufficient for task performance. If sub-sets of subjects engage these degenerate neuronal systems differentially, the consistently activated brain areas might represent only a subset of the regions that are required for task performance. Furthermore, intersubject variability would then contain important information that enables us to identify candidate regions for degenerate neuronal systems. In this paper, we have explored this idea with the help of mixture modeling of the residuals from a second-level analysis. We illustrated our methodology using functional imaging data from a crossmodal priming paradigm.

Semantic decisions on auditory stimuli (e.g., spoken name or sound of a dog) that were preceded by semantically congruent or incongruent visual stimuli (e.g., written name or picture of a dog) consistently activated the superior temporal gyri bilaterally, the left fusiform gyrus and the superior precentral gyrus/inferior frontal sulcus. This suggests that all subjects focussed on the auditory stimulus that determined the response to the crossmodal trial.

Table 2
Comparison of fMRI activations between the two groups

Region	Co-ordinates	*z* score
6 > 11 subjects		
L. superior temporal sulcus ant.	−48, −27, 0	3.1
post	−42, −60, 30	3.4
post.	−33, −72, 33	2.9
R. superior temporal sulcus post	39, −66, 33	3.0
R. angular g.	51, −51, 51	3.4
R. superior frontal sulcus	30, 27, 42	3.1
R. central sulcus	30, −36, 63	2.9
11 > 6 subjects		
Striate cortex	15, −93, −6	3.9
	−21, −63, 12	3.8
	−9, −78, 0	3.5
R. inferior parietal cortex	63, −39, 42	3.5

However, the residuals of the second-level one-sample *t* test exhibited structure that was not modeled or explained by the mean activation across subjects. Bayesian model comparison provided very strong evidence for two groups with distinct activation patterns: 6 subjects exhibited additional activations in the superior temporal sulci bilaterally. 11 subjects showed increased activation (or less activation decreases) in the striate cortex and the right inferior parietal cortex. The few regions that are consistently activated across subjects might therefore form only a subset of the regions that each subject engages for task performance. Importantly, there was no significant difference in behavioural performance across the two groups. Both groups showed comparable mean reaction times and error rates across conditions. Collectively, our results suggest that there might be two partly overlapping degenerate neuronal systems that can each sustain semantic decisions on crossmodal compound trials. Both systems commonly engage the superior temporal gyri bilaterally. However, in the first group, the STG co-activate with the STS bilaterally, while in the second group, they interact with early visual areas in the occipital cortex.

The two degenerate neuronal systems are sufficient for task performance as defined operationally in terms of response accuracy and reaction times. However, it is well recognized in cognitive neuroscience that complex cognitive tasks can often be accomplished in multiple ways, i.e., by engaging different strategies. Many cognitive models even postulate multiple routes that can be used to sustain a task. For instance, the dual route model of reading proposes that reading of familiar, regularly spelt words can be achieved via either spelling–sound relationships or lexical semantic processes (Fiez and Petersen, 1998; Fiez et al., 1999; Hagoort et al., 1999; Marshall and Newcombe, 1973; Shallice, 2003). Similarly, the two degenerate neuronal systems in our experiment might accomplish semantic decisions on crossmodal compound stimuli via similar mechanisms or by implementing different strategies. Given our prior knowledge about STS and occipital cortex (Beauchamp et al., 2004; Calvert, 2001; Gottfried and Dolan, 2003; Malach et al., 1995), we might speculate on potential strategies that the two subject groups engaged in. For instance, subjects with increased occipital cortex activation might have relied more on the visual structural information that in 75% of the trials was sufficient for correct responses. On the contrary, subjects with increased STS activations but relatively decreased activation in the visual cortex might have engaged more in auditory processing, crossmodal integration of visual and auditory information and semantic retrieval (see Shomstein and Yantis, 2004). These hypotheses could be tested in future experiments that degrade visual information or manipulate the percentage of congruent trials in order to modulate the influence of visual information on subjects' judgement.

In this study, the degenerate neuronal systems were revealed using intersubject variability. Degeneracy can therefore be expressed at the level of the individual or the population (Noppeney et al., 2004). In the first case of degenerate functional

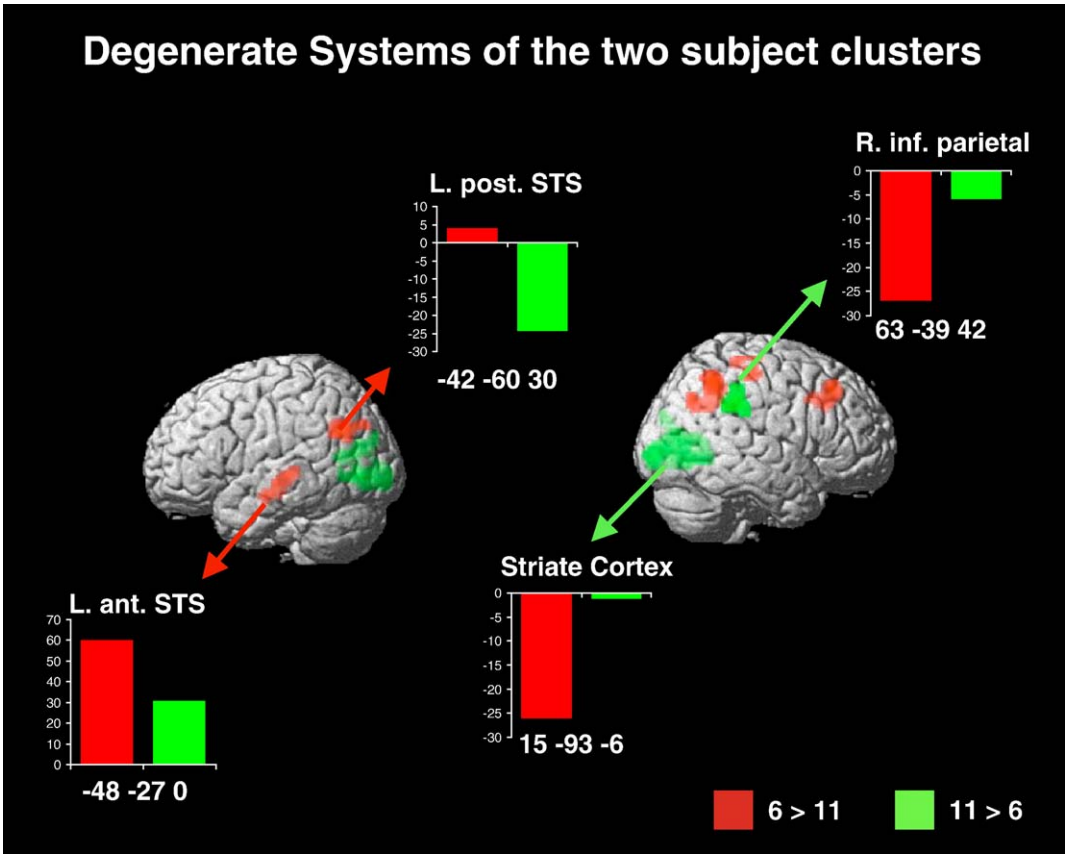

Fig. 2. Semantic decisions on crossmodal compound trials. Differential activation across groups is rendered on an averaged normalized brain. Height threshold: $P < 0.01$ uncorrected. Extent threshold > 50 voxels. Red = 6 > 11 subjects. Green = 11 > 6 subjects. Parameter estimates for 6 subject cluster (red) and 11 subject cluster (green) during semantic decisions on crossmodal stimuli. The bar graphs represent the size of the effect in adimensional units (corresponding to percent whole brain mean).

neuroanatomy (i.e., within-subject degeneracy), semantic decisions on crossmodal stimuli can be performed by the STG co-activating either with the STS or with the occipital cortex within the same subject. In the second case of degeneracy over subjects, there is no degenerate organization within a single brain, but semantic decisions are sustained by STG/STS in one group of subjects and STG/occipital cortex in another subset. Multiple lesion studies – for instance using TMS – may enable us to distinguish between these two cases and confirm that STS and occipital cortex make important functional contributions to semantic decisions. In the case of degenerate functional neuroanatomy, TMS only to both, STS and occipital cortex, will result in a behavioural deficit. In the case of degeneracy over subjects, TMS to any of the two areas alone is associated with impaired performance in one subset of subjects (but not the other one) (Fig. 2).

In conclusion, intersubject variability can help us to identify candidate regions of multiple degenerate neuronal systems that can be investigated using multilesion methods. Moreover, if subsets of subjects engage different neuronal systems, group analyses (e.g., random effects analyses) will not reveal the entire set of regions that sustain a particular cognitive task but only a sub-system that by itself is not sufficient for task performance.

References

Aharonov, R., Segev, L., Meilijson, I., Ruppin, E., 2003. Localization of function via lesion analysis. Neural Comput. 15, 885–913.

Attias, H., 2000. A variational Bayesian framework for graphical models. Neural Inf. Process. Syst. 12, 209–215.

Barlow, H., 2001. Redundancy reduction revisited. Network 12, 241–253.

Beauchamp, M.S., Lee, K.E., Argall, B.D., Martin, A., 2004. Integration of auditory and visual information about objects in superior temporal sulcus. Neuron 41, 809–823.

Beckmann, C.F., Smith, S.M., 2005. Tensorial extensions of independent component analysis for multisubject fMRI analysis. NeuroImage 25, 294–311.

Calvert, G.A., 2001. Crossmodal processing in the human brain: insights from functional neuroimaging studies. Cereb. Cortex 11, 1110–1123.

Edelman, G.M., Gally, J.A., 2001. Degeneracy and complexity in biological systems. Proc. Natl. Acad. Sci. U. S. A. 98, 13763–13768.

Fiez, J.A., Petersen, S.E., 1998. Neuroimaging studies of word reading. Proc. Natl. Acad. Sci. U. S. A. 95, 914–921.

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.

Friston, K.J., Price, C.J., 2003. Degeneracy and redundancy in cognitive anatomy. Trends Cogn. Sci. 7, 151–152.

Friston, K.J., Holmes, A., Worsley, K.J., Poline, J.B., Frith, C.D., Frackowiak, R., 1995. Statistical parametric mapping: a general linear approach. Hum. Brain Mapp. 2, 189–210.

Friston, K.J., Holmes, A.P., Price, C.J., Buchel, C., Worsley, K.J., 1999. Multisubject fMRI studies and conjunction analyses. NeuroImage 10, 385–396.

Friston, K.J., Glaser, D.E., Henson, R.N., Kiebel, S., Phillips, C., Ashburner, J., 2002. Classical and Bayesian inference in neuroimaging: applications. NeuroImage 16, 484–512.

Gottfried, J.A., Dolan, R.J., 2003. The nose smells what the eye sees: crossmodal visual facilitation of human olfactory perception. Neuron 39, 375–386.

Hagoort, P., Indefrey, P., Brown, C., Herzog, H., Steinmetz, H., Seitz, R.J., 1999. The neural circuitry involved in the reading of German words and pseudowords: a PET study. J. Cogn. Neurosci. 11, 383–398.

Josephs, O., Deichmann, R., Turner, R., 2000. Trajectory measurement and generalized reconstruction in rectilinear EPI. ISMRM Meeting 1517.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 773–795.

Kherif, F., Poline, J.B., Meriaux, S., Benali, H., Flandin, G., Brett, M., 2003. Group analysis in functional neuroimaging: selecting subjects using similarity measures. NeuroImage 20, 2197–2208.

Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. U. S. A. 92, 8135–8139.

Marshall, J., Newcombe, F., 1973. Patterns of paralexia: a psycholinguistic approach. J. Psycholinguist. Res. 2, 175–198.

Noppeney, U., Friston, K.J., Price, C.J., 2004. Degenerate neuronal systems sustaining cognitive functions. J. Anat. 205, 433–442.

Penny, W., Kiebel, S., Friston, K., 2003. Variational Bayesian inference for fMRI time series. NeuroImage 19, 727–741.

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

Phillips, J.A., Humphreys, G.W., Noppeney, U., Price, C.J., 2002. The neural substrates of action retrieval: an examination of semantic and visual routes to action. Vis. Cogn. 9 (4), 662–684.

Price, C.J., Friston, K.J., 2002. Degeneracy and cognitive anatomy. Trends Cogn. Sci. 6, 416–421.

Price, C.J., Friston, K.J., 2005. Functional ontologies for cognition: the systematic definition of structure and function. Cogn. Neuropsychol. 22, 262–275.

Rumiati, R.I., Humphreys, G.W., 1998. Recognition by action: dissociating visual and semantic routes to action in normal observers. J. Exp. Psychol. Hum. Percept. Perform. 24, 631–647.

Shallice, T., 2003. Functional imaging and neuropsychology findings: how can they be linked? NeuroImage 20 (Suppl. 1), S146–S154.

Shannon, C.E., Weaver, W., 1949. The Mathematical Theory of Communication. University of Illinois Press.

Shomstein, S., Yantis, S., 2004. Control of attention shifts between vision and audition in human cortex. J. Neurosci. 24, 10702–10706.

Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme, Stuttgart.

Ueda, N., Nakano, R., Gharamani, Z., Hinton, G.E., 2000. A variational Bayesian framework for graphical models. Adv. Neural Inf. 26, 133–140.

Wilkinson, D., Halligan, P., 2004. The relevance of behavioural measures for functional-imaging studies of cognition. Nat. Rev., Neurosci. 5, 67–73.