



**HAL**  
open science

# From Group to Patient-Specific Analysis of Brain Function in Arterial Spin Labelling and BOLD Functional MRI

Camille Maumet

► **To cite this version:**

Camille Maumet. From Group to Patient-Specific Analysis of Brain Function in Arterial Spin Labelling and BOLD Functional MRI. Neuroscience. Université Rennes 1, 2013. English. NNT: . tel-00863908v1

**HAL Id: tel-00863908**

**<https://theses.hal.science/tel-00863908v1>**

Submitted on 19 Sep 2013 (v1), last revised 9 Oct 2013 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE / UNIVERSITÉ DE RENNES 1  
*sous le sceau de l'Université Européenne de Bretagne*

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

*Mention : Informatique*

Ecole doctorale Matisse

présentée par

**Camille MAUMET**

préparée à l'unité de recherche UMR 6074 IRISA / Inria Rennes

Équipe d'accueil: VisAGeS - INSERM U746

Composante universitaire: Université de Rennes 1

---

**From Group to Patient-  
Specific Analysis of  
Brain Function in  
Arterial Spin Labelling  
and BOLD Functional  
MRI**

**Thèse soutenue à Rennes  
le 29/05/2013**

devant le jury composé de :

**Habib BENALI**

Directeur de Recherche à l'INSERM / rapporteur

**Sébastien OURSELIN**

Professeur à University College London / rapporteur

**Patrick BOUTHEMY**

Directeur de Recherche à Inria / examinateur

**Rik ACHTEN**

Responsable du service de radiologie à Ghent University  
Hospital / examinateur

**Jan WARNKING**

Chargé de Recherche au Grenoble Institut des Neuro-  
sciences / examinateur

**Clément DE GUIBERT**

Maître de conférence à l'Université de Rennes 2 / exami-  
nateur

**Christian BARILLOT**

Directeur de Recherche au CNRS / directeur de thèse

**Pierre MAUREL**

Maître de conférence à l'Université de Rennes 1 / co-  
directeur de thèse



## Remerciements

Je tiens à remercier la plateforme Neurinfo et en particulier les acteurs du Contrat de Projet État Région dont elle bénéficie, notamment la Région Bretagne et Rennes Métropole pour le financement de cette thèse.

Je remercie Patrick Bouthémy, président du jury lors de ma soutenance de thèse ainsi que Habib Benali et Sébastien Ourselin, rapporteurs de cette thèse, et Rik Achten, Jan Warnking et Clément de Guibert pour leur participation en tant que membres du jury lors de ma soutenance.

Je remercie Christian Barillot et Pierre Maurel pour leur co-encadrement. Merci à tous deux pour votre bonne humeur et vos encouragements tout au long de cette thèse.

Mes remerciements vont aussi à toute l'équipe du projet Dysphasie : Clément de Guibert, Elisabeth Le Rumeur, Pierre Jannin, Arnaud Biraben, Jean-Christophe Ferré, Catherine Allaire, Catherine Tréguier. Ainsi qu'à Anca Pasnicu et Eduardo Pasqualini que j'ai rencontrés dans le cadre du projet épilepsie. Je remercie aussi Jean-Christophe Ferré et Béatrice Carsin qui ont participé au volet clinique des projets ASL. Merci aussi à Isabelle Corouge et Élise Bannier, pour toutes ces discussions sur nos projets, en cours ou à venir !

Après quasiment 5 années passées dans l'équipe Visages, je remercie tous les membres, anciens et actuels, de l'équipe et en particulier: Tristan, Olivier L., Jan, Adrien, Bacem, Angélique, Odile, Aline, Bernard, Benoit, Marc, Romain, Daniel, Céline, Vincent, Jérémy, Ammar, David, Alexandre, Clément P., Guillaume, Germain, Franck, Romuald, Lei, Pierre H., Arnaud Le C., Emmanuel, Alesandro, Arnaud F., Aurore, Aymeric, Brivael, Elise, Fang, Jean-Christophe, Jean-Yves, Hélène, Guillaume, Hirishikesh, Isabelle, Juan Fran, Justine, Michael, Olivier C., Pierre M., Clément N., Renaud, René-Paul, Sylvain, Yann, Yogesh pour leur bonne humeur et une ambiance de travail à la fois décontractée et extrêmement motivante.

Merci aux cliniciens et aux manipulateurs en radiologie que j'ai pu rencontrer à l'hôpital de Pontchaillou ainsi qu'aux patients et sujets sains qui ont accepté de participer à nos études.

Et bien sûr : merci à ma famille et mes amis de Rennes, et d'ailleurs.



# Contents

---

<b>Résumé en français</b>	<b>ix</b>
<b>Introduction</b>	<b>1</b>
<b>I Context</b>	<b>5</b>
<b>1 The study of brain function</b>	<b>7</b>
1 Brain function . . . . .	9
1.1 Definition . . . . .	9
1.2 An integrative view of brain function . . . . .	10
2 Brain perfusion . . . . .	10
2.1 Blood supply to the brain . . . . .	10
2.2 Definition . . . . .	12
2.3 Parameters . . . . .	13
2.4 Applications . . . . .	13
2.5 Imaging . . . . .	17
3 Task-evoked brain activity . . . . .	26
3.1 Definition and vocabulary . . . . .	26
3.2 Neurovascular coupling . . . . .	27
3.3 Applications . . . . .	28
3.4 Imaging . . . . .	28
4 Conclusion . . . . .	31
<b>2 Pre-processing and statistical analysis</b>	<b>33</b>
1 Pre-processing . . . . .	35
1.1 Pipeline overview . . . . .	35
1.2 Pre-processing steps . . . . .	36
2 Statistical analysis . . . . .	38
2.1 Modelling and inference using the General Linear Model . . . . .	38
2.2 Subject level (First level) . . . . .	41
2.3 Group level (Second level) . . . . .	45
2.4 Correction for multiple comparisons . . . . .	53

3	Conclusion . . . . .	54
<b>3</b>	<b>From group to patient-specific analyses</b>	<b>55</b>
1	The need for patient-specific information . . . . .	57
2	Qualitative analyses . . . . .	57
3	Quantitative analyses . . . . .	57
3.1	Studies without spatial information . . . . .	57
3.2	Comparing one patient to a group of control subjects . . . . .	57
3.3	Multivariate approaches . . . . .	58
3.4	Locally multivariate approaches . . . . .	60
4	Conclusion . . . . .	65
<b>II Language impaired regions in specific language impairment: a group study using BOLD functional MRI</b>		<b>67</b>
<b>4</b>	<b>The “Neuroimaging in Dysphasia” project</b>	<b>69</b>
1	Context . . . . .	71
1.1	Specific Language Impairment . . . . .	71
1.2	Brain and language . . . . .	72
2	The “Neuroimaging in Dysphasia” project . . . . .	75
2.1	A brief chronology . . . . .	75
2.2	Imaging protocol . . . . .	76
<b>5</b>	<b>Functional MRI language mapping in children</b>	<b>77</b>
1	Introduction . . . . .	79
2	Subject and methods . . . . .	81
2.1	Subjects . . . . .	81
2.2	Experimental paradigms . . . . .	82
2.3	Data acquisition . . . . .	83
2.4	Data processing . . . . .	84
3	Results . . . . .	86
3.1	Reference auditory tasks . . . . .	86
3.2	New visual tasks . . . . .	88
3.3	Whole panel analysis . . . . .	89
4	Discussion . . . . .	93
4.1	Methodological issues . . . . .	94
4.2	Reference lexico-semantic auditory tasks . . . . .	95
4.3	New phonological visual tasks . . . . .	96
4.4	Whole panel . . . . .	97
<b>6</b>	<b>Abnormal brain function in Typical-Specific Language Impairment</b>	<b>99</b>
1	Introduction . . . . .	101
2	Methods . . . . .	103
2.1	Participants . . . . .	103

2.2	Neuropsychological and language assessment . . . . .	104
2.3	Functional MRI protocol and task panel . . . . .	104
2.4	Data acquisition . . . . .	106
2.5	Data processing . . . . .	107
3	Results . . . . .	108
3.1	Auditory language tasks . . . . .	108
3.2	Visual language tasks . . . . .	110
3.3	Lateralisation index assessment and comparison using the combined task analysis . . . . .	112
4	Discussion . . . . .	113
4.1	Lack of left lateralisation of core language areas . . . . .	113
4.2	Left hypoactivation of the posterior superior temporal gyrus / supramarginal gyrus junction . . . . .	114
4.3	Right hyperactivation of the anterior insula, adjacent inferior frontal gyrus and head of caudate . . . . .	115
4.4	Comparison to other developmental disorders . . . . .	117
4.5	Methodological considerations . . . . .	118
<b>III Patient-specific analysis of brain function</b>		<b>121</b>
<b>7</b>	<b>Robust Cerebral Blood Flow Maps in ASL</b>	<b>123</b>
1	Introduction . . . . .	125
2	Theory . . . . .	126
3	Material and Methods . . . . .	128
3.1	Data . . . . .	128
3.2	Validation . . . . .	129
4	Results . . . . .	131
4.1	Validation on simulated data . . . . .	131
4.2	Validation on real clinical data . . . . .	133
5	Conclusion . . . . .	135
<b>8</b>	<b>Modelling heterogeneous variances in the detection of patient- specific perfusion abnormalities with ASL</b>	<b>137</b>
1	Introduction . . . . .	139
2	Methods . . . . .	141
2.1	Detection of patient-specific perfusion abnormalities using a mixed-effect hierarchical two-sample t-test . . . . .	141
2.2	Evaluation framework . . . . .	146
3	Materials . . . . .	149
3.1	Data acquisition . . . . .	149
3.2	Pre-processing . . . . .	149
4	Results . . . . .	152
4.1	ASL template: a model of normal perfusion . . . . .	152



4.2	Testing the assumptions of the homoscedastic model . . . . .	152
4.3	Comparison of homoscedastic and heteroscedastic models . . .	155
5	Discussion and conclusion . . . . .	158
<b>9</b>	<b><i>A contrario</i> detection of perfusion abnormalities</b>	<b>161</b>
1	Introduction . . . . .	163
2	Methods . . . . .	164
2.1	The massively univariate General Linear Model . . . . .	164
2.2	The <i>a contrario</i> approach: a locally multivariate procedure .	166
3	Material . . . . .	169
3.1	Data . . . . .	170
3.2	Ground truth . . . . .	171
4	Results . . . . .	172
4.1	Comparison of the <i>a contrario</i> approach with the General Linear Model . . . . .	173
4.2	Distribution of the p-values under the null hypothesis . . . .	176
5	Discussion and conclusion . . . . .	179
<b>10</b>	<b><i>A contrario</i> detection of activation in BOLD fMRI</b>	<b>181</b>
1	Introduction . . . . .	183
2	Materials and methods . . . . .	183
2.1	Data . . . . .	183
2.2	Methods . . . . .	184
3	Results . . . . .	186
3.1	Voxel-wise p-values under the null hypothesis . . . . .	186
3.2	Comparison between the <i>a contrario</i> approach and the GLM	186
4	Discussion and conclusion . . . . .	190
	<b>Conclusion</b>	<b>193</b>
	<b>Bibliography</b>	<b>211</b>
	<b>Software</b>	<b>249</b>
	<b>Publications</b>	<b>251</b>

# Résumé en français

---

Dans ce chapitre, nous proposons un résumé en français du manuscrit rédigé en anglais. Nous conservons, autant que possible, la structure du document.

Ce travail est centré sur l'étude de la fonction cérébrale en Imagerie par Résonance Magnétique (IRM). Le contexte de cette étude est décrit dans la partie 1. Nos contributions sont ensuite présentées en deux parties. Nous proposons tout d'abord une analyse de groupe de la dysphasie chez l'enfant en IRM fonctionnelle BOLD dans la partie 2. Puis, nous nous concentrons sur l'étude de la fonction cérébrale au niveau de l'individu dans la partie 3. Nous concluons dans la partie 4.

In this chapter, we provide the reader with a summary of the work in French. The remaining of the document is written in English. The non-French speakers can easily skip this part.

## Sommaire

<b>1</b>	<b>Contexte</b> . . . . .	<b>xi</b>
1.1	L'étude de la fonction cérébrale . . . . .	xi
1.2	Des analyses de groupe à l'étude de l'individu . . . . .	xii
<b>2</b>	<b>Étude de la dysphasie chez l'enfant en IRMf BOLD</b> . . . . .	<b>xii</b>
2.1	La dysphasie . . . . .	xii
2.2	Données . . . . .	xiii
2.3	Résultats . . . . .	xiii
2.4	Conclusion . . . . .	xv
<b>3</b>	<b>Étude de la fonction cérébrale au niveau de l'individu</b> . . . . .	<b>xv</b>
3.1	Une estimation robuste du débit sanguin cérébral en ASL . . . . .	xv
3.1.1	Introduction . . . . .	xv
3.1.2	Méthode . . . . .	xvi
3.1.3	Résultats . . . . .	xvi
3.2	Modélisation de l'hétérogénéité des variances en ASL . . . . .	xvii
3.2.1	Introduction . . . . .	xvii
3.2.2	Méthode . . . . .	xviii
3.2.3	Validation . . . . .	xviii
3.2.4	Résultats . . . . .	xix
3.2.5	Conclusion . . . . .	xix
3.3	Une approche <i>a contrario</i> pour la détection d'anomalies de perfusion au niveau individuel et d'activations en IRM fonctionnelle BOLD . . . . .	xx
3.3.1	Introduction . . . . .	xx
3.3.2	Méthode . . . . .	xx
3.3.3	Validation . . . . .	xxii
3.3.4	Résultats . . . . .	xxii
3.3.5	Résultats quantitatifs . . . . .	xxii
3.3.6	Exemples de détections . . . . .	xxii
3.3.7	Conclusion . . . . .	xxiii
<b>4</b>	<b>Conclusion</b> . . . . .	<b>xxiv</b>
4.1	Contributions . . . . .	xxiv
4.2	Discussion et perspectives . . . . .	xxv

# 1 Contexte

## 1.1 L'étude de la fonction cérébrale

L'étude du fonctionnement cérébral est depuis longtemps le sujet de recherches soutenues. L'introduction de la neuroimagerie, permettant d'observer l'anatomie et la fonction cérébrale in-vivo, a permis l'émergence de nouvelles problématiques scientifiques.

Dans le domaine de l'étude de la fonction cérébrale, l'Imagerie par Résonance Magnétique (IRM) tient une place particulièrement importante. L'IRM fonctionnelle (IRMf) basée sur la mesure de l'oxygénation du sang, appelée IRMf BOLD (pour "Blood-Oxygen-Level-Dependent") est ainsi la méthode la plus répandue pour analyser le fonctionnement du cerveau. Avec cette méthode, il est possible d'étudier l'activation cérébrale produite en réponse à un stimulus. L'effet observé en IRMf BOLD est la conséquence de mécanismes physiologiques complexes qui ne sont aujourd'hui pas complètement compris. Les zones d'activation sont identifiées en comparant l'activité observée pendant la réalisation d'une tâche, à l'activité au repos. L'IRMf BOLD est donc une technique adaptée à la mesure d'un changement dans l'activité cérébrale, elle fournit une mesure relative et non absolue.

Il est cependant important de remarquer que la fonction cérébrale est un processus qui est principalement spontané, et qu'il serait donc intéressant de pouvoir mesurer l'activité au repos, aussi appelée activité basale. La perfusion cérébrale est le processus biologique qui assure l'approvisionnement du cerveau en oxygène et en nutriments au moyen de la micro-circulation sanguine. En IRM, la technique historique permettant de mesurer la perfusion est l'imagerie dynamique de susceptibilité magnétique (DSC). Cette technique est basée sur l'injection d'un produit de contraste et n'est donc pas adaptée pour les sujets sains ou les populations à risques telles que les femmes enceintes ou les enfants.

L'imagerie de perfusion par marquage de spins (ASL) permet de mesurer la perfusion cérébrale en utilisant un traceur endogène : les protons du sang marqués avec un pulse radiofréquence. Cette technique ne nécessite pas d'injection de produit de contraste et peut donc être appliquée à une population plus large que le DSC. Cet avantage, associé au fait que l'ASL permet d'imager à la fois l'activité basale et l'activation induite par la présentation d'un stimulus, en fait une méthode particulièrement intéressante d'un point de vue scientifique mais aussi clinique. L'utilisation de l'ASL reste cependant assez restreinte et ce, malgré le fait que son principe ait été introduit à peu près au même moment que l'IRMf BOLD [Detre 1992]. Une des explications est sans doute le faible rapport signal sur bruit des images produites par cette séquence ainsi que, initialement, des difficultés techniques à sa mise en oeuvre. Aujourd'hui, l'ASL a gagné en maturité et est maintenant disponible en tant que séquence commerciale sur la plupart des scanners IRM. De nombreuses études ont montré l'utilité de cette séquence, en particulier dans la prise en charge de l'accident vasculaire cérébral ou dans le suivi des tumeurs. Il reste cependant encore beaucoup à faire, notamment dans la défini-

tion de post-traitements standards, pour que cette séquence puisse être utilisable en routine clinique.

## 1.2 Des analyses de groupe à l'étude de l'individu

L'étude du fonctionnement cérébral normal ou des dysfonctions induites par une pathologie est généralement réalisée au niveau du groupe. Ces approches qui visent à produire des conclusions au niveau d'une population se sont montrées particulièrement efficaces par le passé et constituent aujourd'hui la méthode la plus répandue pour analyser la fonction cérébrale. D'un point de vue clinique cependant, c'est le patient, individu unique, qui est le centre d'intérêt. Le diagnostic et le choix du traitement sont des tâches qui sont nécessairement réalisées au niveau de l'individu et non au niveau du groupe. C'est pourquoi il y a aujourd'hui un intérêt grandissant pour les méthodes individuelles qui visent à produire des conclusions au niveau d'un sujet.

Dans ce travail, nous proposons tout d'abord une analyse de groupe de la dysphasie chez l'enfant en IRMf BOLD. Ensuite, nous nous concentrons sur les analyses individuelles. Nous proposons une approche robuste pour calculer la mesure de perfusion en ASL. Ensuite, nous étudions la validité des hypothèses qui sous-tendent les analyses statistiques classiques dans le contexte de l'ASL. Nous proposons ensuite une nouvelle méthode localement multivariée, basée sur l'approche *a contrario* et nous l'appliquons dans deux domaines : la détection d'anomalies de perfusion en ASL et la détection de l'activation cérébrale en IRMf BOLD.

## 2 Mise en évidence de régions cérébrales présentant un fonctionnement atypique dans la dysphasie chez l'enfant : une analyse de groupe en IRM fonctionnelle BOLD

### 2.1 La dysphasie

La dysphasie fait partie des troubles développementaux du langage. Elle est ainsi :

- sévère et persistante : la dysphasie n'est pas un retard de langage ;
- spécifique : bien que la dysphasie soit parfois associée à d'autres troubles, elle n'impacte que le langage (ou des processus liés au langage) ;
- développementale : la dysphasie n'est pas la conséquence d'une lésion cérébrale acquise.

La dysphasie est donc définie par négation (pas un retard de langage, n'impacte pas les capacités intellectuelles, n'est pas acquise) ce qui rend son diagnostic particulièrement complexe. Dans la littérature, la dénomination de dysphasie reste assez floue et représente un sous-ensemble plus ou moins bien défini des troubles

développementaux du langage. La fig. 1 fournit un diagramme des principaux troubles langagiers et les liens qui existent entre eux, en accord avec les définitions utilisées par d'autres auteurs [Brun 2003, Bishop 2006].

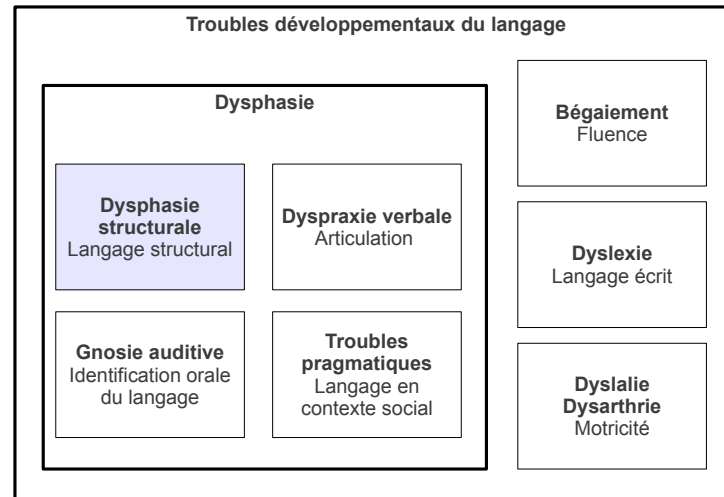


Figure 1: Diagramme des troubles développementaux du langage.

Dans le contexte de ce travail, nous nous sommes restreints à l'étude de la dysphasie structurale qui impacte les aspects structuraux du langage: la phonologie (la maîtrise du système de différenciation phonologique), le lexique (la maîtrise du système de différenciation des mots), la morphologie (formation de mots nouveaux ou dérivés), la syntaxe (la maîtrise de la construction de la phrase dans l'expression et la compréhension). Cette pathologie à une prévalence inférieure à 1% dans la population générale.

## 2.2 Données

21 enfants dysphasiques et 18 enfants présentant un développement caractéristique, âgés de 7 à 18 ans, ont participé à cette étude.

Le protocole d'imagerie incluait deux images anatomiques (T1 et T2 FLAIR), quatre IRMf correspondant aux quatre paradigmes langagiers étudiés et une image de diffusion. La fig. 2 présente une vision d'ensemble des quatre paradigmes étudiés. Chaque paradigme se divisait en six phases : trois phases d'action et trois phases de repos de 27 s chacune en alternance. Deux paradigmes, Catégorie et Dénominations, proviennent de la littérature et sont basés sur l'écoute de stimuli. Les deux paradigmes restants, Phon-diff et Phon-seg ont été développés dans le cadre de cette étude et se basent sur des stimuli visuels.

## 2.3 Résultats

La comparaison inter-groupe des activations (seuil  $p < 0.05$  prenant en compte les tests multiples avec une correction du risque d'erreur par familles (FWE)) a mis en

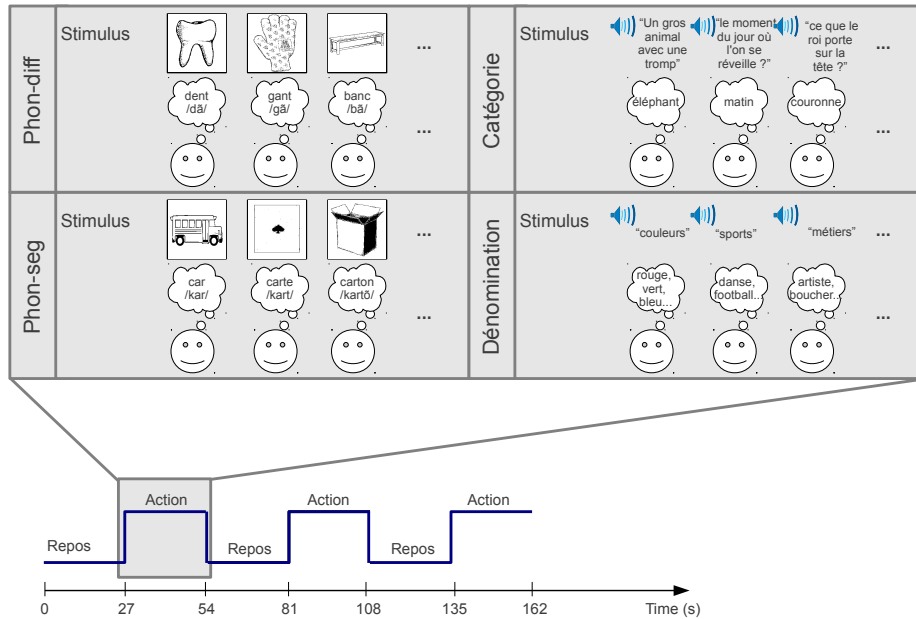


Figure 2: Les quatre paradigmes langagiers inclus dans le protocole Dysphasie.

évidence deux zones de fonctionnement atypique pour deux des quatre paradigmes étudiés.

Lors de la réalisation des tâches du paradigme Dénomination, les sujets dysphasiques ont présenté une activation réduite (hypo-activation) dans la partie postérieure du gyrus temporal supérieur à la jonction avec le gyrus supramarginal à gauche. Cette région, parfois dénommée région de Wernicke, est connue comme étant l'une des aires majeures du traitement du langage dans le cerveau.

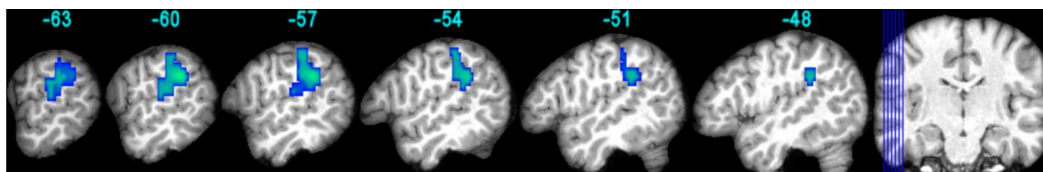


Figure 3: Hypo-activation du groupe dysphasique par rapport au groupe contrôle pour la tâche Définition ( $p < 0.05$  FWE).

Lors de la réalisation des tâches du paradigme Phon-diff, les sujets dysphasiques ont présenté un niveau d'activation plus élevé (hyper-activation) que les enfants à développement caractéristique dans la partie antérieure de l'insula, le gyrus frontal inférieur adjacent et dans la tête du noyau caudé à droite. La partie corticale de cette région est l'homologue à droite de la région parfois dénommée Broca, qui est connue comme étant l'une des aires majeures du traitement du langage dans le cerveau.

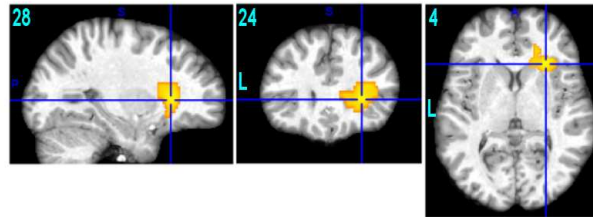


Figure 4: Hyper-activation du groupe dysphasique par rapport au groupe contrôle pour la tâche Phon-diff ( $p < 0.05$  FWE).

## 2.4 Conclusion

Dans ce travail, nous avons étudié la dysphasie structurale, un sous-groupe des troubles développementaux du langage oral affectant les aspects structuraux du langage. En IRMf BOLD, nous avons mis en évidence un fonctionnement atypique de deux régions : une aire temporale dans l'hémisphère gauche connue pour son implication dans le langage et une aire frontale dans l'hémisphère droit, contralatérale à l'aire de Broca qui est elle-aussi impliquée dans le traitement du langage dans le cerveau.

## 3 Étude de la fonction cérébrale au niveau de l'individu

### 3.1 Une estimation robuste du débit sanguin cérébral en ASL

#### 3.1.1 Introduction

L'ASL permet de mesurer le débit sanguin cérébral (CBF) de manière non-invasive. Cette technique souffre cependant d'un faible rapport signal sur bruit qui affecte la qualité des cartographies de perfusion. Afin d'obtenir un niveau de signal satisfaisant, l'acquisition ASL est généralement répétée plusieurs fois. La cartographie de perfusion est ensuite obtenue en moyennant le signal des différentes répétitions. La moyenne est cependant un estimateur qui est très sensible aux valeurs atypiques comme illustré dans la fig. 5. Plusieurs études ont montré que des valeurs atypiques pouvaient être présentes dans la série ASL.

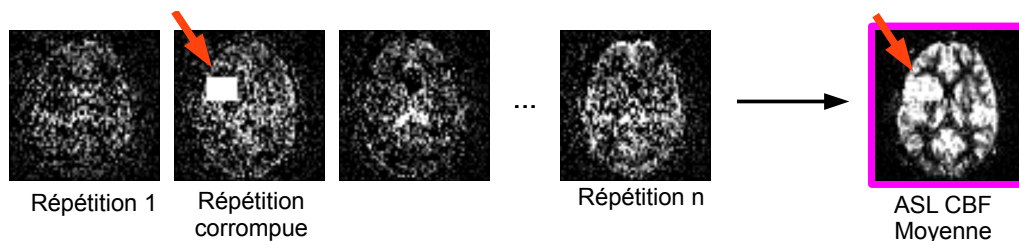


Figure 5: Moyenne : exemple d'estimation du débit sanguin cérébral en présence de valeurs atypiques.



### 3.1.2 Méthode

Dans [Tan 2009], les auteurs ont proposé d'éliminer les volumes dont l'intensité apparaît anormale grâce à une mesure basé sur des scores  $z$ . Un exemple d'estimation est donné en fig. 6. Cette méthode utilise cependant des seuils définis de façon empirique. Par ailleurs, des méthodes statistiques plus robustes que les scores  $z$  sont disponibles dans la littérature.

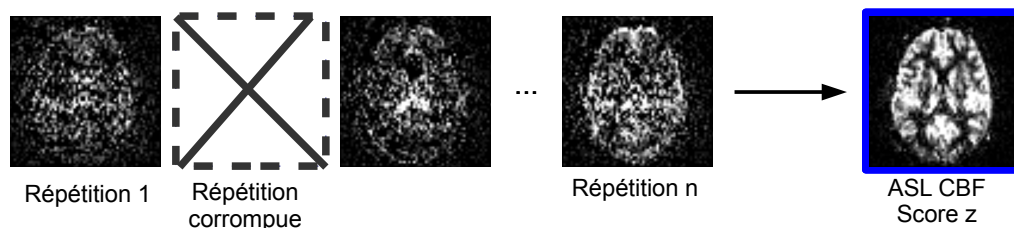


Figure 6: Seuillage par score  $z$  : exemple d'estimation du débit sanguin cérébral en présence de valeurs atypiques.

Nous proposons donc d'estimer la cartographie de perfusion au moyen de statistiques robustes : les M-estimateurs. Nous nous concentrons en particulier sur le M-estimateur de Huber qui est le plus utilisé de cette catégorie. Un exemple d'estimation est présenté en fig 7.

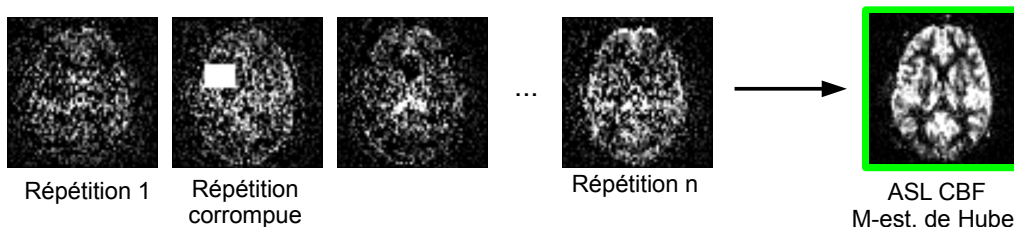


Figure 7: M-estimateur de Huber : exemple d'estimation du débit sanguin cérébral en présence de valeurs atypiques.

### 3.1.3 Résultats

24 patients atteints de tumeurs cérébrales ont participé à cette étude. La validation a été réalisée en mesurant la corrélation voxel-à-voxel de la carte de débit sanguin cérébral produite par chaque estimateur, à partir des données ASL, avec la carte de débit sanguin cérébral acquise avec la méthode d'imagerie de perfusion la plus validée en IRM : le DSC. La fig. 8.a présente le niveau de corrélation obtenu pour chacune des 3 méthodes d'estimation avec des jeux de données de différentes tailles (de 5 volumes à 50 volumes). Le M-estimateur de Huber et la méthode utilisant les scores  $z$  sont tous les deux plus performants que la moyenne. Le M-estimateur fournit une estimation légèrement meilleure que la méthode du score  $z$ . La fig. 8.b donne

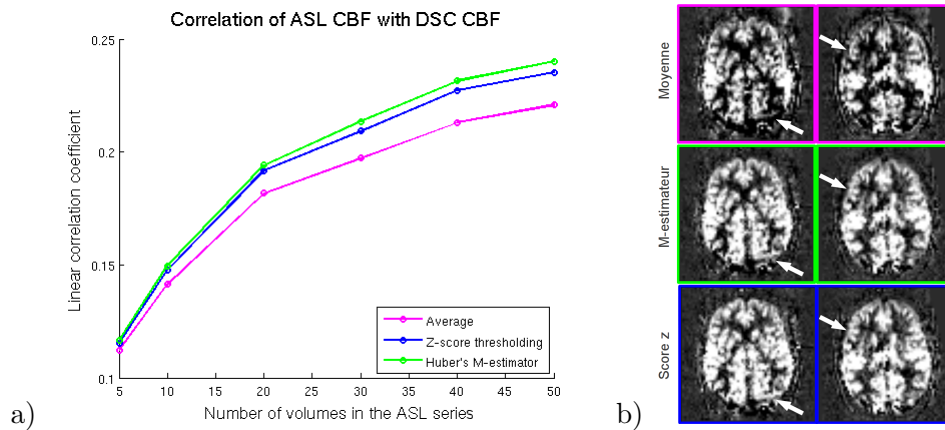


Figure 8: Corrélation de la cartographie ASL CBF avec la carte DSC CBF estimée avec la moyenne, le M-estimateur de Huber ou la méthode des scores z (a). Exemple de résultats sur données réelles (b), les flèches blanches soulignent des artefacts présents dans la carte obtenue avec moyennage qui sont correctement éliminés avec le score z et le M-estimateur.

un exemple sur données réelles pour lequel, la méthode du score z et le M-estimateur corrigent d'importants artefacts visibles sur la carte obtenue par moyennage.

### 3.2 Modélisation de l'hétérogénéité des variances dans la détection individuelle d'anomalies perfusionnelles en ASL

#### 3.2.1 Introduction

Dans cette partie, nous nous proposons de détecter des anomalies de perfusion au niveau individuel en ASL en utilisant la méthode statistique la plus répandue en neuroimagerie : le Modèle linéaire généralisé (GLM). Un exemple de détections est proposé en fig. 9.

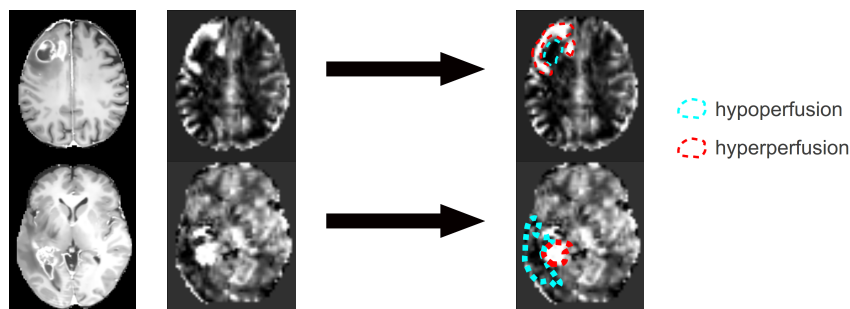


Figure 9: Exemple de détections individuelles d'anomalies de perfusion chez deux patients atteints de tumeurs cérébrales. De gauche à droite : T1-Gd, ASL CBF, détections.

### 3.2.2 Méthode

Dans la détection d'anomalies de perfusion, obtenue en comparant un sujet à un groupe de contrôle, il existe deux composantes qui induisent des variabilités dans la mesure : la variance intra-sujet (aussi appelée erreur de mesure) et la variance inter-sujet (aussi appelée variance intra-groupe). En ASL, la variance intra-sujet mesure la variabilité observée entre les différentes répétitions acquises sur un même sujet comme illustré dans la fig. 10. La variance inter-sujet mesure la variabilité entre sujets d'un même groupe comme décrit dans la fig. 11.

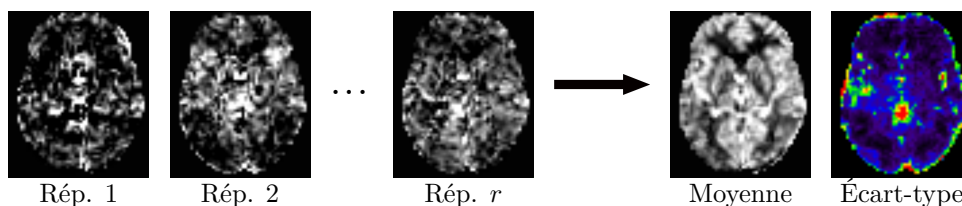


Figure 10: La variance intra-sujet mesure la variabilité entre les répétitions.

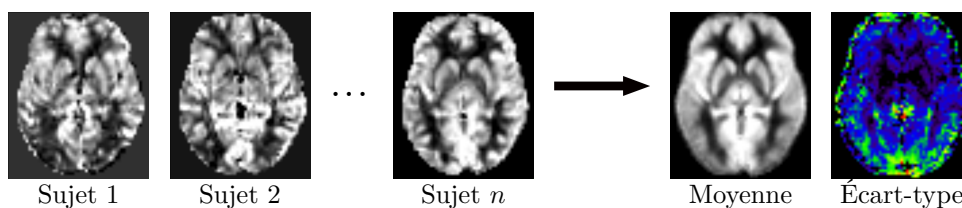


Figure 11: La variance inter-sujets mesure la variabilité entre sujets d'un même groupe.

Deux approches sont actuellement utilisées pour résoudre ce type de problème : l'approche homoscedastique, qui fait l'hypothèse d'une variance intra-sujet constante, et l'approche hétéroscedastique. L'approche hétéroscedastique est théoriquement plus efficace en présence de variances intra-sujet hétérogènes mais algorithmiquement plus demandeuse. Dans la pratique, en IRMf la supériorité du modèle hétéroscedastique est encore remise en question.

En raison du faible rapport signal sur bruit des cartographies produites en ASL, la variance intra-sujet joue un rôle particulièrement important [Viviani 2009]. Nous nous proposons donc d'étudier le comportement des modèles homoscedastique et hétéroscedastique dans la détection d'anomalies de perfusion au niveau individuel en ASL.

### 3.2.3 Validation

36 sujets sains et 25 patients ont participé à cette étude.

La validation des détections est particulièrement complexe dans la mesure où, comme dans de nombreuses applications de neuroimagerie, la vérité terrain n'est pas connue. Nous avons choisi le contexte des pathologies tumorales car, dans ce cadre,

la perfusion est une information qui a été particulièrement étudiée. Par ailleurs, en raison de son intérêt clinique, le DSC fait partie des séquences d'imagerie préconisées chez ces patients. En prenant appui sur les connaissances cliniques nous avons construit notre vérité terrain en combinant les informations anatomiques (T1, T2 FLAIR) et perfusionnelles (DSC). En raison de son faible rapport signal sur bruit, l'ASL n'est pas bien adaptée à la mesure de faibles niveaux de perfusion [Wintermark 2005]. Nous avons donc restreint l'analyse de la sensibilité (détection correcte des anomalies à détecter) aux hyper-perfusions. La fig. 12 fournit un exemple de vérité terrain chez un sujet représentatif.

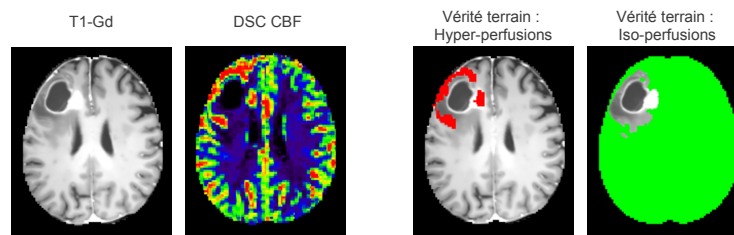


Figure 12: Vérité terrain chez un patient. De gauche à droite : T1-Gd, DSC CBF, Vérité terrain: hyper-perfusions (à détecter) en rouge et iso-perfusion (à ne pas détecter) en vert.

Les performances des modèles homoscédastique et hétéroscédastique ont été comparées à l'aide de courbes ROC (pour "Receiver Operating Characteristic"). En particulier, nous avons calculé l'aire sous la courbe (pour des taux de faux positifs compris entre 0% et 10%) afin d'obtenir une mesure de la qualité de la classification associée à chaque méthode. Le GLM étant généralement appliqué sur des données lissées avec un noyau Gaussien, nous avons étudié différents jeux de paramètres : largeur à mi-hauteur du noyau Gaussien égale à  $0 \text{ mm}^3$  (pas de lissage),  $4 \text{ mm}^3$ ,  $6 \text{ mm}^3$ ,  $8 \text{ mm}^3$ ,  $10 \text{ mm}^3$  et  $12 \text{ mm}^3$ .

### 3.2.4 Résultats

La table 1 résume l'aire sous la courbe ROC obtenue pour chacun des jeux de paramètres testés. Le modèle hétéroscédastique est plus performant que le modèle homoscédastique avec une aire sous la courbe ROC plus importante. Les meilleurs performances sont obtenues avec un noyau Gaussien de largeur à-mi hauteur égale à  $8 \text{ mm}^3$  pour les deux méthodes.

### 3.2.5 Conclusion

Nous avons mis en évidence l'hétérogénéité des variance intra-sujets en ASL et démontré une augmentation du taux de faux positifs si cette hétérogénéité n'est pas correctement prise en compte. Nous conseillons donc l'utilisation du modèle hétéroscédastique dans la détection d'anomalies de perfusion au niveau individuel en ASL.

Table 1: Aire sous la courbe ROC pour des taux de faux positifs compris entre 0% et 10% avec les modèles homoscédastique et hétéroscédastique. Le modèle hétéroscédastique est plus performant que le modèle homoscédastique.

	Modèle homoscédastique						Modèle hétéroscédastique					
	0	4	6	8	10	12	0	4	6	8	10	12
Aire	0.46	<b>0.49</b>	<b>0.49</b>	<b>0.49</b>	0.48	0.48	0.63	0.70	<b>0.72</b>	<b>0.72</b>	0.69	0.65

### 3.3 Une approche *a contrario* pour la détection d’anomalies de perfusion au niveau individuel et d’activations en IRM fonctionnelle BOLD

#### 3.3.1 Introduction

Dans cette partie, nous proposons un modèle localement multivarié pour la détection d’anomalies de perfusion au niveau individuel. Ce modèle est basé sur l’approche *a contrario*, un cadre statistique venant de la communauté de la vision par ordinateur. Cette méthode est comparée au GLM, qui est la méthode standard dans l’analyse des données de neuroimagerie. Le GLM est une approche massivement univariée car elle traite indépendamment chacun des voxels étudiés.

#### 3.3.2 Méthode

L’approche *a contrario* [Desolneux 2003] est généralement définie en trois étapes :

1. La définition d’un modèle de fond ;
2. L’estimation d’une probabilité niveau région ;
3. La correction des comparaisons multiples.

Le terme “*a contrario*” a été initialement choisi dans le but d’illustrer le fait que cette méthode est basée sur la définition d’un modèle de fond (ce que l’on ne souhaite pas détecter) au lieu de la modélisation de l’objet à détecter.

**Définition d’un modèle de fond** La première étape est la définition d’un modèle de fond (cf. fig. 13, 1<sup>ère</sup> colonne). Ce modèle représente le bruit de fond ou bien est construit à partir d’une population contrôle.

Les *événements rares* sont définis comme étant des observations qui ne respectent pas le modèle de fond. Un voxel est associé à un événement rare si la valeur observée à ce voxel a une probabilité (sous le modèle de fond) inférieure à un seuil pré-défini noté  $p_{PRE}$ .

**Estimation d’une probabilité niveau région** Dans une seconde étape, les probabilités calculées au niveau des voxels sont étendues au niveau des régions (cf. fig. 13, 2<sup>ème</sup> colonne). Cette étape permet de prendre en compte les voisins de

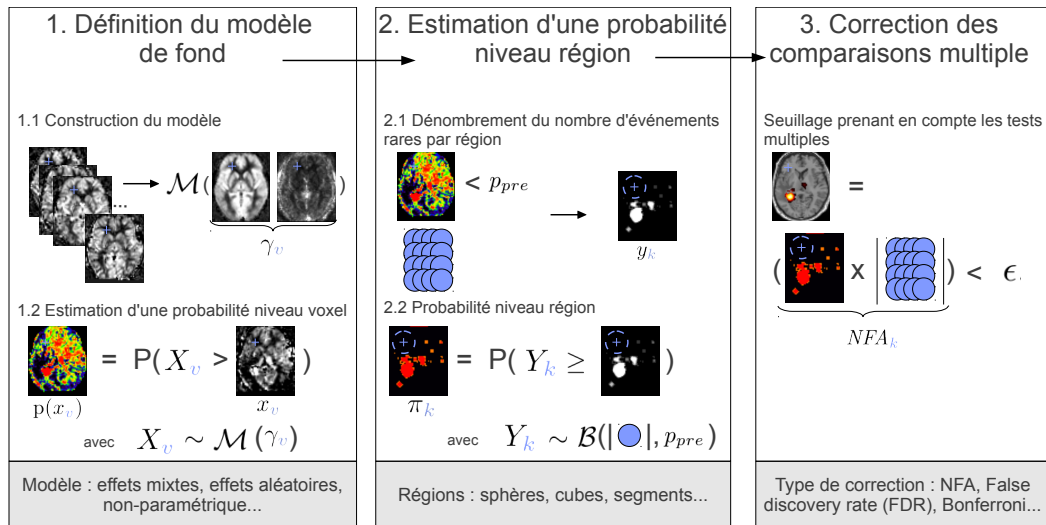


Figure 13: Vue d'ensemble de l'approche *a contrario* : définition du modèle de fond, estimation d'une probabilité niveau région et correction des comparaisons multiples.

chaque voxel avec une métrique (localement) multivariée. Dans ce but, un ensemble de régions est défini *a priori* dans l'image. Si la forme des détections à trouver est connue alors les régions sont choisies en conséquence. Par exemple, dans la première application de la méthode *a contrario*, qui avait pour but la détection de lignes dans des images, des segments de tailles variables avaient été choisis comme régions d'analyse. En l'absence d'*a priori*, deux méthodes ont été proposées : utiliser un algorithme de clustering [Rousseau 2008] ou définir une sphère (ou un cube) centrée en chaque voxel [Aguerrebere 2009, Rousseau 2007].

Le nombre d'événements rares est ensuite calculé pour chaque région d'analyse. Puis, connaissant la loi de probabilité suivie par le nombre d'événements rares par sphère sous le modèle de fond, une probabilité est affectée à chaque région. Dans le cas d'un bruit blanc, la loi de probabilité est une Binomiale paramétrée par le nombre de voxels par sphère et le seuil  $p_{PRE}$ . Dans le cas où les régions d'analyse sont des sphères, la probabilité niveau région peut ensuite être affectée au voxel central de la sphère.

**Correction des comparaisons multiples** La troisième et dernière étape consiste à seuiller la carte de probabilité niveau région pour faire apparaître les détections. Ce seuillage doit être réalisé en prenant en compte les multiples tests réalisés en parallèle (cf. fig. 13, 3<sup>ème</sup> colonne). Dans l'approche *a contrario*, une variante de l'approche de Bonferroni est généralement employée [Rousseau 2008] mais d'autres types de correction sont envisageables (par exemple le FDR pour "False Discovery Rate" de [Benjamini 1995]).

### 3.3.3 Validation

Les performances de la méthode *a contrario* et du GLM ont été comparées à l'aide de courbes ROC. En particulier, nous avons calculé l'aire sous la courbe (pour des taux de faux positifs compris entre 0% et 10%) afin d'obtenir une mesure de la qualité de la classification associée à chaque méthode. Le GLM étant généralement appliqué sur des données lissées avec un noyau Gaussien, nous avons étudié différents jeux de paramètres : largeur à mi-hauteur du noyau Gaussien égale à 0 mm<sup>3</sup> (pas de lissage), 4 mm<sup>3</sup>, 6 mm<sup>3</sup>, 8 mm<sup>3</sup>, 10 mm<sup>3</sup> et 12 mm<sup>3</sup>. Pour la méthode *a contrario*, nous avons étudié 2 tailles de sphère : rayon  $rd \in [1, 2]$  voxel(s) et 3 valeurs de seuils  $p_{PRE} \in [0.01, 0.005, 0.001]$ .

### 3.3.4 Résultats

#### 3.3.5 Résultats quantitatifs

La table 2 résume l'aire sous la courbe ROC obtenue pour chacun des jeux de paramètres testés. La méthode *a contrario* est plus performante que le GLM hétéroscédastique avec une aire sous la courbe ROC plus importante.

Table 2: Aire sous la courbe ROC pour des taux de faux positifs compris entre 0% et 10% avec le GLM hétéroscédastique et la méthode *a contrario*. La méthode *a contrario* est plus performante que le GLM.

$rd/w$	<i>a contrario</i>						GLM						
	1		2		3		0	2	4	6	8	10	12
$p_{PRE}$	0.01	0.005	0.001	0.01	0.005	0.001							
Aire	0.68	0.73	0.79	0.75	0.80	<b>0.87</b>	0.63	0.67	0.70	0.72	<b>0.73</b>	0.70	0.65

### 3.3.6 Exemples de détections

Afin d'illustrer les avantages procurés par la méthode *a contrario*, nous avons choisi 2 sujets représentatifs et comparé les deux méthodes avec un taux fixe de faux ou de vrais positifs. La fig. 14 présente les détections obtenues avec les deux méthodes pour les deux patients sélectionnés. Le premier sujet, qui présente une petite hyper-perfusion, est étudié avec un taux de vrais positifs de 50%. Pour le second patient, les deux méthodes sont comparées avec un taux de faux positifs de 5%. Le patient 15 souffre d'un gliosarcome dans l'hémisphère gauche à proximité de la région parahippocampique. La lésion est caractérisée par de petites hyper-perfusions dans sa partie dorsale, de petites hypo-perfusions sont observées dans l'oedème périphérique. Avec un taux de faux positifs de 50%, le GLM présente un plus grand nombre de faux positifs que l'approche *a contrario*. Le patient 7 souffre d'un méningiome des deux lobes occipitaux caractérisé par une grande hyper-perfusion. Avec un taux de faux positifs de 5%, la plus grande partie d'hyper-perfusion est correcte-

ment détectée avec les deux méthodes. L'approche *a contrario* est cependant plus sensible et détecte une plus large proportion de l'anomalie de perfusion.

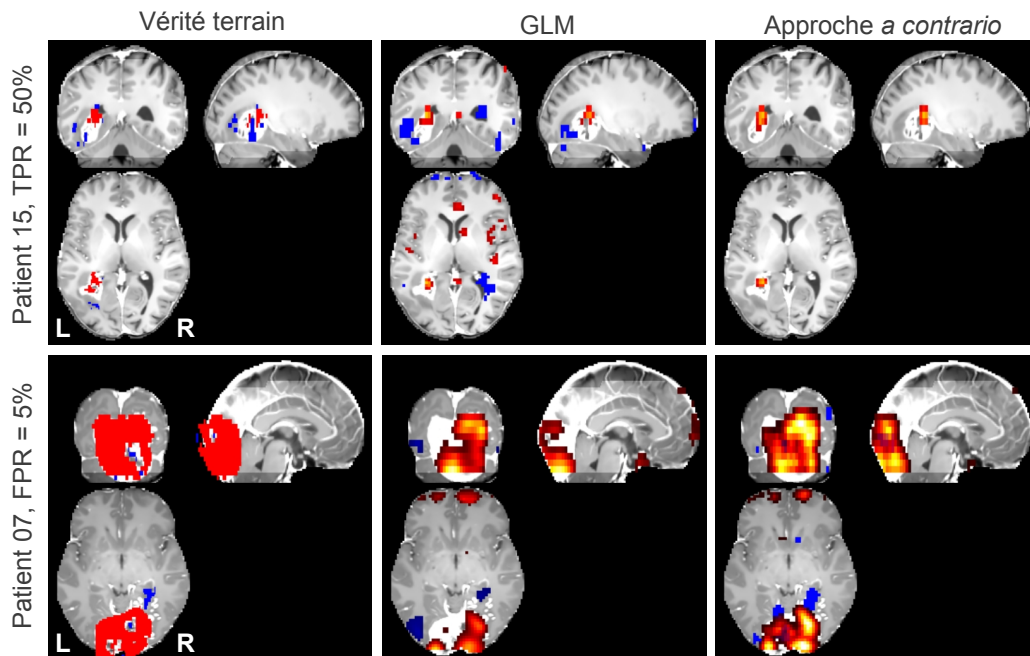


Figure 14: Anomalies de perfusion chez 2 patients avec le GLM et la méthode *a contrario*. De gauche à droite : vérité terrain hyper-perfusions (rouge) et hypo-perfusions (bleu) superposées à la carte T1w-Gd; GLM hyper-perfusions (couleurs chaudes) et hypo-perfusions (bleu); *a contrario* hyper-perfusions (couleurs chaudes) et hypo-perfusions (bleu).

Cette méthode *a contrario* est par ailleurs transposable à d'autres applications comme la détection des zones activées en IRMf. Pour des raisons de place nous ne sommes pas en mesure de traiter cette seconde application en détail dans ce résumé en français. La fig. 15 présente un exemple d'activations détectées avec la méthode *a contrario* et avec un GLM lissé et non-lissé pour un paradigme moteur de la main droite. Dans cet exemple, seule la méthode *a contrario* est en mesure de fournir une détection correcte de l'aire motrice supplémentaire.

### 3.3.7 Conclusion

Nous avons proposé une nouvelle méthode localement multivariée, basée sur une approche *a contrario*, qui présente de meilleures performances dans la détection d'hypo- et hyper-perfusions en ASL ainsi que dans la détection d'aires activées en IRMf BOLD.



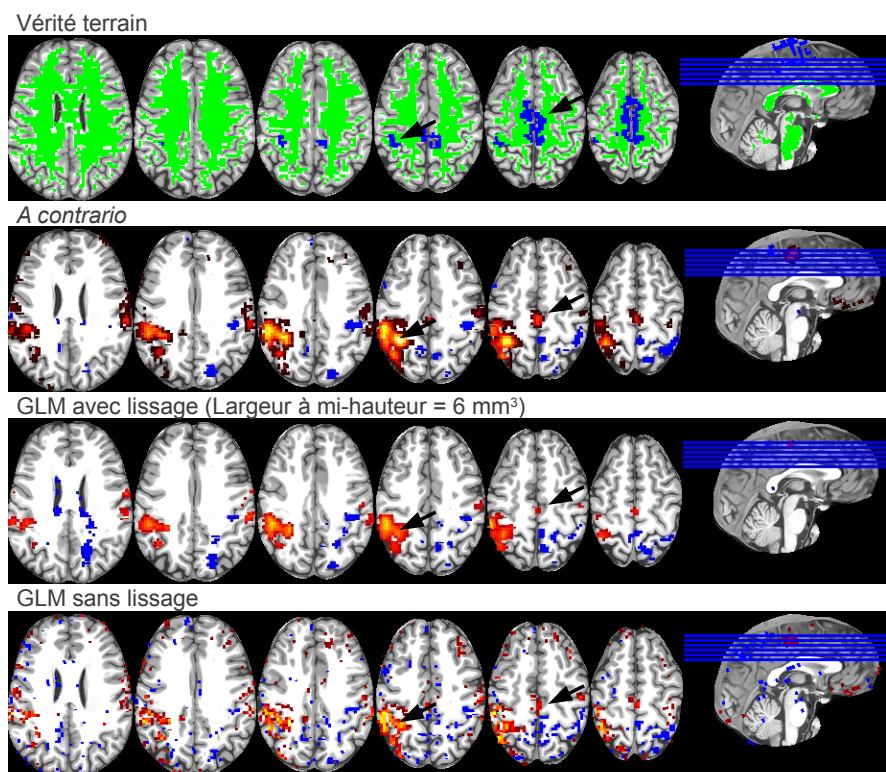


Figure 15: Exemple d'activations obtenues avec un taux de faux positifs de 5%. L'aire motrice primaire de la main (flèche noire sur la 4<sup>ème</sup> coupe) est correctement détectée avec le GLM et l'*a contrario*. L'aire motrice supplémentaire (flèche noire sur la 5<sup>ème</sup> coupe) est mieux détectée avec l'approche *a contrario*.

## 4 Conclusion

Cette thèse traite de l'analyse de la fonction cérébrale en IRM. En conclusion, nous proposons une vue d'ensemble de nos contributions, nous discutons ensuite les limites et perspectives de ces travaux.

### 4.1 Contributions

À l'aide d'une analyse de groupe mettant en oeuvre un modèle linéaire généralisé, nous avons mis en évidence des motifs atypiques d'activation et de latéralité dans la dysphasie chez l'enfant. Nous avons notamment démontré une hypoactivation de la zone postérieure du gyrus temporal supérieur à la jonction avec le gyrus supramarginal dans l'hémisphère gauche ainsi qu'une hyperactivation dans la partie antérieure de l'insula, le gyrus frontal inférieur adjacent et la tête du noyau caudé, dans l'hémisphère droit.

Ensuite, trois contributions méthodologiques ont été proposées dans le cadre des analyses individuelles. Tout d'abord, nous avons proposé l'utilisation de M-

estimateurs pour construire les cartographies de débit sanguin cérébral en ASL. Ensuite, nous avons étudié l'hétéroscédasticité des données ASL et son impact dans la détection d'anomalies perfusionnelles au niveau individuel. Nous avons ainsi montré l'hétérogénéité des variances intra-sujet en ASL. De plus, nous avons mis en évidence une augmentation du taux de faux positifs si cette variabilité n'est pas correctement prise en compte. Enfin, nous avons proposé une approche *a contrario*. En accord avec les récents développements dans le domaine des approches multivariées en neuroimagerie, nous avons souligné la supériorité de notre approche localement multivariée par rapport au GLM standard massivement univarié. Ce gain en performance a été illustré dans deux domaines applicatifs : la détection d'anomalies de perfusion au niveau individuel en ASL et la détection d'activations en IRMf BOLD.

## 4.2 Discussion et perspectives

Nos résultats, mettant en évidence un fonctionnement cérébral atypique dans la dysphasie chez l'enfant, sont particulièrement intéressants dans la mesure où cette pathologie a été très peu analysée en neuroimagerie. Il serait cependant utile de re-discuter ces résultats à la lumière d'informations complémentaires fournies par d'autres séquences ou modalités. Un article centré sur l'anatomie est en cours de rédaction. Les données de diffusion acquises dans le cadre de cette étude ont aussi montré de premiers résultats prometteurs. D'autre part, ces analyses pourraient être reconduites avec une méthode *a contrario* comme celle proposée précédemment.

Notre utilisation des statistiques robustes a été restreinte à la production de cartes de débit sanguin cérébral et n'a pas été étendue à l'analyse statistique. Dans un premier temps, nous avons en effet préféré favoriser le GLM qui est la technique standard d'analyse des données de neuroimagerie. Les statistiques robustes pourraient néanmoins être utiles à des fins de détections dans les analyses individuelles.

Nos expériences en ASL ont été réalisées avec la séquence produit disponible sur notre machine IRM : une PASL PICORE Q2TIPS. Des séquences plus avancées, comme l'ASL pseudo-continu, sont aujourd'hui à même de fournir un meilleur rapport signal sur bruit. Il serait intéressant de mesurer et de comparer le niveau d'hétéroscédasticité observé avec des jeux de données obtenus avec différentes séquences ASL.

Pour finir, une extension possible à ce travail serait l'utilisation d'une approche *a contrario* pour réaliser des statistiques de groupe.



# Introduction

---

For a long time, understanding how the brain works has been the centre of acute researches. The emergence of functional neuroimaging techniques, allowing for in-vivo imaging, has opened the field for new scientific questions and methodological issues.

In particular, [Magnetic Resonance Imaging \(MRI\)](#), thanks to its non-invasiveness, has been widely used to study brain function. In [MRI](#), the preferred imaging technique to measure brain function is functional [MRI](#) based on the [Blood-Oxygen-Level-Dependent \(BOLD\)](#) effect. With this technique, a measure of cerebral activation in response to a stimulus is achieved. However, due to the complex physiological mechanisms underlying the [BOLD](#) effect, this measure is only relative so that basal activity cannot be studied through this sequence.

While [BOLD Functional Magnetic Resonance Imaging \(fMRI\)](#) is useful to measure task-based activation, brain function is mainly an intrinsic process and cannot be confined to task-induced activity. Brain perfusion is the biological process that insures the delivery of oxygen and nutrients to the cerebral tissues by means of microcirculation. In [MRI](#), the historical approach to measure perfusion is the [Dynamic Susceptibility weighted Contrast imaging \(DSC\)](#). This method however requires the injection of a contrast agent necessitating a good venous access and inducing a potential risk of allergy that makes it not widely applicable in sensitive populations such as children and pregnant women or in healthy subjects. [Arterial Spin Labelling \(ASL\)](#) sequences permit the measurement of perfusion with [MRI](#) using an endogenous contrast agent: the protons of blood water. The absence of injection and the possibility to employ this imaging technique to measure task-based activity or basal perfusion is very attractive. However, though the principle of this imaging method were introduced around the same time of [BOLD fMRI](#), the use of [ASL](#) remains much more restricted. The main explanation is probably the low signal-to-noise ratio of this technique and the initial technical difficulties. Nowadays, [ASL](#) has gained in maturity and is available as a commercial product on most clinical scanners. While several studies have shown the potential of [ASL](#) for clinical purposes, in particular in stroke or brain tumours, much still need to be done to provide standardized post-processing.

In order to study the cerebral dysfunctions associated with a pathology or to understand the typical brain function in healthy subjects, group analyses have demon-

strated a great efficiency. Nowadays, group analyses are the standard approach to uncover the, typical or pathological, brain function mechanisms. However, on the clinician viewpoint, the individual patient should be at the centre of the analysis. This is in part driven by the fact that diagnosis and treatment choices are necessarily performed at the patient and not at the group level. This is why there is an increasing interest towards patient-specific analyses that aim at drawing conclusions at the level of a single patient.

In this work, we first introduce a group analysis for the study of specific language impairment, a pathology that was very little studied on the functional viewpoint in the literature. We then focus our discussion to patient-specific analyses. We analyse the validity of the assumptions underlying standard statistical analyses in the context of Arterial Spin Labelling. We propose a new locally multivariate statistical method based on an *a contrario* approach and apply it to the detection of atypical patterns of perfusion in ASL and to activation detection in BOLD functional MRI.

This document is organized in three parts. The first part is dedicated to the presentation of the context and the state of the art. Our contributions are provided in the second and third parts. The former is a standard group analysis applied in an innovative context. The latter presents our methodological contributions to the patient-specific analysis of brain function.

### *Part I: Context*

**Chapter 1** presents the general context of our study. A definition of “brain function” is provided and the focus is made on the two aspects of brain function that are studied in this document: perfusion and task-evoked activity. An overview of the possible applications of each technique is presented along with their interest for clinical purposes. Competing, non Magnetic-Resonance imaging modalities are briefly exposed with their pros and cons.

**Chapter 2** develops the standard pre-processing pipeline employed both for Arterial Spin Labelling and BOLD functional MRI data. The massively univariate general linear model is presented as the standard tool for statistical analysis of neuroimaging data.

**Chapter 3** introduces patient-specific analyses. A brief description of the applications of such studies is provided, followed by the current state of the art.

### *Part II: Language impaired regions in specific language impairment: a group study using BOLD functional MRI*

**Chapter 4** presents the context of this group study on Specific Language Impairment (SLI) entitled the “Neuroimaging of Dysphasia” project. A clinical de-

scription of specific language impairment is provided along with some key-points regarding language processing in the brain.

**Chapter 5** proposes a panel of four language tasks for language mapping in children using **BOLD** functional **MRI**. A group of eighteen typically developing children is involved in this study. Special care is taken to insure the feasibility of each task for young and disordered populations.

**Chapter 6** employs the protocol previously proposed in order to outline atypical patterns of activation and lateralization in children diagnosed with Typical-Specific Language Impairment, a subtype of dysphasia affecting the core aspects of language. A group of twenty-one young patients is involved in this study and compared to typically developing children matched in gender and age.

*Part III: Patient-specific analysis of brain function*

**Chapter 7** presents a robust method based on M-estimators to estimate the cerebral blood flow in arterial spin labelling. This approach is compared to a z-score outlier removal technique previously proposed in the literature and to the sample mean commonly used in arterial spin labelling studies.

**Chapter 8** studies the heteroscedasticity of arterial spin labelling data and its impact on the detection of patient-specific perfusion abnormalities using the massively univariate general linear model.

**Chapter 9** introduces a locally multivariate procedure: the *a contrario* approach to detect patient-specific patterns of perfusion abnormalities. This approach is compared to the standard massively univariate general linear model.

**Chapter 10** expands the use of the *a contrario* approach to the detection of individual activation patterns in **BOLD** functional **MRI**.

**Conclusions**, including a summary of our contributions along with the future work, are discussed in a the last chapter.



**Part I**  
**Context**





## Chapter 1

# The study of brain function

---

In this chapter, we present the biological and experimental context of our study. First, we provide a definition of the term *brain function* in section 1. Then, we focus on the two aspects of brain function that are investigated in the context of this thesis: brain perfusion in section 2 and task-evoked activity in section 3. For each component, we provide an overview of the clinical and research applications, along with the imaging techniques that enable its measurement. We also emphasize the advantages and drawbacks of the two imaging techniques used in this work: Blood-Oxygen-Level-Dependent (BOLD) functional MRI and Arterial Spin Labelling (ASL).

## Contents

<b>1</b>	<b>Brain function</b>	<b>9</b>
1.1	Definition	9
1.2	An integrative view of brain function	10
<b>2</b>	<b>Brain perfusion</b>	<b>10</b>
2.1	Blood supply to the brain	10
2.2	Definition	12
2.3	Parameters	13
2.4	Applications	13
	2.4.1 In clinical practice	13
	2.4.2 Research challenges	15
2.5	Imaging	17
	2.5.1 Overview of imaging techniques for brain perfusion	17
	2.5.2 Arterial Spin Labelling	19
	2.5.3 Pros and Cons of ASL by comparison to other perfusion imaging techniques	25
<b>3</b>	<b>Task-evoked brain activity</b>	<b>26</b>
3.1	Definition and vocabulary	26
3.2	Neurovascular coupling	27
3.3	Applications	28
	3.3.1 In clinical practice: pre-surgical mapping	28
	3.3.2 Research challenges	28
3.4	Imaging	28
	3.4.1 Overview of imaging techniques of task-evoked activity	28
	3.4.2 Functional MRI	30
	3.4.3 Pros and Cons of ASL and BOLD fMRI	31
<b>4</b>	<b>Conclusion</b>	<b>31</b>



duality of brain function, opposing task-evoked activity to intrinsic brain function [Raichle 2010].

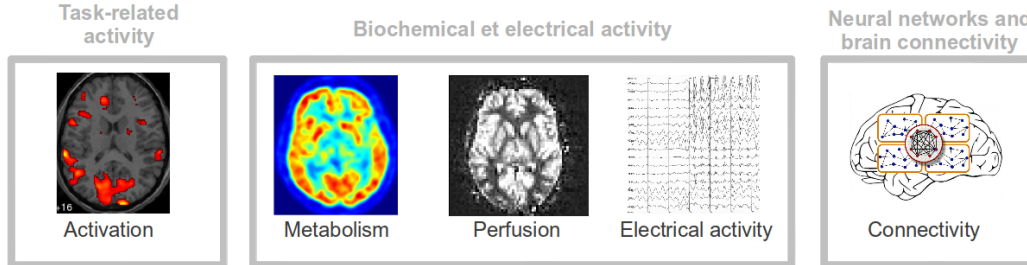


Figure 1.2: The three components of brain function according to [Vik 2005]. Illustration of connectivity reprinted from [Kaiser 2010]

## 1.2 An integrative view of brain function

The different components of brain function are clearly interdependent and nowadays, the study of brain function becomes more and more integrative, targeting several components at the same time. For example, the interest towards resting-state functional techniques has started to underpin the link between brain connectivity and spontaneous biochemical and electrical activity (e.g. [Melzer 2011, Worbe 2012]). In a recent study [Filippini 2012] investigated the relationship between task-evoked and resting brain function. The emergence of new functional sequences, for instance ASL in MRI, opens the field to a better understanding of the relationship between different aspects of brain function, such as basal perfusional activity and task-evoked activity [Tuunanen 2006].

In the context of this thesis, we focus on basal biochemical activity through the measure of brain perfusion and on task-evoked brain activity.

## 2 Brain perfusion

In this section, we review the biological process of brain perfusion. We start with a description of the channels that support blood delivery to the brain in 2.1. The term *brain perfusion* is defined in 2.2 and its parameters are described in 2.3. Then, some examples of how the perfusion information can be useful in research and clinical practice are provided in 2.4. Finally, the different perfusion imaging approaches are described in 2.5

### 2.1 Blood supply to the brain

As other organs of our body, the brain is connected to the heart through an arterial and a venous system. The inflowing blood brings oxygen, nutrients along with hormones while the outflow drains out waste products. Four arteries enter the skull: two carotid and two vertebral arteries as described in figure 1.3.a. These

arteries then ramify in the cerebral arterial circle (also termed circle of Willis) into six main vessels: the anterior, middle and posterior arteries in the left and in the right hemispheres (cf. figure 1.3.b and 1.3.c).

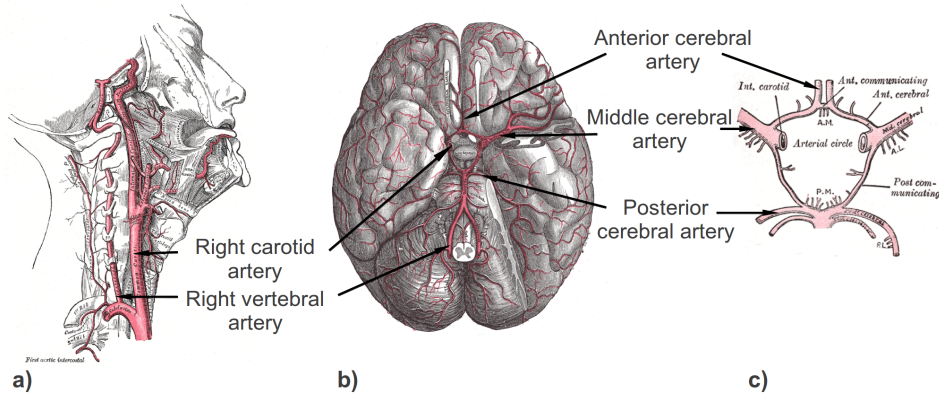


Figure 1.3: Main brain arteries, from [Gray 1918]. Lateral view displaying the carotid and vertebral arteries entering the skull on the right part of the brain (a). Ventral view of the artery ramifications at the basis of the brain (b). Close up on the circle of Willis: the carotid and vertebral arteries ramifies into the anterior, middle and posterior arteries (c).

In healthy subjects, the blood is supplied to each part of the brain by one of these arteries as described in fig 1.4. The anterior cerebral artery is responsible for the blood delivery to the medial part of the frontal and parietal lobes. The posterior cerebral artery supplies the medial posterior portion of the temporal and occipital lobes. The middle cerebral artery supplies the remaining parts of the frontal, temporal, parietal and occipital lobes.

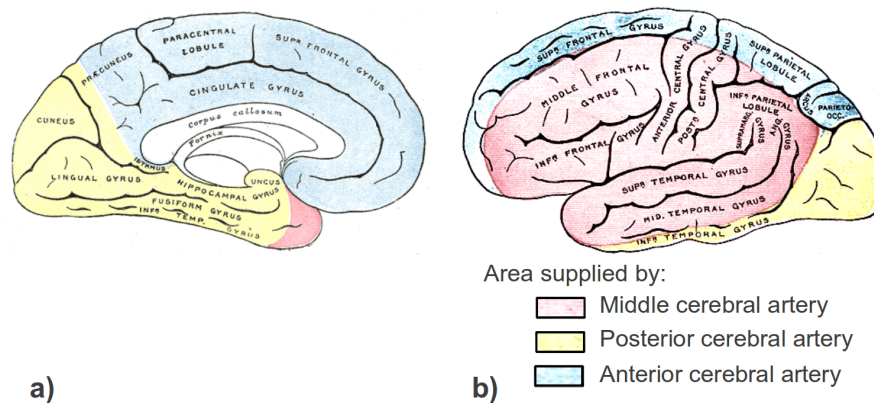


Figure 1.4: Vascular territories of the middle (pink), posterior (yellow) and anterior (blue) brain arteries, from [Gray 1918]. Medial (a) and lateral (b) views.

After exchanging with the brain tissues, the blood circulates into small and then large cerebral veins until it reaches the sinuses. A sinus is a particular type of vein

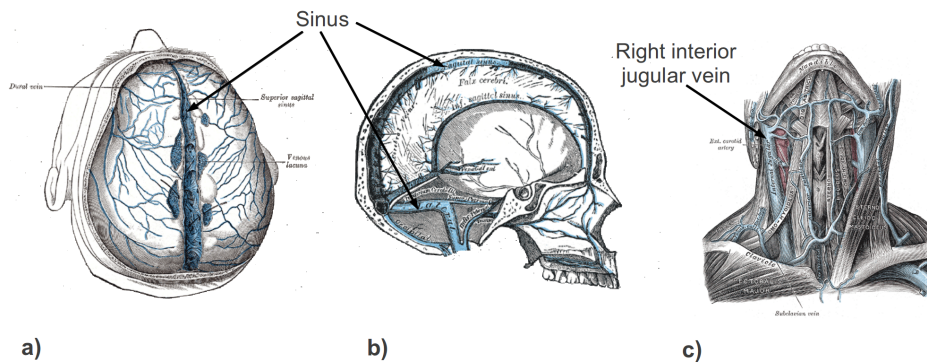


Figure 1.5: Brain veins, from [Gray 1918]. Dorsal (a) and medial (b) views of the sinuses. The interior jugular vein brings the blood out of the skull (c).

with a non-deformable wall made of dura mater. The sinuses are organized as a channel as displayed in fig 1.5.a and fig 1.5.b. Finally, the blood exits the skull through the left and right interior jugular veins fig 1.5.c.

## 2.2 Definition

Brain perfusion is the biological process that ensures the delivery of oxygen and nutrients to the cerebral tissues by means of microcirculation. In the vascular tree, the process of perfusion occurs in small vessels, called the capillaries, as outlined in figure 1.6.

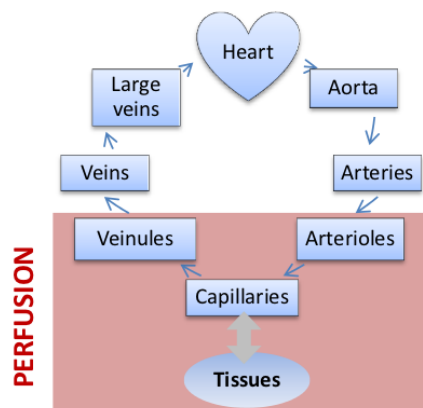


Figure 1.6: Vascular tree and perfusion. Image courtesy of Isabelle Corouge.

The brain is one of the most energy-demanding organs of our body, consuming about 20% of the oxygen, and 15% of total body glucose for a total weight of no more than 3% of the body [Willie 2011]. Simultaneously, there is not much room in the brain for energy storage making the perfusion process particularly important.

### 2.3 Parameters

The perfusion is characterised by a set of complementary parameters. The most common is the **Cerebral Blood Flow (CBF)**, which is a blood flow (i.e. a volume of blood delivered per time unit) per mass of tissue and is usually expressed in millilitre of blood delivered per hundred grams of tissue per minute ( $\text{mL}\cdot 100\text{g}\cdot^{-1}\cdot\text{min}^{-1}$ ). The **Cerebral Blood Volume (CBV)** is a volume of blood per mass of tissue, expressed in millilitre of blood per hundred grams of tissue ( $\text{mL}\cdot 100\text{g}\cdot^{-1}$ ). Finally, various transit times through the vasculature expressed in seconds are also of interest. Depending on the imaging technique, different transit times can be approached. The **Mean Transit Time (MTT)** is defined as the average time spent through the vasculature and calculated by as  $MTT = CBV/CBF$ . The **Bolus Arrival Time (BAT)** represents the time for a bolus of tracer to reach the voxel of interest and is applicable for bolus-based techniques only. The **Time To Peak (TTP)** measures the time necessary to reach a maximum of contrast change in first-pass bolus imaging. Figure 1.7 gives an example of parameter maps obtained with **Dynamic Susceptibility weighted Contrast imaging (DSC)** (more details on this **Magnetic Resonance (MR)** technique are provided in 2.5.1). A clear coupling is visible between the CBF and CBV maps.

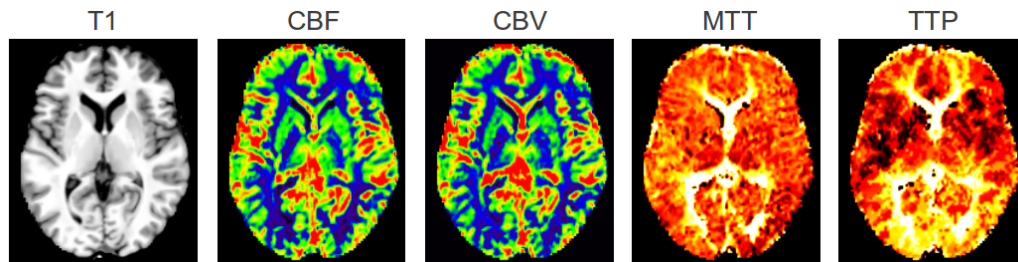


Figure 1.7: Perfusion parameters estimated with DSC in a patient diagnosed with a brain tumour (not visible on the slice displayed): T1w, Cerebral Blood Flow (CBF), Cerebral Blood Volume (CBV), Mean Transit Time (MTT) and Time To Peak (TTP) maps.

## 2.4 Applications

### 2.4.1 In clinical practice

The level of brain perfusion at rest, referred as *basal* perfusion, can be disrupted in the presence of vascular diseases or brain tumours. In this section we present some clinical contexts in which a measure of perfusion is a useful indicator to help the clinician in the diagnosis or monitoring of the treatment efficiency.

#### Diagnosis



**Identifying malignant tumours** In the diagnosis of brain tumours, the level of perfusion is an indicator of the malignancy of the pathology. Malignant high-grade gliomas induce the creation of new brain vessels around the tumour, a phenomenon, referred as neoangiogenesis, that is associated with the presence of hyper-perfusions. In clinics, the hotspot technique [Noguchi 2008] in which the level of *CBV* in the most active part of the tumour is compared to the contralateral perfusion level is common practice. This analysis is usually performed by a clinician who manually locates regions of interest in the tumour and in the contralateral hemisphere. A cut-off ratio of 1.75 has been proposed at 1.5 T [Law 2003] to differentiate benign low-grade from malignant high-grade gliomas. Figure 1.8 illustrates how the complementarity of T1 weighted Gadolinium injected (T1w-Gd) and perfusion imaging can help the clinician differentiating between glioblastomas, lymphomas and low-grade gliomas.

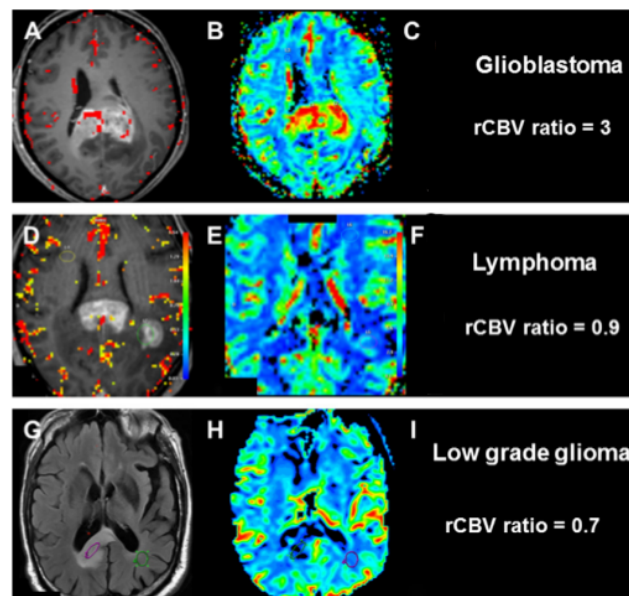


Figure 1.8: Differentiation between 3 types of brain tumours: glioblastoma, lymphoma and low-grade glioma using the complementarity of T1w-Gd and perfusion imaging. T1w-Gd (A,D) or T2w FLAIR (G) anatomical map (first column) and DSC CBV map (second column). The malignant glioblastoma is associated with a contrast enhancement on the T1w-Gd map and a hyper-perfusion on the CBV map. The lymphoma is characterized by a contrast enhancement on the T1w-Gd but a low CBV ratio (tumour/normal). The low-grade glioma has also a low CBV ratio and has no enhancement on the T1w-Gd (T2w FLAIR displayed). Reprinted from [Ferré 2012], Copyright (2013), with permission from Elsevier.

**Indication for thrombolysis in stroke** As described in figure 1.9, the early stage of an ischaemic stroke is characterised by the presence of an ischaemic core in which the brain tissues are already dead and an area of penumbra that is suffering

but can recover if reperfused quickly. The main objective of the early stage treatment is therefore the recanalisation of the penumbra before its necrosis. This can be performed for instance by intra-venous thrombolysis, in which the injection of a pharmacological agent is performed to the aim of destroying the thrombus. Due to the risk of bleeding, this kind of treatment is allowed only within the first hours following the onset of the stroke and when the mismatch between the penumbra and the ischaemic core is large (indicating that most of the tissue is likely to recover). Diffusion imaging is considered as a predictive parameter to measure the extent of the ischaemic core [González 1999] while perfusion imaging is the method of choice to image the penumbra [Huck 2012, Bokkers 2011, Ferré 2012, Wheeler 2013]. The combination of both techniques is advised to check the extent of the mismatch before thrombolysis. This is particularly important when considering a patient more than three hours after the symptom onset or when the onset is unknown.

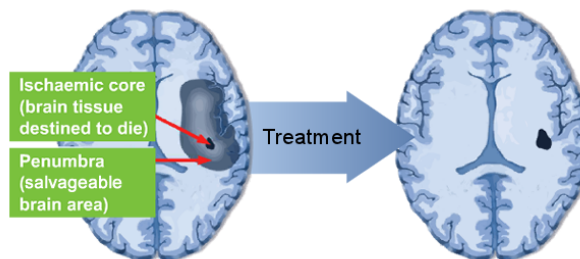


Figure 1.9: Ischaemic core and penumbra in stroke from the [Stroke Forum website](#)<sup>3</sup>. After effective treatment the tissue that was located in the ischaemic penumbra has recovered.

**Post-treatment monitoring** After a radiotherapy, a monitoring of the patient status is performed to ensure that the tumorous process has been stopped. However, on conventional T1w-Gd imaging, a contrast enhancement is either the consequence of the lesions caused by the radiotherapy (radio-necrosis) or an indicator of tumour recurrence. In this context, the perfusion level can help differentiating a pseudo-progression (post-radiation necrosis, characterised by a hypo-perfusion) from a true progression (tumour recurrence, characterised by a hyper-perfusion indicating neoangiogenesis) as described in fig 1.10.

#### 2.4.2 Research challenges

**Biomarkers of Alzheimer’s disease** Reduced perfusion in the temporal lobe is well-documented in Alzheimer’s disease [Du 2006, Alsop 2010, Wolk 2012] and in other types of dementias (e.g. [Corouge 2012] in semantic dementia). Figure 1.11 gives an example of reduced metabolism and CBF in the temporal lobe of a patient

<sup>3</sup>[www.strokeforum.com/stroke-background/pathophysiology.html](http://www.strokeforum.com/stroke-background/pathophysiology.html)

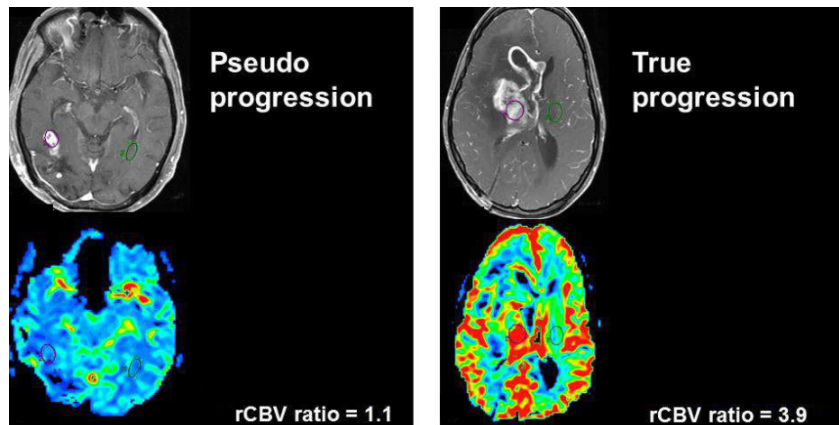


Figure 1.10: Differentiation between post-radiation necrosis (pseudo-progression) and tumour recurrence (true progression) using perfusion imaging. T1w-Gd anatomical map (first line) and DSC CBV map (second line). In pseudo-progression (first column) the presence of radio-necrotic lesions is characterised by a low CBV ratio. In true progression (second column) neoangiogenesis is characterised by hyper-perfusions (high CBV ratio). Reprinted from [Ferré 2012], Copyright (2013), with permission from Elsevier.

diagnosed with semantic dementia [Corouge 2012]. Current research focuses on extracting early biomarkers from the patterns of perfusion and their longitudinal progression. For example, a current issue is the identification of early biomarkers to distinguish patients that will or will not convert to Alzheimer's disease [Klöppel 2012].

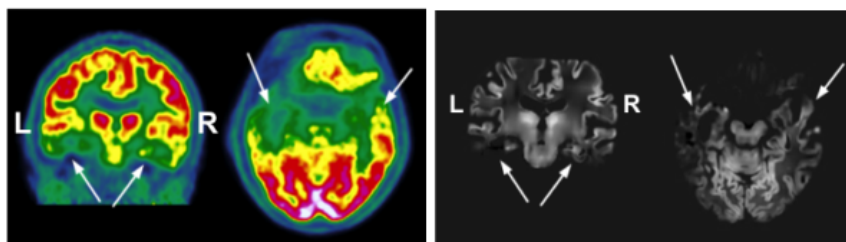


Figure 1.11: Reduced glucose metabolism and perfusion (CBF) of the temporal lobes in a patient suffering from semantic dementia as imaged by FDG-PET (Left) and ASL (Right). Reprinted from [Corouge 2012], with permission.

**Identification of active foci in drug-resistant epilepsy** In drug-resistant epilepsy, where surgical resection is of consideration, locating the source of the seizure is of great interest. In some patients, where the identified anatomical anomaly is large, perfusion could help outlining the core affected area that might be a candidate for surgical resection. Figure 1.12 gives an example of hypoperfusion resulting from a dysplasia in an epileptic patient [Petr 2013].

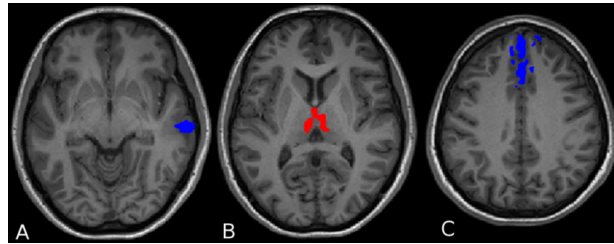


Figure 1.12: Hypoperfusions (in blue) and hyperperfusions (in red) overlaid on T1w-Gd in a patient suffering from drug-resistant epilepsy. A hypoperfusion is co-located with the dysplasia observed in the right temporal lobe of this patient. Reprinted from [Petr 2013], Copyright (2013), with permission from Wiley.

### A better understanding of physiological mechanisms in healthy subjects

Imaging normal brain perfusion in conjunction with other biological parameters can help understanding brain physiological mechanisms. The study of control subjects is also a way to better understand and define pathologies.

## 2.5 Imaging

### 2.5.1 Overview of imaging techniques for brain perfusion

In this section, we briefly describe the main imaging methods allowing the measure of brain perfusion: Positron Emission Tomography (PET), Single Photon Emission Computed Tomography (SPECT), Perfusion Computed Tomography (PCT), Xenon-Computed Tomography (CT), DSC and ASL based on a recent review [Wintermark 2005].

**Nuclear medicine** Two main approaches are currently in use to measure perfusion in the field of nuclear medicine: PET and SPECT

**Positron Emission Tomography** In PET, a radioactive tracer is administered to the subject by injection or inhalation. After waiting a short period, in order to allow for the distribution of the tracer, the radioactive activity of the tissues is recorded. Different imaging tracers can be used in order to image a large range of biological process including metabolism and haemodynamics. The measure of the CBF involves either the injection of  $\text{H}_2^{15}\text{O}$  or the inhalation of  $\text{C}^{15}\text{O}_2$ . The haemodynamic parameters that can be estimated include the CBF, CBV, the regional Oxygen Extraction Fraction (rOEF) and regional cerebral oxygen metabolism ( $\text{rCMRO}_2$ ). But 90% of PET examinations use an FDG molecule as a tracer and therefore measure the metabolism of glucose. The parameter of interest is  $\text{CMRO}_{\text{glu}}$  the consumption rate of glucose. While the metabolism of glucose is not strictly speaking a measure of the haemodynamics, it has been shown to be tightly coupled with regional CBF [Fox 1986]. As the radioactive molecules have to be produced by a cyclotron, the use of PET-scans is limited to specialised centres. PET imaging is

non-invasive but makes use of a radioactive tracer which precludes its use in young or sensitive populations in view of the exposure to ionizing radiation. PET imaging is furthermore impracticable in emergency settings.

***Single Photon Emission Computed Tomography*** Similarly to PET imaging, SPECT is able to image different biological phenomenon depending on the molecule selected as a tracer. Imaging brain haemodynamics in SPECT involves the injection of  $^{133}\text{Xenon}$  (historical tracer) or retention-tracer. The measurement of CBF is relative. A potential advantage of SPECT is the fact that the tracer can be injected before the time of imaging allowing for instance to image seizure for epileptic patients. The resolution of SPECT is lower than PET, however due to its lower cost, SPECT imaging is still quite widespread.

**Computed Tomography** Two main approaches are currently in use to measure perfusion with CT: Xenon-CT and PCT. CT scanners are widely available and therefore usable in the emergency setting. However, CT scans involve the use of ionizing radiation (and intravenous injection of iodinated contrast media for PCT), which makes them not suitable for children or sensitive populations.

***Xenon Computed Tomography*** Measuring perfusion with Xenon-CT involves the inhalation of a mixture of non-radioactive Xenon gas and oxygen. It allows to quantitatively measure the level of CBF. This technique however requires a good cooperation of the patient and specialized equipment [Wintermark 2001].

***Perfusion Computed Tomography*** PCT relies on the monitoring of the first pass of a bolus of iodinated contrast agent. It allows the measurement of several perfusion parameters, both qualitatively and quantitatively: CBF, CBV, MTT [Hoeffner 2004] and is less constraining than a Xenon-CT [Wintermark 2001]. This technique is widely used in clinical centres, in particular for stroke imaging when MRI is not available.

**Magnetic Resonance Imaging** The historical approach to measure perfusion with an MR-scanner is DSC. The increasing robustness and availability of ASL sequences on clinical scanners opens the field for a more widespread use of this competing approach.

***Dynamic Susceptibility weighted Contrast imaging*** Similarly to PCT, DSC relies on the imaging of the first pass of a bolus of contrast agent. Gadolinium chelates are usually used as a contrast agent and their paramagnetic effect on T2 or T2\* imaging is scanned. Gadolinium chelates are non-ionising molecules and allergies to this product are rare. However, the injection requires a good venous access that might not be available in children. Also, the need of injection of a contrast

agent reduces its applicability in pregnant women or in healthy subjects. Quantification is possible with DSC but more complex than with PCT. Also leakage of the blood brain barrier (as often reported in brain tumours) can further complicate the quantification. Perfusion parameters that can be estimated are the TTP, MTT, CBV and CBF.

**Arterial Spin Labelling** ASL is a perfusion imaging technique that uses an endogenous contrast agent: magnetically labelled protons of blood water. The perfusion parameters estimated with this approach vary depending on the imaging sequence but include: CBF and BAT. This method is currently mainly confined to research but its use in the clinical settings is growing rapidly [Detre 2012, Alsop 2012]. This imaging approach will be described in further details in section 2.5.2.

### 2.5.2 Arterial Spin Labelling

**Basics** ASL is an MR technique introduced in the early 1990's [Detre 1992]. This sequence allows the measurement of micro-vascular CBF by means of an endogenous contrast agent: the protons of blood water labelled using radio-frequency pulses. A schematic view of the sequence is presented in figure 1.13. In short, the signal is obtained by subtraction of a control and a labelled image. In the labelled image,

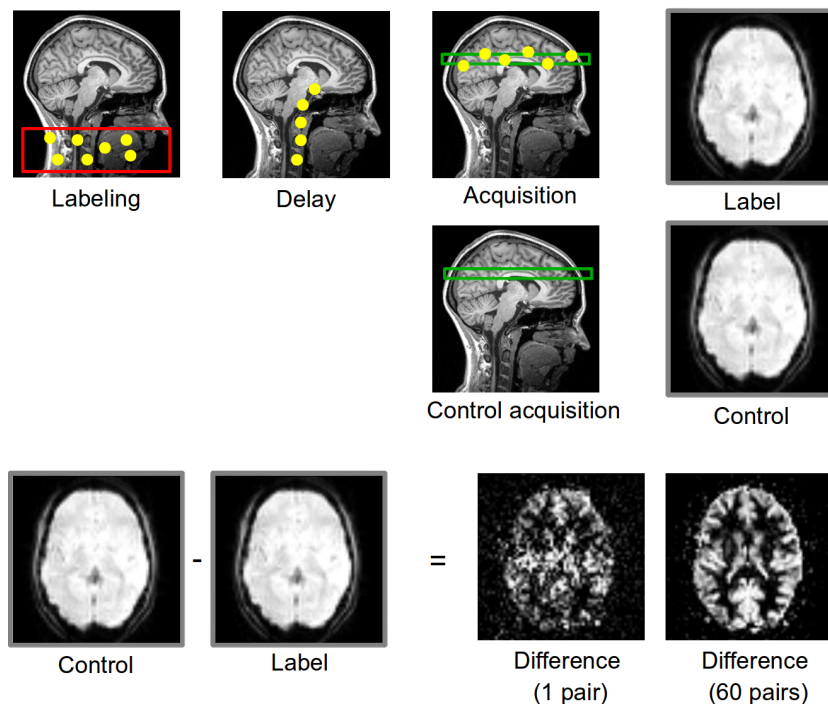


Figure 1.13: Schematic view of an ASL acquisition.

the blood magnetization is inverted, while, in the control, the magnetization of the inflowing blood is not modified. The difference between the two images is

proportional to the microvascular CBF and is called a perfusion-weighted image and usually denoted  $\Delta M$ . As only about 1% of tissue water is replaced by perfusion every second, the Signal to Noise Ratio (SNR) of this imaging technique is inherently low. To increase the SNR of the measurement, the acquisition is usually repeated several times and the perfusion signal averaged across the repetitions.

**Imaging sequence** There is a large number of implementation of ASL imaging sequences but they all rely on three building blocks: the the preparation of tissue magnetization, the labelling scheme and the readout, as displayed in fig 1.14.

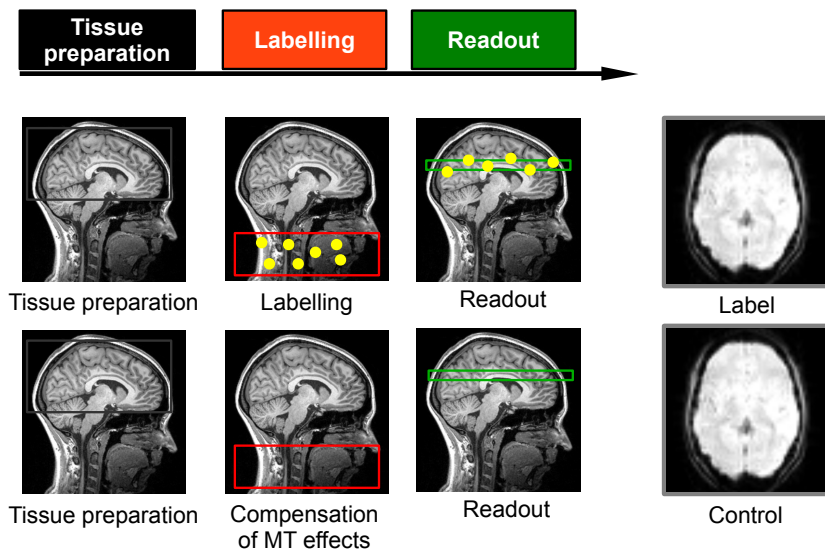


Figure 1.14: Building blocks of an ASL sequence: tissue preparation, labelling and readout.

The tissue preparation focuses on cancelling the magnetization of the static tissues in order to proportionally increase the relative perfusion signal. In practice the tissue preparation can be performed before or after the labelling. Background suppression techniques are sometimes used in order to maintain the tissue magnetization close to zero at the moment of the readout (as otherwise the signal of the tissue recovers).

The labelling aims at inverting (or saturating in the very first ASL sequences nowadays outdated) the magnetization of the inflowing blood in the labelled image while preserving the blood magnetization in the control image. Labelling the blood in the labelled scan may inevitably lead to magnetization transfer effects in the imaged slice. In order for these unwanted effects to be cancelled out in the subtraction of the control and labelled scan, the same magnetization transfer effects must be reproduced in the control scan.

Finally, the readout measures the longitudinal magnetization present in the slices (or volume) of interest.

**Pulsed and Continuous ASL** There are two main groups of labelling approaches that define the two major ASL techniques: **Continuous Arterial Spin Labelling (CASL)** and **Pulsed Arterial Spin Labelling (PASL)**. In **CASL**, the blood is continuously inverted as it passes through a plane while in **PASL** a volume of blood is inverted as described in 1.15. Theoretically, **CASL** is superior to **PASL** in terms of **SNR**. However in practice, due to the historical need of specialized hardware and the high level of energy deposited in the tissues for **CASL** implementation, the **PASL** technique remains largely used.

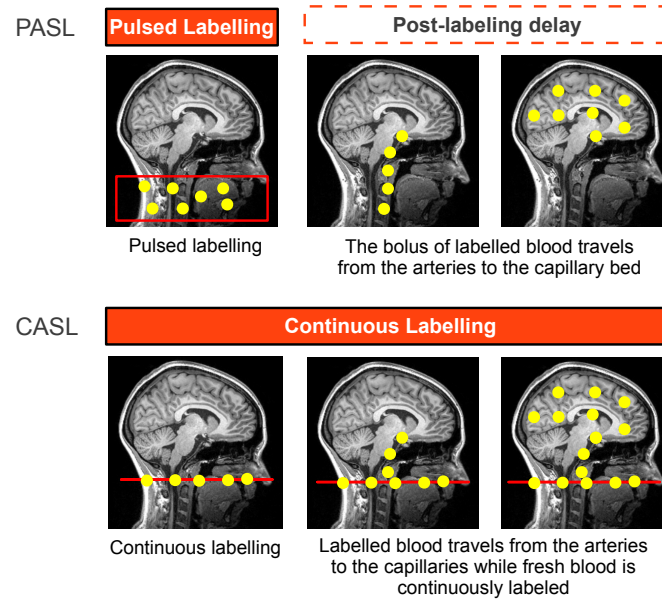


Figure 1.15: Pulsed and Continuous ASL.

**Continuous ASL** Since the introduction of the Pseudo-continuous sequences [Dai 2008] allowing continuous labelling using pulsed radio-frequency pulses, **CASL** has started to be more widespread. **Pseudo-Continuous Arterial Spin Labelling (PCASL)** was designated as one of the preferred ASL sequences during the white paper sessions at the ISMRM workshop on perfusion 2012 held in Amsterdam. This sequence is however not available as a commercial product on all the **MR** scanners.

**Pulsed ASL** Different approaches have been proposed in order to achieve the labelling (and control for magnetization transfer effects) in **PASL**. The three most widespread approaches are: **Flow-sensitive Alternating Inversion Recovery (FAIR)** [Kim 1995], **EPI-Signal Targeting with Alternating Radio-frequency pulses (EPSTAR)** [Edelman 1994] and **Proximal Inversion with a Control for Off-Resonance Effects (PICORE)** [Wong 1997]. For more information on these labelling schemes, the interested reader is referred to the original papers.

In the context of this thesis, we used the commercially available sequence of



our constructor, namely a **PICORE QUIPSS II with Thin-slice Seriodic Saturation (Q2TIPS) PASL** sequence with 2D GRE-EPI readout.

**Quantification of CBF from the ASL signal** In a typical **ASL** study, a perfusion-weighted image is obtained by subtraction of the control and label scan and averaging across the repetitions. Figure 1.16 gives a quick overview of the difference signal  $\Delta M$  observed in a voxel of interest through time. There are basically three phases in which the signal is respectively zero, increasing and decreasing back to zero. At the earliest post-labelling times, none of the bolus has reached the tissue leading to a signal of zero. After a certain period of time, referred as the **BAT** and denoted  $\delta_t$ , the bolus starts entering the voxel of interest and consequently  $\Delta M$  increases. Finally, once all the bolus has reached the voxel of interest,  $\Delta M$  starts decreasing mainly as a consequence of  $T_1$  relaxation.

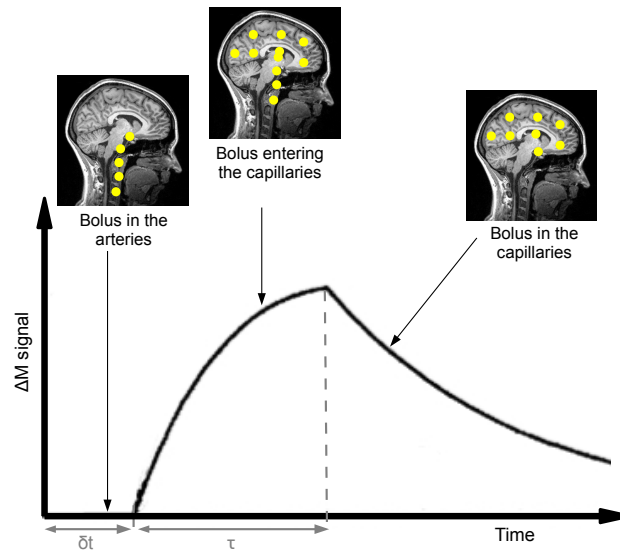


Figure 1.16: Schematic view of the difference signal  $\Delta M$  observed in an **ASL** experiment.

More formally, the signal  $\Delta M$  can be expressed as a convolution between an input function  $c(t)$  and a decaying function  $m(t)$  [Buxton 1998]:

$$\Delta M(t) = c(t) \otimes m(t). \quad (1.1)$$

Assuming that the venous outflow is null (due to the fact that the human  $T_1$  decay is too fast with respect to vascular transit times in humans) and that the bolus has not had the chance to start exchanging with the tissue at the time of

readout we have:

$$c(t) = \begin{cases} 0 & 0 < t < \delta_t \\ \alpha 2f M_{0b} \exp\left(\frac{-t}{T_{1b}}\right) & \delta_t \leq t \leq \delta_t + \tau \text{ for PASL} \\ \alpha 2f M_{0b} \exp\left(\frac{-\delta_t}{T_{1b}}\right) & \delta_t \leq t \leq \delta_t + \tau \text{ for CASL} \\ 0 & t > \delta_t + \tau \end{cases} \quad (1.2)$$

$$m(t) = \exp\left(\frac{-t}{T_{1b}}\right) \quad (1.3)$$

where  $M_{0b}$  is the longitudinal magnetization of blood at equilibrium,  $\alpha$  the labelling efficiency and  $f$  the CBF (parameter of interest),  $T_{1b}$  the  $T_1$  of blood and  $\tau$  the temporal width of the bolus. Figure 1.17 provides a graphical example of  $c(t)$  and  $m(t)$  functions for PASL and CASL.

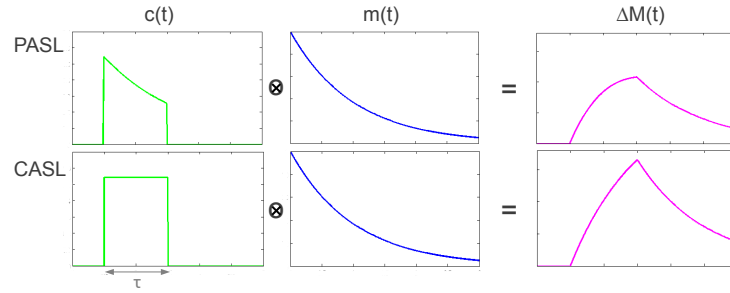


Figure 1.17: Buxton's model for  $\Delta M$ , examples of  $c(t)$  and  $m(t)$  functions for PASL and CASL.

Therefore  $\Delta M_{\text{PASL}}$  and  $\Delta M_{\text{CASL}}$ , the difference signal in a PASL and a CASL experiment, are respectively given by:

$$\Delta M_{\text{PASL}}(t) = \begin{cases} 0 & 0 < t < \delta_t \\ \alpha 2M_{0b}f(t - \delta_t) \exp\left(\frac{-t}{T_{1b}}\right) & \delta_t \leq t \leq \delta_t + \tau \\ \alpha 2M_{0b}f\tau \exp\left(\frac{-t}{T_{1b}}\right) & t > \delta_t + \tau \end{cases} \quad (1.4)$$

and

$$\Delta M_{\text{CASL}}(t) = \begin{cases} 0 & 0 < t < \delta_t \\ \alpha 2M_{0b}f \exp\left(\frac{-\delta_t}{T_{1b}}\right) \left(1 - \exp\left(\frac{\delta_t - t}{T_{1b}}\right)\right) & \delta_t \leq t \leq \delta_t + \tau \\ \alpha 2M_{0b}f \exp\left(\frac{-\tau}{T_{1b}}\right) \left(1 - \exp\left(\frac{-\tau}{T_{1b}}\right)\right) & t > \delta_t + \tau \end{cases} \quad (1.5)$$

**Single-TI quantification** Most of ASL acquisitions are based on the measure of  $\Delta M$  at a single time point, called the **Inversion time (TI)**. If we consider that the  $T_1$  of blood,  $T_{1b}$ , and the labelling efficiency,  $\alpha$ , are known and that the longitudinal magnetization of blood,  $M_{0b}$ , can be estimated from the acquisition, there are three remaining unknown parameters: the **BAT**,  $\delta_t$ , the bolus width  $\tau$  and the **CBF**  $f$ .

In CASL, the temporal width of the bolus corresponds to the duration of the labelling and is a known parameter of the sequence. In the very first applications of PASL, while the spatial width of the labelled region was known, the temporal width was dependent upon the velocity of the flow and the geometry of the vessels. The bolus width was therefore an unknown subject-dependent parameter. To overcome this limitation, a modification of the PASL sequence, known as Quantitative Imaging of Perfusion using a Single Subtraction (QUIPSS II) was proposed [Wong 1998]. The QUIPSS II version of the PASL sequence includes additional saturation pulses in order to “cut” the tail of the labelling and get a sharp shaped bolus. The Q2TIPS variant was then introduced achieving a better saturation with the use of a train of saturation pulses [Luh 1999].

Given the temporal width of the bolus,  $\tau$ , there are two remaining unknown parameters: the CBF,  $f$ , and the BAT,  $\delta_t$ . By looking back at equations (1.4) and (1.5), one can notice that the dependence on the arrival time only affects the ascending part of the  $\Delta M$  curve. The quantification of CBF with a single time point is then achievable as soon as the readout is performed in the decreasing part of the  $\Delta M$  curve, i.e. after all the bolus has entered the voxel of interest. This can be written as:  $TI > \delta_t + \tau$ , where  $TI$  is the readout time. Since the arrival time  $\delta_t$  is unknown,  $TI$  is fixed assuming a reasonable upper threshold for  $\delta_t$ . The choice of the inversion time is therefore a trade-off between the possible range of BATs (the later the inversion time the wider the BATs range) and the SNR (the later the inversion time the smaller the SNR due to  $T_1$  relaxation).

Quantification of CBF,  $f$ , using a single readout time  $TI$  is therefore based on the assumption  $TI > \delta_t + \tau$ , and leads to:

$$f = \frac{\Delta M_{\text{PASL}}}{\alpha 2M_{0b}\tau \exp\left(\frac{-TI}{T_{1b}}\right)}, \quad (1.6)$$

for PASL, and:

$$f = \frac{\Delta M_{\text{CASL}}}{\alpha 2M_{0b} \exp\left(\frac{\tau-TI}{T_{1b}}\right) \left(1 - \exp\left(\frac{-\tau}{T_{1b}}\right)\right)}, \quad (1.7)$$

for CASL.

**Multi-TI quantification** If multiple time points are available then the CBF and the BAT can both be estimated. While the BAT is a parameter of interest, in particular in pathological populations, the availability of multi-TI ASL sequences is still restricted to a few research centres.

**Advanced modelling** The standard quantification approach presented here-above (eq. (1.4) and eq. (1.5)) assumes that the bolus is arriving as a uniform plug flow (no bolus dispersion), that the blood has not exchanged with the tissue at the time of readout and that the labelled water is in a single compartment (in the vasculature). Several advanced approaches have been proposed in which one or several

of these assumptions are relaxed. For instance a model for bolus dispersion was developed in [Hrabe 2004]. Two-compartment models, including micro-vascular and tissue compartments [Parkes 2000], or macro-vascular and micro-vascular compartments [Chappell 2010] have been proposed. More recently, Bayesian and model-free approaches have gained more attention [Chappell 2012]. The possibility to get simultaneous estimates of multiple parameters is quite appealing, however, these approaches are mainly confined to multi-TI ASL sequences. Most ASL studies are thus still relying on the standard quantification approach.

### 2.5.3 Pros and Cons of ASL by comparison to other perfusion imaging techniques

#### Advantages of ASL

***Non-invasiveness*** The absence of radiation and the use of an endogenous tracer, which removes the need of injecting a contrast agent, are clear advantages of ASL over standard perfusion imaging techniques. Contrary to standard perfusion imaging, ASL is completely non-invasive and does not require the injection of an exogenous contrast agent. This non-invasiveness allows the inclusion of sensitive subjects in ASL studies like healthy controls, pregnant women or children. The repeatability is also improved.

***White temporal noise*** An important aspect of ASL time series for post-processing purposes is the fact that they display a noise that is almost white [Agui-re 2002, Wang 2003a]. Indeed, the subtraction process that lead to ASL perfusion-weighted image removes the low frequency noise that is usually observed in other fMRI images [Friston 1994b, Zarahn 1997]. The practical consequences in terms of statistical analysis will be developed in chapter 2.

***Absolute quantification*** Arterial Spin Labelling allows for the measurement of CBF in quantitative units. The quantification is nevertheless dependent upon the precision of the fixed parameters used for the quantification [Parkes 2002] and to the model assumptions (e.g. in single-TI sequences, the readout time should be greater than the time necessary for the full bolus to reach the tissue of interest).

***Multi-modality*** An advantage that is common to all functional MRI techniques, is the possibility to acquire in the same exam session, both anatomical and functional information.

**Drawbacks** While ASL presents important advantages over other perfusion imaging techniques and its first application in human beings date of the early 1990's, its use in clinical practice is still in an early stage.

**Low signal-to-noise ratio and resolution** In ASL, the signal is obtained by subtraction and the perfusion signal of interest is about 1% of the original signal. This makes ASL an inherently low SNR approach. In this work, we usually worked with an in-plane resolution of 3 mm x 3 mm, 7 mm-thick slices with a 0.7 mm gap.

**Impractical quantification for very long arterial transit times** In some pathologies, such as stroke, the time needed for the blood to travel from the labelling region to the imaging region might be much longer than the one typically observed in controls. Since the labelling of the bolus is decaying with the  $T_1$  of blood (approximately equal to 1.5 s at 3 T [Wang 2011]), it is very difficult to measure signal after one  $T_1$  [Kim 2006]. If, due to the pathology the arterial transit time is much longer than the  $T_1$  of arterial blood then quantification using standard ASL approaches is impractical. One future alternative might be the use of velocity selective ASL sequences [Wong 2006].

### 3 Task-evoked brain activity

In this section, we review task-evoked activity. We start with a definition of *task-evoked activity* in 3.1. Then, we describe the neurovascular coupling in 3.2, a biological phenomenon that is at the basis of the main imaging techniques for task-evoked activity. Some examples of how information on task-evoked activity can be useful in research and clinical practice are provided in 3.3. Finally, the different imaging approaches are briefly discussed in 3.4.

#### 3.1 Definition and vocabulary

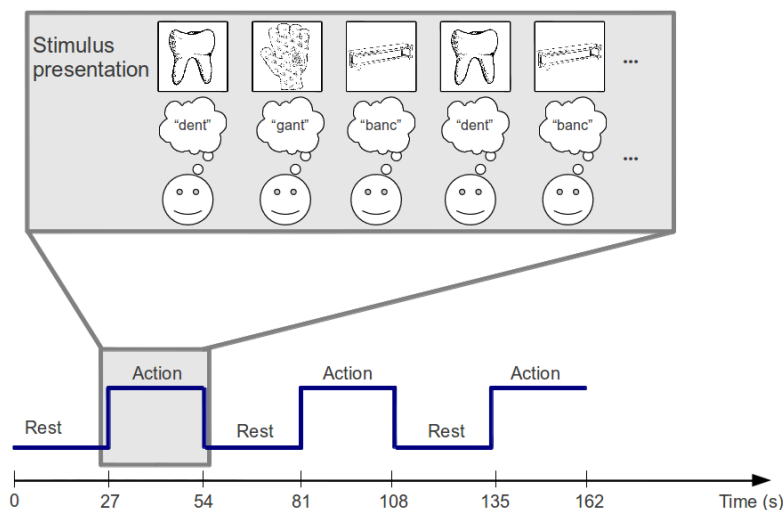
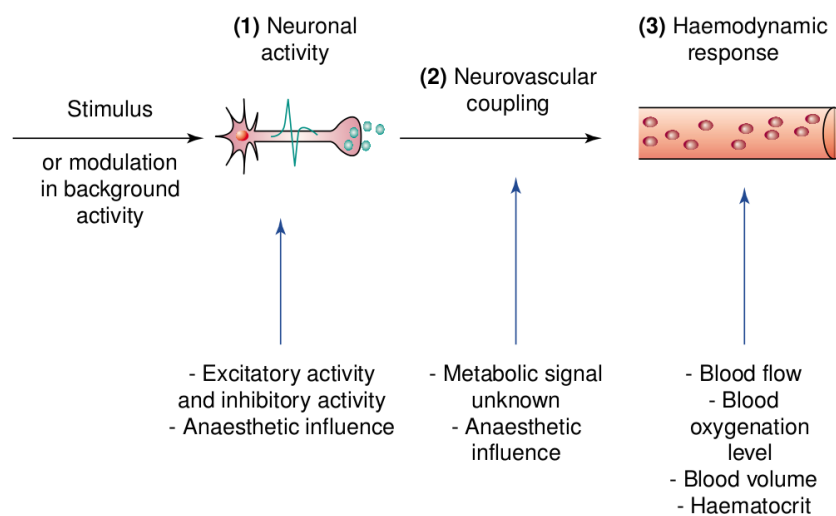


Figure 1.18: PhonoDif paradigm, as described in [de Guibert 2010]. This block paradigm is made of 3 alternating periods of rest and action.

The study of task-evoked activity aims at discovering the spatio-temporal patterns of brain activation changes associated with a particular task or stimulus. The timing of the task and rest periods along with the type of stimuli presented to the subject is referred as the paradigm. An example of paradigm is provided in figure 1.18.

### 3.2 Neurovascular coupling

The neurovascular coupling is a biological process that induces a tight relationship between the electrical activity observed at the neuron level, the local metabolism and the haemodynamics in the surrounding micro-vasculature. An increase in activity in a group of neurons will induce an increased local metabolic activity and in turn increase the local cerebral blood flow. The parameters of task-evoked activation imaging will thus be different depending on the measured biological process.



*TRENDS in Neurosciences*

Figure 1.19: Neurovascular coupling, Reprinted from [Arthurs 2002], Copyright (2013), with permission from Elsevier.

As early as the end of 1800's, Mosso observed blood-flow pulsations related to cognitive activity in the frontal lobe of a patient presenting an injured skull [Mosso 1881]. A decade later, Sherrington and Roy validated these results in dogs [Roy 1890]. On physiological ground, however, this neurovascular coupling is still not completely understood [Obata 2004, Raichle 2010].

### 3.3 Applications

#### 3.3.1 In clinical practice: pre-surgical mapping

The main application of task-evoked functional imaging in clinical practice is the identification of eloquent brain areas during the surgical planning. In view of the surgery, the aim is to identify the brain areas to be avoided during the surgery in order to preserve the patient capacities. Eloquent areas usually correspond to a set of sensorimotor and language regions. Standardized paradigms are used for these examinations including tasks such as picture-naming or word generation from letters to outline the language areas or hand motion for motor areas [Hirsch 2006].

#### 3.3.2 Research challenges

**Studying brain dysfunction** Probably the most widespread application of task-evoked activity imaging in clinical research is the comparison between groups of subjects. The aim is then to identify the brain regions that present different patterns of activation in one group (for instance patients) by comparison to a reference group (for instance healthy subjects). The atypical patterns of activation in pathological populations can indicate the foci of brain dysfunction and help understanding the compensation mechanisms.

**Understanding typical brain function** Imaging task-evoked activity in healthy subjects is of interest for several research topics. First, some studies aim at identifying the brain regions associated with a particular task. Other studies, based on previous results, focus on the identification of the cognitive processes that are involved in a particular task (for example the IOWA gambling task in [Li 2009]). These studies, are based on the assumption that we know that the involvement of a set of brain regions is specific of a particular cognitive process [Aguirre 2006]. Finally identifying the physiology of neurovascular coupling is still a research topic. Imaging using several functional techniques can for example help understanding the physiological mechanism underlying the neurovascular coupling.

### 3.4 Imaging

#### 3.4.1 Overview of imaging techniques of task-evoked activity

While the neural activity observed in reaction to a stimulus is an electrical and biochemical process, in neuroimaging, brain activity is usually imaged through the measurement of a haemodynamic or a metabolic index. This is made possible thanks to the neurovascular coupling of neural activity with metabolism and local haemodynamics as described previously (cf. fig. 1.19). Figure 1.20 presents the three levels of measurement currently in use and the associated imaging modalities.

#### Measurement of neuronal activity

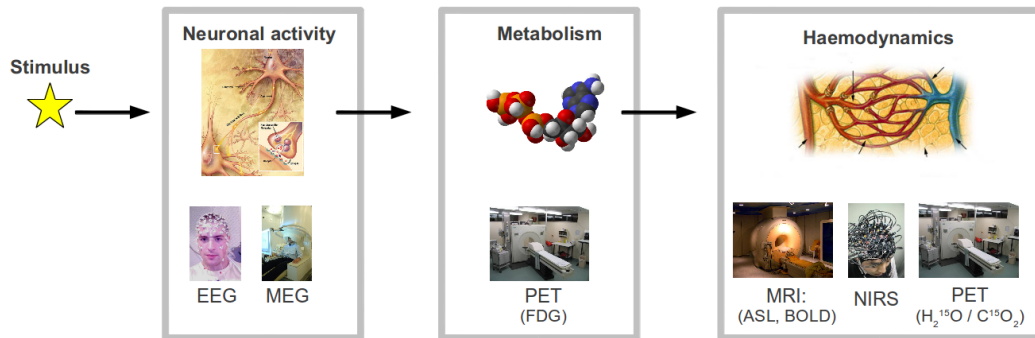


Figure 1.20: Three levels of measurements for brain activity and associated imaging techniques. Studying task-evoked activity can be performed with a measurement of the neuronal electrical activity (with EEG or MEG) or with a measurement of change in metabolism (with PET imaging) or haemodynamics (with MRI or PET).

**Electroencephalography** EEG provides a direct measurement of the electrical activity recorded thanks to a set of electrodes located on the scalp. The recorded signal presents a very high temporal resolution. However this comes at the cost of low spatial resolution.

**Magnetoencephalography** Similarly to EEG, MEG records the magnetic field produced by the neuronal electrical activity. MEG provides a better localisation of the activity than EEG but is much more expensive.

### Measurement of metabolism or haemodynamics

**Positron Emission Tomography** PET-scans can provide measurement of metabolism or haemodynamics depending on the tracer of interest [Fox 1984]. The risk induced by the exposure to ionizing radiation and the existence of alternative non-ionizing approaches have greatly reduced the interest toward PET activation studies.

### Measurement of haemodynamics

**Near Infrared Spectroscopy Imaging** Near Infrared Spectroscopy imaging produces a quantitative measurement of both oxyhaemoglobin and deoxyhaemoglobin. This method provides quantitative measurements but with a low spatial resolution.

**Magnetic Resonance Imaging** The most widespread approach for activation studies is functional MRI using the BOLD effect. The first applications of functional MRI were developed in the early 1990's [Ogawa 1990, Belliveau 1992]. Arterial Spin Labelling is a competing approach based on a more direct measurement



of the haemodynamics [Detre 1992]. More details on BOLD fMRI and functional Arterial Spin Labelling (fASL) are provided in 3.4.2.

### 3.4.2 Functional MRI

**BOLD fMRI** Blood-Oxygen-Level-Dependent (BOLD) fMRI is the most widespread technique for imaging task-evoked activity. It relies on the fact that in presence of activation (electric neuronal activity) the increase in the quantity of oxygen provided to the brain is superior to the increase in oxygen requirements. This leads to an increase of the oxygenation level in the veins and capillaries that can be measured through MRI. The precise relationship between BOLD signal and electrical activity is still under study [Singh 2012].

The Haemodynamic Response Function (HRF) is the temporal response of the BOLD effect to a stimulus. As described in fig. 1.21, the BOLD effect is in fact influenced by a large set of biological parameters including: the CBF, the CBV and the Cerebral Metabolic Rate of Oxygen ( $\text{CMRO}_2$ ).

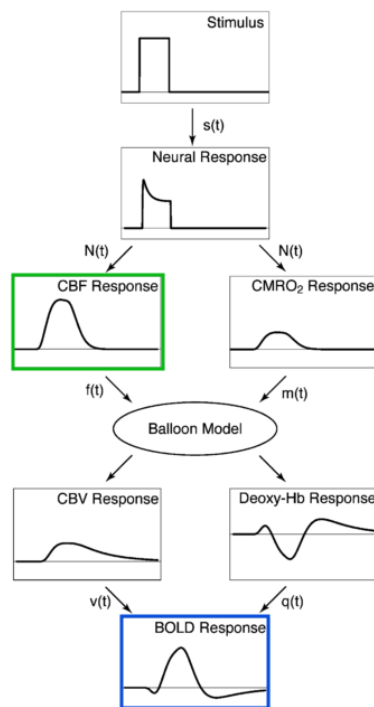


Figure 1.21: The balloon model linking the stimulus time-course with the physiological parameters and the observed BOLD (for BOLD fMRI, circled in blue) and CBF (for fASL, circled in green) effects. Reprinted from [Buxton 2004], Copyright (2013), with permission from Elsevier.

**fASL** Functional Arterial Spin Labelling (fASL) is based on a direct measurement of the CBF as illustrated in fig. 1.21. More details on ASL sequences are provided

in section 2.5.2.

### 3.4.3 Pros and Cons of ASL and BOLD fMRI

**MRI versus other imaging approaches** Since the development of BOLD fMRI in the early 1990's, this method has been the main technique to study task-evoked activity. This can be explained by the fact that competing approaches are either invasive or irradiating (PET) or present low spatial resolution (EEG, MEG, Near Infrared Spectroscopy).

**fASL versus BOLD fMRI** While BOLD fMRI is clearly the current standard for imaging task-evoked activity, fASL presents several advantages that we discuss in the following.

***A quantitative approach*** While relative variations in BOLD signal are of interest to measure activation, i.e. brain activity related to a particular task, baseline BOLD values or BOLD increases are not expressed in quantitative units. On the other hand, fASL can provide a quantitative estimate of CBF. This measure is not only interesting to study task-evoked activity but also in the study of resting perfusion (cf. section 2).

***Imaging of low-frequency tasks*** Temporal fluctuations of the BOLD signal render impossible the measurement of activation related to low-frequency tasks. Since these patterns of temporal noise are not present in ASL [Aguirre 2002], it opens the field to new cognitive tests, such as variations due to sport training [Wang 2003b].

***fASL is more directly coupled with the neuronal activation*** Variations in CBF as imaged by fASL are more directly coupled with neuronal activation than the BOLD effect [Buxton 2004]. Furthermore, while the variability of the haemodynamic response function observed in BOLD is a clear shortcoming [Aguirre 1998], the fASL response might be less variable [Aguirre 2002].

***A better localization of functional activity*** In the context of imaging of task-evoked activity, fASL has been shown to provide a better localization of the foci of activation [Duong 2001, Luh 2000]. The BOLD effect tends to move towards venous side of the circulatory system.

## 4 Conclusion

In this chapter, we presented the context of our study. We provided a general overview of brain function and then focused of the two components that are investigated in this thesis: brain perfusion and task-evoked activity.



## Chapter 2

# Pre-processing and statistical analysis

---

This chapter presents the standard pre-processing pipeline employed both for [ASL](#) and [BOLD](#) functional [MRI](#) data.

We first review the standard pre-processing steps applied to functional [MRI](#) data in order to compensate for undesired effects and to prepare the data for subsequent statistical analysis in [1](#). Then, we describe the most widespread method to analyse voxel-wise data in neuroimaging: the [General Linear Model \(GLM\)](#) in [2](#).

**Contents**

<b>1</b>	<b>Pre-processing</b> . . . . .	<b>35</b>
1.1	Pipeline overview . . . . .	35
1.2	Pre-processing steps . . . . .	36
1.2.1	Slice timing . . . . .	36
1.2.2	Spatial realignment . . . . .	36
1.2.3	Label/Control subtraction in Arterial Spin Labelling . . . . .	37
1.2.4	Registration onto anatomical space . . . . .	37
1.2.5	Spatial normalisation . . . . .	37
1.2.6	Smoothing . . . . .	38
<b>2</b>	<b>Statistical analysis</b> . . . . .	<b>38</b>
2.1	Modelling and inference using the General Linear Model . . . . .	38
2.1.1	Introduction to the GLM . . . . .	38
2.1.2	Modelling . . . . .	39
2.1.3	Estimation of the parameters . . . . .	40
2.1.4	Inference . . . . .	40
2.2	Subject level (First level) . . . . .	41
2.2.1	Single subject analysis for task-evoked activation . . . . .	41
2.2.2	Single subject analysis for basal perfusion . . . . .	44
2.3	Group level (Second level) . . . . .	45
2.3.1	Group analysis . . . . .	46
2.3.2	Between-group analysis . . . . .	49
2.3.3	Fixed effects . . . . .	53
2.4	Correction for multiple comparisons . . . . .	53
<b>3</b>	<b>Conclusion</b> . . . . .	<b>54</b>

## 1 Pre-processing

In this section, we present the pre-processing applied to raw **ASL** or **BOLD fMRI** data in order to prepare the data for the statistical analysis. We start with an overview of the pre-processing pipeline in 1.1 and then present each step in further details in 1.2. These typical pre-processing steps are implemented in various software including SPM<sup>1</sup> and FSL<sup>2</sup> and reviewed in greater details in [Poldrack 2011].

### 1.1 Pipeline overview

An example of pre-processing pipeline is provided in figure 2.1. In this pipeline, slice timing is first applied to the data followed by 3D spatial realignment. The order of these first two steps (slice timing and spatial realignment) might be reversed. For each subject, registration of the functional data (perfusion or task-evoked imaging) onto the anatomical image is then performed. Registration to a template space is optional and can be performed in order to compare or merge data from different subjects. Finally smoothing is often performed.

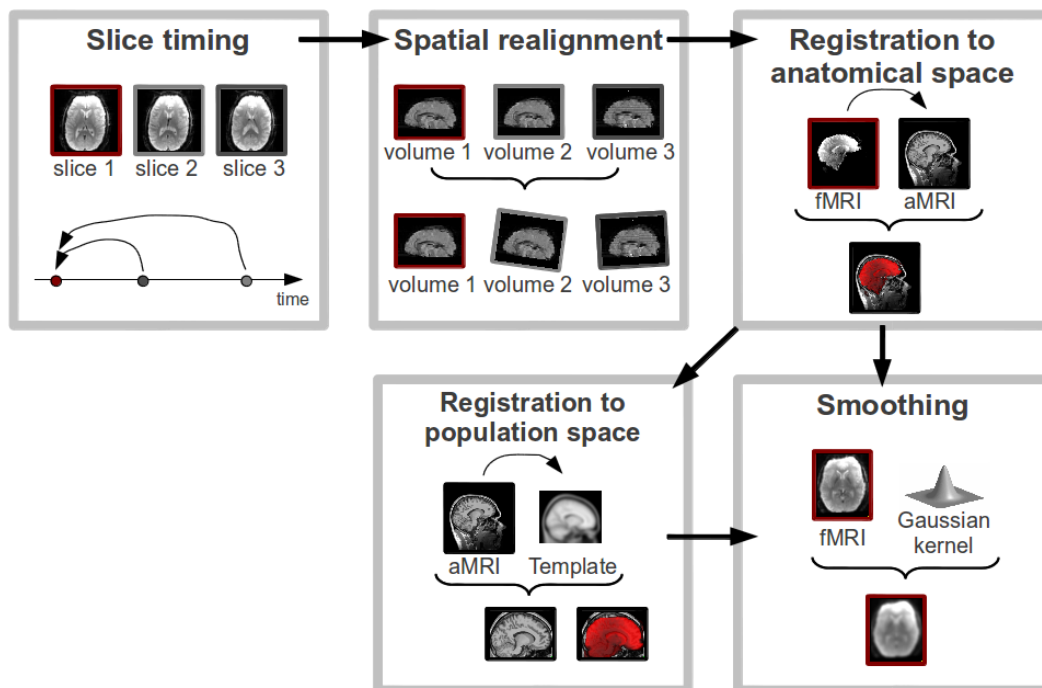


Figure 2.1: An example of pre-processing pipeline for **BOLD fMRI** or **ASL** data including: slice timing, spatial realignment, registration to the anatomical space, an optional registration into a population space (for group studies) and finally smoothing.

<sup>1</sup><http://www.fil.ion.ucl.ac.uk/spm>

<sup>2</sup><http://fsl.fmrib.ox.ac.uk/fsl/fslwiki>

## 1.2 Pre-processing steps

### 1.2.1 Slice timing

In the subsequent statistical analysis, the same model is usually employed throughout the brain, implicitly considering that all voxels were acquired at a single time-point. The aim of slice timing is to compensate for differences in acquisition times between slices [Henson 1999].

In **BOLD fMRI**, the typical readout time necessary to acquire a slice is in the order of 100 ms. It is easy to see that the first and last acquired slices will not correspond to the same time point. Fig. 2.2 illustrates the shifts on the haemodynamic function for three different slices located at the bottom, middle and top of the brain in a sequential acquisition (assuming that the same activation is present in the three observations). In practice, slice timing is performed by interpolating the available data with sinc interpolation.

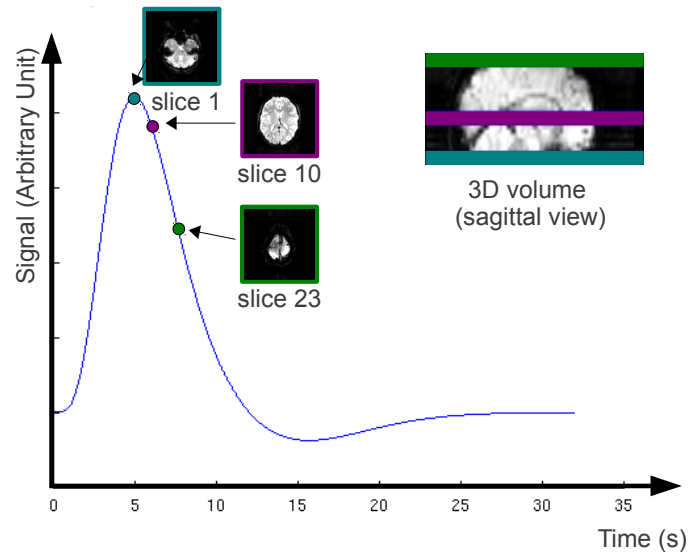


Figure 2.2: Example of slice delays in a sequential acquisition.

In 2D **ASL** sequences, the readout time per slice is in the order of 50 ms, which is twice as small as the usual **BOLD fMRI** slice readout time. In **ASL**, interpolation between labelled and control scans must be avoided. The shifting in time between slices is therefore taken into account in the quantification model (cf. eq. 1.6 and eq. 1.7) by setting a different readout time  $TI$  for each slice (e.g. in [Wang 2011]).

### 1.2.2 Spatial realignment

A **BOLD fMRI** or **fASL** experiment can last from about 5 min to 30 min for more complex and demanding paradigms. In resting **ASL**, the acquisition time is around 5 min. During this period of time, though foam padding is usually employed to

maintain the subject still, subject motion cannot be completely avoided. To compensate for motion between volumes, a rigid realignment is usually performed.

### 1.2.3 Label/Control subtraction in Arterial Spin Labelling

In basal ASL or in fASL, the perfusion signal comes from the difference between a control and a label scan. After realignment, a subtraction scheme is applied to the ASL series in order to obtain the perfusion-weighted series. Several subtraction procedures have been proposed, the three main variants are: pair-wise subtraction, surround subtraction and sinc-interpolation.

With pair-wise subtraction, the simplest approach, each labelled image is subtracted to the preceding control image. Surround subtraction and sinc-interpolation are two alternatives that takes into account the difference in readout times of the control and labelled scans. This is particularly important in fASL where two successive volumes might not be in the same state relatively to the paradigm (task-rest). With surround subtraction, a linear interpolation is performed between two subsequent labelled volumes in order to generate a virtual labelled volume with the same readout time as the control volume lying in the middle. With sinc-interpolation, similarly to slice timing in BOLD fMRI, the control scans are shifted by one Repetition time (TR) and then pair-wise subtracted with the labels. In fASL, pair-wise subtraction should not be used as otherwise the perfusion-weighted series is contaminated by BOLD contrast [Liu 2005].

In the study of perfusion at rest, the perfusion-weighted series is usually averaged in order to get a single perfusion-weighted map. While this is common practice, we will see later, that the standard deviation across the repeated acquisitions can also be of interest. In fASL, the complete perfusion-weighted series is always kept for further analysis.

Following the subtraction, the quantification is performed in order to convert the perfusion-weighted series into a CBF series as described in chapter 1, section 2.5.2. In single-TI studies, the quantification step is just a rescaling of the perfusion-weighted series.

### 1.2.4 Registration onto anatomical space

In order to be able to locate the functional information into the anatomy, a registration of the functional volumes (usually the transformation is estimated using the average of all volumes) onto the anatomical volume of the subject is performed. This is done using a cross-modality cost function such as normalised mutual information [Studholme 1999].

### 1.2.5 Spatial normalisation

When comparison between- or combination across-subjects is of interest, the functional maps must be registered into a common space. Several algorithms have been proposed for this purpose and this is still an open issue investigated in research.



In the context of this work, we applied publicly available registration schemes from SPM: the standard unified segmentation and registration scheme [Ashburner 2005] and DARTEL [Ashburner 2007]. While the latter is promising in terms of better cross-subject registration, the registration algorithms with the largest degrees of freedom are not well-suited for the registration of maps presenting large set of unexpected values, such as the ones of damaged brains [Crinion 2007a].

### 1.2.6 Smoothing

The last step before the statistical analysis is usually to smooth the data with a Gaussian kernel. While this clearly reduces the spatial resolution of the output results and cancels out high-frequency information, it provides several advantages [Poldrack 2011]. First, in multi-subject analysis, smoothing the data is a way to correct for small misregistrations following the spatial normalisation. Second, smoothing will inherently increase the SNR of large-scale signal and potentially the sensitivity of the statistical analysis. Third, by increasing the spatial coherency, smoothing reduces the burden of multiple comparisons via random-field theory [Nichols 2003]. Fourth, by the central limit theorem, smoothed data will tend to be more Gaussian than raw data and therefore better complies with the GLM assumptions. Depending on the expected improvement, the optimal size of the smoothing kernel differs. In practice, smoothing kernels of Full Width at Half Maximum (FWHM) from 6 to 12 mm<sup>3</sup> are commonly used.

## 2 Statistical analysis

In 2.1, we review the General Linear Model (GLM) in its general form. Then, in 2.2 and 2.3, we focus on the inference at the subject and group levels and provide some details on the most common models (one-sample t-test, two-sample t-test). Finally in 2.4 we discuss the issue of multiple testing. For more general information about the GLM, the interested reader is referred to a recent book [Poldrack 2011], online materials: SPM course<sup>3</sup>, FSL course<sup>4</sup>, MRC Cognition and Brain Sciences Unit Wiki<sup>5</sup> and to a set of references focusing on the application of the GLM in neuroimaging [Holmes 1998, Woolrich 2004, Cox 1996, Worsley 2002, Mumford 2006, Mumford 2009, Beckmann 2003].

### 2.1 Modelling and inference using the General Linear Model

#### 2.1.1 Introduction to the GLM

The GLM is a convenient and versatile way to model a dataset of interest as a linear combination of a set of pre-defined parameters representing the experimentally controlled factors and the potential confounds. Fig. 2.3 provide an example in

<sup>3</sup><http://www.translationalneuromodeling.org/spm-course-2013-presentation-slides>

<sup>4</sup><http://fsl.fmrib.ox.ac.uk/fslcourse>

<sup>5</sup><http://imaging.mrc-cbu.cam.ac.uk/imaging/PrinciplesStatistics>

which the time-course of a BOLD acquisition is modelled by a linear combination of a regressor matching the paradigm blocks and a constant term.

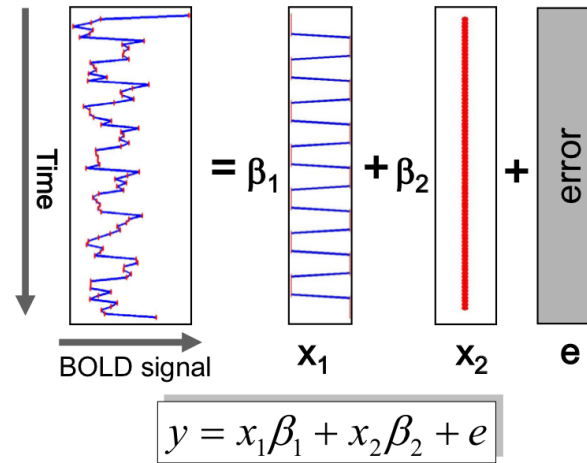


Figure 2.3: Example of use of the GLM to model the time-course of a BOLD acquisition related to a block paradigm, reprinted from “The General Linear Model for fMRI analyses” by Klaas Enno Stephan, SPM 2013 Course<sup>6</sup>.

The aim of any GLM experiment is to provide an estimate of the model parameters ( $\beta_1$  and  $\beta_2$  in the example of fig. 2.3) and, most usually, to draw inference about these estimates.

### 2.1.2 Modelling

In its more general form, the GLM can be defined by:

$$\mathbf{Y}_{(v)} = \mathbf{X}_{(v)}\boldsymbol{\beta}_{(v)} + \boldsymbol{\varepsilon}_{(v)}, \quad (2.1)$$

where  $\mathbf{Y}_{(v)}$  is a vector of length  $n$  containing the observations at voxel  $v$ . The  $p$  elements of  $\boldsymbol{\beta}_{(v)}$  represent the underlying effects.  $\mathbf{X}_{(v)}$  is the  $n \times p$  design matrix that defines the relationship between the observations and the predicting parameters defined in  $\boldsymbol{\beta}_{(v)}$ . Finally  $\boldsymbol{\varepsilon}_{(v)}$ , a vector of length  $n$ , is the error term that captures the variability that cannot be explained by the model. Fig. 2.4 provides an illustration of this matrix representation. The design matrix  $\mathbf{X}_{(v)}$  is known *a priori*. The aim is to estimate the model parameters stored in  $\boldsymbol{\beta}_{(v)}$ .

The main assumption behind the GLM is the linearity of the effects. Gaussian noise is usually assumed. In the standard approach, the GLM is defined independently for each voxel,  $v$ , which is why this approach is often termed “massively univariate”.

<sup>6</sup>[http://www.translationalneuromodeling.org/spm2013/slides/Klaas\\_GLM.pptx](http://www.translationalneuromodeling.org/spm2013/slides/Klaas_GLM.pptx)

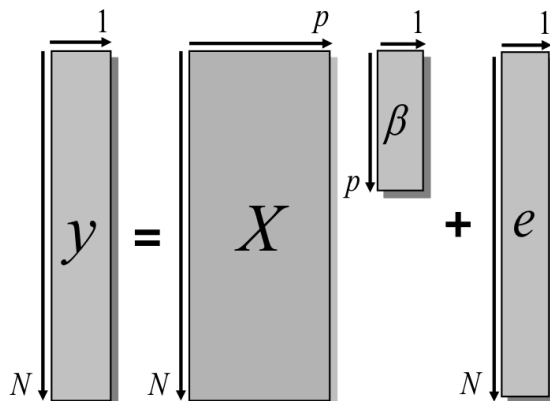


Figure 2.4: Matrix representation of the GLM, reprinted from “The General Linear Model for fMRI analyses” by Klaas Enno Stephan, SPM 2013 Course<sup>6</sup>.

### 2.1.3 Estimation of the parameters

If the errors are independent and identically distributed, i.e.  $\varepsilon_{(v)} \sim \mathcal{N}(0, \sigma_{(v)}^2 \mathbf{I})$  where  $\mathbf{I}$  is the identity matrix, then the GLM can be solved by Ordinary Least Squares (OLS):

$$\hat{\beta}_{(v)}^{OLS} = (\mathbf{X}_{(v)}^T \mathbf{X}_{(v)})^{-1} \mathbf{X}_{(v)}^T \mathbf{Y}_{(v)}. \quad (2.2)$$

If the errors are correlated and/or not identically distributed, i.e.  $\varepsilon_{(v)} \sim \mathcal{N}(0, \mathbf{V}_{(v)})$ , then the GLM can be solved by Weighted Least Squares (WLS):

$$\hat{\beta}_{(v)}^{WLS} = (\mathbf{X}_{(v)}^T \mathbf{W}_{(v)} \mathbf{X}_{(v)})^{-1} \mathbf{X}_{(v)}^T \mathbf{W}_{(v)} \mathbf{Y}_{(v)} \quad (2.3)$$

where  $\mathbf{W}_{(v)} = \mathbf{V}_{(v)}^{-1}$  is the pre-whitening matrix.

### 2.1.4 Inference

Now, since we are able to estimate the underlying parameters of our model, we would like to be able to draw conclusions about these parameters. This part is known as inference. Inference can be done for a single parameter (e.g. testing if the first parameter  $\beta_0$  is significantly different from zero using:  $H_0 : \beta_0 = 0$ ) or for a combination of parameters (e.g. testing if the second and third parameters  $\beta_1$  and  $\beta_2$  are significantly different from one another:  $H_0 : \beta_1 = \beta_2$ ). Both cases can be described by a contrast vector  $\mathbf{c}$  defining a linear combination in the set of parameters. Then the associated null hypothesis is  $H_0 : \mathbf{c}\beta_{(v)} = 0$ .

In order to infer whether  $H_0$  is rejected in the current observations, we need to define the distribution of  $\mathbf{c}\hat{\beta}_{(v)}$  under the null hypothesis with:

$$\mathbf{c}\hat{\beta}_{(v)} \sim \mathcal{N}(0, \text{Var}(\mathbf{c}\hat{\beta}_{(v)})). \quad (2.4)$$

In OLS, the variance is:

$$\text{Var}(\mathbf{c}\hat{\beta}_{(v)}^{OLS}) = \mathbf{c}(\mathbf{X}_{(v)}^T \mathbf{X}_{(v)})^{-1} \mathbf{c}^T \sigma_{(v)}^2 \quad (2.5)$$

where  $\sigma_{(v)}^2$  is estimated by:

$$\hat{\sigma}_{(v)}^2 = \frac{\mathbf{e}_{(v)}^T \mathbf{e}_{(v)}}{n-p} \quad \text{where } \mathbf{e}_{(v)} = \mathbf{Y}_{(v)} - \mathbf{X}_{(v)} \hat{\boldsymbol{\beta}}_{(v)}. \quad (2.6)$$

In WLS, the variance is:

$$\text{Var}(\mathbf{c} \hat{\boldsymbol{\beta}}_{(v)}^{OLS}) = \mathbf{c} (\mathbf{X}_{(v)}^T \mathbf{W}_{(v)} \mathbf{X}_{(v)})^{-1} \mathbf{c}^T \quad (2.7)$$

where  $\mathbf{W}_{(v)}$  will be estimated differently depending on the structure of the noise at hand.

Since  $\text{Var}(\mathbf{c} \hat{\boldsymbol{\beta}}_{(v)})$  is estimated from the data (and is *a priori* unknown), we have:

$$t_{(v)} = \frac{\mathbf{c} \hat{\boldsymbol{\beta}}_{(v)}}{\widehat{\text{Var}}(\mathbf{c} \hat{\boldsymbol{\beta}}_{(v)})} \sim \mathcal{T}_{n-p}, \quad (2.8)$$

where  $t_{(v)}$  is the value of the statistics. Then one-sided or two-sided tests can be performed by looking at  $\Pr(\mathcal{T}_{n-p} \geq t_{(v)})$  or  $\Pr(\mathcal{T}_{n-p} \geq |t_{(v)}|)$ .

## 2.2 Subject level (First level)

At the subject level, the GLM can be defined with the following notations. Given a voxel  $v$ , for a subject  $s$  we have:

$$\mathbf{Y}_{s(v)} = \mathbf{X}_{s(v)} \boldsymbol{\beta}_{s(v)} + \boldsymbol{\varepsilon}_{s(v)}, \quad (2.9)$$

where  $\mathbf{Y}_{s(v)}$  is a vector containing the  $n$  observations (usually along the time),  $\boldsymbol{\beta}_{s(v)}$  is the set of subject parameters to be estimated,  $\mathbf{X}_{s(v)}$  is the subject-level design matrix and  $\boldsymbol{\varepsilon}_{s(v)}$  contains the residual errors.

### 2.2.1 Single subject analysis for task-evoked activation

**Objective** In the analysis of task-evoked activity in a single subject, the aim is to outline activated (or deactivated) brain areas associated with a given task. At the end of the analysis, the activated regions are usually overlaid on the anatomical MRI of the subject under study as illustrated in fig. 2.5.

**Modelling** In BOLD fMRI at the subject level, the data under study is a 4D image of  $n$  3D-volumes acquired during the accomplishment of the paradigm. In fASL at the subject level, the data under study is a 4D perfusion-weighted series obtained after pair-wise subtraction of the 3D-volumes acquired during the accomplishment of the paradigm. The design matrix,  $\mathbf{X}_{s(v)}$ , is therefore closely related to the time-course of the paradigm under study. Several tasks (or actions) might be performed in a single session and modelled by different regressors in the GLM.

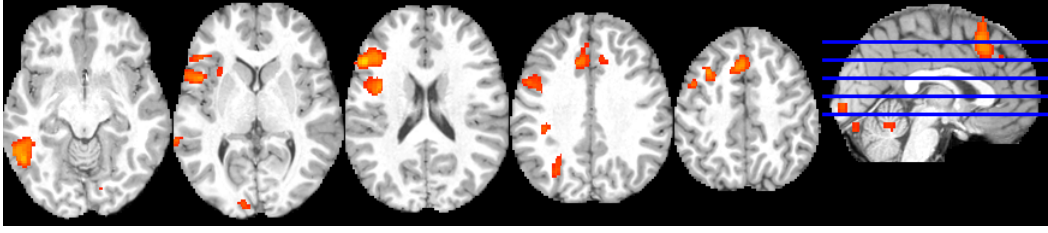


Figure 2.5: Example of subject activations for a language paradigm.

Each task-related regressor is built by convolution of the task time-course with the Haemodynamic Response Function (HRF) as illustrated in fig. 2.6. While the HRF is supposed to be different in BOLD fMRI and fASL, for convenience most fASL analyses rely on the BOLD response function. Including the derivative (temporal and dispersion) is also an option to allow for more flexibility in the response function [Friston 1998]. Also, a canonical HRF without undershoot is sometimes used in fASL studies.

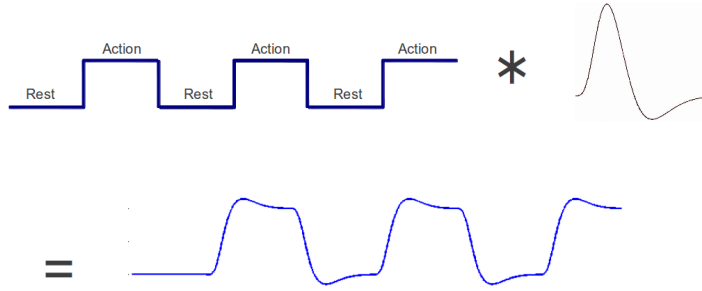


Figure 2.6: Creation of a task-related regressor (one column of the design matrix  $\mathbf{X}_{s(v)}$ ) by convolution of the paradigm time-course with the HRF. In this example, a block paradigm with 3 action and rest blocks is presented.

Other regressors, such as the motion parameters estimated in 1.2.2, might be included to model nuisance effects [Johnstone 2006]. The GLM model for task-invoked activity of a single subject is therefore defined by:

$$\mathbf{Y}_{s(v)} = \begin{bmatrix} x_s^{1,1} & \dots & x_s^{k,1} & a_s^{1,1} & \dots & a_s^{p-k,1} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ x_s^{1,n} & \dots & x_s^{k,n} & a_s^{1,n} & \dots & a_s^{p-k,n} \end{bmatrix} \begin{bmatrix} \beta_{s(v)}^1 \\ \vdots \\ \beta_{s(v)}^k \\ \alpha_{s(v)}^1 \\ \vdots \\ \alpha_{s(v)}^{p-k} \end{bmatrix} + \boldsymbol{\varepsilon}_{s(v)} \quad (2.10)$$

where  $(\beta_{s(v)}^i)_{1 \leq i \leq k}$  are the parameters of interest for each of the  $k$  paradigm-related

effects and  $(\alpha_{s(v)}^i)_{1 \leq i \leq p-k}$  are the nuisance parameters. An example of design matrix with  $k = 1$  is displayed in fig. 2.7. More complex paradigms, can require more than one paradigm-related effect. For example, the IOWA Gambling task, a paradigm in which the subject is required to play cards and receive or loose money at each run, can be modelled with two paradigm-related effects corresponding to the trials in which the subject won (or lost) money during the game [Lawrence 2009].

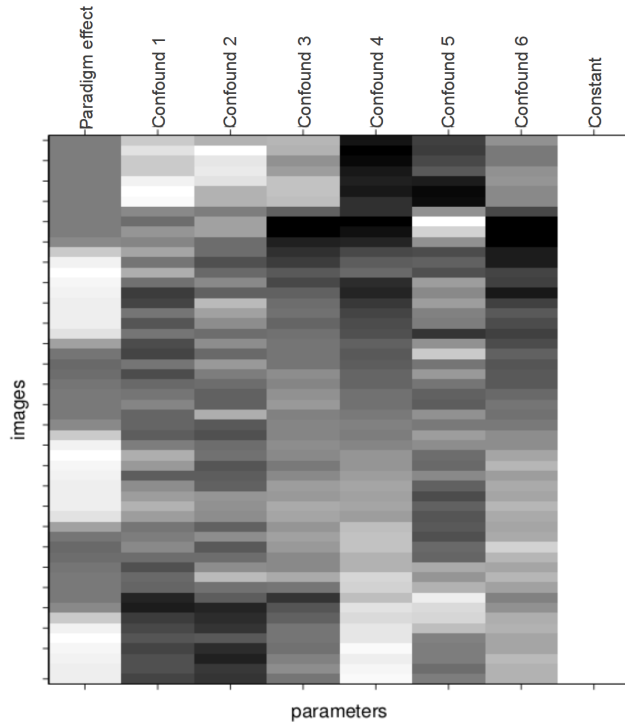


Figure 2.7: Example of design matrix,  $\mathbf{X}_{s(v)}$ , for a subject analysis of task-evoked activation (computed with SPM8). The matrix presents a total of 8 regressors, from left to right: 1 task-related regressor modelling a block paradigm (from fig. 2.6), 6 nuisance motion covariates (3 translations and 3 rotations parameters estimated during the registration step in the pre-processing) and 1 constant regressor.

In fASL, the pre-processed ASL data before subtraction can also be used as observations in the  $\mathbf{Y}_{s(v)}$  vector [Mumford 2006]. As BOLD effect is also present in fASL series [Wong 1997], the model is then modified so that each paradigm-related effect is represented by an ASL and a BOLD regressor. An additional regressor represents basal perfusion. While this modelling approach offers interesting perspective, in the context of this thesis we focused on the simpler model, which is applicable for both BOLD fMRI and fASL.

## Noise properties and estimation

**BOLD fMRI** In BOLD fMRI, the temporal noise is not a white noise as it is characterized by autocorrelation and slow frequency drifts [Zarahn 1997, Aguirre 1997]. Both phenomena must be taken into account in the statistical analysis. The reason for low-frequency drifts is still not completely understood but seems to be inherent to BOLD imaging as it was also observed in phantoms [Smith 1999]. Slow frequency drifts are usually removed from the data by high pass filtering. This will however affect the potential designs that can be conducted in BOLD fMRI experiments. In practice, due to the noise structure, high-pass filtering of 128 s can be performed (default value in SPM) making it impossible to detect variations that occurs more slowly than every two minutes.

Temporal autocorrelation is usually dealt with by pre-whitening the data. This relies on the fact that the temporal autocorrelation is estimated accurately. Once the pre-whitening matrix  $\mathbf{W}_{s(v)}$  has been estimated, given the original observations and design matrix,  $\mathbf{Y}_{s(v)}$  and  $\mathbf{X}_{s(v)}$ , the data can be updated using:

$$\begin{aligned}\mathbf{Y}_{s(v)}^* &= \mathbf{W}_{s(v)}\mathbf{Y}_{s(v)} \\ \mathbf{X}_{s(v)}^* &= \mathbf{W}_{s(v)}\mathbf{X}_{s(v)}\end{aligned}\quad (2.11)$$

and the updated system:

$$\mathbf{Y}_{s(v)}^* = \mathbf{X}_{s(v)}^* \boldsymbol{\beta}_{s(v)} + \boldsymbol{\varepsilon}_{s(v)}^*, \quad (2.12)$$

can be solved by OLS where  $\boldsymbol{\varepsilon}_{s(v)}^* \sim \mathcal{N}(0, \sigma_{s(v)}^2 \mathbf{I})$  (direct solving by WLS is equivalent).

**fASL** Thanks to the subtraction process between label and control scans, it can be reasonably assumed that ASL sequences are not contaminated by slow frequency drifts and temporal autocorrelation [Aguirre 2002]. The system (2.10) can therefore be directly solved by OLS where:  $\boldsymbol{\varepsilon}_{s(v)} \sim \mathcal{N}(0, \sigma_{s(v)}^2 \mathbf{I})$ .

## 2.2.2 Single subject analysis for basal perfusion

**Objective** In the analysis of basal perfusion in single subjects, the aim is to get a map of the CBF for the subject of interest. At the end of the analysis, the CBF map is usually displayed in grey or rainramp colormap as illustrated in fig 2.8.

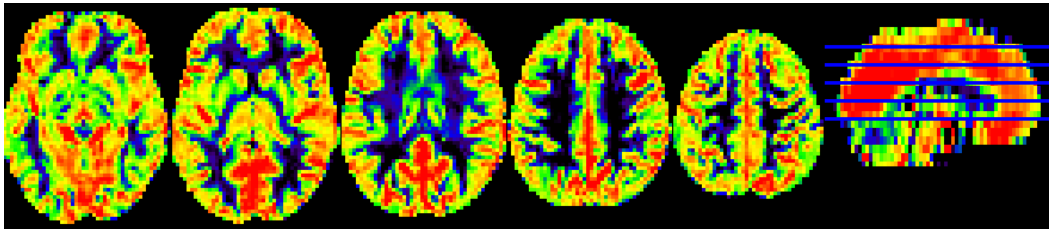


Figure 2.8: Example of CBF map for a control subject.

**Modelling** In ASL, at the subject level, the data under study is a 4D volume of observed CBF maps containing  $n$  volumes (1 per repeated acquisition). The observations  $\mathbf{Y}_{s(v)}$  are therefore repeated measurements of a same underlying value (the CBF) and the  $\mathbf{X}_{s(v)}$  matrix is a vector of ones tabulated by nuisance regressors. The model is therefore:

$$\mathbf{Y}_{s(v)} = \begin{bmatrix} 1 & a_s^{1,1} & \dots & a_s^{p-1,1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & a_s^{1,n} & \dots & a_s^{p-1,n} \end{bmatrix} \begin{bmatrix} \beta_{s(v)} \\ \alpha_{s(v)}^1 \\ \vdots \\ \alpha_{s(v)}^{p-1} \end{bmatrix} + \boldsymbol{\varepsilon}_{s(v)} \quad (2.13)$$

where  $\beta_{s(v)}$  is the parameter of interest (the CBF) and  $(\alpha_{s(v)}^i)_{1 \leq i \leq p-1}$  are the nuisance parameters.

**Noise properties and estimation** Again, in ASL, thanks to the subtraction process between control and labelled scans, we can reasonably assume white noise [Agui-re 2002, Mumford 2006]. The system (2.13) can therefore be solved by OLS given  $\boldsymbol{\varepsilon}_{s(v)} \sim \mathcal{N}(0, \sigma_{s(v)}^2 \mathbf{I})$ .

**Example** If no nuisance regressors are defined in the model, the OLS estimate of the subject parameter  $\beta_{s(v)}$  is the sample average across the ASL repeats:

$$\hat{\beta}_{s(v)} = \frac{1}{n} \sum_{i=1}^n y_{s(v)}^i, \quad (2.14)$$

where  $y_{s(v)}^i$  is the  $i^{\text{th}}$  element of the vector  $\mathbf{Y}_{s(v)}$ . Similarly, the sampling variance of  $\hat{\beta}_{s(v)}$  is estimated by:

$$\widehat{\text{Var}}(\hat{\beta}_{s(v)}) = \frac{\hat{\sigma}_{s(v)}^2}{n} \quad \text{where} \quad \hat{\sigma}_{s(v)}^2 = \frac{1}{n-1} \sum_{i=1}^n (y_{s(v)}^i - \hat{\beta}_{s(v)})^2. \quad (2.15)$$

### 2.3 Group level (Second level)

We limit our discussion to the case where a single estimate per subject (from the subject level) is brought to the group level. In a group of  $n$  subjects, the subject parameters  $(\beta_{s(v)})_{1 \leq s \leq n}$  can be combined using the following model:

$$\begin{bmatrix} \beta_{1(v)} \\ \vdots \\ \beta_{n(v)} \end{bmatrix} = \mathbf{X}_{G(v)} \boldsymbol{\beta}_{G(v)} + \boldsymbol{\gamma}_{G(v)}, \quad (2.16)$$

where  $\mathbf{X}_{G(v)}$  is the group-level design matrix,  $\boldsymbol{\beta}_{G(v)}$  the group parameters and  $\boldsymbol{\gamma}_{G(v)}$  the residual error term (outlining the intra-group variabilities).



However, the true subject parameters are in fact unknown and, in practice, their estimates from the first level (i.e. the subject level) are used in the second-level leading to a slightly different model:

$$\begin{bmatrix} \hat{\beta}_{1(v)} \\ \vdots \\ \hat{\beta}_{n(v)} \end{bmatrix} = \mathbf{X}_{G(v)} \boldsymbol{\beta}_{G(v)} + \boldsymbol{\gamma}_{G_C(v)}. \quad (2.17)$$

The new error term  $\boldsymbol{\gamma}_{G_C(v)}$  is impacted by two combined sources of variations: the measurement error on the subject parameters (also termed within-subject variance) and the between-subject variance. Let  $\mathbf{Y}_{G(v)} = [\hat{\beta}_{1(v)} \dots \hat{\beta}_{n(v)}]$  denote the vector of subject-level parameters brought to the second level.

### 2.3.1 Group analysis

**Objective** In single group analyses, the aim is to produce a map of the functional activity representative of the group of subjects under study. In task-evoked activation, a map of brain areas involved for a given task across the group is produced. This activation map is obtained through the inference step, by thresholding the estimated parameter map to keep only the areas displaying a significant task-related signal. At the end of the analysis, the activated regions are usually overlaid on the average anatomical MRI of the subjects under study as illustrated in fig. 2.9. In the

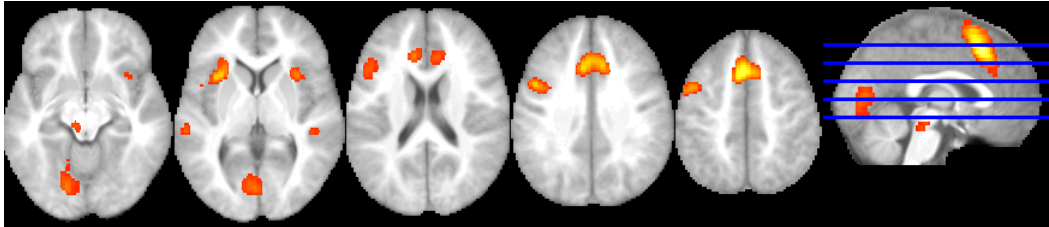


Figure 2.9: Example of group activations for a language paradigm.

study of basal perfusion, the complete (unthresholded) parameter map is of interest and is a representation of the perfusion information across the group. At the end of the analysis, the group perfusion map is usually displayed in grey or rainramp colormap as illustrated in fig 2.10.

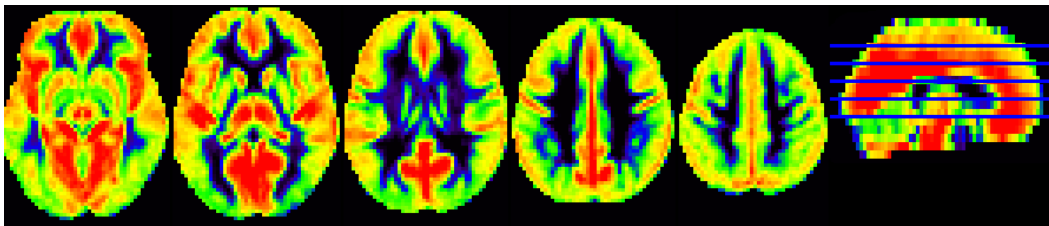


Figure 2.10: Example of group perfusion map.

Group task-evoked activity and group basal perfusion can be estimated with the same model, as described hereafter.

**Modelling** In **BOLD fMRI** and **fASL** at the group level, the data under study is a 4D image of  $n$  (1 per subject) 3D-volumes representing an estimated task-related effect. In **ASL** at the group level, the data under study is a 4D image of  $n$  (1 per subject) estimated **CBF** maps. The observations  $\mathbf{Y}_{G(v)}$  are therefore seen as samples of a same underlying effect (i.e. the activation level or the **CBF**) and the  $\mathbf{X}_{G(v)}$  matrix is a vector of ones tabulated with nuisance regressors. The model is therefore:

$$\begin{bmatrix} \hat{\beta}_{1(v)} \\ \vdots \\ \hat{\beta}_{n(v)} \end{bmatrix} = \begin{bmatrix} 1 & a_G^{1,1} & \dots & a_G^{p-1,1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & a_G^{1,t} & \dots & a_G^{p-1,t} \end{bmatrix} \begin{bmatrix} \beta_{G(v)} \\ \alpha_{1(v)} \\ \vdots \\ \alpha_{p-1(v)} \end{bmatrix} + \gamma_{G_C(v)}, \quad (2.18)$$

where  $\beta_{G(v)}$  is the parameter of interest and  $(\alpha_{i(v)})_{1 \leq i \leq p-1}$  are the nuisance parameters.

**Noise properties and estimation** The error term  $\gamma_{G_C(v)}$  is impacted by two combined sources of variations: the measurement error on the subject parameters (also termed within-subject variance) and the between-subject variance (also termed intra-group or group variance).  $\gamma_{G_C(v)}$  therefore follows a normal distribution:

$$\gamma_{G_C(v)} \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_{G(v)}^2 + \frac{\sigma_{1(v)}^2}{v_1} & 0 & \dots & \dots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \sigma_{G(v)}^2 + \frac{\sigma_{s(v)}^2}{v_s} & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \dots & \dots & 0 & \sigma_{G(v)}^2 + \frac{\sigma_{n(v)}^2}{v_n} \end{bmatrix}\right) \quad (2.19)$$

where  $v_s$  is the number of repeated volume at the subject level for subject  $s$  and  $\sigma_{s(v)}^2$  is the within-subject variance (from the subject level).

Depending on the assumptions made on the within-subjects variances, two methods have been proposed to solve this hierarchical **GLM**: the random-effects and mixed-effects approaches.

**Random-effects** If the subject-level variance is considered either negligible by comparison to the intra-group variance:

$$\frac{\sigma_{s(v)}^2}{v_{s(v)}} \ll \sigma_{G(v)}^2, \quad \forall 1 \leq s \leq n \quad (2.20)$$

or constant across subjects:

$$\frac{\sigma_{s(v)}^2}{v_{s(v)}} \approx \sigma_{SUB(v)}^2, \quad \forall 1 \leq s \leq n \quad (2.21)$$

then the data is homoscedastic and  $\gamma_{G_C(v)}$  follows a normal distribution:

$$\gamma_{G_C(v)} \sim \mathcal{N}(0, \sigma_{G_C(v)}^2 \mathbf{I}) \quad (2.22)$$

where  $\sigma_{G_C(v)}^2$  is the combined within- and between-subject variance. Depending on the assumption, we have  $\sigma_{G_C(v)}^2 = \sigma_{G(v)}^2 + \sigma_{SUB(v)}^2$  or  $\sigma_{G_C(v)}^2 \approx \sigma_{G(v)}^2$ . The sphericity assumption (no heteroscedasticity and no autocorrelation) hence holds and the system (2.18) can be solved by OLS.

**Mixed-effects** In case of heteroscedasticity (i.e. when the assumptions of the random-effects estimation does not hold), the system (2.18) is solved by WLS. This type of estimation is known as mixed-effects in the neuroimaging community.

**Example** In a one-sample t-test (i.e. in the absence of nuisance covariates),  $\mathbf{X}_{G(v)}$  reduces to a vector of ones and  $\beta_{G(v)}$  contains a single estimate.

$$\begin{bmatrix} \hat{\beta}_{1(v)} \\ \vdots \\ \hat{\beta}_{n(v)} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \beta_{G(v)} + \gamma_{G_C(v)}, \quad (2.23)$$

**Random-effects** In case of homoscedasticity, the OLS estimate of the group parameter  $\beta_{G(v)}$  is the cross-subject sample average:

$$\hat{\beta}_{G(v)}^{RFX} = \frac{1}{n} \sum_{s=1}^n \hat{\beta}_{s(v)}, \quad (2.24)$$

and the sampling variance of  $\hat{\beta}_G$  is estimated by:

$$\widehat{\text{Var}}(\hat{\beta}_{G(v)}^{RFX}) = \frac{\hat{\sigma}_{G(v)}^2}{n} \quad \text{where} \quad \hat{\sigma}_{G(v)}^2 = \frac{1}{n-1} \sum_{i=1}^n (\hat{\beta}_{i(v)} - \hat{\beta}_{G(v)})^2. \quad (2.25)$$

**Mixed-effects** In case of heteroscedasticity, the WLS estimate of the group parameter  $\beta_{G(v)}$  is a weighted average of the subject parameters:

$$\hat{\beta}_{G(v)}^{MFX} = \frac{1}{\sum_{i=1}^n \kappa_{i(v)}} \sum_{i=1}^n \kappa_{i(v)} \hat{\beta}_{i(v)}, \quad (2.26)$$

where  $\kappa_{s(v)}^{-1} = \sigma_{G(v)}^2 + \frac{\sigma_{s(v)}^2}{v_{s(v)}}$

and the sampling variance of  $\hat{\beta}_G$  is estimated by:

$$\widehat{\text{Var}}(\hat{\beta}_{G(v)}^{MFX}) = \frac{1}{\sum_{i=1}^n \kappa_{i(v)}} \quad (2.27)$$

### 2.3.2 Between-group analysis

**Objective** In between-group analyses, the aim is to find brain areas that have a different pattern of functional activity in a group by comparison to another group of subjects. At the end of the analysis, the hyper-activations and hypo-activations (in the patient group by comparison to the control group) are usually overlaid on the mean anatomical MRI of the subjects under study as illustrated in fig. 2.11.

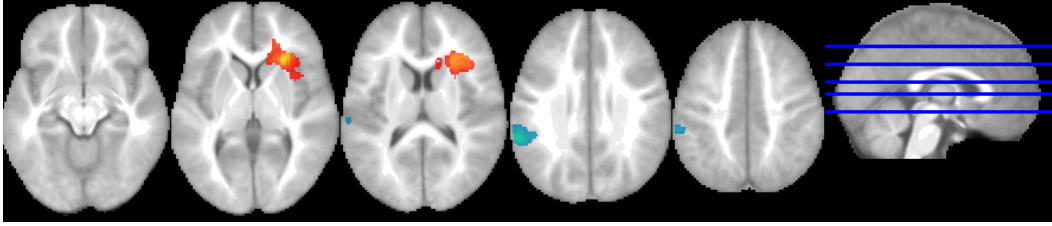


Figure 2.11: Example of task-evoked hyper-activations (hot colormap) and hypo-activations (winter colormap) in a group of children diagnosed with typical specific language impairment by comparison to typically developing children.

**Modelling** In the context of between-group analyses, the subjects are part of two groups:  $n_1$  subjects (controls) in the first group and  $n_2$  subjects (patients) in the second. Let  $\beta_{controls(v)}$  and  $\beta_{patients(v)}$  be the first and second group parameters respectively. Without loss of generality, we assume that subjects 1 to  $n_1$  are part of the first group and subject  $n_1 + 1$  to  $n = n_1 + n_2$  are part of the second group. Then the second-level model is:

$$\begin{bmatrix} \hat{\beta}_{1(v)} \\ \vdots \\ \hat{\beta}_{n_1(v)} \\ \hat{\beta}_{n_1+1(v)} \\ \vdots \\ \hat{\beta}_{n(v)} \end{bmatrix} = \begin{bmatrix} 1 & 0 & a^{1,1} & \dots & a^{1,p-2} \\ \vdots & \vdots & \dots & \dots & \dots \\ 1 & 0 & a^{n_1,1} & \dots & a^{n_1,p-2} \\ 0 & 1 & a^{n_1+1,1} & \dots & a^{n_1+1,p-2} \\ \vdots & \vdots & \dots & \dots & \dots \\ 0 & 1 & a^{n,1} & \dots & a^{n,p-2} \end{bmatrix} \begin{bmatrix} \beta_{controls(v)} \\ \beta_{patients(v)} \\ \alpha_{1(v)} \\ \vdots \\ \alpha_{p-2(v)} \end{bmatrix} + \gamma_{GC(v)} \quad (2.28)$$

where  $\beta_{controls(v)}$  and  $\beta_{patients(v)}$  are the parameters of interest and  $(\alpha_{s(v)}^i)_{1 \leq i \leq p-2}$  are the nuisance parameters. An example of design matrix is provided in fig. 2.12.

**Noise properties and estimation** Similarly to single-group analyses, the error term  $\gamma_{GC(v)}$  is impacted by two combined sources of variations: the measurement error on the subject parameters (also termed within-subject variance) and the between-subject variance (also termed intra-group or group variance).  $\gamma_{GC(v)}$

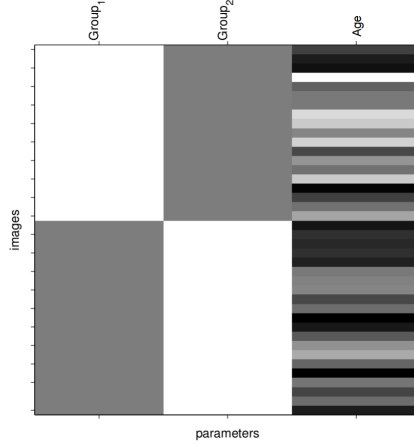


Figure 2.12: Example of design matrix,  $\mathbf{X}_{G(v)}$ , for a between-group analysis (computed with SPM8). The matrix presents a total of 3 regressors, from left to right: 2 regressors modelling the group effects and 1 nuisance regressor (age of the subjects).

therefore follows a normal distribution:

$$\gamma_{GC(v)} \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_{G_1(v)}^2 + \frac{\sigma_{1(v)}^2}{v_{1(v)}} & 0 & \dots & \dots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \sigma_{G_1(v)}^2 + \frac{\sigma_{n_1(v)}^2}{v_{n_1(v)}} & & \vdots \\ \vdots & & & \sigma_{G_2(v)}^2 + \frac{\sigma_{n_1+1(v)}^2}{v_{n_1+1(v)}} & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \dots & \dots & 0 & \sigma_{G_2(v)}^2 + \frac{\sigma_{n(v)}^2}{v_{n(v)}} \end{bmatrix}\right) \quad (2.29)$$

where  $v_{s(v)}$  is the number of repeated volume at the subject level for subject  $s$ .

**Random-effects** If the subject-level variances are considered either negligible by comparison to the intra-group variances:

$$\begin{aligned} \frac{\sigma_{s(v)}^2}{v_{s(v)}} &\ll \sigma_{G_1(v)}^2, & \forall 1 \leq s \leq n_1 & \quad \text{and} \\ \frac{\sigma_{s(v)}^2}{v_{s(v)}} &\ll \sigma_{G_2(v)}^2, & \forall n_1 + 1 \leq s \leq n \end{aligned} \quad (2.30)$$

or constant across subjects of the same group

$$\begin{aligned} \frac{\sigma_{s(v)}^2}{v_{s(v)}} &\approx \sigma_{SUB_1(v)}^2, & \forall 1 \leq s \leq n_1 & \quad \text{and} \\ \frac{\sigma_{s(v)}^2}{v_{s(v)}} &\approx \sigma_{SUB_2(v)}^2, & \forall n_1 + 1 \leq s \leq n \end{aligned} \quad (2.31)$$

then  $\gamma_{G_C(v)}$  follows a normal distribution:

$$\gamma_{G_C(v)} \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_{G_c1}^2 & 0 & \dots & 0 \\ 0 & \ddots & & \\ & & \sigma_{G_c1}^2 & \vdots \\ \vdots & & \sigma_{G_c2}^2 & \\ & & & \ddots & 0 \\ 0 & \dots & 0 & \sigma_{G_c2}^2 \end{bmatrix}\right) \quad (2.32)$$

where  $\sigma_{G_Cj(v)}^2$  is the combined within- and between-subject variance in group  $j$ . Depending on the assumption, we have  $\sigma_{G_Cj(v)}^2 = \sigma_{G_j(v)}^2 + \sigma_{SUB_j(v)}^2$  or  $\sigma_{G_Cj(v)}^2 \approx \sigma_{G_j(v)}^2$ . The system (2.28) can be solved by **WLS**. This type of estimation is known as random-effects in the neuroimaging community.

**Mixed-effects** When the assumptions of the random-effects estimation does not hold, the system (2.28) is solved by **WLS** using the complete variance-covariance matrix defined in eq. (2.29). This type of estimation is known as mixed-effects in the neuroimaging community.

**Example** In a two-sample t-test (*i.e.* in the absence of covariates),  $\mathbf{X}_{G(v)}$  reduces to a  $n \times 2$  matrix and  $\beta_{G(v)}$  contains two parameters:

$$\begin{bmatrix} \hat{\beta}_{1(v)} \\ \vdots \\ \hat{\beta}_{n_1(v)} \\ \hat{\beta}_{n_1+1(v)} \\ \vdots \\ \hat{\beta}_{n(v)} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_{controls(v)} \\ \beta_{patients(v)} \end{bmatrix} + \gamma_{G_C(v)}. \quad (2.33)$$

**Random-effects** If the measurement error on the subject parameter estimates can be considered as constant across subjects or negligible by comparison to the cross-subject variability, then the system (2.33) can be solved by **WLS**, giving:

$$\begin{aligned} \hat{\beta}_{controls}^{RFX} &= \frac{1}{n_1} \sum_{s=1}^{n_1} \hat{\beta}_s \\ \hat{\beta}_{patients}^{RFX} &= \frac{1}{n_2} \sum_{s=n_1+1}^n \hat{\beta}_s \end{aligned} \quad (2.34)$$

and their sample variance:

$$\begin{aligned} \widehat{\text{Var}}(\hat{\beta}_{controls}^{RFX}) &= \frac{\hat{\sigma}_{G_c1}^2}{n_1} \\ \widehat{\text{Var}}(\hat{\beta}_{patients}^{RFX}) &= \frac{\hat{\sigma}_{G_c2}^2}{n_2}. \end{aligned} \quad (2.35)$$

The parameter estimate associated with the between-group contrast:  $c = [-1 \ 1]$  is therefore:

$$\begin{aligned}\hat{b}^{RFX} &= \hat{\beta}_{controls}^{RFX} - \hat{\beta}_{patients}^{RFX} \\ &= \frac{1}{n_1} \sum_{s=1}^{n_1} \hat{\beta}_s - \frac{1}{n_2} \sum_{s=n_1+1}^n \hat{\beta}_s\end{aligned}\quad (2.36)$$

and its sampling variance:

$$\begin{aligned}\widehat{\text{Var}}(\hat{b}^{RFX}) &= \widehat{\text{Var}}(\hat{\beta}_{controls}^{RFX}) + \widehat{\text{Var}}(\hat{\beta}_{patients}^{RFX}) \\ &= \frac{\hat{\sigma}_{Gc1}^2}{n_1} + \frac{\hat{\sigma}_{Gc2}^2}{n_2}\end{aligned}\quad (2.37)$$

**Mixed-effects** If the measurement error on the subject parameter estimates is varying across subjects and not negligible by comparison to the cross-subject variability, then the homoscedasticity assumption is violated and the system (2.33) can be solved by WLS giving:

$$\begin{aligned}\hat{\beta}_{controls(v)}^{MFX} &= \frac{1}{\sum_{i=1}^{n_1} \kappa_{i(v)}} \sum_{s=1}^{n_1} \kappa_{s(v)} \hat{\beta}_{s(v)}, \\ \hat{\beta}_{patients(v)}^{MFX} &= \frac{1}{\sum_{i=n_1+1}^n \kappa_{i(v)}} \sum_{s=n_1+1}^n \kappa_{s(v)} \hat{\beta}_{s(v)},\end{aligned}\quad (2.38)$$

where  $\kappa_{s(v)}^{-1} = \begin{cases} \sigma_{G_1(v)}^2 + \frac{\sigma_{s(v)}^2}{v_{s(v)}} & \forall 1 < s \leq n_1 \\ \sigma_{G_2(v)}^2 + \frac{\sigma_{s(v)}^2}{v_{s(v)}} & \forall n_1 < s \leq n \end{cases}$

and their sample variance:

$$\begin{aligned}\widehat{\text{Var}}(\hat{\beta}_{controls}^{MFX}) &= \frac{1}{\sum_{s=1}^{n_1} \kappa_{s(v)}} \\ \widehat{\text{Var}}(\hat{\beta}_{patients}^{MFX}) &= \frac{1}{\sum_{s=n_1+1}^n \kappa_{s(v)}}\end{aligned}\quad (2.39)$$

The parameter estimate associated with the between-group contrast:  $c = [1 \ -1]$  is therefore:

$$\begin{aligned}\hat{b}^{MFX} &= \hat{\beta}_{controls}^{MFX} - \hat{\beta}_{patients}^{MFX} \\ &= \frac{1}{\sum_{i=1}^{n_1} \kappa_{i(v)}} \sum_{s=1}^{n_1} \kappa_{s(v)} \hat{\beta}_{s(v)} - \frac{1}{\sum_{i=n_1+1}^n \kappa_{i(v)}} \sum_{s=n_1+1}^n \kappa_{s(v)} \hat{\beta}_{s(v)}\end{aligned}\quad (2.40)$$

and its sampling variance:

$$\begin{aligned}\widehat{\text{Var}}(\hat{b}^{MFX}) &= \widehat{\text{Var}}(\hat{\beta}_{controls}^{MFX}) + \widehat{\text{Var}}(\hat{\beta}_{patients}^{MFX}) \\ &= \frac{1}{\sum_{s=1}^{n_1} \kappa_{s(v)}} + \frac{1}{\sum_{s=n_1+1}^n \kappa_{s(v)}}\end{aligned}\quad (2.41)$$

### 2.3.3 Fixed effects

In the preceding sections we have presented the difference between mixed-effect analyses and random-effect group analyses. Both methods are currently in use in the neuroimaging community and available in different software packages. In the early years of functional neuroimaging, another type of analysis, known as fixed-effect GLM, was common practice. In a fixed-effect analysis, several single-subject analyses were concatenated to build the group-level design matrix. This method was very sensitive but a given subject would potentially overly influence the overall group results. Nowadays, it is acknowledged that fixed-effects analyses lack generalisation and should therefore be confined to analyses in which inference is drawn for the subjects under study [Penny 2003]. In practice, this fixed-effect group analyses are almost always avoided.

## 2.4 Correction for multiple comparisons

In the massively univariate GLM, a test is performed at each voxel. In a typical GLM study, there is usually no less than 30 000 voxels and the objective is to draw inference over all voxels. Let's consider that we are thresholding the tests with a p-value such as  $p < 0.05$ , then we will have  $0.05 \times 30\,000 = 1\,500$  voxels falsely detected on average. While a threshold of  $p < 0.05$  is commonly employed in hypothesis testing, when it comes to multiple tests, such a threshold is no more appropriate. This is known as the multiple testing (also termed multiple comparisons) problem.

To overcome this limitation, several corrections have been proposed. Family Wise Error (FWE) controls the probability to have one false positive per image. The historical approach was the Bonferroni correction, but this approach is overly conservative in the presence of spatial autocorrelation of the noise such as the one observed in MRI (for a complete review of the potential causes of such correlation, see [Chumbley 2009], Table 1). Random field theory was therefore proposed as an alternative for FWE correction [Friston 1991, Worsley 1992, Friston 1994c]. While taking into account the spatial autocorrelation should increase the sensitivity (by alleviating the correction), it has however been argued that FWE correction based on random-field theory is usually equivalent to Bonferroni's correction [Logan 2004].

Also, the conservativeness of FWE correction is sometimes discussed. And a strong control on the false positive rate, as performed with FWE corrections, often leads to a reduced sensitivity. To overcome this limitation, the False Discovery Rate (FDR) has been introduced in [Benjamini 1995] as an alternative correction method in the context of multiple tests. The FDR controls the percentage of false positives in the detections.

Both FWE and FDR corrections can be applied at the voxel or the cluster level, providing a range of concurrent method to threshold a probability map while handling the multiple testing issue (for a review cf. [Nichols 2003]).



### 3 Conclusion

This chapter was dedicated to the description of the standard processing pipeline commonly used in the analysis of neuroimaging data. We first described the data preparations steps, known as the pre-processing. Then we focused on the statistical analysis with a complete description of the general model. We reviewed the main models at the subject and at the group level.

## Chapter 3

# From group to patient-specific analyses

---

This chapter introduces patient-specific analyses that aim at drawing inference at the subject level.

We first provide a brief description of the applications of such studies and outline their importance in the clinical setting in [1](#). Then, the current state-of-the-art approaches are presented divided in two parts: qualitative analyses in [2](#) and quantitative methods in [3](#).

**Contents**

<b>1</b>	<b>The need for patient-specific information . . . . .</b>	<b>57</b>
<b>2</b>	<b>Qualitative analyses . . . . .</b>	<b>57</b>
<b>3</b>	<b>Quantitative analyses . . . . .</b>	<b>57</b>
3.1	Studies without spatial information . . . . .	57
3.2	Comparing one patient to a group of control subjects . . . . .	57
3.3	Multivariate approaches . . . . .	58
3.4	Locally multivariate approaches . . . . .	60
3.4.1	The searchlight approach . . . . .	61
3.4.2	The <i>a contrario</i> approach . . . . .	62
<b>4</b>	<b>Conclusion . . . . .</b>	<b>65</b>

## 1 The need for patient-specific information

While group statistical analyses are of tremendous importance in order to understand the general mechanisms underlying a pathology, they can only reach conclusions at the population level. There is today an increasing interest towards patient-specific analyses [Commowick 2009, Maumet 2012b], also referred as *single-subject* studies [Viviani 2007a], *single case studies* [Scarpazza 2013], *individual* analyses [Colliot 2006] or *single image tests* [Viviani 2007b], that aim to draw conclusions at the patient level.

Inference at the patient level is highly desirable in order to perform a diagnosis or provide a personalised treatment, both of these operations being necessarily made at the patient level. On top of this, it is also important to note that some pathologies are intrinsically not well-suited to perform group voxel-wise analyses. It is for instance the case of brain tumours [Warmuth 2003, Chawla 2007, Weber 2006, Sugahara 2000], stroke [Huck 2012, Bokkers 2011, Ferré 2012, Wheeler 2013], refractory epilepsy characterised by dysplasia and heterotopia [Colliot 2006, Wilke 2003b, Focke 2008, Huppertz 2005], or multiple sclerosis lesion [Commowick 2008], that exhibit a different pattern of spatial abnormalities for each patient.

## 2 Qualitative analyses

Probably the simplest method to deal with single-subject data, the qualitative approach is mainly employed by clinicians in case reports and deals with global spatial patterns (e.g. recent ASL studies: [Chen 2012b, Zaharchuk 2012]). In this type of analysis, an expert, usually a clinician, discusses the spatial patterns of hyper- and hypo-signals and links his/her observations with external clinical parameters, such as the patient outcome. While qualitative observations are interesting as a first global information, quantitative indices are usually preferred.

## 3 Quantitative analyses

### 3.1 Studies without spatial information

The first quantitative approach that can be performed in order to analyse patient-specific data is to measure a certain index per subject and compare it to a clinical parameter. For example in brain tumours, the level of CBF or CBV in the most malignant part of the tumour (manually identified by an expert clinician) has been proven to be predictive of the tumour grade [Gaa 1996, Weber 2006, Sugahara 2000].

### 3.2 Comparing one patient to a group of control subjects

The simplest approach to perform a patient-specific analysis while retaining spatial information is to perform a *one-versus-many* parametric test in which a single

patient map is compared to a set of maps from a reference group. The *one-versus-many* GLM is simply a sub-type of between-group analysis (described in chapter 2, section 2.3.2), in which one of the group is made of one subject only (the patient of interest). This particular model will be developed in greater details in chapter 8.

This approach has mainly been used for the detection of focal cortical dysplasias in epileptic patients using voxel-based morphometry [Wilke 2003b, Huppertz 2005, Colliot 2006, Focke 2008, Bannier 2012]. In ASL, [Petr 2013] used a z-score (which is equivalent to a GLM in which the control group estimate and the variances of the model are considered as known *a priori*) to outline perfusion abnormalities in an epileptic patient.

While, in principle, performing a *one-versus-many* analysis is no more complex than dealing with a GLM, special care has to be taken when interpreting the results of such analyses. Indeed, an inflation of the false positive rate, sometimes leading to invalidity of the tests has been described [Scarpazza 2013, Viviani 2007a]. This phenomenon is explained by the fact that unbalanced designs, and in particular *one-versus-many* analyses, tend to be more sensitive to deviations from the Gaussianity assumption [Viviani 2007a]. Indeed, in between-group analyses featuring a large number of subjects, the averaging across subjects in each group tends to make the data more Gaussian (thanks to the central limit theorem). Semi-parametric tests, with a recalibration of the probabilities have been proposed to overcome this limitation [Viviani 2007b].

In conclusion, the massively univariate comparison of a patient to a group of reference subjects is an option to study patient-specific atypical patterns. When using this type of design, and in particular with the GLM, the false positive rate must be carefully assessed, as violations of the model assumptions are more prone to induce invalidity.

### 3.3 Multivariate approaches

As opposed to the mainstream massively univariate statistics approach presented in the previous section, an emerging field tries to make use of multivariate patterns to draw inference. Fig. 3.1, from [Klöppel 2012], illustrates the concept of multivariate statistics. In this example, the aim is to distinguish the patients suffering of Alzheimer's disease from the healthy controls. In this artificial sample, the two groups are perfectly separable in the 2-dimensional space. It would however be impossible to differentiate patients from control based on 1-dimensional analyses (i.e. using only the anterior or posterior volume of the hippocampus). Multivariate approaches can therefore be more sensitive than univariate approaches. But, this comes at the cost of specific methodological and technical issues that will be discussed in the following.

Machine learning techniques are more and more used in the field of brain imaging (for a review, see [Poldrack 2008, Yang 2012]). These methods are particularly interesting in the context of patient-specific analyses as they aim at maximising the prediction accuracy for new observations.

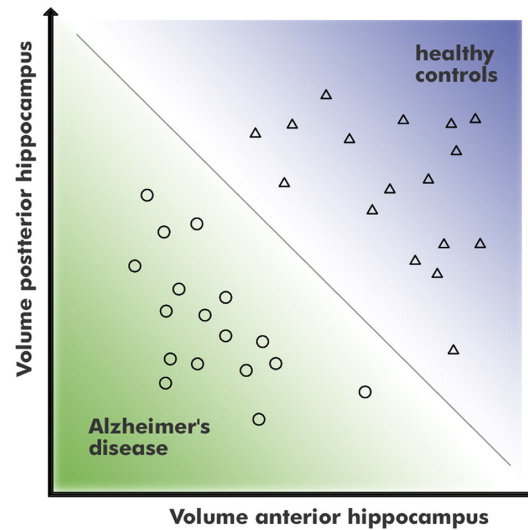


Figure 3.1: Concept of multivariate classification in two dimensions. Reprinted from [Klöppel 2012], Copyright (2013), with permission from Elsevier.

In multivariate analyses of neuroimaging data, each voxel represents a dimension and there is usually less than one hundred observations for each dimension (voxel). Standard multivariate statistics are impractical in whole-brain analyses due to the large number of dimensions (from 10 000 to 100 000 voxels) by comparison to the number of observations (from 10 to 100 subjects) [Poldrack 2011, Mourão Miranda 2011]. On the other hand, classifiers from the field of machine learning are mathematical tools that are particularly suited to deal with a large number of dimensions and few observations. In pattern recognition, the dimensions are referred to as features. Machine learning has been applied to neuroimaging data, and in particular to fMRI experiments for various purposes, including mental state decoding, stimulus prediction, group memberships prediction [Poldrack 2008, Yang 2012]. In the context of this section, we will restrain our discussion to the identification of group memberships.

Briefly, in a typical machine learning experiment, the data is first split into a test set and a training set. The parameters of the selected classifier, most usually a Support Vector Machine [Yang 2012, Poldrack 2011], are fitted on the training sets so that the accuracy is maximised in this dataset. The trained classifier is then used to predict the group labels of the data in the test set. The training and testing operations are iterated several times using different test and training set and the final error rate associated with the classifier is averaged across the iterations. The process of training and testing the classifier on separate datasets is known as cross-validation and is required in order to achieve a correct estimate of the prediction accuracy [Kriegeskorte 2009]. A leave-one-out cross-validation refers to the case where the test set is made of one subject only, and N-fold cross-validation to the case in which the test set contains  $1/N$  of the initial observation. An illustration of

a 10-fold cross-validation is displayed in fig. 3.2.

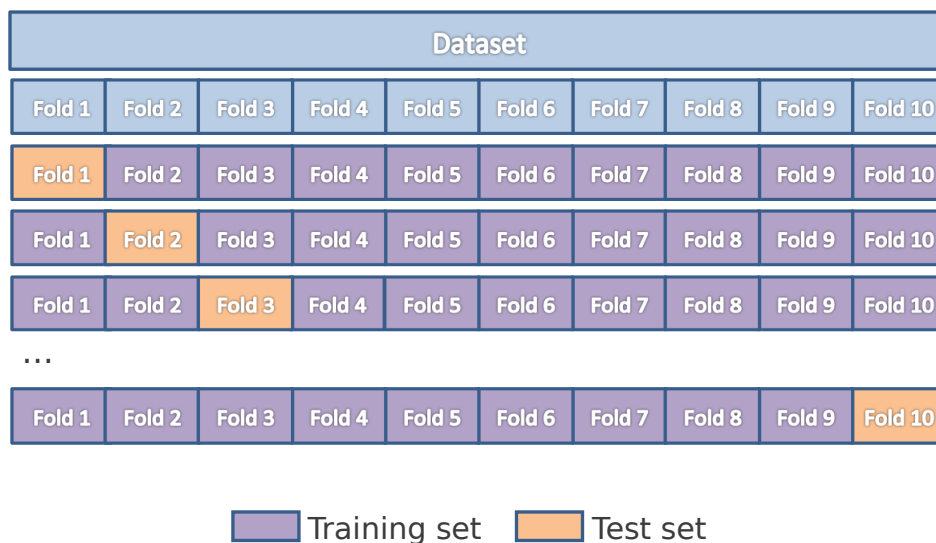


Figure 3.2: 10-fold cross-validation.

In practice, there is usually a feature selection step that aims at reducing the dimensionality of the feature space. If the feature selection step is based on an *a priori* knowledge, it can be performed prior to the cross-validation. On the other hand, the whole brain can be considered as the feature space, and a feature reduction techniques can be embedded with the classifier (for example recursive feature elimination [Guyon 2002] with the support vector machine) [Poldrack 2011]. With this latter approach, it is very important that the test set is not included in the feature selection process to avoid a selection bias [Kriegeskorte 2009], an issue that is not specific to neuroimaging (e.g. in microarray analysis [Ambroise 2002]).

Machine learning has been used in the context of neuroimaging based on different type of data, including anatomical and functional MRI. Classifiers have outlined interesting patterns predicting the patient diagnosis [Mourão Miranda 2011, Fu 2008] or future response to treatment [Costafreda 2009]. For more applicative examples, the interested reader is referred to a recent review: [Klöppel 2012].

While machine learning classifiers have a great potential for identifying interesting new patterns, the massively univariate approach is still prevailing. Some methodological issue, in particular, how to appropriately select the features in fMRI are still under discussion [Mumford 2012]. On top of this, these approaches remain computationally intensive, in particular when one is interested in multimodal data analysis, where the number of features is multiplied by the number of modalities.

### 3.4 Locally multivariate approaches

Lying in between the massively univariate and the fully multivariate approaches, locally multivariate methods focus on multivariate patterns found in the neigh-

bourhood of each voxel. The searchlight approach [Kriegeskorte 2006] is the most recognised in the neuroimaging literature, while the *a contrario* approach [Desolneux 2003] comes from the computer vision community. We will see that these two approaches have a lot in common.

### 3.4.1 The searchlight approach

The searchlight approach introduced in [Kriegeskorte 2006] works by computing a multivariate analysis at each brain location based on its neighbourhood. Fig. 3.3, from [Kriegeskorte 2007a], illustrates the concept of *searchlight*. A spherical region of interest is centred sequentially at each location in the brain in order to compute a locally multivariate metric that is then affected to the centre voxel.

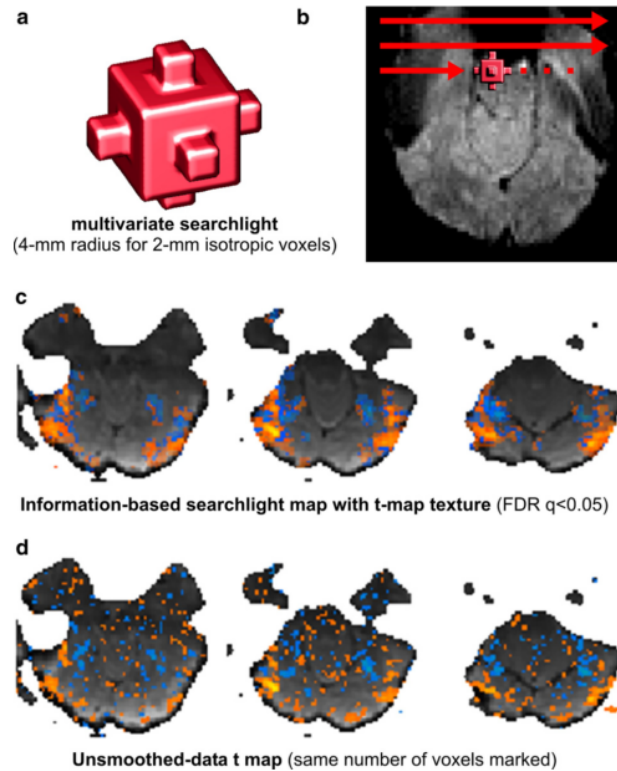


Figure 3.3: Principle of a searchlight analysis (a, b) and examples of activations obtained with the searchlight approach (c) and the massively univariate GLM (d). Reprinted from [Kriegeskorte 2007a], Copyright (2013), with permission from Elsevier.

At its inception, two multivariate metrics were considered: a Euclidean distance (computed as the sum of the squared t-statistic in the searchlight) and a Mahalanobis distance (taking into account the spatial autocorrelation of the noise) [Kriegeskorte 2006, Kriegeskorte 2007b]. Then this approach was extended to machine learning classifiers such as the support vector machine (for recent ex-



amples, see [Uddin 2011, Weygandt 2012]). Recently, the searchlight was also used to provide regularised estimates [Feng 2013].

As illustrated in fig. 3.3, the searchlight has the potential to provide detection maps that are spatially more coherent than an unsmoothed massively univariate approach. The authors of the searchlight also make the emphasis on the fact that this analysis is information-based as opposed to standard activation-based approaches [Kriegeskorte 2007a, Kriegeskorte 2007b]. The main idea is that smoothing the data, as usually done in a pre-processing step in massively univariate analyses, is optimal to increase the SNR of large activation patterns but is detrimental for fine-activity patterns.

### 3.4.2 The *a contrario* approach

The *a contrario* approach is a statistical framework that comes from the field of computer vision [Desolneux 2003]. In this context, the *a contrario* approach is usually defined by three components:

1. The definition of a *background model*;
2. The estimation of a region-based probability (under the background model);
3. The correction for multiple comparisons (using the *number of false alarms*).

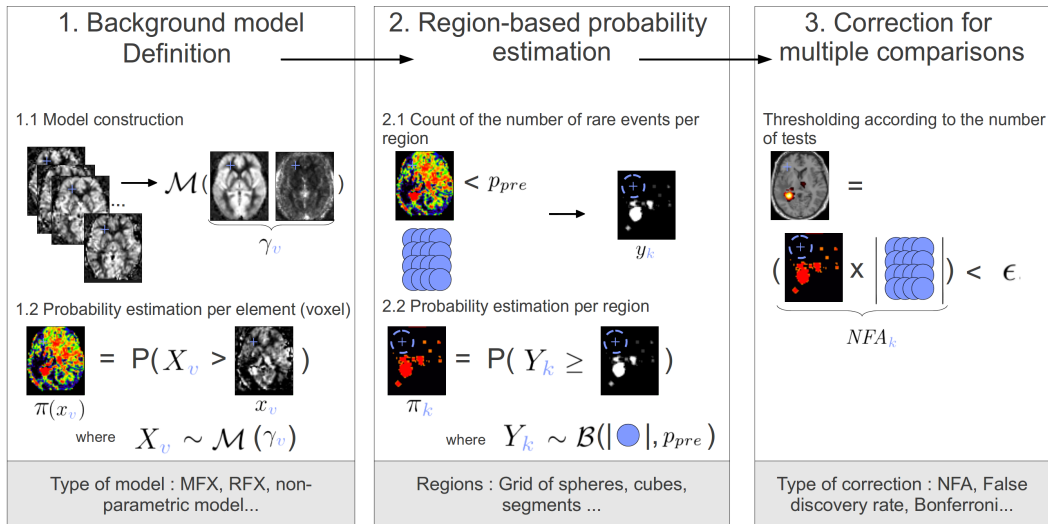


Figure 3.4: Overview of the *a contrario* approach: background model definition, region-based probability estimation and correction for multiple comparisons.

The term *a contrario* was chosen in order to outline that the detections are performed thanks to the definition of a model of the background, i.e. areas that should not be detected, instead of modelling the object of interest, which is common practice in the computer vision community.

**Definition of a background model** In the first step, a model of the background is defined (cf. fig. 3.4, 1<sup>st</sup> panel). This model can represent the background noise or be learned from a control population. The structures of interest are later detected as voxels deviating from this model. Let  $X_v$  be the random variable representing the value observed at the voxel  $v$  of the image of interest. Under the background model  $\mathcal{M}$ , depending on the set of parameters  $\gamma$ , we have:

$$X_v \sim \mathcal{M}(\gamma_v). \quad (3.1)$$

Given  $x_v$ , the value observed at voxel  $v$ , each voxel of the image can be associated with a probability under the background model:

$$\pi(x_v) = \Pr(X_v \geq x_v), \text{ where } X_v \sim \mathcal{M}(\gamma_v) \quad (3.2)$$

Here, we focused on increases with respect to the background model, if, instead, decreases are of interest  $\Pr(X_v > x_v)$  is substituted by  $\Pr(X_v < x_v)$  in (3.2).

*Rare events* are defined as observations contradicting the background model. A rare event occurs at voxel  $v$  if the probability (under the background model) to observe such value, or a more extreme, is smaller than a pre-defined threshold  $p_{PRE}$ . The initial voxel-wise probability map is thus thresholded:

$$k_v = \begin{cases} 1 & \text{if } \pi_v \leq p_{PRE} \\ 0 & \text{otherwise} \end{cases} \quad (3.3)$$

to produce a binary map.

**Estimation of a region-based probability** In the second step, the probability estimation is extended from voxels to regions of the image (cf. fig. 3.4, 2<sup>nd</sup> panel). This allows to take into account the spatial neighbourhood of each voxel in the statistical analysis. To this aim, a grid of regions is specified in the image. If an *a priori* on the shape of the structures of interest is known the regions are selected accordingly. For instance, in the first *a contrario* application, which aimed at detecting alignments, segments of variable sizes were chosen as regions of analysis. In the absence of *a priori*, two methods have been proposed: either using a non-supervised clustering algorithm [Rousseau 2008] or defining a sphere (or cube) centred at each voxel [Aguerrebere 2009, Rousseau 2007]. In the case of spheres, this set of regions can be viewed as a searchlight.

Then, the number of rare events in region  $r$ ,  $l_{(r)}$ , is determined by:

$$l_{(r)} = \sum_{v \in R_r} k_v \quad (3.4)$$

where  $R_r$  is the set of voxels belonging to region  $r$ . Let  $L_{(r)}$  be the random variable representing the number of rare events observed in region  $r$ . Then, we have:

$$L_{(r)} = \sum_{v \in R_r} K_v \text{ where } K_v \sim \text{Bern}(p_{PRE}) \quad (3.5)$$

where  $\text{Bern}(j)$  is a Bernoulli distribution with probability  $j$ .

Each region  $R_k$  is made of  $e$  voxels. While some attempts have been made in order to model coloured noise [Grosjean 2008, Myaskouvskey 2013], a white noise is usually assumed [Desolneux 2003, Veit 2006, Aguerrebere 2009]. In case of independence, the probability of  $L_{(r)}$  is given by the Binomial distribution  $\mathcal{B}$  with parameters:  $e$ , the number of tests, and  $p_{pre}$  the rare event probability. Given  $y_k$  the observed number of rare events in region  $R_k$ , this region is associated with the probability:

$$\pi_{R_r} = \Pr(L_{(r)} \geq l_{(r)}), \text{ where } L_{(r)} \sim \text{B}(e, p_{PRE}). \quad (3.6)$$

In a few papers, a variant was proposed in which the region-based probabilities  $\pi_k$  were directly estimated without the preliminary rare event count computation [Aguerreberere 2009, Grosjean 2008].

At the end of the second step, each region of the image is associated with a probability (under the background model). As in the searchlight approach, the probability associated to each region can be affected to its centre voxel to produce a voxel-wise probability map.

**Correction for multiple comparisons** The third and last step aims at outlining the detections while taking into account multiple comparisons (cf. fig. 3.4, 3<sup>rd</sup> panel). Indeed, since a test is performed for each region leading to  $numReg$  tests ( $numReg$  being the number of regions), we need to control for the overall number of false positives. Originally, the *a contrario* approach controls the expectation of the number of regions falsely detected in the background. The control is done by computing the Number of False Alarms (NFA) that corresponds to the average number of false detections per image, according to the background model:

$$NFA_k = numReg \pi_{R_r}. \quad (3.7)$$

Given  $\varepsilon$ , the average number of false detections tolerated by the user, the regions which verify

$$NFA_k < \varepsilon \quad (3.8)$$

are finally outlined as detections.

In practice, to alleviate the influence of the preselected  $p_{pre}$  parameter in the second step, a set of p-values  $\{p_1, p_2, \dots, p_\alpha\}$  is usually tested. Then a probability  $\pi_{R_r, j}$  is associated to each region  $R_r$  and each p-value  $p_j$  and the NFA is computed as:

$$NFA_r = \alpha numReg \min_j(\pi_{R_r, j}) \quad (3.9)$$

At the end of the third step, each region of the image is detected or undetected (under the background model). As in the searchlight approach, the detection of each region can be affected to its centre voxel to produce a voxel-wise detection map.

**Link with standard hypothesis testing and searchlight approach** As outlined by [Rousseau 2007], the *a contrario* approach can be related to standard hypothesis testing by noting that the definition of the background model is equivalent to the selection of a null hypothesis  $H_0$ . On top of this, the thresholding on the NFA is in fact equivalent to a Bonferroni correction where the expected number of false alarms is expressed in number of regions per image. In the *a contrario* approach, it makes sense to set  $\varepsilon > 1$  whereas this is impossible with Bonferroni correction [Rousseau 2007].

The originality of the *a contrario* approach stands in its capability to integrate the multivariate information in pre-defined regions of analysis. When spheres are used as regions of analysis and the region-based probability is affected to its centre voxel, the *a contrario* approach can be viewed as a special type of searchlight in which the multivariate metrics is the size of an excursion set.

## 4 Conclusion

In this chapter, we introduced patient-specific analyses. We presented the context that justifies such analyses and we provided a review of the state-of-art methods to draw inference at the subject level. We outlined the importance of checking the false positive rate when performing a *one-versus-many* GLM. Then, we focused on multivariate approaches and presented machine learning classifiers and locally multivariate procedures.



## Part II

# Language impaired regions in specific language impairment: a group study using BOLD functional MRI



## Chapter 4

# The “Neuroimaging in Dysphasia” project

---

Typical Specific Language Impairment, or structural Dysphasia, is a developmental language disorder affecting the core aspects of language. The “Neuroimaging of Dysphasia” project was initiated in 2005, in Rennes taking advantage of a multidisciplinary group made of clinicians, research scientists in medical imaging and linguistic and a methodologist. The primary goal of this project was to investigate the functional aspects of structural dysphasia based on functional [fMRI](#). In this work, I contributed to the methodological choices, programmed the processing pipelines, applied these tools to the available data, performed the statistical analysis and participated in the writing of the papers. The secondary goal was to study the anatomy with standard anatomical [MRI](#) as well as diffusion [MRI](#).

In this introductory chapter, we present the context of the “Neuroimaging in Dysphasia” project. The following two chapters present the results obtained on functional [fMRI](#).



**Contents**

<b>1</b>	<b>Context</b> . . . . .	<b>71</b>
1.1	Specific Language Impairment . . . . .	71
1.2	Brain and language . . . . .	72
1.2.1	Cortex . . . . .	72
1.2.2	Lateralisation of language . . . . .	74
1.2.3	Communication networks . . . . .	74
<b>2</b>	<b>The “Neuroimaging in Dysphasia” project</b> . . . . .	<b>75</b>
2.1	A brief chronology . . . . .	75
2.2	Imaging protocol . . . . .	76

## 1 Context

### 1.1 Specific Language Impairment

**Specific Language Impairment (SLI)**, also known as dysphasia, is part of the developmental language disorders and, as such, is:

- severe and persistent: **SLI** is not a language delay;
- specific: although **SLI** may be associated with other disorders, it impacts oral language (or related processes) only;
- developmental: **SLI** is not related to an acquired brain injury.

Specific Language Impairment is therefore defined by negation (not a language delay, not impacting intellectual capacities, not acquired. . . ) which makes it difficult to diagnose.

In the literature, the denomination *Specific Language Impairment* is somewhat loosely defined and used to name a subclass of the developmental language disorders that can be more or less broad. In order to define the terminology that will be later used in this manuscript, fig. 4.1 provides a diagram of the main language disorders and their relationship along with the aspect of language that is affected by each trouble. This view is in line with the vocabulary employed by other authors including [Brun 2003, Bishop 2006].

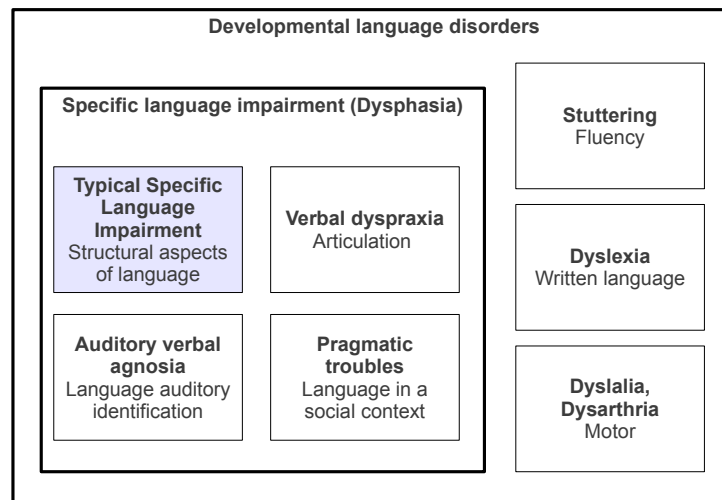


Figure 4.1: Developmental language disorders diagram.

Developmental language disorders form therefore a broad and heterogeneous group that includes: affections of the oral language (specific language impairment), of the fluency (stuttering), of the written language (dyslexia) and motor affections (dyslalia, dysarthria). Specific Language Impairment forms itself an heterogeneous

subgroup including affection of the core aspects of language (typical specific language impairment), articulatory troubles (dyspraxia), disorders related to the social aspects of language (pragmatic troubles) and affection of auditory language identification (verbal agnosia).

In the context of our study, we focused on typical specific language impairment, also termed structural language impairment. This pathology has a prevalence of less than 1% in the general population. **Typical Specific Language Impairment (T-SLI)** is characterised by an affection of the core aspects of language: phonology (managing phonological differentiations), lexicon (managing words differentiation), morphology (formation of derivative words) and syntax (managing sentences construction both for expression and comprehension).

## 1.2 Brain and language

### 1.2.1 Cortex

Fig. 4.2 gives an overview of the three major cortex components involved in language: Broca’s area, Wernicke’s area and Geschwind’s territory. In his pioneering work, Broca outlined the importance of the inferior frontal gyrus in speech production [Broca 1861]. A decade later, Wernicke’s outlined the importance of the superior temporal gyrus in speech understanding along with the major role of the connexions between the so-called Broca’s and Wernicke’s areas [Wernicke 1874].

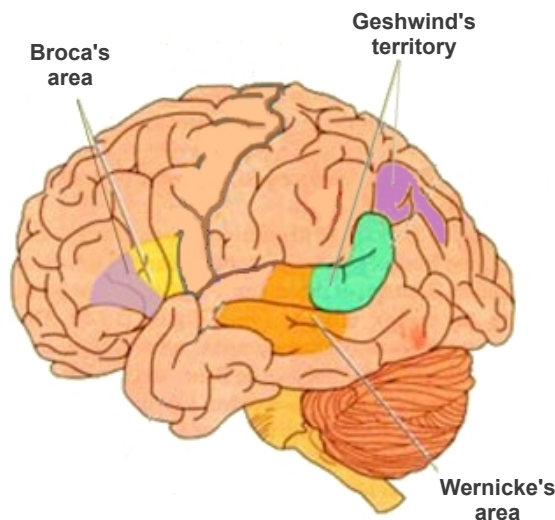


Figure 4.2: Main brain areas involved in language, adapted from *The Brain From Top To Bottom* website<sup>1</sup>

<sup>1</sup>[http://thebrain.mcgill.ca/flash/i/i\\_10/i\\_10\\_cr/i\\_10\\_cr\\_lan/i\\_10\\_cr\\_lan.html](http://thebrain.mcgill.ca/flash/i/i_10/i_10_cr/i_10_cr_lan/i_10_cr_lan.html)

About a century later, Geschwind revisited these works and proposed the Geschwind-Wernicke's model as illustrated in fig. 4.3.

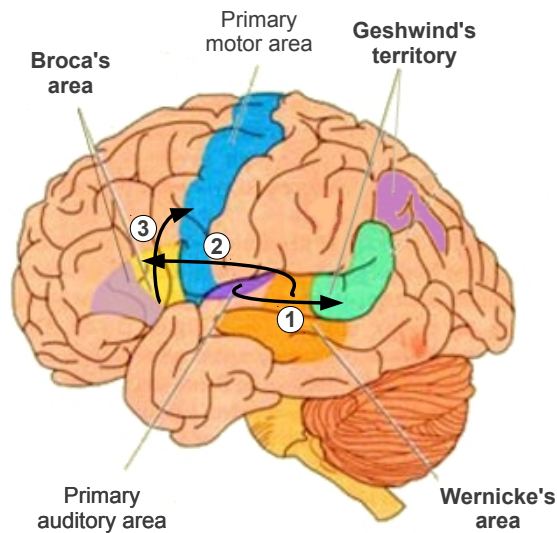


Figure 4.3: Geschwind-Wernicke's model of language production following an auditory stimulus, adapted from [The Brain From Top To Bottom](#) website<sup>1</sup>

While this pioneering work clearly demonstrated the involvement of the frontal and parieto-temporal regions in language, the Geschwind-Wernicke's model was a simplification and is nowadays outdated. In particular the very serial view proposed by the Geschwind-Wernicke's model is now replaced by cooperative models proposing a more parallel view (e.g. [Abrams 2012, Radanovic 2011]).

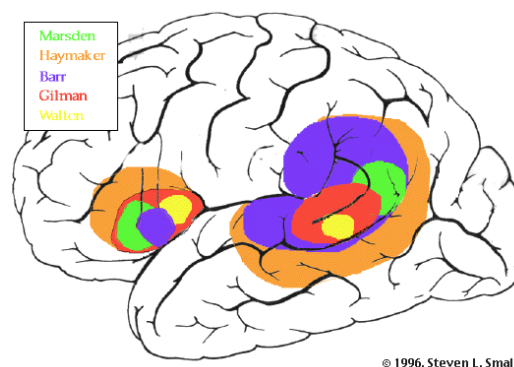


Figure 4.4: Variability of Broca's (in the frontal lobe) and Wernicke's (in the temporal and parietal lobes) areas from [The Brain From Top To Bottom](#) website<sup>1</sup>.

It is also worth noting that the precise anatomical location of these three cortical areas is highly variable across studies (cf. fig 4.4). In particular, Geschwind's

territory and Wernicke’s area are often merged under the denomination Wernicke’s area. This anatomical variability is probably partly explained by the fact that these three cortical entities are, in fact, compounded of several sub-units in charge of different processing (e.g. [Radanovic 2011]).

### 1.2.2 Lateralisation of language

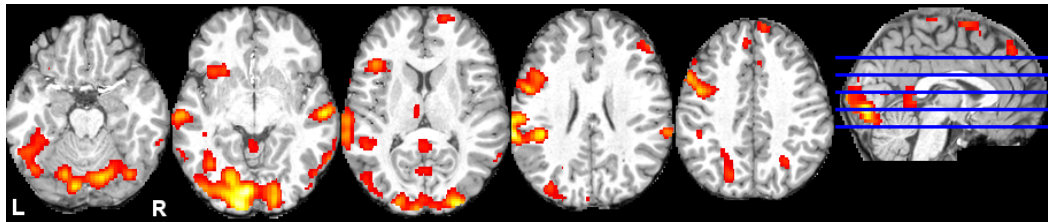


Figure 4.5: Activation in one subject during a language paradigm (Phon-seg paradigm described hereafter). The language activation patterns observed in the IFG (Broca’s area) and STG (Wernicke’s area) are clearly left-lateralised.

Brain lateralisation (or brain dominance) refers to the fact that a particular function is predominantly taken in charge by one of the two brain hemispheres. In the case of language, the dominant hemisphere is left in 95% of right-handers and 60% in left-handers [Taylor 1990]. Fig. 4.5 provides an example of left-lateralised patterns of activation. The degree of lateralisation is however known to be dependent upon the task.

### 1.2.3 Communication networks

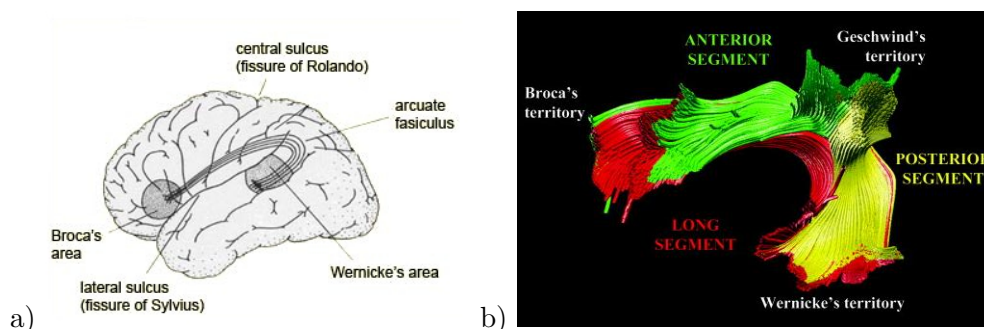


Figure 4.6: The arcuate fasciculus: historical view (a) from [The Brain From Top To Bottom](#) website<sup>2</sup> and modern three-segment view (b), reprinted from [Catani 2005], Copyright (2013), with permission from Wiley.

In his early works, Wernicke’s identified the importance of the communication pathways between brain regions involved in language [Wernicke 1874]. The historical fasciculus of white matter tracts identified as language-related is the arcuate fasciculus. This fibre bundle connects the inferior frontal gyrus (Broca’s region) to

the superior temporal gyrus (Wernicke’s region) passing through the supramarginal region (Geschwind’s territory) as illustrated in fig. 4.6.

Thanks to modern neuroimaging techniques, the arcuate fasciculus has been divided in three parts: the anterior (Broca’s area  $\leftrightarrow$  Geschwind’s territory), posterior (Wernicke’s area  $\leftrightarrow$  Geschwind’s territory) and long (Wernicke’s area  $\leftrightarrow$  Broca’s area) segments [Catani 2005].

The involvement of an additional ventral pathway is nowadays widely accepted [Mandonnet 2007], and the exact contribution of each pathway in language processing is still under investigation [Rolheiser 2011].

## 2 The “Neuroimaging in Dysphasia” project

### 2.1 A brief chronology

Based on the observation that SLI had been very little studied in neuroimaging and taking advantage of the multidisciplinary group created in Rennes, the “Neuroimaging in Dysphasia” project started in 2005. This multidisciplinary group, created in Rennes, was composed of four clinicians: Catherine Allaire (neuropediatrics Department, Centre Référent des Troubles Sévères du Langage et des Apprentissages, Pontchaillou Hospital), Jean-Christophe Ferré (Neuroradiology Department, Pontchaillou Hospital), Catherine Tréguier (Radiopediatrics Department, South Hospital), Arnaud Biraben (Neurology Department, Pontchaillou Hospital), three research scientists: Elisabeth le Rumeur (Radiology Department, Pontchaillou Hospital), Christian Barillot (VisAGeS team, CNRS, University of Rennes 1, Inria, INSERM) and Pierre Jannin (VisAGeS team, INSERM, University of Rennes 1, Inria), one methodologist: David Veillard (Pontchaillou Hospital) and two engineers: Alexandre Abadie (VisAGeS team, University of Rennes 1, Inria, INSERM) and myself, under the supervision of a neuropsychologist, research associate in linguistic: Clément de Guibert (University of Rennes 2).

In 2006-2008, Clément de Guibert, the principal investigator on this project, was appointed on Inria secondment to the VisAGeS team, for two years. During this period, the imaging protocol was built in collaboration with the other members of the project. The acquisitions started in the first months of 2008.

Two journal publications, reported in chapters 5 and 6 have arisen from this fruitful collaboration. These first two communications focus on functional data, one journal paper based on anatomical data is in preparation. My contribution in this work was to participate in the methodological choices, program the processing pipelines, apply these tools to the available data, perform the statistical analysis and participate in the writing of the papers.

Analysis of the diffusion data was performed during two internships. First, in 2009, Arnaud Le Cavernec in a three-months project, under my supervision, employed the tool available in the VisAGeS team to extract the arcuate fasciculus

---

<sup>2</sup>[http://thebrain.mcgill.ca/flash/d/d\\_10/d\\_10\\_cr/d\\_10\\_cr\\_lan/d\\_10\\_cr\\_lan.html](http://thebrain.mcgill.ca/flash/d/d_10/d_10_cr/d_10_cr_lan/d_10_cr_lan.html)

and looked at different metrics along these tracts. Then in 2012, Emmanuel Vallée did his Masters project on the analysis of the diffusion data. He built a pipeline to automatically extract diffusion bundles from a set of atlas-based **Region of Interests (ROIs)**. He outlined interesting between-group differences. I was involved in his supervision along with Olivier Commowick, Christian Barillot and Aymeric Stamm.

## 2.2 Imaging protocol

The imaging protocol included a **T1 weighted (T1w)** and a **T2w FLAIR** anatomical images, a set of four **fMRI** images corresponding to four distinct language paradigms and a sixteen-direction diffusion image.

Fig. 4.7 provides an overview of the four language paradigms. Each paradigm was implemented with six blocks (three action blocks / three resting blocks) of 27 s each. Two paradigms, Category and Denomination, were previously described in the literature and based on auditory stimuli. The two remaining paradigms were originally developed in the context of the “Neuroimaging in Dysphasia” project and based on visual stimuli.

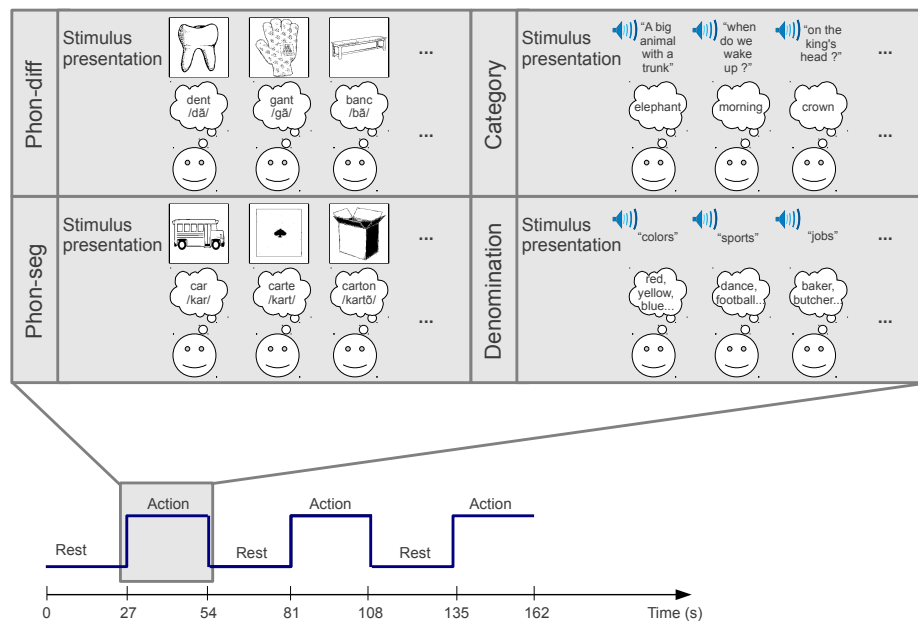


Figure 4.7: The four language paradigms used in the “Neuroimaging in Dysphasia” project.

More details on the implementation of this task panel for language mapping in children are provided in chapter 5. The patterns of activations and language lateralisation in **T-SLI** patients by comparison to a group of typically developing children are studied, using the same protocol, in chapter 6.

## Chapter 5

# Functional **fMRI** language mapping in children

---

This chapter is presented in the form of a paper. This work was realised in collaboration with Clément de Guibert (cf. chapter 4) and was published in Neuroimage [de Guibert 2010].

### Abstract

In the context of presurgical mapping or investigation of neurological and developmental disorders in children, language **fMRI** raises the issue of the design of a tasks panel achievable by young disordered children. Most language tasks shown to be efficient with healthy children require metalinguistic or reading abilities, therefore adding attentional, cognitive and academic constraints that may be problematic in this context. This study experimented a panel of four language tasks that did not require high attentional skills, reading, or metalinguistic abilities. Two reference tasks involving auditory stimulation (words generation from category, “category”; auditory responsive naming, “definition”) were compared with two new tasks involving visual stimulation. The latter were designed to tap spontaneous phonological production, in which the names of pictures to be named involve a phonological difference (e.g. in French poule/boule/moule; “phon-diff”) or change of segmentation (e.g. in French car/car-te/car-t-on; “phon-seg”). Eighteen healthy children participated (mean age:  $12.7 \pm 3$  years). Data processing involved normalizing the data via a matched pairs paediatric template, and inter-task and region of interest analyses with laterality assessment. The reference tasks predominantly activated the left frontal and temporal core language regions, respectively. The new tasks activated these two regions simultaneously, more strongly for the phon-seg task. The union and intersection of all tasks provided more sensitive or specific maps. The study demonstrates that both reference and new tasks highlight core language regions in children, and that the latter are useful for the mapping of spontaneous phonological processing.



**Contents**

<b>1</b>	<b>Introduction</b> . . . . .	<b>79</b>
<b>2</b>	<b>Subject and methods</b> . . . . .	<b>81</b>
2.1	Subjects . . . . .	81
2.2	Experimental paradigms . . . . .	82
2.2.1	General technical implementation . . . . .	82
2.2.2	Auditory lexico-semantic tasks (reference tasks) . . . . .	82
2.2.3	Visual phonological tasks (new tasks) . . . . .	83
2.2.4	Preparation before the scanner . . . . .	83
2.3	Data acquisition . . . . .	83
2.4	Data processing . . . . .	84
<b>3</b>	<b>Results</b> . . . . .	<b>86</b>
3.1	Reference auditory tasks . . . . .	86
3.2	New visual tasks . . . . .	88
3.3	Whole panel analysis . . . . .	89
<b>4</b>	<b>Discussion</b> . . . . .	<b>93</b>
4.1	Methodological issues . . . . .	94
4.2	Reference lexico-semantic auditory tasks . . . . .	95
4.3	New phonological visual tasks . . . . .	96
4.4	Whole panel . . . . .	97

## 1 Introduction

fMRI is a safe and non-invasive method for determining the brain functional localisation and lateralisation of language in children, which is an important issue both for paediatric clinical applications and research purposes [Gaillard 2001a, O’Shaughnessy 2008, Wilke 2003a]. From a clinical perspective, it is considered that fMRI may replace or serve as an important adjunct to the invasive intracarotid amobarbital (Wada test) or direct cortical stimulation mapping procedures, in order to delineate the eloquent cortex to be spared in children who are candidates for surgical resection [O’Shaughnessy 2008]. In line with studies showing the utility of presurgical fMRI language mapping in adults [Binder 1996, Gaillard 2002, Roux 2003, Rutten 2002, Tie 2008, Tie 2009], studies have reported the potential utility of this procedure in the even more crucial context of childhood [Anderson 2006, Gaillard 2000, Gaillard 2001b, Hertz-Pannier 1997, Holland 2001, Wilke 2005, Wilke 2006]. For research purposes, fMRI allows to specify the normal functional development of the brain [Durstun 2006] and, particularly, the functional development of language during childhood [Holland 2007, Sachs 2008, Sachs 2003, Gaillard 2006]. One promising perspective is the investigation of the neural bases of developmental disorders, whose aetiology remains largely unknown [Berl 2006, Frith 2006, O’Shaughnessy 2008, Wilke 2003a], including childhood developmental language disorders, [Dick 2008, Friederici 2006a, Rapin 2003].

A major issue of fMRI, in this context as in others, is its sensitivity and specificity, i.e. the ability to draw a comprehensive as well as selective picture of the essential language brain network [Medina 2007, Tie 2008, Tie 2009]. The “ideal” procedure may thus highlight all core language areas but only core language areas, by minimizing the false negatives and false positives. Therefore, owing to the complexity of language, the choice of activation tasks is crucial, as any single language task is unlikely to engage all aspects of language and exclusively involve language processing [Gaillard 2004, Ramsey 2001, Tie 2008]. One way of bypassing this difficulty is to use a panel of different tasks targeting distinct aspects of language. First, the union of activations from several tasks increases the sensitivity by providing a more comprehensive picture of the overall network [Deblaere 2002, Gaillard 2001b, Gaillard 2004, Holland 2007, Ramsey 2001, Roux 2003, Tie 2008, Wilke 2005, Wilke 2006]. Secondly, the intersection of activations across several tasks, as obtained by conjunction analysis [Friston 2005a, Nichols 2005], increases the specificity by neglecting non language-specific brain areas that only participate in, but are not essential to language [Tie 2008]. Furthermore, the use of several tasks allows to focus separately on particular parts of the language network [Gaillard 2004, Wilke 2005, Wilke 2006].

In the context of paediatric fMRI, another major issue is that many language paradigms shown to be useful in adult studies are not well suited for children, and that child-specific tasks have to be specially designed [O’Shaughnessy 2008, Wilke 2003a, Wilke 2006]. This may be even more problematic for children suffering from neurological or developmental disorders, especially language disorders. For example, classical tasks such as verbs-to-word, words-to-letter, and words-to-

phoneme generation, involve an understanding of what is an “action word”, a basic knowledge of written language, and the ability to explicitly segment oral utterances, respectively. Such tasks may not be achievable for young and/or disordered children. Similar difficulties arise with metalinguistic tasks such as rhyme, syntactic, or semantic decision tasks. These latter tasks, in addition, require explicit forced-choice analysis and judgement, as opposed to spontaneous discourse, and are likely to involve undesired attentional effects even in adults (e.g. [Crinion 2003]).

Following previous fMRI studies using a panel of language tasks with children [Gaillard 2001b, Gaillard 2004, Holland 2007, Wilke 2005, Wilke 2006], the aim of our study was to carry out an fMRI investigation of the essential language brain network that may be accessible to young disordered children by avoiding reading, high attentional, or metalinguistic requirements. In addition, we aimed to use auditory and visual stimuli delivery, to solicit language comprehension and production, and to target lexico-semantic and phonological processing within the whole procedure.

We first choose from the literature two reference lexico-semantic tasks with auditory stimulation that have been shown to elicit in children distinct and selective activations within the language network. In the first task (words generation from category, hereafter “category”; see [Gaillard 2003]), children have to name several examples of a given category. Gaillard et al. reported, as similarly found in adults, an activation in the left IFG without consistent activation in the STG and Middle Temporal Gyrus (MTG). In the second reference task (responsive naming task, hereafter “definition”; see [Balsamo 2002]), which resembles riddles, children have to name the concept corresponding to a short verbal description. Balsamo et al. reported a highly left-lateralised activation of the STG and MTG without consistent activation of the IFG.

In addition, we aimed to design two new tasks with visual stimulation able to involve spontaneous phonological processing and highlight distinct parts of the phonological brain network. According to current knowledge, phonological processing implies a left distributed network encompassing frontal and temporo-parietal language areas [Buchsbaum 2001, Burton 2001, Démonet 2005, Hickok 2007, Indefrey 2004, Vigneau 2006], but some authors have reported that the left IFG may be involved in phonological segmentation (i.e. sublexical phoneme isolation), as opposed to phoneme identification or storage [Paulesu 1993, Burton 2000, Burton 2006]. In this context, [Burton 2000] have reported that same/different judgements about the first phoneme in pairs of phonologically close words (e.g. dip-tip) and distant words (e.g. dip-ten) both involved the left STG. However, only the latter condition – requiring word segmentation in order to isolate the whole phoneme from the word – involved the left IFG (see also [Gandour 2003, Heim 2003, Zatorre 1996]).

When designing our new tasks, to avoid reading and metalinguistic requirements, we first based both tasks on a picture-naming paradigm in which the names of three familiar objects to be successively named are phonologically close. Globally, this procedure is inspired by the so-called minimal pairs in linguistics [Jakobson 1951, Chomsky 1968] and by procedures largely used in the assessment (e.g.

in French: [Piérart 2005, Chevrie-Muller 1985]) and remediation (e.g. [Barlow 2002, Moore 2005]) of phonological disorders in children. This repetitive evocation of only three familiar but phonologically close words attenuates the lexico-semantic requirement, while stressing the phonological constraints. Secondly, to involve distinct phonological brain areas, the two tasks were partially different. In the first task, the three names differ from each others only in one phonological feature (e.g. in French: poule/boule/moule [hen, ball, tin]; for English equivalent: batch/patch/match). In the second task, they differ by the number of phonemes (e.g. in French: car/carte/car-t-on [car, card, cardboard box]; for English equivalent: car/car-t/car-t-on). Thus, while the first task (hereafter called “phon-diff”) implies only a difference of feature in one phoneme (e.g. the voicing between /p/ and /b/ in poule and boule), the second task (hereafter “phon-seg”) requires a change of segmentation, i.e. the subtraction or addition of whole phonemes (e.g. from carte to car and inversely). We expected from both new tasks the inferotemporal activation found in classical picture naming tasks in adults [Démonet 2005] (see also [DeLeon 2007]) and children [Gaillard 2006]. Furthermore, we predicted an activation of the left posterior language areas for both tasks, but a stronger activation of the left IFG for the task implying a change of segmentation [Burton 2000, Burton 2001].

Our study tested these four tasks with a group of healthy children. To optimize the feasibility of the paradigms by children, all tasks were contrasted with rest condition in four identical block designs. To reduce the bias due to the normalization of children’s brains with respect to adult standard [Wilke 2002, Wilke 2003d], we created a customized and matched pairs paediatric template [Wilke 2008]. Task comparisons, intersections and unions were carried out to highlight the specificity and sensitivity of the tasks as well as the whole panel [Tie 2008]. ROI analyses were performed using ROIs adapted to our template from the Anatomic Automatic Labelling atlas (AAL atlas) [Tzourio-Mazoyer 2002]. Lateralization Indexes (LIs) in the ROIs were assessed using a recent dedicated toolbox [Wilke 2007].

## 2 Subject and methods

### 2.1 Subjects

The study was approved by the local ethics committee (Consultative Committee for Protection of Persons in Biomedical Research) of the University Hospital (Rennes, France). Healthy children aged from 8 to 18 were recruited by word of mouth in the context of a larger study of developmental language disorders. Exclusion criteria included non-French native speakers, previous or current neurological, developmental or psychiatric illness, learning disability and abnormal academic performance, as well as language delay, MRI contraindication and the presence of orthodontic braces. Handedness was assessed by a child-modified version of the Edinburgh Handedness Inventory [Oldfield 1971]. All subjects were pre-screened for any conditions which would prevent an MRI scan from being acquired. A group of 18 children was recruited (age range = 8.7–17.7, mean age =  $12.7 \pm 3$ ), with 9 boys (mean age =

12.3  $\pm$  3.2) and 9 girls (mean age = 13  $\pm$  3). Sixteen children were right-handed and two children were left-handed (11%), which is within the estimated range of 8–15% left-handers present in the general population [Hardyck 1977]. All parents and children were informed about the experiment and procedure; parents signed the informed consent and children gave their verbal assent.

## 2.2 Experimental paradigms

### 2.2.1 General technical implementation

A single scanner session included the four paradigms separately implemented with the same parameters: a simple block design alternated a rest condition as control and the language task, starting with rest, with a preliminary period of signal acquisition for MRI signal stabilization which was later discarded during data processing. Each paradigm included three 27 s blocks of each condition and had a total duration of 2 min 48 s. The scanner session, including the anatomical acquisition and the four language paradigms, had a duration of about 30-35 min. All subjects performed the tasks in the same order, as during the preparation step, in order to avoid the mix of auditory and visual tasks and the resulting complication for the child. Words required by the tasks were one-to-three-syllable words highly frequent in the lexicon of French 8 years old children [Lambert 2001].

During the rest condition, a red cross was displayed on the projection screen and children were asked “not to work”, to “think about nothing” and, because of the complexity of this instruction, to listen to the noise of the scanner and fix attention on the red cross. Visual stimuli were delivered through a screen placed within the head-coil (IFIS-SA fMRI system, Invivo, Orlando, FL) just in front of the face, and synchronised with the scanner. In cases of poor eyesight, the children wore corrective glasses compatible with the high-magnetic- field environment. Auditory verbal stimuli were delivered by an experimented member of the staff using the machine microphone, via specially converted high-fidelity stereo headphones.

### 2.2.2 Auditory lexico-semantic tasks (reference tasks)

**Category task (words generation from category).** In this task adapted from [Gaillard 2003], children heard category names (e.g. animals, colours, things to eat) and had to silently generate as many as possible verbal examples of these categories. A category name was delivered every 9 s, with three categories per block and nine categories for the whole paradigm.

**Definition task (auditory responsive naming).** In this task adapted from [Balsamo 2002], children heard descriptions of concepts (e.g. “a big animal with a trunk”; “the moment of the day when one wakes up”) and had to find and silently name the corresponding word (e.g. elephant; morning). Descriptions were delivered every 9 s, with three definitions per block and nine definitions for the whole paradigm.

### 2.2.3 Visual phonological tasks (new tasks)

The two new tasks are based on picture naming and used black-and-white line drawings of familiar objects. Children had to silently name successively three pictures one by one (i.e. triplets) whose names are semantically unrelated but exhibit a close phonological composition. The pictures of the triplets were presented successively and randomly (without any picture being delivered twice successively) every 1.4 s, resulting in 19 stimulations within each block, so that the child could not predict the upcoming picture. One distinct triplet was used for each language block, resulting in three distinct triplets for the whole task.

**Phon-diff task.** The names of the objects to be named present a minimal difference in the phonological distinctive features of the initial phoneme. In the triplet poule/boule/moule (/pul/-/bul/-/mul/ [hen, ball, tin]), the difference between poule and boule is the voicing feature of /p/ and /b/ (voiceless vs. voiced). A similar reasoning applies for the distinctions poule–moule and boule–moule (stop vs. nasal consonant). Concretely, children successively named for example: “poule, moule, boule, moule, poule...” for the first block, then: “banc, dent, gant, dent, banc...” (/bã/, /dã/, /gã/ [bench, tooth, glove]) for the second block, and so on.

**Phon-seg task.** The names of the objects to be named present a small change in their phonological length, resulting in phoneme addition or subtraction. For example, in the triplet car/car-te/car-t-on (/kar/-/kart/-/kartõ/ [car, card, cardboard box]), there may be an addition (car towards carte or carton) or a subtraction (carton towards carte or car) of phonemes. Concretely, children successively named for example: “car, carte, carton, carte, car...” for the first block, then: “croix, roi, oie, roi, croix...” (/krwa/, /rwa/, /wa/ [cross, king, goose]) for the second block, and so on.

### 2.2.4 Preparation before the scanner

Children were prepared extensively just before the scanning session. Each task was thoroughly explained and practised prior to entering the scanner, using original task material. Each task was performed several times with the clinician, both aloud and silently to check for the comprehension of the tasks and to prevent mouth movements during silent responses. For the phonological visual tasks, the clinician made sure that the child would use the expected names, e.g. not naming carton as “boîte” [box], which would be inconsistent with the logic of the triplet car–carte–carton in French.

## 2.3 Data acquisition

Acquisitions were performed on a 3 T whole-body scanner (Achieva, Philips Medical Systems, Best, The Netherlands) using a 8-channel head coil. Anatomical 3D T1w images were acquired with a Fast Field Echo sequence. The acquisition parameters

were as follows: Echo time (TE)/TR/Flip angle: 4.6 ms/9.9 ms/8°; acquired matrix size: 256 x 256 mm; Field of View (FOV): 256 mm; voxel size: 1 x 1 x 1 mm; volume: 160 sagittal 1 mm thickness slices; acquisition time: 3 min 56 s. Functional images were acquired using a single-shot T2\* weighted (T2\*w) gradient-echo echo planar imaging sequence. Twenty-four 4 mm slices were acquired with the following parameters: TE/TR/Flip angle: 35 ms/3000 ms/90°; acquired matrix size: 80 x 80; reconstructed matrix size: 128 x 128; FOV: 230 x 230; acquired voxel size: 2.9 x 2.9 x 4 mm; reconstructed voxel size: 1.8 x 1.8 x 4 mm. Slices were positioned parallel to the anterior commissure-posterior commissure line, with no gap, and were interleaved from bottom to top. Each functional run consisted of 56 series of images acquisition for the 24 slices covering the entire brain volume separated by a 3000 ms delay for a total acquisition time of 2 min 48 s. Children were positioned supine in the system. The subject's head motion was minimized using straps and foam padding.

## 2.4 Data processing

MRI data preprocessing and analysis were performed using the GLM [Friston 1994a], as implemented in SPM5<sup>1</sup> (Statistical Parametric Mapping 5, Wellcome Department of Imaging Neuroscience, University College, London). The first two volumes of fMRI data were discarded to allow for signal stabilization. Slice timing and motion correction were applied to the remaining 54 volumes. To prevent bias caused by the normalization of paediatric data on adult templates [Wilke 2002, Wilke 2003d], we used the Template-O-Matic toolbox [Wilke 2008] to generate a customized paediatric template based on the age and sex of our 18 subjects. The matched pair option of the toolbox creates a reference map for each subject, based on the Paediatric MRI Data Repository funded by the National Institute of Health (n = 404, age range = [5;18]), with a final averaging of these individual reference maps. Structural MRI were rigidly realigned with this template, segmented using unified segmentation [Ashburner 2005], and then normalized. FMRI data were registered on segmented grey matter, and then normalized and smoothed using an isotropic 8 mm FWHM 3D Gaussian kernel.

Statistical activation maps were obtained using a mixed effects analysis. At the subject level, a high-pass filter was applied to fMRI data so as to remove slow signal drifts due to undesired effects. To model possible delay and dispersion of the canonical HRF, we used the Informed Basis Set [Friston 1998], including temporal and dispersion derivatives, to model the haemodynamic response. For each task, individual and group activations were identified by contrasting out the effect of temporal and dispersion derivative, focusing on the canonical variable, at a threshold of  $p < 0.05$  FWE corrected at the voxel level, and an extent threshold ( $k$ ) of 5 voxels was chosen as the minimal cluster size to reduce the effect of noise. In addition, since they had never been used before, a prospective less conservative threshold of  $p < 0.001$  uncorrected, at the voxel level was also applied to the new visual phono-

---

<sup>1</sup>[www.fil.ion.ucl.ac.uk](http://www.fil.ion.ucl.ac.uk)

logical paradigms. Exclusion criteria of individual data included motion artefacts associated with a movement exceeding 3 mm in translation or 3° in rotation.

Additional statistical comparisons between the paradigms were carried out by entering individual contrast files into a three (basis)-by-four (paradigm) **Analysis Of Variance (ANOVA)**, with a significance threshold set at  $p < 0.001$  (uncorrected). Comparison and conjunction (intersection) analyses [Friston 2005a, Nichols 2005] were performed to study the specificity of each paradigm and the whole panel. Furthermore, to address the sensitivity of the whole protocol, we performed a union analysis to select the voxels activated by any of the paradigms (logical OR). A conjunction analysis of auditory tasks union and visual tasks union ([category OR definition] AND [phon-diff OR phon-seg]) was also performed.

For ROI analyses, thirteen ROIs covering brain areas involved in language were selected from the literature: the pars opercularis (IFG-oper) and triangularis (IFG-tri) of the IFG; the precentral gyrus; the rolandic operculum; the STG, MTG, and Inferior Temporal Gyrus (ITG); the Heschl's, lingual, fusiform, Supramarginal Gyrus (SMG) and angular gyri; and the insula. Left and right ROIs as delineated in the AAL atlas [Tzourio-Mazoyer 2002] were adapted to our customized paediatric template using an approach suggested by [Wilke 2003c]. Firstly, to match our template, we performed a non-linear deformation of the structural image on which the AAL atlas was delineated. The deformation parameters previously determined were then applied. Finally, each region was smoothed with an isotropic Gaussian kernel of 6 mm to partially correct for the registration inaccuracy.

LIs were estimated using the LI toolbox [Wilke 2007], based on unsmoothed regions. For each subject the average t-value within each ROI was measured and voxels smaller than this threshold were discarded. The LI was then calculated with the remaining voxels as follows:

$$li = \frac{\sum \text{Activation}_L - \sum \text{Activation}_R}{\sum \text{Activation}_L + \sum \text{Activation}_R}, \quad (5.1)$$

where  $\sum \text{Activation}_L$  and  $\sum \text{Activation}_R$  denote the sum of the remaining voxels in the left and right parts of the ROI, respectively.

Boxplots based on these values were created. Similarly weighted mean activations within a ROI were estimated using the Marsbar toolbox [Brett 2002] for each smoothed ROI. Wilcoxon signed rank tests ( $p < 0.05$ ) were performed on each ROI to determine significant activations (for all subjects) and group lateralisation (i.e. left if  $li$  significantly greater than zero, right if  $li$  significantly smaller than zero, otherwise bilateral). To highlight laterality differences between paradigms, the Kruskal-Wallis test was performed on each ROI. A post-hoc non-parametric Mann-Whitney U-test ( $p < 0.05$ ) was performed to determine between-paradigm differences when the Kruskal-Wallis test was significant ( $p < 0.05$ ).



### 3 Results

We first carried out group analysis including statistical comparisons and conjunctions of the tasks within each category (i.e. reference auditory tasks; new visual tasks). Then, we carried out a comparison of the two categories as well as conjunction and union analysis of all tasks (i.e. whole panel). The statistical comparisons aimed to assess the specificity of each task or category. The conjunction (intersection) aimed to reveal more language-specific activations, while the union analysis aimed to combine the results of all tasks and to improve the sensitivity of the procedure.

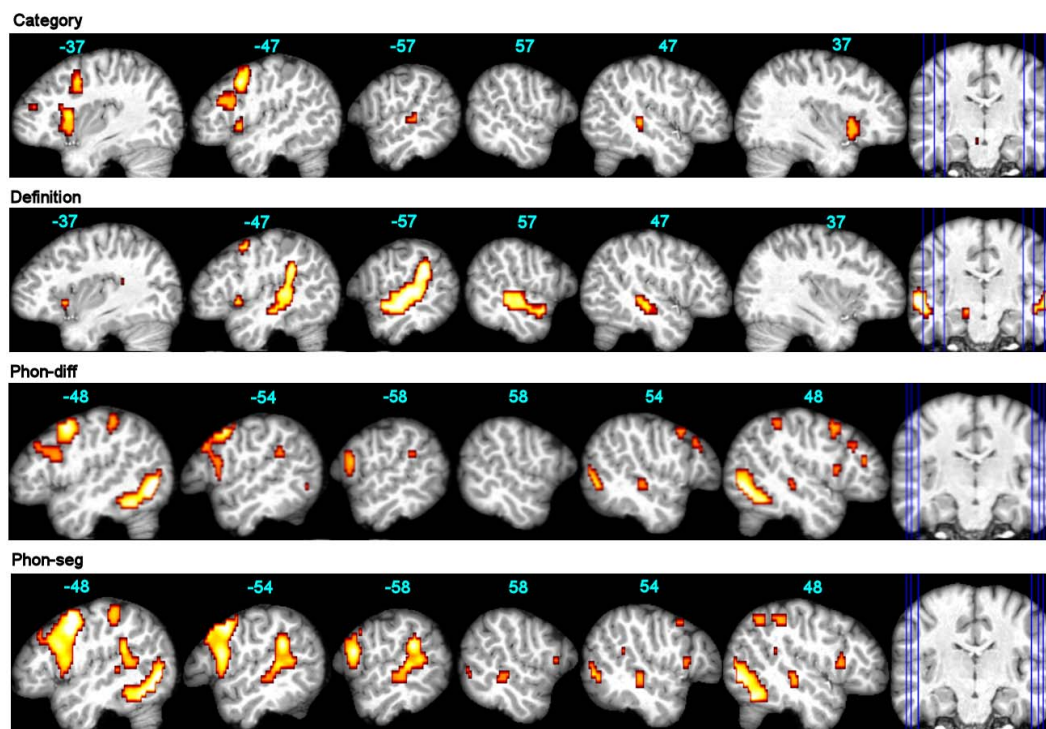


Figure 5.1: fMRI group effects for each language task. Results are displayed at  $p < 0.05$  FWE for category and definition tasks, and at  $p < 0.001$  uncorrected for phon-diff and phon-seg tasks. The functional maps are superimposed onto an individual brain normalized with respect to our customized paediatric template, with x-coordinates in MNI space. Left panels display the left hemisphere.

#### 3.1 Reference auditory tasks

These two lexico-semantic tasks with auditory stimulation were chosen from the literature to highlight a predominant activation of the left IFG (for the category task) or the left STG (for the definition task). The data from one subject had to

<i>Left hemisphere</i>	<i>Auditory language tasks</i>		<i>Visual language tasks</i>	
	Category	Definition	Phon-Diff	Phon-Seg
Inf frontal–Oper	—	—	—	5–4.78 <sup>(4)</sup>
			(509–4.02) <sup>(4)</sup>	(1295–5.18) <sup>(4)</sup>
Inf frontal–Tri	346–5.31 <sup>(4)</sup>	—	—	—
			(509–3.91) <sup>(4)</sup>	
Precentral	346–7.02	36–5.38	52–5.74	102–5.67
			(509–5.74)	(1295–5.67)
Insula / Inf frontal	357–6.77 <sup>(2)</sup>	29–4.74 <sup>(2)</sup>	—	—
SMA	1114–7.72	195–5.76	37–5.30	18–5.16
Ant cingulate	1114–5.76	—	—	—
Mid frontal	23–5.07	—	—	—
Supramarginal	—	1038–6.35	—	—
Sup temporal	—	1038–6.56 <sup>(2)</sup>	—	15–5.14 <sup>(1)</sup>
		1038–6.43 <sup>(3)</sup>	(22–3.46) <sup>(1)</sup>	(421–5.14) <sup>(1)</sup>
Mid temporal	24–5.32 <sup>(2)</sup>	—	—	—
				(421–4.42) <sup>(2)</sup>
Hippocampus	—	28–5.32	—	—
Parahippocampus	—	24–4.70	—	—
Inf parietal cortex	—	—	21–4.83	—
Postcentral	—	—	—	18–5.43
Sup parietal	—	—	—	249–6.32
Lingual	529–5.67	402–6.09	—	1296–7.72
		24–4.91		
Fusiform	—	—	890–6.53	1296–7.45
Inf occipital	—	—	890–>8	1296–7.53
Mid occipital	—	22–5.02	890–6.66	—
<i>Right hemisphere</i>				
Inf frontal–Oper	—	—	—	—
			(13–3.50)	(228–4.28) <sup>(4)</sup>
Inf frontal–Tri	—	—	—	—
			(43–3.37)	
Ant Cingulate	—	195–4.83	—	—
Med sup frontal	—	195–5.13	—	—
Insula	137–5.96	—	—	—
Mid Cingulate	1114–6.42	—	—	—
Sup temporal	—	437–6.11 <sup>(2,3)</sup>	—	—
				(90–3.31) <sup>(1)</sup>
Mid temporal	28–5.27 <sup>(2)</sup>	437–5.60 <sup>(2)</sup>	—	—
			(26–3.50)	(90–3.86) <sup>(2)</sup>
Parahippocampus	—	20–4.83	—	—
Angular	—	—	67–5.16	—
Sup parietal	—	—	67–5.06	24–5.20
Lingual	—	402–5.69	—	1129–7.46
Fusiform	—	—	760–7.40	1129–7.26
Mid occipital	—	—	760–6.89	1129–7.26
Sup occipital	—	14–5.04	—	—
Cerebellum	19–5.02	—	760–5.94	—

Table 5.1: fMRI results for each language task: peak location, cluster extent – Z-score ( $p < 0.05$  FWE;  $k = 5$ ). For visual language tasks, additional results among language areas are reported at  $p < 0.001$  uncorrected in brackets. Clusters in italics correspond to sub-clusters belonging to cluster of the same extent given in the column. <sup>(1)</sup>Posterior part; <sup>(2)</sup>Middle part; <sup>(3)</sup>Anterior part; <sup>(4)</sup>Dorsal part; <sup>(5)</sup>Ventral part; <sup>(6)</sup>Extending in the adjacent IFG. Oper = opercularis; Tri = triangularis; Inf = inferior; Mid = middle; Sup = superior; Med = medial; Ant = anterior.

be discarded because of a technical problem with sound delivery.

According to group analysis (Table 5.1; Fig. 5.1), the category task elicited a left-only activation in the caudal and dorsal IFG-triangularis, extending into the

IFG-opercularis, and a bilateral activation of the insula extending only on the left in the ventral IFG-opercularis. Only a small and bilateral activation was seen in the middle part of the Superior Temporal Sulcus (STS), and other activations appeared in the left precentral and lingual gyri. ROI analyses and LIs (Figs. 5.3 and 5.4) confirmed the significant left-lateralised activation of the IFG-triangularis and opercularis, as well as of the precentral and lingual gyri, while the activation of the STG did not appear to be significantly left lateralised.

The definition task yielded a strong bilateral activation along the STS, with a large and left-only sub-cluster in the posterior STG and adjacent SMG. Small left-only clusters occurred in the insula / ventral IFG-opercularis and in the precentral gyrus, and other clusters were located in the left and right lingual and parahippocampus region. ROI analyses and LIs confirmed a significant and left-lateralised activation of the MTG as well as the SMG and angular gyri, along with the IFG-opercularis and precentral gyrus.

The statistical comparison of these two tasks (Table 5.2; Fig. 5.2) showed specific activations in the left precentral gyrus / IFG-opercularis junction for the category task compared to the definition task. The reverse comparison (definition > category) highlighted a strong left-dominant activation extending from the ventral inferior parietal cortex to the anterior STS. This significant difference is corroborated by ROI analyses. The conjunction analysis between the two tasks (Table 5.2; Fig. 5.2) revealed common activations along the left STS, extending posteriorly in the STG / SMG junction, and in the right middle STS. Left-only common activations were also found in the dorsal IFG-opercularis and the insula, extending into the adjacent ventral IFG.

Therefore, as expected, the category task elicited a predominant activation of the left IFG compared to the left STG, while results from the definition task showed the inverse pattern. More precisely, in the group analysis, the category task highlighted the caudal and dorsal left IFG-triangularis, although the location is distinct in the statistical comparison. By contrast, the definition task specifically involved the left posterior STG/SMG region. Furthermore, the conjunction of both tasks detected both left temporal and frontal core language areas.

### 3.2 New visual tasks

These new phonological tasks with visual stimulation were designed to focus on the phonological brain network without any requirement for either reading or metalinguistic skills. In addition to occipital and inferotemporal activation due to the picture-naming condition, we expected both tasks to elicit an activation of the left posterior STG. Furthermore, in the case of the task involving a phonological change of segmentation, we expected a higher activation of the left IFG. For the phon-seg task, the last paradigm of the session, the data from two subjects had to be discarded because of excessive movement.

According to the group analysis (Table 5.1), the phon-diff task elicited activations centred on the left and right fusiform, the left precentral and the right angular

gyri ( $p < 0.05$  FWE). At the threshold of  $p < 0.001$  uncorrected (Fig. 5.1), large clusters appeared in the left precentral gyrus and IFG (opercularis and triangularis), with smaller similar clusters on the right; a small cluster is located at the left posterior STG / SMG junction. ROI analyses (Fig. 5.3) confirmed a significant and left-dominant activation for the precentral gyrus and the IFG-opercularis, but this did not reach significance for lateralisation. LIs (Fig. 4) showed a significant left lateralisation for the rolandic operculum, a region displaying a significant deactivation during the task ( $p < 0.001$  uncorrected).

The phon-seg task elicited a strong bilateral activation centred on the lingual gyri, and a left activation along the precentral gyrus / IFG-opercularis junction; two small left clusters appeared in the dorsal IFG-opercularis and posterior STG ( $p < 0.05$  FWE). At the threshold of  $p < 0.001$  uncorrected (Fig. 5.1), results showed a large cluster encompassing the precentral gyrus and the ventral and dorsal IFG-opercularis, as well as a large cluster centred on the posterior STG / SMG junction; only small counterparts were seen on the right. ROI analyses confirmed a significant activation of the left and right IFG and MTG, and LIs revealed a significant left lateralisation for the precentral and lingual gyri.

The statistical comparison (Table 5.2; Fig. 5.2) showed that the phon-diff task did not elicit specific activation compared to the phon-seg task. The inverse comparison (phon-seg > phon-diff) highlighted the left lingual gyrus, as confirmed by ROI analyses, as well as the left and right fusiform gyri, the left rolandic operculum and a small cluster on the right STS. By analysing the conjunction of the two tasks (Table 5.2; Fig. 5.2), we found a significant common activation in the left precentral gyrus, as well as in the adjacent IFG-opercularis into its more ventral and caudal part, associated with smaller counterparts in the right hemisphere. Another left cluster appeared in the posterior STG.

Therefore, these two new tasks, taken individually, were able to activate core language areas, namely the left IFG-opercularis and the posterior temporal / SMG region, even though the LIs did not show significant lateralisation. In the group analyses, the phon-seg task compared to the other task activated more the left IFG, as expected. But, contrary to expectations, it also elicited higher activation in the left posterior temporal/SMG region. This may reflect an overall superiority of the phon-seg task for the language regions, which nevertheless did not appear in the statistical comparisons. Finally, the conjunction of the two tasks, as observed with previous auditory tasks, was able to detect both left temporal and frontal core language areas.

### 3.3 Whole panel analysis

When statistically comparing the auditory reference tasks and the new visual tasks (Table 5.2; Fig. 5.2), the former elicited specific strong and bilateral activations centred on the middle STG and the Heschl's gyri, with both areas also showing significant interparadigm differences in the ROI analysis (Fig. 5.3). Another specific activation was found bilaterally in the lingual gyri. The reverse comparison (visual

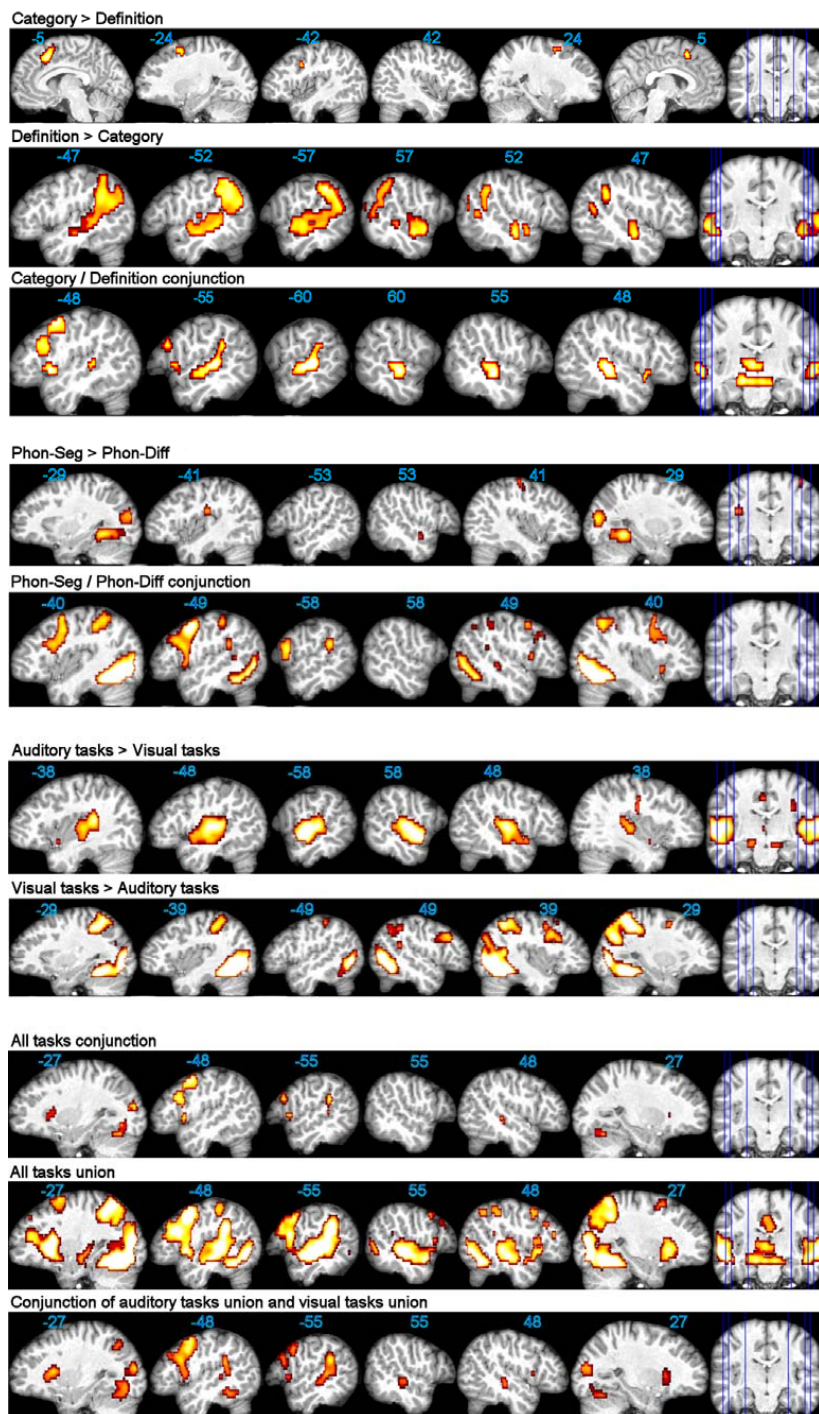


Figure 5.2: Task comparisons, conjunctions, and union ( $p < 0.001$  uncorrected). Functional maps are superimposed onto an individual brain normalized with respect to our customized paediatric template, with x-coordinates in MNI space. Left panels display the left hemisphere.

<i>Left Hemisphere</i>	<i>Auditory language tasks</i>			<i>Visual language tasks</i>			<i>All tasks</i>	
	Categ >Def	Def > Categ	Categ C Def	Ph-s >Ph-d	Ph-s C Ph-d	Audit >Vis	Vis >Audit	All tasks C
Inf frontal-Oper	—	—	<i>348–4.10<sup>(4)</sup></i>	—	<i>825–4.83<sup>(4)</sup></i>	—	—	<i>286–4.10<sup>(4)</sup></i> <i>33–3.64<sup>(7)</sup></i>
Precentral	18–3.38 <sup>(5)</sup>	—	348–5.09	—	825–6.96	—	—	286–5.09
Mid frontal	33–3.66	—	—	—	—	—	—	—
SMA	—	—	1433–5.48	—	357–6.32	—	—	289–5.37
Cingulate	—	—	<i>1433–5.08<sup>(3)</sup></i>	—	—	55–3.92 <sup>(2)</sup>	—	—
Med sup frontal	174–4.69	—	—	—	—	—	—	—
Rol operculum	—	—	—	36–4.31	—	—	—	—
Insula	—	—	396–4.87 <sup>(8)</sup>	—	58–3.81	—	—	51–3.81
Sup temporal	—	—	<i>351–3.81<sup>(1)</sup></i>	—	91–4.54 <sup>(1)</sup>	1588–7.26 <sup>(123)</sup>	—	37–3.8 <sup>(1)</sup>
Mid temporal	—	1658–4.67 <sup>(3)</sup>	351–5.61 <sup>(2)</sup>	—	10–3.26 <sup>(2)</sup>	—	—	—
Inf parietal	—	<i>1658–5.18<sup>(6)</sup></i>	—	—	—	—	—	—
Sup parietal	—	—	—	—	976–6.50	—	821–6.50	—
Postcentral	—	—	—	—	<i>976–4.65</i>	—	—	—
Sup occipital	—	—	—	—	—	—	—	<i>41–3.44</i>
Mid occipital	—	—	—	146–4.43	<i>1465-&gt;8</i>	—	<i>1366-&gt;8</i>	41–4.11
Inf occipital	—	—	—	—	<i>1465-&gt;8</i>	—	<i>1366-&gt;8</i>	—
Fusiform	—	—	—	397–5.44	<i>1465-&gt;8</i>	—	<i>1366-&gt;8</i>	—
Lingual	—	—	3632–6.53	397–4.60	—	4559->8	—	116–5.55
Precuneus	—	1060–4.69	—	—	—	—	—	—
Thalamus	—	—	—	—	—	21–3.43	—	—
Caudate	—	—	—	—	—	14–3.73	—	—
Cerebellum	—	—	—	—	—	—	—	<i>116–3.76</i>
Hippocampus	—	—	—	—	—	—	—	<i>11–3.42</i>
<i>Right Hemisphere</i>								
Inf frontal-Oper	—	—	—	—	14–3.48 <sup>(5)</sup>	—	332–3.98 <sup>(4)</sup>	—
Mid frontal	34–3.66	—	—	—	11–3.98	19–3.89	332–4.31	—
Precentral	—	—	—	—	252–4.13	—	—	—
Cingulate	—	—	—	—	—	2037–5.95 <sup>(3)</sup>	—	—
Insula	—	—	200–3.88 <sup>(8)</sup>	—	61–3.76	—	—	38–3.45
Rol operculum	—	—	—	—	—	50–4.16	—	—
Sup temporal	—	<i>436–4.67<sup>(3)</sup></i>	364–5.08 <sup>(2)</sup>	11–3.43 <sup>(3)</sup>	—	1139–7.66 <sup>(2)</sup>	14–3.28 <sup>(1)</sup>	—
Mid temporal	—	<i>436–4.71<sup>(3)</sup></i>	<i>364–5.04<sup>(2)</sup></i>	—	10–3.29 <sup>(2)</sup>	—	—	10–3.29
Postcentral	—	—	—	10–3.27	—	—	—	—
Sup parietal	—	—	—	—	800–5.50	—	—	—
Sup Occipital	—	—	—	—	<i>800–4.46</i>	—	—	—
Angular / SMG	—	388–4.22	—	—	—	—	—	—
Fusiform	—	—	—	284–6.33	1266->8	—	3118->8	—
Inf occipital	—	—	—	—	<i>1266-&gt;8</i>	—	<i>3118-&gt;8</i>	—
Mid occipital	—	—	—	225–4.91	—	—	—	—
(Pre)cuneus	—	<i>1060–3.99</i>	—	225–3.76	—	—	—	—
Lingual gyrus	—	—	3632–5.26	18–3.74	—	4559–6.99	—	67–3.68
Hippocampus	—	—	—	—	14–3.84	—	—	—
Caudate	—	—	—	—	—	26–3.86	—	—
Cerebellum	—	—	—	—	—	—	—	<i>67–5.06</i>

Table 5.2: Task comparisons (N) and conjunctions (C). Peak locations, cluster extent–Z-score ( $p < 0.001$  uncorrected;  $k = 10$ ). Clusters in italics correspond to sub-clusters belonging to cluster of the same extent given in the column. <sup>(1)</sup>Posterior part; <sup>(2)</sup>Middle part; <sup>(3)</sup>Anterior part; <sup>(4)</sup>Dorsal part; <sup>(5)</sup>Ventral part; <sup>(6)</sup>Ventral part, extending into the SMG, angular, superior temporal gyri; <sup>(7)</sup>Ventral part, extending into the adjacent insula; <sup>(8)</sup>Extending into the adjacent IFG; <sup>(9)</sup>Overlapping most of the adjacent dorsal and ventral IFG-Opercularis. Ph-s = Phonol-Seg task; Ph-d = Phonol-Diff task; Oper = opercularis; Tri = triangularis; Rol = rolandic; Inf = inferior; Mid = middle; Sup = superior; Med = medial; Ant = anterior; SMG = SMG.

tasks > auditory tasks) revealed specific clusters in the right middle frontal gyrus, in the right dorsal IFG-opercularis, and in the right posterior and medial STG. In

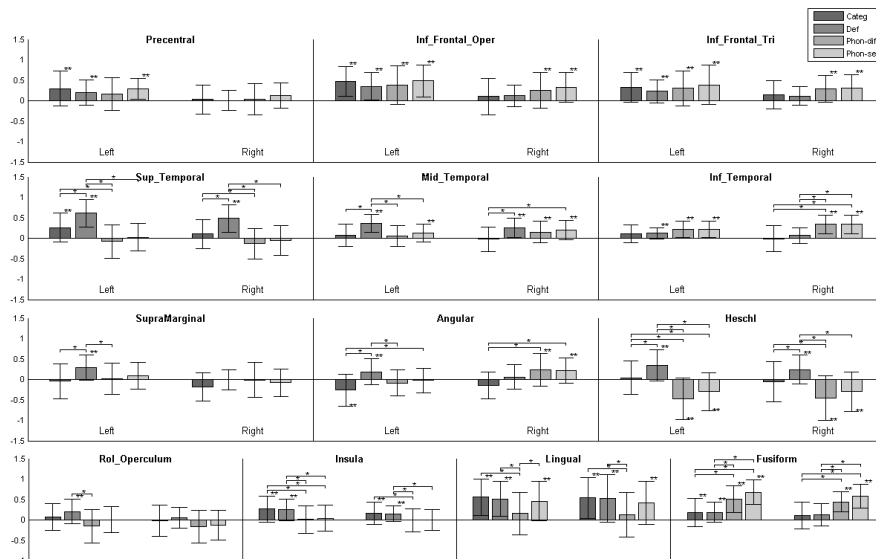


Figure 5.3: ROI analyses for each paradigm. Bar plots depict the effect size of activation (mean  $\pm$  standard deviation) in each left and right ROI for each paradigm. Significant activation or deactivation ( $p < 0.05$ ) are indicated by a double star. In addition, significant interparadigm differences ( $p < 0.05$ ) are indicated by brackets. Oper = opercularis; Tri = triangularis; Rol = rolandic; Inf = inferior; Mid = middle; Sup = superior.

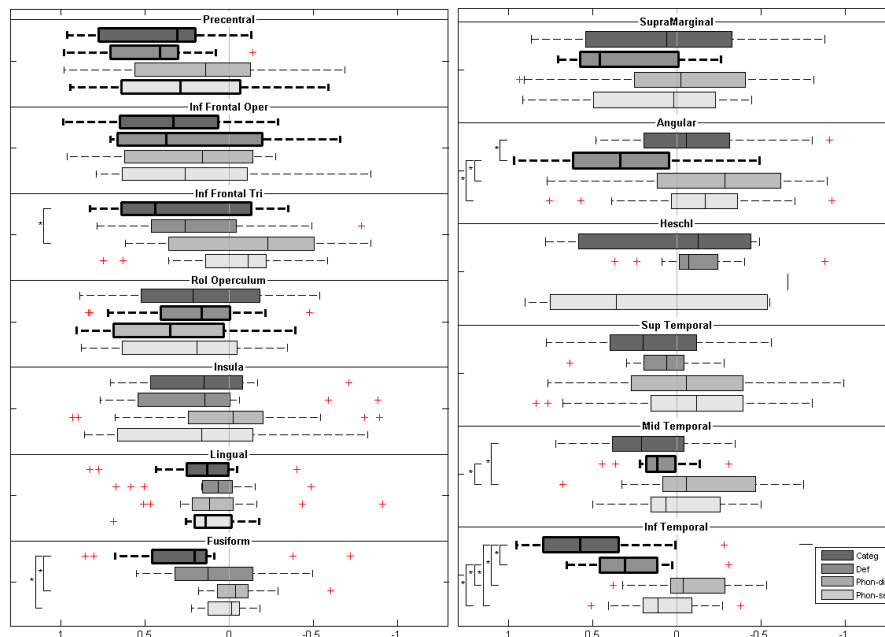


Figure 5.4: LIs and significant lateralisations within the ROIs for each paradigm. The box plots depict lateralisation for each ROI, with positive LIs (resp. negative) reflecting left (resp. right). Significant lateralisations ( $p < 0.05$ ) are indicated by bold lines. Significant interparadigm differences of LIs are indicated by a bracket. Oper = opercularis; Tri = triangularis; Rol = rolandic; Inf = inferior; Mid = middle; Sup = superior.

addition, strong specific activations appeared in the fusiform gyri, as also confirmed by significant interparadigm differences in the ROI analysis (Fig. 5.3).

The conjunction analysis of the four tasks (Table 5.2; Fig. 5.2) highlighted the dorsal and ventral parts of the left IFG-opercularis, as well as the left posterior STG/SMG junction; on the right, only one small cluster appeared in the middle of the MTG. Additional common clusters were located in the left precentral gyrus, and, bilaterally, in the lingual gyri and the anterior insula. The conjunction analysis of the unions of auditory and visual tasks (Fig. 5.2) showed an enlargement of the same clusters with an extension of the left temporal cluster onto the middle STS.

The union analysis of all tasks (Fig. 5.2) highlighted, on the left, a large cluster encompassing the precentral gyrus, the dorsal IFG-opercularis/triangularis junction, the insula and the immediately adjacent ventral IFG-opercularis, as well as a strong activation extending from the SMG to the anterior STS. Smaller counterparts were seen on the right.

Therefore, the auditory lexico-semantic tasks showed more activation of the auditory receptive language region (i.e. middle STG, Heschl's gyrus), whereas the visual phonological tasks involved more the occipitotemporal visual region (i.e. fusiform gyrus). For these latter tasks, the stronger contribution of the right IFG and STG, compared to the reference tasks, reflects a less left-lateralised activation of these regions, as suggested also by the LIs. Finally, the conjunction of all tasks highlighted more focused activations in the left language areas, while their union showed enlarged activations, especially for the posterior STG/SMG region.

## 4 Discussion

fMRI has an expanding role in the localisation and lateralisation of language in children, which is an important issue in clinical and research applications such as presurgical mapping [Anderson 2006, O'Shaughnessy 2008, Wilke 2006] and the investigation of developmental language disorders [Dick 2008, Friederici 2006a, Rapin 2003]. A major issue of language fMRI is due to the complexity of language, which means that it is crucial to choose a panel of language activation tasks able to detect all, but only, language brain areas (e.g. [Tie 2008]). In the paediatric context, another specific issue is the need for child-adapted tasks [O'Shaughnessy 2008, Wilke 2003a, Wilke 2006], which is even more important with impaired children whose achievement of the task may be compromised by high attentional, reading or metalinguistic requirements.

In this study, we tested with a group of healthy children a panel of four fMRI language tasks that could be used with young impaired children. One important constraint was to avoid reading and metalinguistic requirements, with the aim of increasing the feasibility and efficiency of the procedure, while reducing the attentional and academic demands. Two reference auditory lexico-semantic tasks were chosen from the literature, and two new tasks with visual stimulation were designed to focus on spontaneous (i.e. non-metalinguistic) phonological processing. When



taken individually, the tasks of each modality aimed to stress distinct language areas, and the whole panel aimed to be sensitive and specific in the detection of the language network.

#### 4.1 Methodological issues

FMRI in children implies special methodological precautions due to the risk of movement, attentional constraints, task design, task preparation and achievement, as well as appropriate reference brain data [Gaillard 2001a, O’Shaughnessy 2008, Wilke 2003a].

To optimize the feasibility of the procedure for young disordered children, and minimize the movement artefacts and the attentional complications, we implemented four identical block-designed paradigms of equal periods, without control active tasks or motor responses required from the child. These choices reduced the heterogeneity and complexity of the protocol, as the child did not have to understand and achieve distinct control tasks, or give motor responses in addition to understanding and achieving the four language tasks themselves. This may be particularly crucial for disordered children with lower attentional, cognitive or language abilities. Furthermore, while requiring motor responses is well suited for metalinguistic tasks (i.e. judgement tasks), this is precluded in the investigation of more natural condition such as word production, whether covert or overt. Finally, the simplicity and identity of the paradigm parameters for all tasks facilitated the comparison and combination of the results.

Nevertheless, these choices have some drawbacks. First, the achievement of the tasks could not be directly checked. Requiring overt responses would have allowed online performance monitoring, but aloud speech increases the risk of movement, which is crucial with children and all the more with disordered children [O’Shaughnessy 2008]. Therefore, children were intensively prepared before the scanner session using the same order of tasks and stimuli, allowing to check that they understood and were able to achieve successfully the task, and they were questioned after the session. Secondly, the use of a low-level control condition (listening to the noise of the scanner and fixing a cross) may involve more non-language-specific coactivations than a more specified control task (e.g. [Holland 2007, Wilke 2006]). Therefore, we used conjunction analysis between the various tasks to highlight specific language activations.

Moreover, to select a sample representative of the general population and close to the clinical context for disordered children, we did not only recruit right-handed children. In this study, the proportion of left-handed children (11%), is within the normal range estimate (8–15%). To investigate the effect of left-handedness on our results, we carried out supplementary ROIs and LIs analysis of the data from the sixteen right-handed children only, focusing on the IFG, STG and SMG. For the four tasks, results showed no difference concerning the pattern of activation and lateralisation within these ROIs compared to the whole group.

In data processing, to avoid distortions due to normalisation of the children’s

data on an adult standard template [Wilke 2002, Wilke 2003d], we used a recent tool dedicated to the creation of pair- and group-matched normalized templates based on normative brain data [Wilke 2008]. This enabled to avoid an age-related bias in the normalization steps by using a customized paediatric template based on the age and sex of our 18 subjects. To carry out local analyses, ROIs were based on the non-linear deformation of the AAL atlas [Tzourio-Mazoyer 2002] on our customized template, as suggested by [Wilke 2003c]. An inherent limitation of ROIs generated from an atlas is that they cannot adequately model subject variability, and hence do not allow to focus on sub-regions that might be of interest. Furthermore, activations may overlap several ROIs (e.g. the posterior STG and the SMG), so that the actual activations and lateralisations may, in fact, be minimized within each separated ROI. However, given the context and our objectives, this approach remains instructive by confirming significant activation and lateralisation when they are located within the ROIs.

## 4.2 Reference lexico-semantic auditory tasks

As expected, the words generation from category task (category), when contrasted with rest, highlighted left frontal rather than temporal language areas. More precisely, the paradigm elicited a left activation of the dorsal and caudal IFG-triangularis, at the junction with the pars opercularis, with another cluster of interest extending from the insula to the adjacent ventral IFG-opercularis. By contrast, there was a weaker and bilateral activation of the middle STS, without activation in the posterior STG. Thus, our study confirms that this paradigm, easier than other verbal fluency paradigms (e.g. [Riva 2000, Warburton 1996]), is able to induce a relatively specific activation of Broca's area in children.

This result is in line with the study by [Gaillard 2003], who used this paradigm with a group of 16 children (mean age: 10.2) and reported a left activation of the IFG without consistent activation of the STG. However, the location within the IFG is somewhat different from the present results, as these authors reported a main activation in the ventral pars orbitalis and a weaker activation in the anterior pars triangularis. Furthermore, no activation in the middle STS was reported. The task used in our study included three categories in each 27 s active block whereas Gaillard et al.'s task delivered one category within each 32 s block. Therefore, it is likely that our task results in a weaker executive demand and a greater receptive component. Moreover, in adults, the location of semantic processing within the left IFG remains to be specified. In their systematic review of studies using words generation from category, [Costafreda 2006] reported a ventral location. However, other studies of semantic processing have reported a dorsal location (e.g. [Wagner 2001]), and [Vigneau 2006], in their meta-analysis of studies using various semantic contrasts, reported semantic clusters in the dorsal pars opercularis.

In contrast with the category task, the auditory-responsive task (definition) elicited a strong and left-dominant activation of the temporal language regions without significant activation in the left IFG. Left and right activations appeared

along the middle and anterior *STS*, extending specifically on the left in the posterior *STG* and adjacent *SMG*. Again, one may note a left-only cluster extending from the insula into the adjacent ventral *IFG*-opercularis.

This is in line with the study by [Balsamo 2002], who used a similar paradigm with a group of 11 children (mean age: 8.5) and reported a left-dominant activation centred on the middle *STG* and *MTG*, and including the primary auditory cortex. Furthermore, these results are similar to those obtained in other studies with children using different language comprehension tasks such as picture/verbal-description matching [Wilke 2006], story listening [Ahmad 2003], or sentence listening with correction judgement [Brauer 2007]. Interestingly, a reading variant of the definition task (i.e. read response naming) has been shown to elicit activations in the left middle *MTG* without activation in the posterior *STG* [Gaillard 2001b]. Thus, the definition task appears to be able to cause a relatively specific activation in the left temporal region and especially in Wernicke's area. In line with current knowledge [Hickok 2007, Vigneau 2006], the bilateral activation of the middle *STS* may reflect the phonological-level processing of auditory speech (or written text) input, whereas the activation of the posterior *STG* / *SMG* may reflect the mapping of phonemes onto articulatory representation.

Our statistical comparison between these two reference tasks confirmed the specificity of the definition task and highlighted a specific activation for the category task in the left precentral/*IFG* junction. Moreover, the conjunction of both tasks revealed specific core language areas on the left, focusing on the posterior *STG*, the middle *STS*, the insula along with the adjacent ventral *IFG*-opercularis, and the dorsal *IFG*-opercularis.

### 4.3 New phonological visual tasks

The two new tasks with visual stimulation were designed to investigate spontaneous phonological processing. In contrast with most studies using metalinguistic tasks (e.g. rhyme judgements), these tasks are based on a picture-naming condition in which familiar objects to be named exhibit a close phonological composition.

The first new task (phon-diff) was designed to assess the brain basis of spontaneous production of minimal phonological differences (e.g. /pul/-/bul/-/mul/), and was expected to induce activations in the left *IFG* and temporo-parietal areas, as suggested by current knowledge about the brain areas involved in phonological production (e.g. [Vigneau 2006]). Interestingly, at the threshold of  $p < 0.001$  uncorrected, this task, when compared to rest, yielded left-dominant activations in the whole dorsal and ventral *IFG*-opercularis along the precentral sulcus, in a dorsal and more medial part of the *IFG*-triangularis, and, to a lesser extent, in the posterior *STG*.

The second new task (phon-seg) was designed to assess the brain basis of spontaneous phonological change of segmentation involved in the subtraction or addition of phonemes (e.g. /kar/-/kart/-/kartõ/). Compared to the previous task, a higher activation of the left *IFG* was expected according to current hypotheses about the

role of this structure in phonological segmentation [Burton 2001]. In fact, this paradigm elicited an interesting activation not only in the left IFG-opercularis, lying along the precentral sulcus, but also an equivalent activation in the left posterior STG. At the threshold of  $p < 0.001$  uncorrected, our results showed an enlargement of these clusters, with an extension of the posterior temporal activation into the middle STS.

Statistical comparisons between the two tasks did not show task-specific activations in language areas, reflecting the similarity of their global activation patterns. Promising results are provided by the conjunction analysis, which showed left-dominant common activations in core language areas, namely the dorsal and ventral IFG-opercularis, lying along the precentral sulcus, and the posterior STG.

Thus, both new visual tasks yielded an interesting and similar activation pattern of core left language areas, in agreement with current knowledge about the brain basis of phonological processing [Buchsbaum 2001, Burton 2001, Heim 2003, Vigneau 2006]. The location within the left IFG, along the precentral sulcus, is convergent with previous studies of phonological tasks involving phonemes isolation or sequencing (e.g. [Vigneau 2006]). In addition, according to the group analysis, the task involving a phonological change of segmentation showed as expected, a greater activation of the IFG than the other task, in agreement with current hypotheses [Burton 2001]. Nevertheless, this task also yielded a higher activation of the left STG, which requires further explanation (for discussion, see also [Gandour 2003, Heim 2009]). The next step of the work needs to use these new tasks for disordered children population investigation. In particular, although the phon-diff task appeared to be less efficient in this study of healthy children, further study may show more efficiency with disordered children.

#### 4.4 Whole panel

When statistically contrasted with the new visual tasks, the two auditory reference tasks yielded specific activations of the left and right middle STG, including the primary auditory cortex, and lingual gyri. The involvement of the middle STG may reflect auditory verbal processing. The recruitment of the lingual gyri is consistent with the involvement of ventromedial temporo-occipital regions during semantic processing, even in non-visual tasks, suggesting mental imagery or visualization strategies (see e.g. [Abel 2009, Sachs 2008, Vitali 2005, Wise 2000]).

Compared to the reference tasks, the two new visual tasks activated slightly more the right IFG-opercularis and posterior STG, which suggests less lateralisation associated with the new tasks in these regions. Moreover, these new tasks yielded greater activation of the bilateral fusiform gyrus, whose function has been the subject of much debate and which has been shown to be involved in a number of tasks, such as picture naming, object processing, reading and amodal conceptual processing (for discussion, see [Cohen 2004, Price 2003, Hillis 2005, Karnath 2009]).

Although this contrast between the auditory and visual tasks did not show differences in the left IFG and posterior STG / SMG, the separated conjunction

analyses of the two groups of tasks provide distinctive results, with the former leading to a wider activation in the left posterior **STG** and the latter in the left **IFG**.

The dissociation between the lingual and fusiform regions, which are more intensely activated by the auditory and visual language tasks, respectively, may be surprising. Using a picture-naming task with verbal semantic distracters (interference paradigm), [Abel 2009] reported an activation of the left and right lingual gyri. By contrast, [Balsamo 2006] reported an activation of the left fusiform gyrus in children during an auditory semantic decision task. Further work is needed to clarify the respective contributions of these two regions in object and language processing.

The conjunction analysis across all four tasks was assumed to reveal more specific and essential language areas. Interestingly, it highlighted left-only common and focal activations in core language regions, i.e. the dorsal and ventral **IFG**-opercularis and the posterior **STG**. Furthermore, the union analysis of all the paradigms, assumed to be more sensitive for detecting a more comprehensive language network, showed clusters of slightly distinct but close locations in the left **IFG**, and an extended left-dominant parietotemporal activation from the **SMG** to the anterior **STS**. The conjunction of the auditory tasks union and the visual tasks union showed an intermediate picture more informative than the conjunction and more specific than the union. Thus, in agreement with previous authors [Gaillard 2004, Ramsey 2001, Roux 2003, Tie 2008, Wilke 2006], our study confirms the usefulness of using a number of language tasks in an **fMRI** procedure.

In conclusion, out of the four language tasks in our panel, the two reference tasks (category and definition) demonstrated good abilities to yield selective left activations in the **IFG** and **STG**, respectively. The two new tasks studied here (phon-diff and phon-seg), which targeted phonological processing without requiring any metalinguistic or reading abilities, also yielded left-dominant activations in the dorsal and ventral **IFG**-opercularis, as well as the posterior **STG**, with an overall superiority of the phon-seg task. Compared to the reference tasks, the new tasks activated simultaneously both left frontal and temporal language regions, but less strongly and more bilaterally than the category task for the left **IFG** and than the definition task for the left posterior **STG**. When all tasks are taken together, conjunction and union analyses yielded interesting delineations of similar core language regions, with greater sensitivity being obtained from union analysis. This study confirms that a combination of several tasks tapping different aspects of language is useful for language brain mapping in children, and provides new tasks for the investigation of the brain basis of spontaneous phonological processing. We believe that such an **fMRI** panel could be efficient and useful with young children in the context of presurgical mapping as well as the investigation of acquired or developmental childhood disorders.

## Chapter 6

# Abnormal brain function in Typical-Specific Language Impairment

---

This chapter is presented in the form of a paper. This work was realised in collaboration with Clément de Guibert (cf. chapter 4) and was originally published in *Brain* [de Guibert 2011].

### Abstract

Atypical functional lateralisation and specialisation for language have been proposed to account for developmental language disorders, yet results from functional neuroimaging studies are sparse and inconsistent. This functional MRI study compared children with a specific subtype of Specific Language Impairment affecting structural language (n=21), to a matched group of typically-developing children using a panel of four language tasks neither requiring reading nor metalinguistic skills, including two auditory lexico-semantic tasks (category fluency and responsive naming) and two visual phonological tasks based on picture naming. Data processing involved normalizing the data with respect to a matched pairs paediatric template, groups and between-groups analyses, and laterality indices assessment within regions of interest using single and combined task analysis. Children with Specific Language Impairment exhibited a significant lack of left lateralisation in all core language regions (inferior frontal gyrus-opercularis, inferior frontal gyrus-triangularis, supramarginal gyrus, superior temporal gyrus), across single or combined task analysis, but no difference of lateralisation for the rest of the brain. Between-group comparisons revealed a left hypoactivation of Wernicke's area at the posterior superior temporal/supramarginal junction during the responsive naming task, and a right hyperactivation encompassing the anterior insula with adjacent inferior frontal gyrus and the head of the caudate nucleus during the first phonological task. This study thus provides evidence that this specific subtype of Specific Language Impairment is associated with atypical lateralisation and functioning of core language areas.

**Contents**

<b>1</b>	<b>Introduction</b> . . . . .	<b>101</b>
<b>2</b>	<b>Methods</b> . . . . .	<b>103</b>
2.1	Participants . . . . .	103
2.2	Neuropsychological and language assessment . . . . .	104
2.3	Functional MRI protocol and task panel . . . . .	104
2.4	Data acquisition . . . . .	106
2.5	Data processing . . . . .	107
<b>3</b>	<b>Results</b> . . . . .	<b>108</b>
3.1	Auditory language tasks . . . . .	108
3.2	Visual language tasks . . . . .	110
3.3	Lateralisation index assessment and comparison using the combined task analysis . . . . .	112
<b>4</b>	<b>Discussion</b> . . . . .	<b>113</b>
4.1	Lack of left lateralisation of core language areas . . . . .	113
4.2	Left hypoactivation of the posterior superior temporal gyrus / supramarginal gyrus junction . . . . .	114
4.3	Right hyperactivation of the anterior insula, adjacent inferior frontal gyrus and head of caudate . . . . .	115
4.4	Comparison to other developmental disorders . . . . .	117
4.5	Methodological considerations . . . . .	118

## 1 Introduction

Some children fail to develop typical language for no obvious reason. Their impairment typically reflects difficulties with producing and understanding oral language, and cannot be attributed to sensory-motor or intellectual deficits or other developmental impairment, especially autistic spectrum disorders, and is not the consequence of evident brain lesion or socio-affective deprivation [Bishop 1997, Rapin 2003]. It usually leads to literacy difficulties [Bishop 2004b] and, in a substantial number of cases, to persistent language difficulties through adolescence [Stothard 1998]. This condition is known as developmental dysphasia [Pari-  
sise 2009], developmental language disorder [Rapin 2003], or SLI [Bishop 1997].

The current optimal diagnostic procedure of SLI involves both psychometric assessment and clinical appraisal and expertise [Rapin 2003, Bishop 2008], since psychometric criteria alone do not provide sufficient sensitivity, specificity, and clinical congruence [Dunn 1996, Bishop 2004a]. Psychometric criteria may include a normal nonverbal intelligence quotient and abnormal language scores [Tomblin 1996]. Nonword repetition, sentence repetition, and syntactic manipulation are known to be more sensitive and specific than other tests [Conti-Ramsden 2003, Bishop 2004a].

SLI is known to be heterogeneous, encompassing distinct clinical profiles that may reflect distinct underlying deficits. A convergent trend distinguishes structural impairments affecting the core components of language (especially morphosyntax and phonology), from those affecting language reception, articulation or social use (auditory verbal processing, oro-motor verbal gesture, and pragmatics) [Rapin 1983, Conti-Ramsden 1997, Rapin 2003, Bishop 2004a]. Structural impairments are referred to as grammatical SLI [van der Lely 2005], linguistic dysphasia [Pari-  
sise 2009], or T-SLI [Bishop 2004a].

The aetiology of SLI remains largely unknown. There is growing evidence that SLI has a genetic component and can be inherited, but a complex picture is emerging of interaction between several genes and environmental risk factors [Fisher 2003, Bishop 2006]. This interaction may affect the anatomo-functional development and organization of the brain language network [Bates 1997, Bishop 2000, Rapin 2003, Friederici 2006a]. In addition, given the left lateralisation of language in most typically-developing individuals, atypical language lateralisation has been hypothesized to explain literacy and language developmental disorders [Annett 1985, Gesch-  
wind 1985]. Moreover, as acquired childhood aphasia show that unilateral lesion of the dominant hemisphere in childhood rarely leads to a persistent deficit [Van Hout 1997], the abnormality of brain development in SLI is suspected to be bilateral [VarghaKhadem 1998, Rapin 2003].

Although some studies have reported morphometric and functional brain anomalies in SLI, the results remain inconsistent and heterogeneous [Webster 2004, Friederici 2006a, Herbert 2007].

Volumetric studies have reported abnormal volume and/or asymmetries of the perisylvian fronto-temporo-parietal region [Plante 1991], inferior frontal region [Jerni-  
gan 1991] and posterior perisylvian region or Planum Temporale (PT) [Jernigan 1991,



Leonard 2002]. However, other studies have reported normal [Gauger 1997, Preis 1998] and even exaggerated [De Fossé 2004, Herbert 2005] asymmetry of the PT. Furthermore, voxel-based morphometry studies have reported no differences of grey matter for core language areas [Jäncke 2007] and grey matter increase in the right posterior Superior Temporal Gyrus (pSTG) [Soriano-Mas 2009].

Studies using SPECT have also detected perisylvian functional anomalies, including hypoactivation of the anterior perisylvian region at rest [Lou 1984] and the left inferior parietal region during phoneme discrimination [Tzourio 1994], as well as Broca's area during a dichotic verbal task [Chiron 1999]. However, [Ors 2005] reported hypoactivation at rest affecting the right parietal region and a right hyperactivation with lower left asymmetry involving the temporal lobes.

To our knowledge, three studies have investigated SLI using fMRI. In a study of 8 adolescents with SLI, [Weismer 2005] reported hypoactivation in the left parietal region and precentral gyrus during sentence comprehension, as well as in the left "insular portion of the IFG" during final word recognition, without any difference of lateralisation. [Dibbets 2006], in a study of 4 adolescents with SLI, found hyperactivations in frontal, temporal and angular regions during a non-verbal executive paradigm. Finally, based on listening to speech sounds in 5 members of the same family, [Hugdahl 2004] reported a leftward but smaller and weaker temporal activation than in controls. Thus, the results from neuroimaging studies of SLI remain inconsistent, which may be partly due to the heterogeneity of SLI [Rapin 2003, Friederici 2006a, Herbert 2007, Whitehouse 2008]. In addition, although fMRI is promising for investigating developmental language disorders [Friederici 2006a, Gaillard 2006], there are as yet very few fMRI studies.

One issue of fMRI language mapping is its sensitivity and specificity, i.e. the ability to draw a comprehensive and selective picture of the essential language network [Medina 2007, Tie 2008]. Since any single language task is unlikely to engage all aspects of language, and be limited to language processing alone, one strategy is to use several tasks targeting distinct aspects of language [Deblaere 2002, Roux 2003, Gaillard 2004]. This makes it possible to focus separately on parts of the network that are distinctly recruited by the tasks, while combined tasks analysis provides a more robust laterality assessment [Ramsey 2001, Rutten 2002].

fMRI in children requires special precautions due to the risk of movement, attentional constraints, task design, and preparation of the child [O'Shaughnessy 2008, Leach 2010]. The procedure may be even more problematic with young and language-disordered children, who may not be able to do classical tasks involving reading or verbs-to-words and words-to-letter generation, which have been used with typically-developing children. Similar difficulty arises with metalinguistic tasks, which, in addition, require explicit forced-choice analysis and are likely to involve undesired executive effects [Blank 2002, Crinion 2003].

Following fMRI studies that developed task panels for language mapping in children [Gaillard 2004, Wilke 2006, Holland 2007], we developed and tested with typically-developing children a four-task panel that was specifically designed to be feasible for use with young and language-disordered children [de Guibert 2010]. The

whole procedure avoids reading, metalinguistic or complex executive requirements. It makes use of auditory and visual stimuli, and involves language comprehension and production.

In the present **fMRI** study, we use this task panel to compare a group of children with **T-SLI** to a matched group of typically-developing children. All **T-SLI** children were referred to a hospital specialized centre and diagnosed on both psychometric and clinical grounds. Data processing includes single task, group and between-group analysis, as well as **LI** assessment and comparison within **ROIs** using single tasks and combined tasks analysis.

## 2 Methods

### 2.1 Participants

The study was approved by the regional ethics committee of the University Hospital. Parents and children were informed about the experiment; parents signed the informed consent and children gave their verbal assent.

Children with **T-SLI** were native French speakers recruited among children referred to the Centre for Language and Learning Disorders (University Hospital). They were diagnosed on psychometric and clinical grounds by the interdisciplinary team, after neuropsychiatric, neuroradiological, neuropsychological and language examinations. Out of the twenty-five children initially recruited, four were finally excluded, because of associated attention-deficit hyperactivity disorder ( $n = 1$ ), non-verbal index within the deficit range (i.e.  $< 70$ ;  $n = 1$ ), and fear ( $n = 1$ ) or teeth braces ( $n = 1$ ) preventing the **MRI** session. This resulted in a group of 21 **T-SLI** children aged from 7 to 18 years (mean age =  $11.4 \pm 3.3$  years old), with 9 boys (mean age =  $11.4 \pm 3.7$  years old) and 12 girls (mean age =  $11.4 \pm 3.1$  years old). Three children were left-handed (14.3%), as assessed using the Edinburgh inventory [Oldfield 1971], which is within the estimation of 8–15% left-handers for the general population [Hardyck 1977]. None exhibited any neurological anomalies or auditory deficit, or was affected by communication, behavioural or attentional disorders. The visual inspection of anatomical 3D **T1w** and **FLAIR** images by an experienced neuroradiologist showed no significant abnormalities.

A matched group of typically-developing children was recruited excluding non-French native language speakers, previous or current neurological, developmental or psychiatric illness, as well as learning disability or abnormal academic performance. They did not undergo psychometric testing. This resulted in a group of 18 children aged from 8.7 to 17.7 years (mean age =  $12.7 \pm 3$  years old), with 9 boys (mean age =  $12.3 \pm 3.2$  years old) and 9 girls (mean age =  $13.0 \pm 3.0$  years old). Two children (11.1%) were left-handed.

The **T-SLI** and control groups were similar for sex and handedness, and no significant between-group difference was found for age (two-sample t-test,  $p = 0.22$ ).

## 2.2 Neuropsychological and language assessment

Neuropsychological assessment includes achievement of the full *Wechsler Intelligence Scale for Children, Fourth version* (WISC-IV; [Wechsler 2003]). This version provides a Verbal Comprehension Index (VCI) and a Perceptual Reasoning Index (PRI) replacing the previous verbal and nonverbal intelligence quotients. Two subjects older than 16 years performed the full *Wechsler Adult Intelligence Scale, Third version* (WAIS-III; [Wechsler 1997]).

T-SLI children as a group showed a discrepancy between low verbal (VCI) and higher nonverbal (PRI) indexes (Table 6.1). Individually, nonverbal scores were all above the intellectual deficit range (i.e.  $PRI \geq 70$ ).

The language assessment included subtests from the *Langage oral, Langage écrit, Mémoire, Attention* battery (L2MA; [Chevrie-Muller 1997]). Since the L2MA is standardized from 8 years 7 months, four children below this age performed analogous subtests from the *Nouvelles Epreuves pour l'Examen du Langage*, standardized for younger children (N-EEL; [Chevrie-Muller 2001]). All children performed the 6 following subtests: phonology (repetition of complex unfamiliar words); vocabulary (picture naming); morphosyntactic integration and comprehension (sentence completion; sentence-picture matching); comprehension of complex instructions; sentence repetition.

Children with T-SLI as a group showed scores lower than 1 standard-deviation below the normal mean for phonology (repetition of complex unfamiliar words), sentence repetition, and morpho-syntactic integration (sentence completion) (Table 6.1). Individually, all children performed less than 1.5 standard deviation below the normal mean for at least one of these subtests. Thus, all children demonstrated impairments of the phonological or morpho-syntactic components of language, or both, which is characteristic of T-SLI. No child was diagnosed with developmental verbal dyspraxia, verbal auditory agnosia, or pragmatic language impairment.

## 2.3 Functional MRI protocol and task panel

The protocol has been previously described [de Guibert 2010] and is summarized here. The session includes four language tasks separately implemented with the same parameters, aiming to minimize attentional complications: a simple block design involves alternating a rest condition with a task, starting with rest. Each paradigm comprises three 27 s blocks of each condition and has a total duration of 2 min 48 s. The session has a duration of 30-35 min, including the anatomical acquisition and the four tasks. All subjects perform the tasks in the same order, as during the preparation step, to avoid the mixing of auditory and visual tasks. During the rest condition, children are asked “not to work”, to listen to the noise of the scanner, while fixing their attention on a red cross displayed on the screen.

The panel of tasks does not involve reading, metalinguistic (i.e. explicit analysis of language) or high-executive skills, and targets anterior and posterior core language areas. It uses auditory and visual stimuli delivery, solicits language com-

Table 6.1: Groups characteristics. Ranges are reported in brackets. Bold font indicates language scores smaller than one standard deviation below the normal mean. (a) Including scores of the four youngest children who performed equivalent subtests from the Nouvelles Epreuves pour l’Evaluation du Langage (N-EEL; see text). L2MA = battery Langage oral, Langage écrit, Mémoire, Attention; WISC-IV= Wechsler Intelligence Scale for Children-Fourth version; VCI and PRI= Verbal Comprehension and Perceptual Reasoning Indexes.

Control group ( $n = 18$ )	
Age (years)	12.7±3 (8.7–17.7)
Gender (male:female)	9:9
Handedness (left:right)	2:16
T-SLI Group ( $n = 21$ )	
Age (years)	11.4±3.3 (7–18)
Gender (male:female)	9:12
Handedness (left:right)	3:18
Intellectual indexes (WISC-IV)	
Verbal index (VCI)	77.0±15.7 (45 to 112)
Non-verbal index (PRI)	90.0 ± 13.1 (73 to 121)
Language z-scores (L2MA) <sup>a</sup>	
Phonology (complex unfamiliar word repetition)	-2.4 ± 1.7 (-6.9 to 0.4)
Vocabulary (picture naming)	-0.3 ± 1.0 (-2.6 to 1.3)
Morphosyntactic integration (sentence completion)	-1.2 ± 1.5 (-5.0 to 1.3)
Complex instructions comprehension	-0.2 ± 1.1 (-2.2 to 1.7)
Morphosyntax-Comprehension (sentence-picture match)	-0.5 ± 1.0 (-2.5 to 1.1)
Sentence repetition	-1.4 ± 0.8 (-2.5 to 0.8)

prehension and production, and involves lexico-semantic and phonologic processing.

Two auditory lexico-semantic tasks were chosen from the literature because of their distinct and selective activations of either the left IFG (word generation from categories, [Gaillard 2003] hereafter category task) or the left STG (auditory responsive naming, [Balsamo 2002]; hereafter definition task).

In the category task adapted from [Gaillard 2003], children hear category names (e.g. animals, colours) and have to silently generate examples of these categories. A category name is delivered every 9 s, with three categories per block and nine categories for the whole task.

In the definition task adapted from [Balsamo 2002], children hear definitions of concepts (e.g. “a big animal with a trunk”) and have to find and silently name the corresponding word (e.g. elephant). Definitions are delivered every 9 s, with three definitions per block and nine definitions for the whole task.

Two new visual phonological tasks are based on picture naming and require the child to silently name three objects repetitively (i.e. triplets) one-by-one. Within the triplets, the names are semantically unrelated, but exhibit a minimal phonological change, either, for the first task, a phonological minimal difference (hereafter

phon-diff task), or, for the second task, a change in segmentation (hereafter phon-seg task). These tasks are adapted from so-called minimal pairs in linguistics and from procedures used in the assessment and remediation of phonological disorders. The repetitive evocation of just three familiar but phonologically close words attenuates the lexico-semantic requirements, while stressing the phonological constraints.

In both tasks, line-drawings of objects from each triplet are displayed every 1.4 s, successively and randomly (without any picture being delivered twice in succession), so that the child could not predict the upcoming picture. Three distinct triplets (one per block) are used for each task.

In the phon-diff task, children name objects such as “poule, boule and moule” (/pul-/bul-/mul/; [hen, ball, tin]). In the triplets, a difference of distinctive feature occurs in the first phoneme, e.g. the voicing feature of /p/ and /b/ (voiceless vs. voiced) for poule and moule. Concretely, children may successively name, for example: “poule, moule, boule, moule, poule...” in a first block, then: “banc, dent, gant, dent, banc...” (/bã/, /dã/, /gã/...; [bench, tooth, glove]) in a second block, and so on.

In the phon-seg task, children name objects such as “car, car-te and car-ton” (/kar-/kart-/kartõ/; [car, card, cardboard box]). In the triplets, there is a change of segmentation, with phoneme addition (car and then carte or carton) or subtraction (carton and then carte or car). Concretely, children may successively name, for example: “car, carte, carton, carte, car...” in a first block, then: “croix, roi, oie, roi, croix...” (/krwa/, /rwa/, /wa/...; [cross, king, goose]) in a second block, and so on.

For all tasks, children were prepared extensively prior to entering the scanner, with each task being thoroughly explained and practised several times, using original task material and both aloud and silently, to check for comprehension and achievement by the child.

## 2.4 Data acquisition

Acquisitions were performed on a 3 T whole-body scanner (Achieva, Philips Medical Systems, Best, The Netherlands) using a 8-channel head coil. Anatomical 3D T1w images were acquired with a Fast Field Echo sequence. The acquisition parameters were as follows: TE/TR/Flip angle: 4.6 ms/9.9 ms/8°; acquired matrix size: 256 mm x 256 mm; FOV: 256 mm; voxel size: 1 mm x 1 mm x 1 mm; volume: 160 sagittal 1 mm thickness slices; acquisition time: 3 min 56 s. Functional images were acquired using a single-shot T2\*w gradient-echo echo planar imaging sequence. Twenty-four 4 mm slices were acquired with the following parameters: TE/TR/Flip angle: 35 ms/3000 ms/90°; acquired matrix size: 80 x 80; reconstructed matrix size: 128 x 128; FOV: 230 x 230; acquired voxel size: 2.9 mm x 2.9 mm x 4 mm; reconstructed voxel size: 1.8 mm x 1.8 mm x 4 mm. Slices were positioned parallel to the anterior commissure-posterior commissure line, with no gap, and were interleaved from bottom to top. Each functional run consisted of 56 series of image acquisitions for the 24 slices covering the entire brain volume separated by a 3000 ms delay,

with a total acquisition time of 2 min 48 s. Children were positioned supine in the system. The subject’s head motion was minimized using straps and foam padding.

Visual stimuli were delivered through a screen placed within the head-coil (IFIS-SA fMRI system, Invivo, Orlando, FL) just in front of the face, and synchronised with the scanner. If necessary, the children wore corrective glasses compatible with the high-magnetic-field environment. Auditory verbal stimuli were delivered by an experienced member of the staff using the machine microphone, via specially converted high-fidelity stereo headphones.

## 2.5 Data processing

MRI data were pre-processed and analysed using the GLM [Friston 1994a] with SPM5<sup>1</sup> (Statistical Parametric Mapping 5; Wellcome Department of Imaging Neuroscience, University College, London). The first two volumes of fMRI data were discarded to allow for signal stabilization. Slice timing and motion correction were applied to the remaining 54 volumes. Data were excluded if associated with excessive motion (more than 3 mm of translation in any direction or 3° of rotation throughout the session). To prevent bias caused by the normalization of paediatric data on adult templates [Wilke 2003d], we used the match pair option of the Template-O-Matic toolbox [Wilke 2008] to generate a customized paediatric template based on the age and sex of our 39 subjects. Structural MRI were segmented using unified segmentation [Ashburner 2005], and then normalized. FMRI data were registered on structural images, normalized and then smoothed using an isotropic 8-mm FWHM 3D Gaussian kernel.

Statistical activation maps were obtained using a mixed effects analysis. At the subject level, a high-pass filter was applied to fMRI data to remove slow signal drifts due to undesired effects. The haemodynamic response was modelled by the Informed Basis Set [Friston 1998] to account for possible delay and dispersion of the response from the canonical haemodynamic response function. For each task, group activations were identified by contrasting out the effect of temporal and dispersion derivative, focusing on the canonical variable, at a threshold of  $p < 0.05$  FWE at the cluster level with a cluster-defining threshold of  $p < 0.001$ . Between-group comparisons were considered significant at the cluster level at a threshold of  $p < 0.05$  FWE, with a cluster-defining threshold of  $p < 0.005$ .

For regional analyses, ROIs covering brain areas involved in language were selected from the literature: the IFG-opercularis, IFG-triangularis, STG, SMG, and insula. For laterality assessment, the following extended ROIs were also computed: “frontal language” (IFG-opercularis, IFG-triangularis, insula), “temporo-parietal language” (STG, SMG), “language” (i.e. combining the two previous regions) and “non-language” (i.e. all brain ROIs except the “language ROI”). Left and right ROIs as delineated in the AAL atlas [Tzourio-Mazoyer 2002] were adapted to our customized paediatric template using an approach by [Wilke 2003c]. To match our template, we performed a non-linear deformation of the structural image on which

---

<sup>1</sup>[www.fil.ion.ucl.ac.uk](http://www.fil.ion.ucl.ac.uk)

the AAL atlas was delineated. The deformation parameters were then applied to the ROIs.

LIs were estimated using the LI toolbox [Wilke 2007]. For each subject, the average t-value within each ROI was measured and voxels smaller than this threshold were discarded. The LI was then calculated with the remaining voxels as follows:

$$li = \frac{\sum \text{Activation}_L - \sum \text{Activation}_R}{\sum \text{Activation}_L + \sum \text{Activation}_R}, \quad (6.1)$$

where  $\sum \text{Activation}_L$  and  $\sum \text{Activation}_R$  denote the sum of the remaining voxels in the left and right parts of the ROI, respectively.

LIs were calculated for each single task and also using combined task analysis [Ramsey 2001], the latter being known to yield more robust LIs when dealing with a panel of tasks [Rutten 2002]. Boxplots based on these values were created for each region and each task. T-tests were performed on each ROI to determine significant group lateralisation (i.e. left or right if  $li$  was significantly greater or smaller than zero, respectively; otherwise bilateral). Correlation with age was assessed using covariance analysis including the factors group, age and interaction. When the effect of interaction and age were non-significant, and based on current hypotheses about SLI, a one-tailed two-sample t-test was performed to highlight laterality differences in language ROIs, while a two-sided two-sample t-test was performed in non-language ROIs.

### 3 Results

The results are reported successively for each single task, focusing on language areas, including group analysis (fig. 6.1), between-group comparisons (fig. 6.2), as well as LI measurements and comparison (fig. 6.3). Finally, LIs comparison using the combined task analysis is reported (fig. 6.4).

#### 3.1 Auditory language tasks

**Category task** The control group shows left-only activations in the dorsal IFG as well as in the pSTG with, as expected, a predominance of the former. A left-dominant activation is situated in the anterior insula and extends into the ventral IFG. According to the LIs, the left lateralisation is significant in the IFG-opercularis and -triangularis.

In contrast, the T-SLI group shows small activations in the left dorsal IFG, and no activation in the most posterior STG. The activation of the anterior insula is right-dominant and extends into the ventral IFG on the right. According to the LIs, no ROI is significantly lateralised.

The between-group analysis does not reveal any significant between-group differences. However, the LI comparison reveals a significant lack of left lateralisation of the IFG-opercularis in the T-SLI group (fig. 6.3).

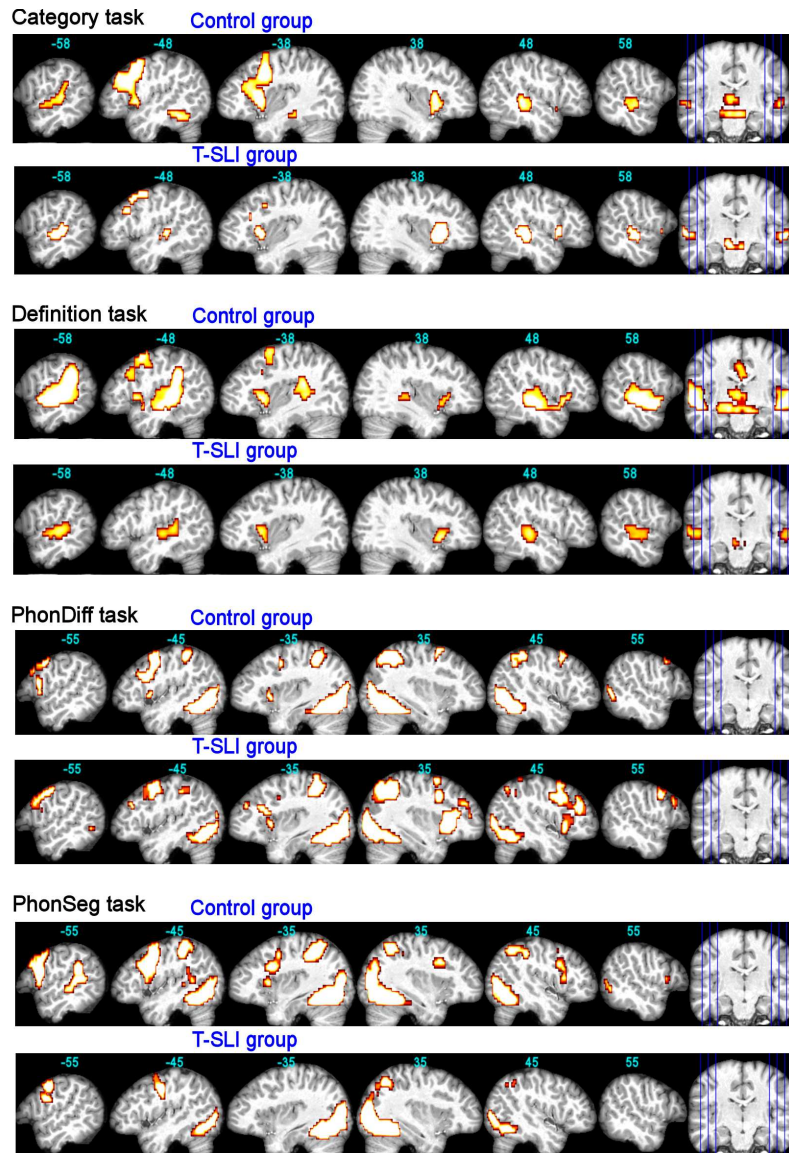


Figure 6.1: FMRI group effects for each language task ( $p < 0.05$  FWE). The functional maps are super-imposed onto an individual brain normalized with respect to our customized paediatric template, with x-coordinates in MNI space. Left slices are left hemisphere.

**Definition task** The control group shows left-only activations in the pSTG/adjacent SMG, and in the dorsal IFG, with an expected predominance of the former. A left-dominant activation occurs in the anterior insula and extends into the adjacent ventral IFG. According to the LIs, the SMG and IFG-opercularis are significantly left-lateralised.

In contrast, the T-SLI group shows no activation in the pSTG/adjacent SMG or in the dorsal IFG. A bilateral activation is centred on the anterior insula and



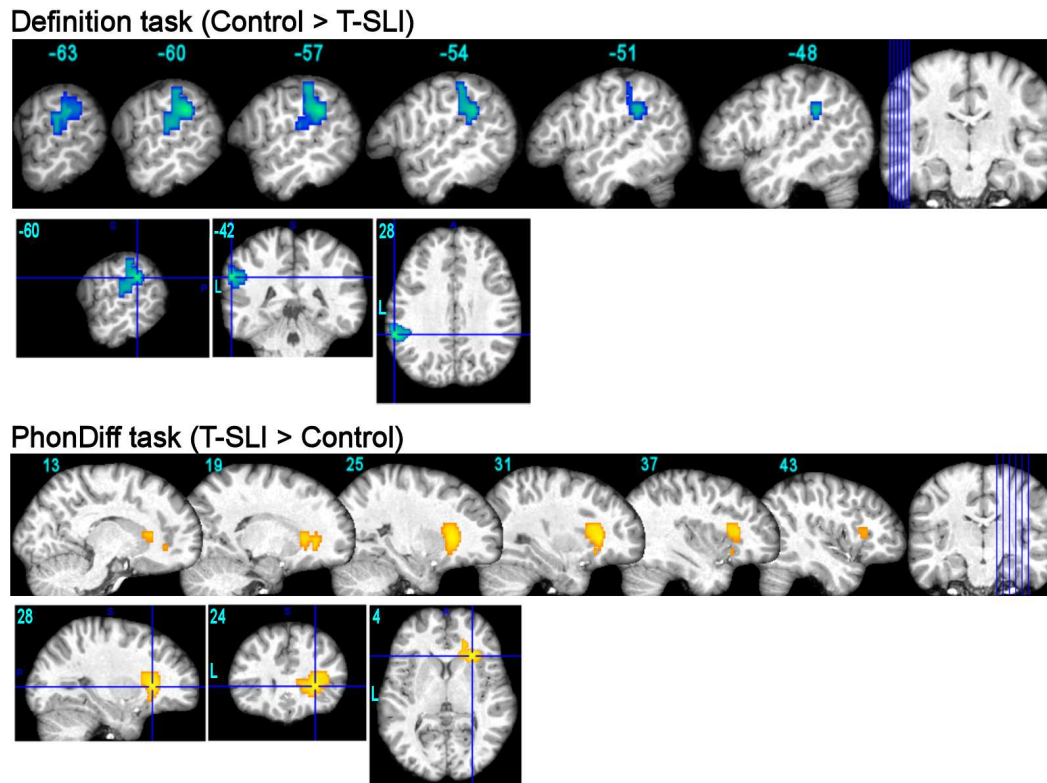


Figure 6.2: FMRI between-group comparisons for definition and phon-diff tasks ( $p < 0.05$  FWE). Blue and yellow colours indicate the hypo- and hyperactivations, respectively, for the T-SLI group compared with the control group. The 3D view focuses on the peak contrasts. Functional maps are superimposed on an individual brain normalized with respect to our customized paediatric template. Coordinates are in Montreal Neurological Institute space.

extends, superiorly on the right, into the ventral IFG. According to the LIs, no ROI is significantly lateralised, although the IFG and the SMG tend towards the right.

The between-group analysis highlights a left hypoactivation centred on the pSTG/SMG junction in the T-SLI group (fig. 6.2;  $k = 304$ ;  $T = 4.42$ ;  $p = 0.03$ ). According to the LI comparison, there is a significant lack of left lateralisation of the SMG, IFG-opercularis and IFG-triangularis in the T-SLI group (fig. 6.3).

### 3.2 Visual language tasks

**Phonological minimal difference task** The control group shows left-only activations in the ventral and dorsal IFG and in the anterior insula, without activation in the STG. According to the LIs, the IFG-opercularis is significantly left-lateralised.

In contrast, the T-SLI group shows a right-only activation in the IFG and a right-dominant activation in the anterior insula. According to the LIs, the IFG-triangularis is significantly right-lateralised.

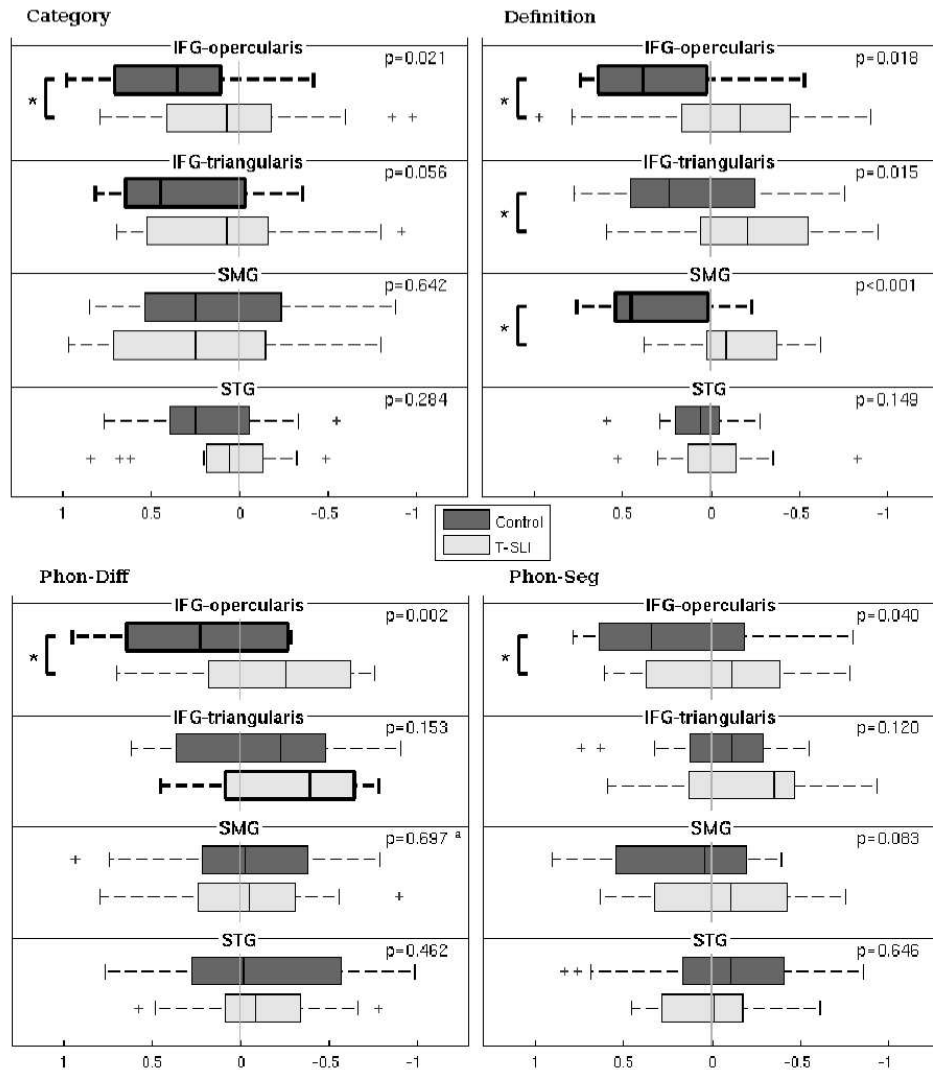


Figure 6.3: LIs, significant group lateralisations and between-group comparisons within the ROIs for each single language task. The box plots depict group lateralisation for each ROI, with positive and negative lateralisation indices reflecting left and right, respectively, and with significant left or right lateralisations ( $p < 0.05$ ) outlined by bold lines. The p-values of between-group comparisons are indicated, with significant between-group differences outlined by a square bracket with an asterisk. (a) P-value of the group factor from analysis of covariance with factors age and group.

Between-group analysis highlights a right hyperactivation centred on the anterior insula, extending into the adjacent ventral IFG (opercularis and triangularis) and into the head of the caudate nucleus in the T-SLI group (fig. 6.2;  $k = 362$ ;  $T = 4.34$ ;  $p = 0.02$ ). According to the LIs comparison, there is a significant lack of left lateralisation of the IFG-opercularis in the T-SLI group (fig. 6.3). A positive

effect of age on LI is found in the SMG ( $p = 0.024$ ), but no group difference is detected in this region.

**Phonological change in segmentation task** The control group shows a left-dominant activation in the ventral and dorsal IFG-opercularis, as well as left-only activations in the pSTG/adjacent SMG and in the anterior insula. According to the LIs, no region is significantly lateralised.

In contrast, the T-SLI group shows no activation in the IFG, the STG, the SMG or the insula. According to the LIs, no ROI is significantly lateralised.

The between-group analysis does not reveal any significant between-group differences. However, the LI comparison highlights a significant lack of left lateralisation of the IFG-opercularis in the T-SLI group (fig. 6.3).

### 3.3 Lateralisation index assessment and comparison using the combined task analysis

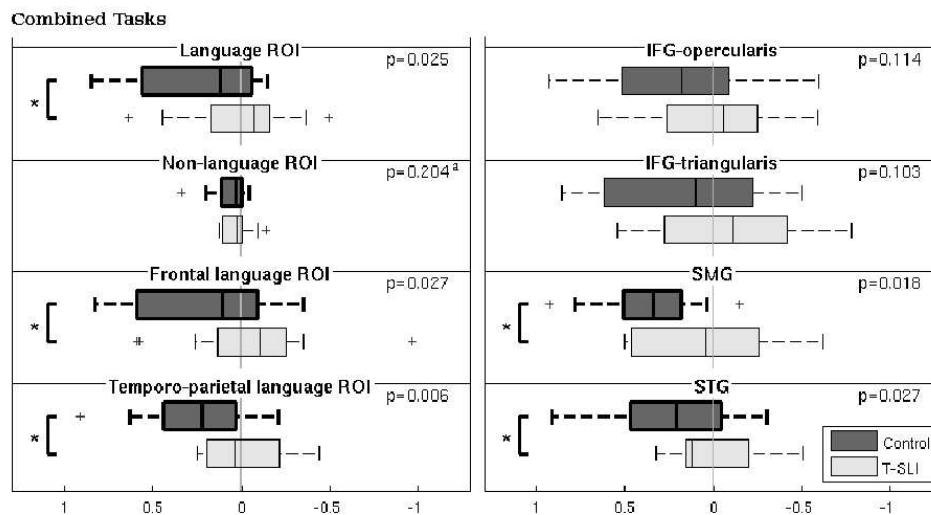


Figure 6.4: LIs, significant group lateralisations and between-group comparisons using combined tasks analysis within single (right) and extended (left) ROIs. The box plots depict group lateralisation for each ROI, with positive and negative lateralisation indices reflecting left and right, respectively, and with significant left or right lateralisations ( $p < 0.05$ ) outlined by bold lines. The P-values of between-group comparisons are indicated, with significant between-group differences outlined by a square bracket with an asterisk. (a) P-value from two-sided two sample t-test. Temporoparietal language ROI = STG and SMG; frontal language ROI = IFG (opercularis and triangularis) and insula; language ROI = frontal language and temporoparietal language ROIs; non-language ROI = whole brain (i.e. all AAL atlas ROIs) except the language ROI.

According to the assessment of LIs using the combined task analysis, the control group exhibits a left lateralisation of the SMG and STG. This is not observed with the T-SLI group, where no ROI is lateralised. According to the between-group comparison, the lack of left lateralisation in the T-SLI group is significant for the SMG and the STG (fig. 6.4).

Subsequently, the combined task analysis was carried out using extended ROIs: “frontal language”, “temporo-parietal language”, “language” (i.e. combining the two latter) and “non-language ROI” (i.e. all AAL atlas ROIs except the latter). In the control group, LIs from the combined task analysis show a left lateralisation in all these extended ROIs. This is not the case for the T-SLI group, where no lateralisation appears, despite a right trend for the “frontal language” and “language” ROIs.

The between-group comparison of the LIs highlights a significant lack of left lateralisation in all extended language ROIs (i.e. “temporo-parietal language”, “frontal language” and “language”), while, inversely, there is no significant difference for the “non-language” ROI (fig. 6.4).

In summary, our main results highlight a left hypoactivation centred on the pSTG/SMG junction (definition task), a right hyperactivation of the anterior insula including the adjacent IFG and extending into the head of caudate (phon-diff task), and a lack of left lateralisation of core language areas in the T-SLI group. The lack of left lateralisation is found for the IFG-opercularis (all tasks), the IFG-triangularis (definition task), the SMG (definition task and combined tasks), the STG (combined tasks), and in all extended language ROIs when using combined tasks. On the contrary, there is no difference of lateralisation for the rest of the brain when the language regions are excluded.

## 4 Discussion

Although functional neuroimaging may have an expanding role in the investigation of development language disorders, fMRI studies of SLI are sparse and available results from functional studies remain inconsistent, in parallel with heterogeneous morphometric findings. Based on a comparison with typically-developing children, we studied a group of 21 children with T-SLI, a main form of SLI affecting structural aspects of language, which was diagnosed on psychometric and clinical grounds. To apply an appropriate procedure and to improve the mapping by using several tasks, we set up a panel of tasks without reading, metalinguistic or high attentional requirements. Three main interesting results arise from our study:

1. The lack of left lateralisation of core language areas;
2. The left hypoactivation centred on the pSTG/SMG junction (definition task);
3. The right hyperactivation of the anterior insula including the adjacent IFG and extending into the head of caudate (phon-diff task).

### 4.1 Lack of left lateralisation of core language areas

The study reveals a lack of left functional lateralisation for all single language ROIs across single or combined tasks, i.e. in the IFG-opercularis (all tasks), the IFG-triangularis (definition task), the SMG (definition and combined tasks), and the STG (combined tasks). In addition, when using combined tasks analysis, this lack

also applies to larger frontal and temporo-parietal language ROIs, while, interestingly, no significant difference appears for the whole brain when language ROIs are excluded. Thus, our study provides evidence that T-SLI is associated with atypical lateralisation of language function in core language areas.

No anomaly of lateralisation was found or reported in previous fMRI studies of SLI [Hugdahl 2004, Weismer 2005, Dibbets 2006], including a study of oro-facial verbal dyspraxia [Liégeois 2003], which could be due to reduced sample sizes, distinct activation tasks, and/or distinct clinical subtypes. However, SPECT studies at rest have reported a reversed asymmetry in Wernicke’s area [Chiron 1999] and a more symmetric activation in the temporal lobe [Ors 2005]. Furthermore, one study using functional transcranial Doppler ultrasonography during words-to-letter generation reported a lack of leftward dominance of blood flow in adults with persisting SLI, a condition more associated with structural language impairment than transient SLI [Whitehouse 2008].

Our results support the hypothesis of atypical cerebral dominance for literacy and language developmental disorders, although atypical lateralisation is not in itself either abnormal or specific. Atypical functional lateralisation has been reported in 5% of right-handed and 73–80% of non-right-handed normal subjects [Szaflarski 2002] and in other developmental clinical conditions including speech delay [Bernal 2003], stuttering [Brown 2005], autism spectrum disorder [Knaus 2008, Kleinhans 2008], and dyslexia [Maisog 2008, Heim 2010].

Therefore, our study shows that a well-defined form of SLI affecting structural aspects of language is more associated with atypical functional lateralisation of core language areas, but we cannot yet affirm that it is a specific marker of T-SLI. As argued by [Whitehouse 2008], atypical cerebral lateralisation may be an indicator, albeit imperfect, of some causal factor that leads together to atypical cerebral lateralisation and language impairment.

#### 4.2 Left hypoactivation of the posterior superior temporal gyrus / supramarginal gyrus junction

The second result, provided by the definition task, is the left hypoactivation centred on the junction of the posterior supratemporal plane (STG) and the SMG, extending laterally, deeply into the Sylvian fissure, and superiorly in the parietal operculum and parietal inferior lobule. This region is crucial for language, belonging to the so-called “Wernicke’s area” or “territory” [Blank 2002, Catani 2005].

The central location of the hypoactivation corresponds to the posterior PT/ventral SMG region [Price 2010] or “area sylvian parietal-temporal” [Hickok 2007] which may be a sensorimotor interface translating acoustic speech signals from the posterior temporal sulcus into articulatory representation for the premotor cortex and posterior IFG. This region is involved in both complex speech perception and production (e.g. [Hickok 2003, Price 2010]), and its lesion may be associated with conduction aphasia, which exhibits phonemic paraphasias with better preserved comprehension [Hillis 2007]. Therefore, on the whole, the putative function of this

region is in agreement with the task used here involving both speech reception and (covert) production, and with the linguistic deficit inherent to T-SLI.

In line with our result, SPECT studies of SLI have reported bilateral posterior perisylvian hypoactivation at rest [Lou 1984], and no activation of the left inferior parietal region during phonological discrimination [Tzourio 1994]. One fMRI study has also reported a hypoactivation in the left parietal lobe during sentence comprehension [Weismer 2005]. In contrast, using SPECT at rest, [Ors 2005] reported a hypoactivation of the right parietal region.

One question is whether our result can be linked to morphological anomalies. Volumetric studies of SLI have shown a reduced volume of the perisylvian temporoparietal region [Jernigan 1991] and reduced left asymmetry of the PT when compared with reading disability [Leonard 2002]. However, other authors report normal volume and asymmetry for the PT (leftward) and for the parietal ascending ramus (rightward) [Gauger 1997, Preis 1998] and even an exaggerated leftward asymmetry for the PT [De Fossé 2004, Herbert 2005], without any difference for the parietal opercule and the SMG [De Fossé 2004]. Using voxel-based morphometry, [Jäncke 2007] found no grey matter differences for core language areas, and [Soriano-Mas 2009] reported an increase in grey matter at the right temporoparietal junction. Apart from methodological differences, this heterogeneity may result from the heterogeneity of SLI.

In summary, the left hypoactivation of the pSTG/SMG junction found in T-SLI children during the auditory responsive naming task could reflect a dysfunction of a core region considered as an interface between complex language reception and production [Hickok 2007, Price 2010]. This result converges with a SPECT study of SLI using a discrimination task involving phonologically-close words [Tzourio 1994].

### 4.3 Right hyperactivation of the anterior insula, adjacent inferior frontal gyrus and head of caudate

The third main result of this study is the right hyperactivation centred on the anterior insula and extending into the adjacent IFG (opercularis and triangularis), as well as into the head of caudate, during the phon-diff task. The anterior insula and adjacent IFG is an important region for language, in continuity with the frontal operculum and, at the left side, Broca's area [Keller 2009]. The head of caudate has already been highlighted in a developmental speech disorder, oro-facial verbal dyspraxia [VarghaKhadem 2005].

The anterior insula is involved in motor aspects of speech, although its specific role for speech remains unclear. Overall, this region has been linked with coordination and motor control of speech articulation, vocal tract and mimic muscles, as well as non-speech orofacial gestures, swallowing and respiratory voluntary regulation [Brown 2005, Brown 2009, Ackermann 2010, Price 2010]. Clinical studies have yielded some controversial results, with one study based on lesion overlap reporting a correlation of left insular lesion with deficits in motor programming of speech [Dronkers 1996], which was not replicated using neuroimaging at

stroke onset [Hillis 2004]. Single clinical cases and brain stimulation studies involving the insula have reported aphasic, dysarthric, speech initiation and/or non-speech oro-motor disturbances, and functional neuroimaging studies have revealed the involvement of the anterior insula, predominantly on the left, in motor aspects of speech [Ackermann 2010, Price 2010]. Interestingly, [Bohland 2006] found increased bilateral activation of the insula/IFG junction in proportion to phonological complexity by requiring triads of syllables of varying complexity (“ta-ta-ta”/“karu-ti”/“stra-stra-stra”/“kla-stri-splu”), which suggests a function of integration of low-level motor aspects, abstract speech sounds and prosodic components in speech planning. The right anterior insula/frontal operculum may specifically mediate supra-segmental aspects of speech (prosody, intonation contour), as well as vocal imitation and musical melodies [Brown 2005, Ackermann 2010].

Therefore, the hyperactivation in the right anterior insula/adjacent IFG during our phonological task could reflect an articulatory and/or prosodic compensatory mechanism of defective structural phonological function in T-SLI. Such a compensation could be inter-hemispheric (i.e. left to right) and possibly intra-hemispheric (i.e. lateral IFG to insula). Compensatory interhemispheric recruitment of the right inferior frontal region is well-known after left acquired lesions associated with aphasia [Crinion 2007b], and the recruitment of the right insula may compensate dysfunction of the left counterpart [Duffau 2001]. An intrahemispheric shift of frontal response towards the anterior insula/frontal operculum, at the right side, has also been reported in subjects with left temporal lobe epilepsy [Voets 2006]. If further corroborated, this suggests that the insular structure might be involved in the compensation of speech/language function [Ackermann 2010].

Previous functional neuroimaging studies of SLI yield results that are heterogeneous with respect to the insula. A right hyperactivation has been highlighted in the anterior part during speech sound listening [Hugdahl 2004] and in the posterior part during a non-verbal executive paradigm [Dibbets 2006]. This contrasts with hypoactivation on the left during word recognition [Weismer 2005]. While one volumetric study reports a volume reduction of the left insula [Jernigan 1991], no anomaly of the insula has been observed using voxel-based morphometry [Jäncke 2007, Soriano-Mas 2009].

In our study, the right hyperactivation extends into the head of caudate. The caudate nucleus participates in sensorimotor coordination including response selection and initiation, in executive-related processes, and may support the planning and execution of correct strategies required for complex goals [Grahn 2008]. With regards to speech, the caudate nucleus is involved in the control and selection of articulatory motor sequences, and may initiate cortical phonological and controlled processes when automatic processes are not well-suited [Friederici 2006b, Booth 2007]. Furthermore, the bilateral head of caudate is involved in language-based conflict, suggesting that it participates in the suppression of inappropriate responses in a competitive context [Price 2010]. Similarly, the hyperactivation of the right head of caudate during the phon-diff task in our study could reflect higher compensatory attempts of initiation of phonological processes in the con-

text of phonological conflict (e.g. the minimal difference between pain/bain/main).

Regarding previous neuroimaging studies of SLI, the right caudate has been found to be hyperactive during a non-verbal switch paradigm in four children [Dibbets 2006]. Moreover, it was bilaterally reduced in the volumetric study by [Jernigan 1991], but increased on the left in the voxel-based morphometry study by [Soriano-Mas 2009].

In summary, the right hyperactivation of the anterior insula/adjacent IFG and the head of caudate in T-SLI, which is highlighted when requiring phonological differentiation, could reflect compensatory recruitments of non language-specific functions resulting from the structural phonological deficit. This could be associated with higher recruitment of orofacial and intonative motor functions for the anterior insula/adjacent IFG, and response initiation and selection in the context of interferences for the head of caudate.

#### 4.4 Comparison to other developmental disorders

Another question is whether the functional abnormalities reported in our study are specific to structural language disorder (i.e. T-SLI) compared to disorders affecting other aspects of language such as communication, reading or speech.

With regards to the left temporoparietal hypoactivation, in autistic spectrum disorder, on the contrary, hyperactivation has been detected in the posterior temporal region during a responsive naming task [Knaus 2008] and other language tasks [Just 2004, Harris 2006]. On the other hand, left temporoparietal hypoactivation appears to be a “neural signature” of dyslexia [Shaywitz 2008], as obtained during reading, rhyme or semantic tasks [Paulesu 1996, Schulz 2008, Richlan 2009]. Phonological disturbances in dyslexia nevertheless concern metalinguistic tasks (i.e. phonological awareness) rather than direct oral language, likely reflecting a phonological-access deficit [Ramus 2008], and the differentiation from SLI remains clinically and aetiologically justified [Bishop 2004b]. Finally, with regards to speech disorders, developmental stuttering has been associated with hypoactivation of the auditory cortices, but not of the temporoparietal junction [Brown 2005, Watkins 2008], while oro-facial verbal dyspraxia in the KE family was associated with hypoactivation near the left pSTG/SMG junction during covert verb generation [Liégeois 2003]. Although the core deficit in this family is dyspraxic, affected members also exhibit phonological and grammatical impairments, so linguistic deficits cannot be ruled out [VarghaKhadem 2005].

With regards to the right hyperactivation of the anterior insula/adjacent IFG, in autistic spectrum disorder, the insula has been found to be hypoactive during tasks involving social processing [Uddin 2009]. In dyslexia, hyperactivation of the anterior insula has been observed during reading tasks, either on the left or on the right [Maisog 2008, Richlan 2009]. However, one study reports a bilateral hypoactivation in parallel to higher activation of the adjacent frontal operculum during click and speech sound listening [Steinbrink 2009]. With regards to speech disorders, the anterior insula has been found to be hyperactive at the left side during overt words



repetition in oro-facial verbal dyspraxia [Liégeois 2003]. Nevertheless, it is noteworthy that hyperactivation of the right frontal operculum/anterior insula during speech production is considered as a “neural signature” of developmental stuttering [Brown 2005, Watkins 2008]. This suggests that lower skills for speech result in compensatory hyperactivation of vocal-motor areas, which are right- lateralised because of left dysfunction of the normal dedicated regions [Preibisch 2003]

With regards to the right hyperactivation of the caudate, no functional anomaly of the caudate has been found either in autistic spectrum disorder during language tasks [Harris 2006, Knaus 2008], or in dyslexia [Maisog 2008, Richlan 2009]. Nevertheless, in studies of oro-facial verbal dyspraxia in the KE family, the caudate was found to be both bilaterally morphologically reduced [VarghaKhadem 1998, Watkins 2002, Belton 2003], and functionally hyperactive on the left [VarghaKhadem 1998]. In this context, these anomalies come under the hypothesis of a dysfunction of the frontostriatal network [VarghaKhadem 2005]. Finally, the activity of the caudate on both sides has been positively correlated with the severity of developmental stuttering [Giraud 2008], and a dysfunction of the basal ganglia has been speculated to underlie the deficit in the timing of speech motor initiation [Alm 2004], or to reflect a secondary dysfunction resulting from a left inferior frontal anomaly [Kell 2009].

In summary, the left temporoparietal hypoactivation associated with T-SLI is similar to results obtained by studies of dyslexia [Richlan 2009] and by one study of oro-facial verbal dyspraxia [Liégeois 2003]. Moreover, in developmental stuttering, the hyperactivation of the right anterior insula is regarded as a “neural signature” [Brown 2005] and the activity of the right caudate is correlated with the severity of impairment [Giraud 2008]. As these disorders and the activation tasks used are distinct, further studies are needed to elucidate whether these similarities reflect common dysfunctions, common atypical compensatory modes of resolution of language tasks, and/or more task-specific effects.

#### 4.5 Methodological considerations

As described previously [de Guibert 2010], to optimize the feasibility of the procedure for young disordered children, and minimize motion artefacts as well as attentional complications, we implemented four identical block-designed paradigms with a low-level condition as baseline (listening to the noise and fixing a red cross) and without requiring motor responses. These choices reduced the heterogeneity and complexity of the protocol, as the child did not have to understand and achieve supplementary control tasks and also give motor responses. Furthermore, although requiring motor responses is well suited for metalinguistic tasks (i.e. judgement tasks with explicit analysis), which may involve additional non-language functions [Blank 2002, Crinion 2003], it is not appropriate for investigations under more natural conditions such as word production.

However, these choices have some drawbacks. First, the achievement of the tasks cannot be directly assessed. Requiring overt responses would allow online perfor-

mance monitoring, but speech increases the risk of movement, which is crucial in the case of children, and especially with disordered children [O'Shaughnessy 2008]. Therefore, children were intensively prepared before the scanner session using the same order of tasks and stimuli, allowing us to check that they understood and were able to achieve the tasks, and were also questioned after the session. Thus, the design of the tasks and the preparation aimed to ensure high performance for each child during the scanner session. Secondly, low-level control conditions involve more non-language-specific coactivations [Wilke 2006, Holland 2007] and, in some research contexts, it may be important to use high-level control tasks to target more specific functions. Nevertheless, together with the use of several tasks, a low-level control condition makes it possible to map a comprehensive and specific left-lateralised language network in normal children or adults [Ramsey 2001, Tie 2008, de Guibert 2010], and was appropriate in our study for the comparison with language-impaired children. Furthermore, we used a combined tasks analysis to provide a robust assessment of language lateralisation.

Moreover, to select a representative sample of the general population, close to the clinical context for language-disordered children, we did not solely recruit right-handed children. In this study, the proportions of left-handed in the T-SLI and control groups are similar and within the normal range estimate. Furthermore, additional analyses excluding left-handed children ( $n=5$ ) were carried out, which showed no major change of results (e.g. out of 11 LI differences for the whole group, 8 remained significant and 2 were nearly significant ( $p = 0.057$ ) when eliminating left-handers). Moreover, to avoid distortions due to normalization of the children's data with respect to an adult template, we used a tool dedicated to the creation of pair- and group-matched normalized templates based on normative brain data [Wilke 2008].

Finally, a well-known issue with the category of SLI is its clinical heterogeneity, since it encompasses impairments reflecting structural rule-like deficits (i.e. mainly phonology and morphosyntax), as well as impairments of articulatory, auditory receptive or pragmatic aspects of language. Since this heterogeneity may be a crucial source of inconsistency in the neuroimaging results, we focus here on structural language impairments, known as T-SLI or linguistic dysphasia. Future studies need to investigate whether distinct subtypes of SLI are associated with distinct brain functional anomalies. Secondly, since the current psychometric diagnostic tools are not totally adequate when used in isolation, because they lack clinical congruence, we selected T-SLI children on both psychometric and clinical grounds. The T-SLI children, as a group, failed three subtests (repetition of unfamiliar words, sentence completion, and sentence repetition) that are acknowledged as being especially sensitive to SLI, and all of the subjects had an early history of selective language impairment and had been diagnosed in our specialized hospital Centre.

In conclusion, by using fMRI with a panel of distinct language tasks, this study provides evidence that a well-defined type of SLI affecting structural components of language is associated with a lack of left functional lateralisation in core language areas (pars opercularis and triangularis of the IFG; STG; SMG), with hypoactiva-

tion of the left superior temporoparietal junction, within Wernicke's area, as well as with hyperactivation of the right anterior insula, adjacent inferior frontal gyrus and head of caudate. These results are similar to some findings from studies of developmental disorders involving other aspects of language such as dyslexia, stuttering or orofacial verbal dyspraxia, which will require further comparisons.

## **Part III**

# **Patient-specific analysis of brain function**



## Chapter 7

# Robust Cerebral Blood Flow Maps in Arterial Spin Labelling

---

The introduction of [ASL MRI](#) techniques has made feasible a non-invasive measurement of the [CBF](#). However, to date, the low signal-to-noise ratio of [ASL](#) gives us no option but to repeat the acquisition in order to accumulate enough data to get a reliable signal. The perfusion signal is then usually extracted by averaging across the repetitions. However, the sample mean is very sensitive to outliers as a single incorrectly large observation can cause strong detrimental effects on the sample mean estimate.

We propose to estimate robust [ASL CBF](#) maps with M-estimators to overcome the deleterious effects of outliers. The behaviour of this method is compared to z-score thresholding as recommended in [\[Tan 2009\]](#). Validation on simulated and real data is provided. Quantitative validation is undertaken by measuring the correlation with the most widespread technique to measure perfusion with [MRI: DSC](#).

Part of this work was published in the Multimodal Brain Image Analysis (MBIA) workshop of the International conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) in 2012 [\[Maumet 2012c\]](#). An extension of this first analysis was then presented at the 21st Annual Meeting & Exhibition of the International Society for Magnetic Resonance in Medicine (ISMRM) in 2013 [\[Maumet 2013c\]](#).

**Contents**

<b>1</b>	<b>Introduction</b> . . . . .	<b>125</b>
<b>2</b>	<b>Theory</b> . . . . .	<b>126</b>
	2.0.1    Z-score thresholding . . . . .	127
	2.0.2    M-estimators . . . . .	127
<b>3</b>	<b>Material and Methods</b> . . . . .	<b>128</b>
	3.1    Data . . . . .	128
	3.1.1    Acquisition protocol . . . . .	128
	3.1.2    Pre-processing . . . . .	128
	3.2    Validation . . . . .	129
	3.2.1    Simulated corruption . . . . .	129
	3.2.2    Experiments on real clinical data sets . . . . .	130
<b>4</b>	<b>Results</b> . . . . .	<b>131</b>
	4.1    Validation on simulated data . . . . .	131
	4.1.1    Dataset with 250 repetitions: . . . . .	131
	4.1.2    Simulation based on pathological data: . . . . .	131
	4.2    Validation on real clinical data . . . . .	133
<b>5</b>	<b>Conclusion</b> . . . . .	<b>135</b>

## 1 Introduction

ASL allows a non-invasive quantification of the CBF [Detre 1992]. Due to the low SNR of the ASL sequence, a single pair of control and label image is not sufficient to measure perfusion. The acquisition is therefore repeated several times, leading to  $r$  pairs of images (usually  $r \geq 30$ ). Perfusion information is then usually extracted by pair-wise subtracting the control and label images and averaging across the repetitions.

Though sample average, as an unbiased estimate of mean, ensures convergence as  $r$  grows, it has a breakdown point of 0% (i.e. a single arbitrary large value can induce an arbitrary large estimate) and is thus very sensitive to outliers as illustrated in fig. 7.1. And yet it is well-known that instabilities during the acquisition and

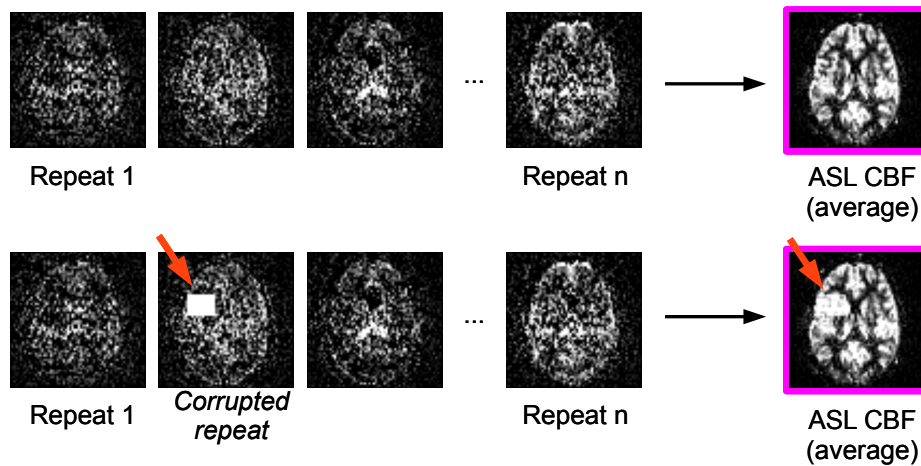


Figure 7.1: Sensitivity of the sample mean to outliers. A single corrupted volume can induce strong artefact in the final CBF maps computed with the sample mean.

improperly corrected patient motion can cause artefactual values [Tan 2009]. In particular, sudden subject motion often induces strong corolla-shaped artefacts.

To avoid the detrimental effects that a few abnormal repetitions could have in the final perfusion map, it is often suggested to ignore the volumes corresponding to the motion peaks using an appropriate threshold [Sidaros 2005]. Volumes with (estimated) motion parameters greater than  $[1-3]^\circ$  or  $[1-3]$  mm are thus sometimes discarded before averaging. However the choice of these thresholds is empirical and there is no common rule across studies or automatic methods to tune these ad-hoc parameters. In [Tan 2009], the authors proposed an automatic algorithm for outlier rejection in ASL perfusion series based on z-score thresholding at the volume (or slice) level as illustrated in fig. 7.2. Their method produced satisfactory results on a qualitative validation based on ratings made by medical experts. However, their approach is based on z-scores, while more robust statistical measures are known to be better suited to deal with outliers. Also, they rely on empirically tuned parameters that might limit the generalisation of their procedure to new datasets.



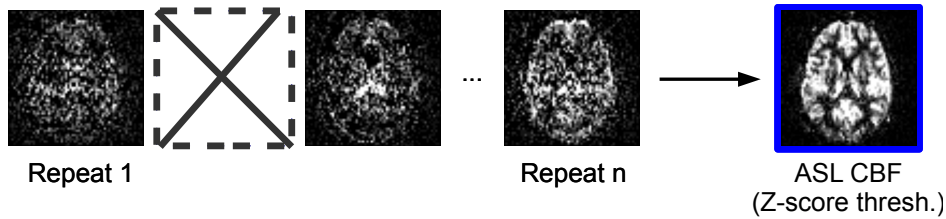


Figure 7.2: Volume exclusion via z-score thresholding as proposed in [Tan 2009].

How to appropriately deal with outliers has been widely studied in the statistical literature and a large range of methods has emerged. Z-score is known to be sensitive upon sample size and is suffering from masking effects when more than one outliers is present in the series [Shiffler 1988]. Indeed, in a dataset containing more than one outlier, the standard deviation estimate will be artificially inflated which may prevent z-score based outlier detection. On the other hand, M-estimators are robust techniques to estimate location and scale in the presence of outliers [Rousseeuw 2003]. We focus on Huber’s M-estimator [Huber 1964], as it is the most widely used. An illustration of this method is provided in fig. 7.3.

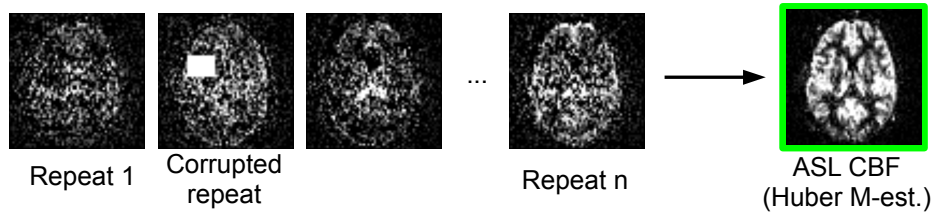


Figure 7.3: Robust ASL CBF map via Huber’s M-estimator.

In this chapter, we propose to estimate robust ASL CBF maps with Huber’s M-estimator. This method is compared to z-thresholding as proposed in [Tan 2009]. Validation is undertaken by measuring the voxel-to-voxel correlation between ASL CBF maps and DSC CBF maps as an affine relationship is expected between these estimates of CBF [Warmuth 2003].

Section 3 presents the statistical methods and the validation procedure. Section 4 presents the results on simulated data and on real datasets from patients diagnosed with brain tumours.

## 2 Theory

Starting from a perfusion-weighted series, namely a 4D volume made of the  $r$  repetitions obtained after pair-wise subtracting the control and label scans, the objective is to compute a single perfusion-weighted volume. This section presents z-score thresholding and M-estimators as statistical method to compute robust CBF maps.

### 2.0.1 Z-score thresholding

In [Tan 2009], an outlier rejection algorithm based on z-scores is proposed in order to remove outliers from the perfusion-weighted series. The outlier rejection is performed both on a volume-by-volume and a slice-by-slice basis. For each volume (respectively slices)  $v$ , the mean  $m_v$  and standard deviation  $s_v$  of in-brain voxel intensities is computed. Assuming a Gaussian distribution of  $m_v$  and  $s_v$ , a volume is then rejected if:

$$|m_v| > \hat{\mu}_m + 2.5 \hat{\sigma}_m, \text{ or } s_v > \hat{\mu}_s + 1.5 \hat{\sigma}_s \quad (7.1)$$

where  $\hat{\mu}_m = \frac{1}{r} \sum_{i=1}^r m_i$   $\hat{\sigma}_m^2 = \frac{1}{r-1} \sum_{i=1}^r (m_i - \hat{\mu}_m)^2$

and  $\hat{\mu}_s = \frac{1}{r} \sum_{i=1}^r s_i$   $\hat{\sigma}_s^2 = \frac{1}{r-1} \sum_{i=1}^r (s_i - \hat{\mu}_s)^2$ .

The parameters 1.5 and 2.5 were determined empirically. To avoid over-filtering, a heuristic is added saying that series verifying  $\ln(\max_i(s_i) - \min_i(s_i)) < 1$  are not searched for outliers. Once the outliers are identified, the perfusion map is then computed by averaging the remaining repetitions.

### 2.0.2 M-estimators

Another solution to deal with outliers is to employ robust statistics such as M-estimators, which will not be overly influenced by outliers. In [Huber 1964], M-estimators are defined, given a function  $\rho$ , as solutions  $\hat{\theta}$  of:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \left( \sum_{i=1}^r \rho(x_i, \theta) \right). \quad (7.2)$$

If  $\rho$  is differentiable, and  $\psi$  is its derivative then eq. (7.2) can be solved by finding the root of:

$$\sum_{i=1}^r \psi(x_i, \theta) = 0. \quad (7.3)$$

The sample average can be seen as an M-estimator with  $\rho(x_i, \theta) = (x_i - \theta)^2$  and  $\psi(x_i, \theta) = 2(x_i - \theta)$  leading to  $\hat{\theta} = \frac{1}{r} \sum_{i=1}^r x_i$ . Fig. 7.4 illustrates the  $\psi(x_i, \theta)$  functions for the three estimators of interest.

The M-estimator of location proposed by Huber in [Huber 1964] is defined by:

$$\psi(x_i, \theta) = \gamma\left(\frac{x_i - \theta}{\sigma}\right) \quad \text{where } \gamma(x) = \begin{cases} -k, & x < -k \\ x, & -k < x < k \\ k, & x > k \end{cases} \quad (7.4)$$

$k$  will be set to 1.345 throughout this chapter corresponding to 95% efficiency in Gaussian data [Krasker 1982]. Likewise,  $\sigma$  is estimated by a robust estimator:

the median absolute deviation divided by 0.6745 [Rousseeuw 2003]. Huber’s M-estimator is applied voxel by voxel on the perfusion-weighted series to obtain the robust perfusion-weighted map.

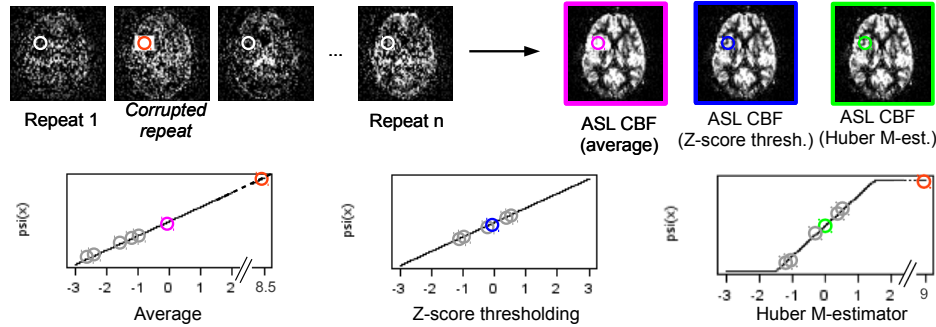


Figure 7.4:  $\Psi$  functions for the three estimators: average, z-score thresholding and Huber’s M estimator.

### 3 Material and Methods

#### 3.1 Data

##### 3.1.1 Acquisition protocol

26 patients diagnosed with brain tumours were involved in this study. Data acquisition was performed on a 3T Siemens Verio MR scanner with a 32-channel head-coil. Patients were scanned in the context of clinical practice. The imaging protocol included a 3D T1w anatomical sequence (TR: 1900 ms, TE: 2.27 ms, FOV: 256 mm x 256 mm x 176 mm, flip angle:  $9^\circ$ , resolution: 1 mm x 1 mm x 1 mm), a PICORE Q2TIPS sequence with crusher gradients (TR: 3000 ms, TE: 18 ms, FOV: 192 mm x 192 mm, flip angle:  $90^\circ$ , in plane resolution: 3 mm x 3 mm, slice thickness: 7 mm, inter-slice gap: 0.7 mm, TI: 1700 ms, bolus width: 700 ms,  $r = 60$ ), a DSC sequence (GRE EPI, TR: 1500 ms, TE: 30 ms, FOV: 230 mm x 230 mm, flip angle:  $90^\circ$ , in plane resolution: 1.8 mm x 1.8 mm, slice thickness: 4 mm, inter-slice gap: 1.2 mm) and 3D T1-weighted post gadolinium sequence (TR: 1900 ms, TE: 2.27 ms, flip angle:  $9^\circ$ , FOV: 250 mm x 250 mm x 176 mm, resolution: 1 mm x 1 mm x 1 mm).

1 healthy subject was involved in this study. The imaging protocol included a 3D T1-weighted anatomical sequence (same parameters as above) and a PICORE Q2TIPS sequence with crusher gradients (TR: 2500 ms, TE: 19 ms, flip angle:  $90^\circ$ , in plane resolution: 3 mm x 3 mm, slice thickness: 7 mm, inter-slice gap: 0.7 mm, TI: 1800 ms, bolus width: 700 ms,  $r = 250$ ).

##### 3.1.2 Pre-processing

Image pre-processing was performed using SPM8 (Wellcome Department of Imaging Neuroscience, University College, London) Matlab toolbox. A six-parameter

rigid-body registration of the ASL volumes was carried out in order to reduce undesired effects due to subject motion. Coregistration on grey matter map was then performed based on normalised mutual information. The average of unlabelled volumes was used to estimate the geometrical transformation to apply to each volume.

The 60 unlabelled and labelled ASL volumes were pair-wise subtracted in order to obtain a perfusion-weighted series per subject. Robust ASL perfusion-weighted map was then carried out as described in section 3. A standard kinetic model [Buxton 1998] was then applied in order to obtain quantitative ASL CBF maps.

The DSC images were processed using MR manufacturer software by manually choosing an arterial input function to calculate CBF and mean transit time maps. Similarly to ASL, DSC CBF maps were coregistered to grey matter maps.

## 3.2 Validation

### 3.2.1 Simulated corruption

In order to assess the efficiency of each technique, we generated simulated data with a known quantity of outliers based on two real datasets. Outliers were drawn from a uniform distribution with extrema (-100;100). These values were determined empirically. Indeed, in an uncorrupted perfusion-weighted map, the values usually range between -10 and +10 and voxel standard deviation can in fact be up to 50. Also, by looking at the values of identified outliers in a real dataset, we found values as big as 300 in absolute value.

As data corruption usually affects multiple voxels per volume [Tan 2009], outlier simulation was undertaken by corrupting from 0% to 50% of the volumes. We will later refer to these volumes as outlier volumes. Then, 2%, 20% or 50% of the voxels in each outlier volume were replaced by random outliers leading to low, medium and high level of volume corruption respectively. A brief description of the corruption simulation is provided in fig. 7.5. Each simulation was repeated 30 times in order to get estimates of the standard deviation.

Simulated data were based on two real datasets as described hereafter. The first dataset was a perfusion-weighted series with a large number of repetitions,  $r = 250$ , from a healthy subject. The perfusion-weighted map obtained by averaging the 250 repetitions was considered as the ground truth. The 60 first volumes of the series were extracted and used as dataset for robust CBF map estimation. The quality of the maps produced by each method was measured in terms of sum of squared differences with the estimated ground truth.

The second dataset was built on the perfusion-weighted map of one patient diagnosed with a brain tumour. The original ASL CBF map of this subject presented few artefacts identified by visual inspection and a very low level of motion ( $< 0.5$  mm and  $< 0.2^\circ$  in all directions). As DSC is currently the reference method to estimate perfusion with MRI, the quality of the maps produced by each method was measured by computing the Pearson linear correlation coefficient with the DSC CBF map. This assumes an affine relationship between CBF maps produced by

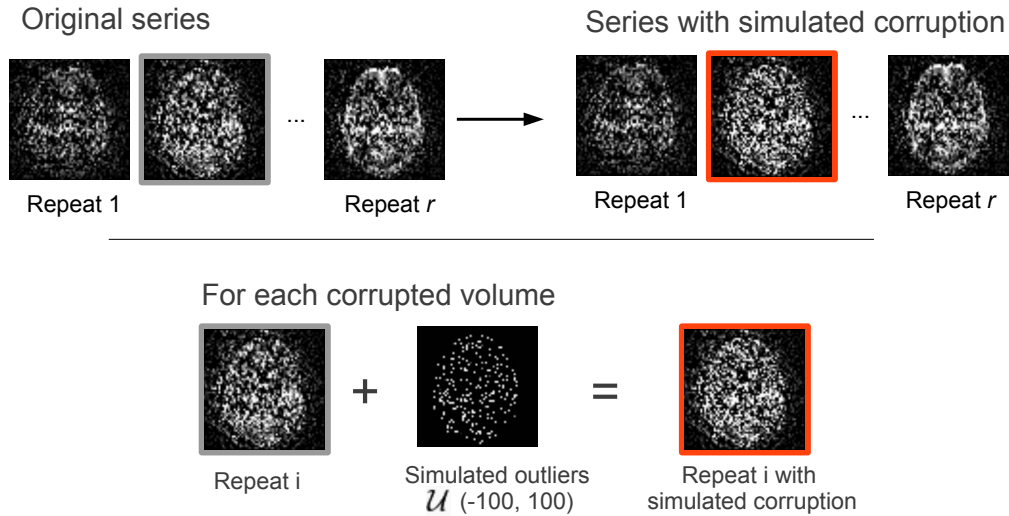


Figure 7.5: Simulation: corruption of a given percentage of ASL volumes (from 0% to 50%) by adding samples drawn from a uniform distribution to a given percentage of the voxels (2%, 20% or 50%) per volume.

ASL and DSC [Warmuth 2003].

### 3.2.2 Experiments on real clinical data sets

We compared the performances of the 3 methods on real clinical datasets. The quality of the estimated ASL CBF map was assessed by computing a linear correlation index with a CBF map extracted from the most validated MR technique for perfusion: DSC.

**Full dataset comparison** The efficiency of both algorithms was estimated on a dataset of 14 perfusion-weighted maps of patients diagnosed with brain tumours. The quality of the ASL CBF map was assessed by voxel-to-voxel correlation with the DSC CBF map.

**Comparison based on subsets of the available data** To investigate the behaviour of the different CBF estimators on datasets of various sizes, we further decimated the data (to get from 5 to 50 volumes) by randomly picking a subset of the available volumes. The experiment was repeated 20 times.

## 4 Results

### 4.1 Validation on simulated data

#### 4.1.1 Dataset with 250 repetitions:

Fig. 7.6 presents the simulation study based on a healthy subject data. The performances of sample average, z-score thresholding [Tan 2009], and Huber's M-estimator are assessed by measuring the sum of squared differences of the ASL CBF map with the ground truth estimated by averaging a large number of repetitions.

As described in fig. 7.6, with a medium or a high level of corruption, z-score thresholding and Huber's M-estimator perform equally and better than averaging until 20% of volumes are corrupted. If more than 20% of the volumes are affected by outliers, then M-estimators provide better estimates than both z-score thresholding and averaging. The robust M-estimator CBF map is closer to the ground truth and less sensitive to an increase in the number of outliers. The same behaviour is observed with a low number of corrupted voxels per volume except that the separation point is at 5% of corrupted volumes instead of 20%. The lower performances of z-thresholding when the number of corrupted volumes exceeds 20% (or 5% with low corruption) is a consequence of the masking effect which penalize this estimator when several outliers are present in the series. Moreover, the performance of Huber's M-estimator always depicts a smaller variance than z-thresholding.

Both Huber's M-estimator and z-score thresholding provide better estimates than the sample average. As the level of corruption per volume decreases, the separation point between Huber's M-estimator and z-score thresholding tends to become lower. This can probably be explained by the fact that the method proposed in [Tan 2009] is based on a global mean and standard deviation estimate per volume (or slice) and is therefore less suited to detect sparsely corrupted volumes.

#### 4.1.2 Simulation based on pathological data:

Fig. 7.7 presents the simulation study based on pathological data of a subject suffering from a brain tumour. The performances of sample average, z-score thresholding [Tan 2009], and Huber's M-estimator are assessed by measuring the correlation coefficient of ASL CBF with DSC CBF.

The simulation involving a high level of volume corruption leads to very similar results than the one obtained in the previous section on healthy subject data. Both Huber's M-estimator and z-thresholding perform better than averaging until 20% of corrupted volumes. After this threshold, z-thresholding performances drop until reaching the same correlation as the sample average for 30% of outlier volumes. This result suggests that correlation with DSC is a good measure of ASL CBF map quality. For medium level of volume corruption, the same tendency is observable.

With a low level of volume corruption, the trend is less clear. Overall the correlation coefficient seems much less affected by the increasing number of outliers. Z-score thresholding and Huber's M-estimator are both better estimator of the mean

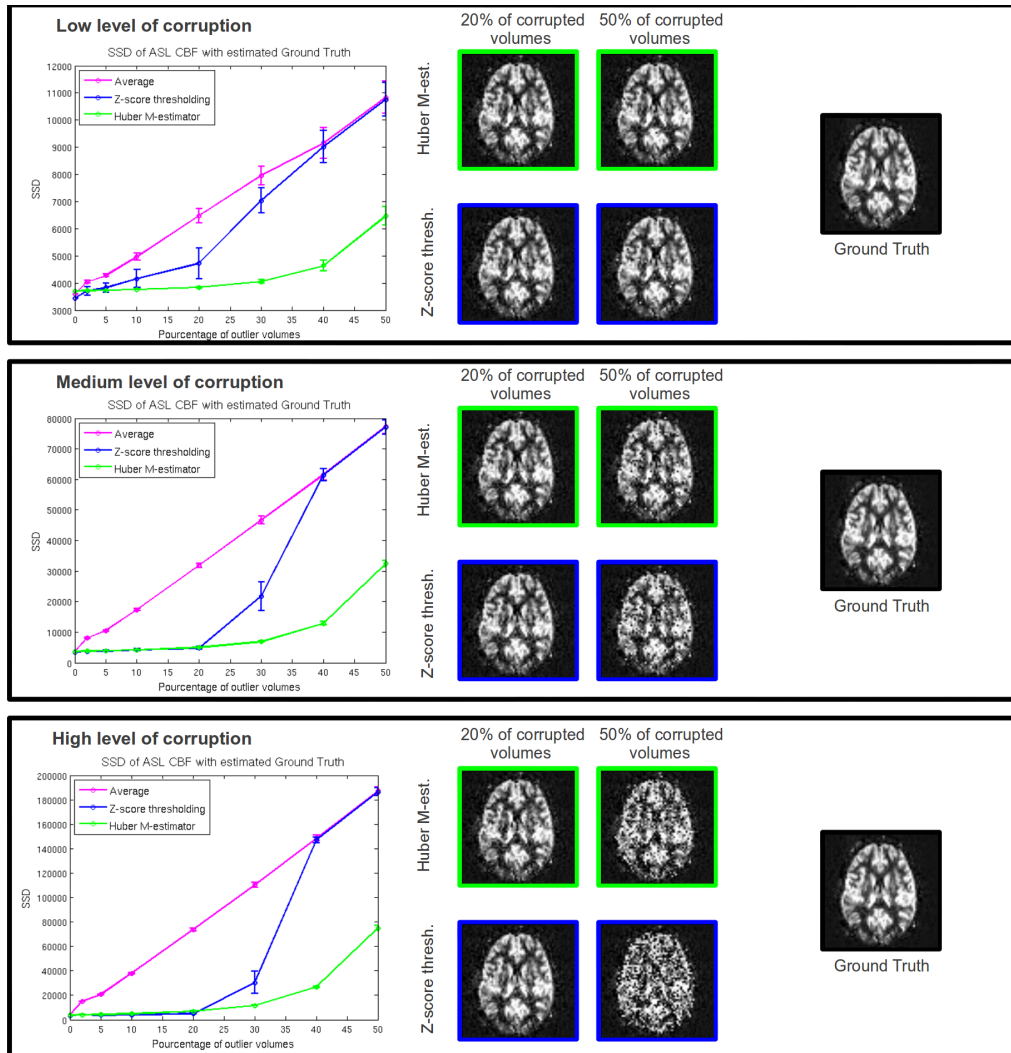


Figure 7.6: Healthy subject dataset with simulated outliers: sum of squared differences (SSD) of ASL CBF map, computed by M-estimator, z-score thresholding [Tan 2009] and sample average, with the estimated ground truth. Low, medium and high level of volume corruption, from 0% to 50% of corrupted volumes. In all configuration Huber’s M-estimators is either better or as good as z-thresholding to estimate robust CBF maps. In the presence of outliers, Huber’s M-estimator is always more accurate than the sample average.

than the sample average. Z-score thresholding however displays a higher variance in its performance estimates. In comparison with the previous simulation study, there is probably a higher level of noise in the so-called “uncorrupted” pathological data than in the “uncorrupted” healthy subject data. The inherent higher level of noise in pathological data might prevent the correct detection of low level of volume corruption.

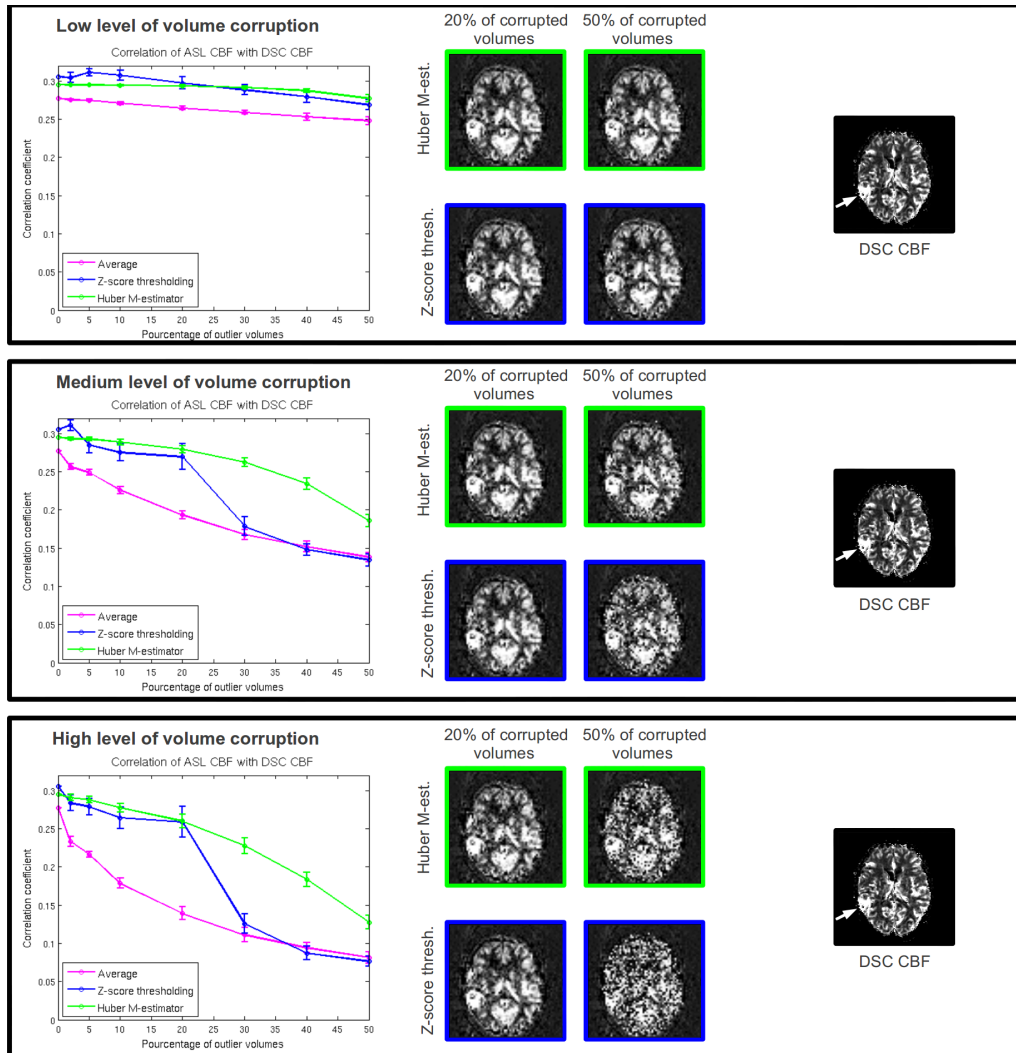


Figure 7.7: Patient dataset with simulated outliers: correlation of ASL CBF map, computed by M-estimator, z-score thresholding [Tan 2009] and sample average, with the DSC CBF map. Low, medium and high level of volume corruption, from 0% to 50% of corrupted volumes. The white arrow points the tumour site. A similar behaviour as for healthy subject simulation (fig. 7.6) is observed outlining that correlation with DSC CBF is a valuable indicator to measure the quality of the ASL CBF estimates.

## 4.2 Validation on real clinical data

**Full dataset** A subset of 14 patients was studied, as the full dataset (26 patients) was not available at the moment of the analysis. Table 7.1 presents the correlation coefficient obtained for the 14 patients diagnosed with brain tumours. Overall, there is a significant improvement of both Huber's M-estimator ( $p=0.007$ ) and z-score thresholding ( $p=0.010$ ) over the sample average (paired two sample t-test). In



this dataset, there was no significant difference between the two filtering methods (paired t-test  $p=0.84$ ).

Table 7.1: Real clinical dataset: correlation coefficient with DSC CBF map of ASL CBF map computed by M-estimator, z-score thresholding [Tan 2009] and sample average in 14 patients diagnosed with brain tumours. Last column: mean and standard deviation across subjects.

Patients	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Mean $\pm$ std.
Huber M-est.	.45	.32	.29	.51	.52	.34	.28	.12	.14	.27	.35	.16	.17	.17	.29 $\pm$ .13
z-score thresh.	.45	.24	.27	.53	.51	.35	.28	.15	.14	.30	.35	.18	.16	.20	.29 $\pm$ .13
Average	.46	.25	.20	.42	.52	.31	.25	.12	.14	.25	.32	.13	.17	.12	.26 $\pm$ .13

Fig. 7.8 presents an example of robust ASL CBF maps in which motion artefacts are significantly reduced by both Huber's M-estimator and z-thresholding.

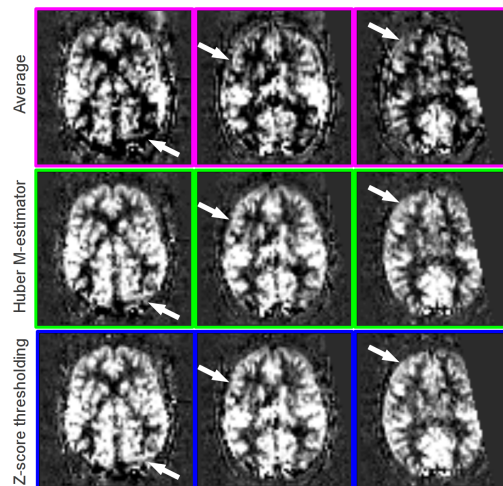


Figure 7.8: Example of robust CBF map in one patient: three contiguous axial slices are depicted. White arrows outlines large artefacts presents in the averaged perfusion-weighted map and correctly corrected by both z-score thresholding and M-estimator.

**Comparison based on subsets of the data** As displayed in fig. 7.9, both z-thresholding and Huber's M-estimator outperform the sample mean. Huber's M-estimator appears marginally better than z-thresholding. Fig. 7.10 presents an example in which we can clearly see the improved quality of both robust ASL CBF maps.

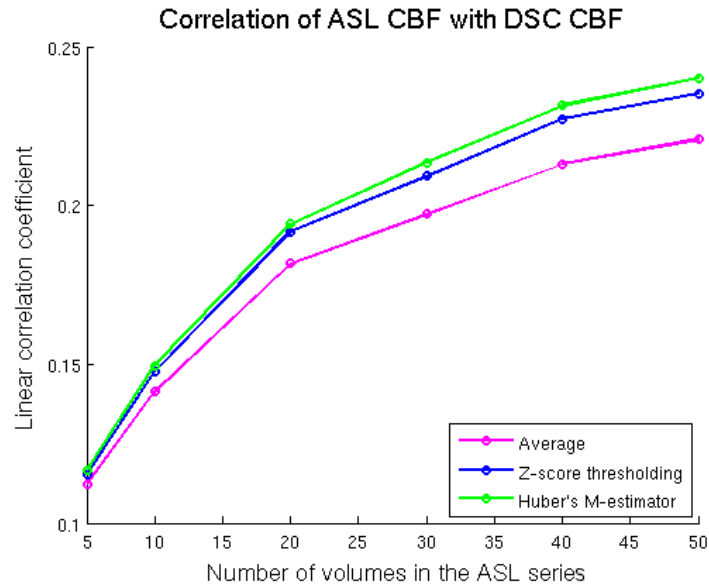


Figure 7.9: Real clinical dataset: correlation of ASL CBF map with DSC CBF for different number of repeats  $r$ .

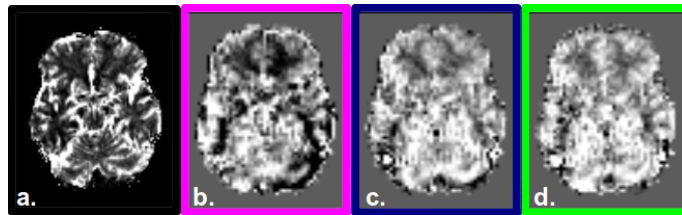


Figure 7.10: DSC CBF estimate (a) and ASL CBF estimates based on 40 repetitions: average (b), z-thresholding (c), Huber's M-estimator (d).

## 5 Conclusion

We studied the ability of Huber's M-estimator to compute robust CBF maps in ASL. The behaviour of this estimator was studied in both simulated and real clinical datasets and compared to an outlier removal technique based on z-thresholding previously introduced in the ASL literature [Tan 2009].

Out of this study, it is confirmed that outlier filtering, either via outlier removal or M-estimation, provides more robust CBF maps than the sample average. The simulation study clearly confirmed the superior robustness of M-estimators over z-score thresholding in the presence of outliers. On the tested clinical cases, the superiority was less pronounced but still present. Overall, both robust methods outperform the sample mean on simulated and real data and Huber's M-estimates are either as good as or better than z-thresholding and are always less variable.

As M-estimators are able to deal with a broader range of outliers, we recommend

the use of M-estimators as robust method to compute ASL CBF maps. This study focused on patients diagnosed with brain tumours, as DSC sequence is part of their routine clinical protocol. Other pathologies might be related with different outlier patterns and a larger validation study on real datasets is therefore needed in order to outline the cases in which M-estimator will have a significantly better behaviour than z-thresholding.

## Chapter 8

# Modelling heterogeneous variances in the detection of patient-specific perfusion abnormalities with ASL

---

This chapter is presented in the form of a paper. A shorter version of this work is currently in press at NeuroImage [Maumet 2013d].

### Abstract

In this paper, patient-specific perfusion abnormalities in ASL were identified by comparing a single patient to a group of healthy controls using a mixed-effect hierarchical GLM. Two approaches are currently in use to solve hierarchical GLMs: (1) the homoscedastic approach assumes homogeneous variances across subjects and (2) the heteroscedastic approach is theoretically more efficient in the presence of heterogeneous variances but algorithmically more demanding. In practice, in functional magnetic resonance imaging studies, the superiority of the heteroscedastic approach is still under debate. Due to the low signal-to-noise ratio of ASL sequences, within-subject variances have a significant impact on the estimated perfusion maps and the heteroscedastic model might be better suited in this context.

In this paper we studied how the homoscedastic and heteroscedastic approaches behave in terms of specificity and sensitivity in the detection of patient-specific ASL perfusion abnormalities. Validation was undertaken on a dataset of 25 patients diagnosed with brain tumours and 36 healthy volunteers. We showed evidence of heterogeneous within-subject variances in ASL and pointed out an increased false positive rate of the homoscedastic model. In the detection of patient-specific brain perfusion abnormalities with ASL, modelling heterogeneous variances increases the sensitivity at the same specificity level.

## Contents

<b>1</b>	<b>Introduction</b>	<b>139</b>
<b>2</b>	<b>Methods</b>	<b>141</b>
2.1	Detection of patient-specific perfusion abnormalities using a mixed-effect hierarchical two-sample t-test	141
2.1.1	Subject level (First level)	142
2.1.2	Group level (Second level)	143
2.1.3	Hypothesis testing	146
2.2	Evaluation framework	146
2.2.1	Ground truth	146
2.2.2	Validity of the assumptions of the homoscedastic model	147
2.2.3	Comparison of the homoscedastic and heteroscedastic models	148
<b>3</b>	<b>Materials</b>	<b>149</b>
3.1	Data acquisition	149
3.2	Pre-processing	149
3.2.1	CBF estimation with ASL	149
3.2.2	Intensity normalization	150
3.2.3	CBF estimation with DSC	151
<b>4</b>	<b>Results</b>	<b>152</b>
4.1	ASL template: a model of normal perfusion	152
4.2	Testing the assumptions of the homoscedastic model	152
4.2.1	Relative weights of between-subject and within-subject variances in the control group	153
4.2.2	Cross-subject comparison of within-subject variances	153
4.3	Comparison of homoscedastic and heteroscedastic models	155
4.3.1	Fixed threshold	155
4.3.2	Receiver-Operating-Characteristics (ROC) analysis	157
<b>5</b>	<b>Discussion and conclusion</b>	<b>158</b>

## 1 Introduction

Brain perfusion is the biological process that ensures the delivery of oxygen and nutrients to the cerebral tissues by means of micro-circulation. As an indicator of the well-being of the tissues, perfusion is a useful measurement for diagnosis in clinical practice. Identifying regions of abnormal perfusion, either hypo-perfusions or hyper-perfusions, can help understanding the mechanism of a disease and impact patient care. For instance, for patients diagnosed with tumours, the clinician is interested in hyper-perfusions that would reveal the grade of the tumour [Weber 2006], or help differentiating between post-radiation necrosis and tumour recurrence [Sugahara 2000]. In strokes, the extent of hypo-perfusion and mismatch with diffusion imaging provide an insight on the possible recovery of the tissue [Huck 2012].

ASL, a MRI technique introduced in the early 1990's, allows to measure the level of perfusion through a quantitative index: the cerebral blood flow (CBF). Contrary to standard perfusion imaging, including PET and SPECT in nuclear medicine, or DSC in MRI, ASL is completely non-invasive and does not require the injection of an exogenous contrast agent. In ASL, blood water, used as an endogenous tracer, is labelled with a radio-frequency pulse. After a delay called *inversion time*, a labelled image of the brain is acquired. A control image is also acquired without prior labelling. The difference between the labelled and the control image leads to a perfusion-weighted map. To increase the SNR of the measurement, the acquisition is repeated several times. The absence of radiation and the use of an endogenous tracer, which removes the need of injecting a contrast agent, are clear advantages of ASL over standard perfusion imaging techniques. The ASL sequence however suffers from a low SNR, which is still a serious obstacle for its use in clinical practice.

Since the introduction of ASL, and despite its low SNR, a large number of studies have demonstrated its usefulness in identifying patterns of abnormal perfusion at the group level (e.g. [Pinkham 2011]). To date, identification of individual patterns of hypo- and hyper-perfusions in patients with ASL mainly relies on visual analysis [Zaharchuk 2012, Chen 2012b]. Very few ASL studies focused on voxelwise quantitative perfusion abnormality detections at the individual level. In [Petr 2013], a template-based analysis was presented in order to detect individual activation patterns in functional ASL data. Interestingly, they also applied their method to a pathological case in order to detect a hypo-perfusion co-located with a dysplasia in an epileptic patient. While their approach opens the field to the detection of patient-specific perfusion abnormalities with ASL, a single pathological case was presented and no quantitative validation performed. Their detections relied on z-scores, but in a more general setting, the most widespread approach to compare voxel-wise maps in neuroimaging is the massively univariate GLM. To detect differential patterns between two groups with repeated measurements a subtype of GLM is employed: a mixed-effect hierarchical two-sample t-test with two levels: subject and group. In this context, two variance components are of interest: the within-subject variance (or the measurement error, estimated from the repeated ASL acquisitions of a single subject) and the between-subject variance.

In the functional MRI community, where similar statistical models are applied, two approaches are currently in use to solve hierarchical GLMs. On the one hand, the homoscedastic approach, also termed “summary statistics” [Penny 2003], or referred as “ordinary least square estimation” [Mumford 2006, Mumford 2009] or “conventional group analysis” [Chen 2012a], assumes homogeneous within-subject variances or negligible within-subject variances by comparison to between-subject variance. On the other hand, the heteroscedastic approach, also referred as “full mixed-effect” [Friston 2005b, Thirion 2007, Poldrack 2011], “mixed-effect model” [Chen 2012a] or “weighted least square estimation” [Mumford 2006, Mumford 2009] models heterogeneous within-subject variances. There is indeed a large panel of homonyms to refer to these two approaches, in current practice they are also sometimes referred as “mixed-effect” and “random-effect” approaches [Lindquist 2012]. However, in a statistical sense both the homoscedastic and the heteroscedastic approaches are mixed-effect models. That is why, in the context of this paper, we chose to term the two approaches “homoscedastic model” and “heteroscedastic model” to identify them according to their main difference: the homoscedasticity (constant within-subject variance across subjects) assumption.

The homoscedastic model is theoretically less efficient in the presence of heterogeneous within-subject variances. However, in practice, the true variance components are unknown and the superiority of the heteroscedastic model is therefore questioned. Whether modelling heterogeneous variances should be preferred over the homoscedastic model is still under discussion in the fMRI community. Several authors outlined the benefits of heteroscedastic models including [Worsley 2002, Beckmann 2003, Woolrich 2004, Mumford 2006, Mériaux 2006, Thirion 2007, Chen 2012a]. However, [Mumford 2009] showed that the homoscedastic approach is still valid with near optimal sensitivity in the context of one-sample t-tests. Nonetheless, the same authors acknowledged that, in two-sample t-tests, appropriate modelling of heterogeneous within-subject variances might be crucial [Mumford 2009, Poldrack 2011]. Recently the conclusions regarding one-sample t-tests were revisited leading to opposite conclusions [Chen 2012a]. In the different software packages currently available to deal with fMRI data both approaches are represented: SPM<sup>1</sup> favours the homoscedastic approach [Holmes 1998] while FSL<sup>2</sup> [Woolrich 2004], AFNI<sup>3</sup> [Cox 1996] and fmristat<sup>4</sup> [Worsley 2002] use the heteroscedastic model.

In ASL, the importance of the within-subject variance has been outlined by [Viviani 2009]. While we previously investigated the ability to detect hypo- and hyper-perfused regions using a template of normal perfusion in two conference papers [Maumet 2012a, Maumet 2012b], here, we focus on the comparison of the homoscedastic and heteroscedastic approaches. We further improve on these papers by providing quantitative measurements of heterogeneity. Also, a larger database

<sup>1</sup><http://www.fil.ion.ucl.ac.uk/spm>

<sup>2</sup><http://fsl.fmrib.ox.ac.uk/fsl/fslwiki>

<sup>3</sup><http://afni.nimh.nih.gov/afni>

<sup>4</sup><http://www.math.mcgill.ca/keith/fmristat>

is studied and a novel quantitative validation is performed.

In this paper, we focus on quantitative detections of pathological brain perfusion abnormalities at the individual level. To this aim, we employ and compare two GLM-based models: the homoscedastic and heteroscedastic approaches. We test whether the assumptions underlying the homoscedastic approach are verified in pulsed ASL datasets. We furthermore study how the homoscedastic and heteroscedastic approaches behave in terms of specificity and sensitivity in the detection of patient-specific perfusion abnormalities.

A quantitative validation is performed on a dataset of 25 patients diagnosed with brain tumours. Though there is no well-defined ground truth, this pathology was selected since patterns of hypo- and hyper-perfusions have been widely studied in this clinical context. The model of normal perfusion is computed out of the data of 36 healthy volunteers.

Section 2 begins with a presentation of the homoscedastic and heteroscedastic models employed in the detection of patient-specific brain perfusion abnormalities with ASL. Then, the experiments designed to test the assumptions of the homoscedastic model and measure sensitivity and specificity are presented. Section 3 presents the datasets under study: the acquisition and pre-processing are described. The results are presented in section 4. Section 5 gives a discussion and concludes.

## 2 Methods

In 2.1, the homoscedastic and heteroscedastic models and their implementation in the context of patient-specific detection of perfusion abnormalities in ASL are presented. Then, in 2.2, the experiments undertaken in order to test the assumptions of the homoscedastic model and to compare the homoscedastic and heteroscedastic models are described.

### 2.1 Detection of patient-specific perfusion abnormalities using a mixed-effect hierarchical two-sample t-test

This section presents the common massively univariate GLM usually employed for voxel-based analysis in the neuroimaging community. The main assumption behind the GLM is the linearity of the effects. Moreover, Gaussian noise is usually assumed. In the standard approach, the GLM is defined for each voxel, which is why this approach is often termed “massively univariate”. For ease of notation the voxel index is omitted in the remainder of the paper.

In the context of this paper, we aim at outlining areas of abnormal perfusion in the perfusion map of a subject of interest by comparison to a group of healthy controls. We hence focus on a subtype of GLM: a *one-versus-many* mixed-effect two-sample t-test. As several measurements are available for each subject, a hierarchical model is defined with 2 levels: subject and group.

Since the subject under study is usually an individual suffering of a pathological condition, the term *patient* will be used in the following. This approach is never-



theless suited for the comparison of any single subject (control, patient) to a group of individuals.

A description of the subject level is proposed in 2.1.1. In 2.1.2 the group level is presented along with a description of the two main related approaches: the homoscedastic and heteroscedastic models. Lastly, hypothesis testing performed in order to outline the perfusion abnormalities is discussed in 2.1.3.

### 2.1.1 Subject level (First level)

Given a voxel, for each subject  $s$  we have:

$$Y_s = X_s \beta_s + \varepsilon_s, \quad (8.1)$$

where  $Y_s$  is a vector containing the observations at the given voxel,  $\beta_s$  is the set of subject parameters to be estimated,  $X_s$  is the subject-level design matrix and  $\varepsilon_s$  contains the residual errors. In the following, we assume that the data can be described with a single parameter per subject,  $\beta_s$ , however more parameters (e.g. nuisance covariates as in [Wang 2012]) could be considered. In fMRI,  $X_s$  is closely related to the time-course of the paradigm under study. In ASL, at the subject level, the data under study is a 4D volume of CBF maps containing  $V$  volumes (1 per repeated acquisition). The observations  $Y_s$  are therefore repeated measurements of a same underlying value and the  $X_s$  matrix is a vector of ones. Then, the model reduces to:

$$Y_s = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \beta_s + \varepsilon_s.$$

Assuming Gaussian noise of the errors,  $\varepsilon_s$  follows a normal distribution:  $\varepsilon_s \sim \mathcal{N}(0, \sigma_s^2)$ . While, in fMRI, temporal autocorrelation must be accounted for [Aguirre 1997], in ASL, thanks to the subtraction process between control and labelled scans, we can reasonably assume white noise [Aguirre 2002, Mumford 2006]. The subject parameter  $\beta_s$  is thus estimated by OLS:

$$\hat{\beta}_s = \frac{1}{V} \sum_{i=1}^V y_{s,i}, \quad (8.2)$$

where  $y_{s,i}$  is the  $i^{\text{th}}$  element of vector  $Y_s$ . Similarly, the sampling variance of  $\hat{\beta}_s$  is estimated by:

$$\widehat{\text{Var}}(\hat{\beta}_s) = \frac{\hat{\sigma}_s^2}{V} \quad \text{where} \quad \hat{\sigma}_s^2 = \frac{1}{V-1} \sum_{i=1}^V (y_{s,i} - \hat{\beta}_s)^2. \quad (8.3)$$

### 2.1.2 Group level (Second level)

In a general setting, in a group of  $n$  subjects, the subject parameters  $(\beta_s)_{1 \leq s \leq n}$  can be combined using the following model:

$$\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = X_G \beta_G + \gamma_G, \quad (8.4)$$

where  $X_G$  is the group-level design matrix,  $\beta_G$  the group parameters and  $\gamma_G$  the residual error term (outlining the intra-group variability). In a one-sample t-test,  $X_G$  would be a vector of ones and  $\beta_G$  contain a single estimate. In our context, the subjects are part of two groups ( $n - 1$  controls in the first group and 1 patient in the second) and  $\beta_G$  is a vector with 2 elements. Let  $\beta_{controls}$  and  $\beta_{patient}$  be the control group and patient parameters. Without loss of generality, we assume that subjects 1 to  $n - 1$  are part of the control group and subject  $n$  is the patient of interest. Then the second-level model reduces to:

$$\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{n-1} \\ \beta_n \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_{controls} \\ \beta_{patient} \end{bmatrix} + \gamma_G. \quad (8.5)$$

Assuming Gaussian errors,  $\gamma_G$  follows a normal distribution  $\gamma_G \sim \mathcal{N}(0, \sigma_G^2)$ , where  $\sigma_G^2$  is the common between-subject (also termed ‘‘within-group’’ or even ‘‘group’’) variance of both groups. Indeed, since in a *one-versus-many* study, a single subject is available in the patient group, estimating a different variance for each group (control, patient) is impractical.

However, the true subject parameters are in fact unknown and, in practice, their estimates from the first level are used in the second-level leading to a slightly different model:

$$\begin{bmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_{n-1} \\ \hat{\beta}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_{controls} \\ \beta_{patient} \end{bmatrix} + \gamma_{G_C}. \quad (8.6)$$

The new error term  $\gamma_{G_C}$  is impacted by two combined sources of variations: the error measurement on the subject parameters (also termed within-subject variance) and the between-subject variance. Each element of  $\gamma_{G_C}$  therefore follows a normal distribution:  $\gamma_{G_C}^s \sim \mathcal{N}\left(0, \sigma_G^2 + \frac{\sigma_s^2}{V}\right)$ .

Linear combinations of the group parameters can be calculated using a particular contrast. Here we are interested in the patient versus control group contrast:

$$b = [1 \ -1] \begin{bmatrix} \beta_{controls} \\ \beta_{patient} \end{bmatrix} = \beta_{controls} - \beta_{patient}. \quad (8.7)$$

Two approaches have been proposed in the neuroimaging literature to solve the system (8.6) and find an estimate of the patient versus control group contrast,  $\hat{b}$ , and its associated sampling variance,  $\widehat{\text{Var}}(\hat{b})$ . The homoscedastic approach assumes homoscedasticity while the heteroscedastic approach explicitly models heterogeneous within-subject variances. Both models are described hereafter.

**Homoscedastic model** The homoscedastic model is based on the assumption that the within-subject variance is either negligible by comparison to the between-subject variance (i.e.  $\sigma_s^2 \ll \sigma_G^2$ ,  $\forall 1 \leq s \leq n$ ) or, roughly constant across subjects (i.e.  $\sigma_s^2 \approx \sigma_{SUB}^2$ ,  $\forall 1 \leq s \leq n$ ). Then each element of  $\gamma_{G_C}$  follows a normal distribution:

$$\gamma_{G_C}^s \sim \mathcal{N}(0, \sigma_{G_C}^2) \quad (8.8)$$

where  $\sigma_{G_C}^2$  is the combined within- and between-subject variance. Depending on the assumption, we have  $\sigma_{G_C}^2 = \sigma_G^2 + \frac{\sigma_{SUB}^2}{V}$  or  $\sigma_{G_C}^2 \approx \sigma_G^2$  and the combined within- and between-subject variance is therefore constant across subjects. The sphericity assumption (no heteroscedasticity and no autocorrelation) hence holds and the system (8.6) is then solved by OLS to get:

$$\begin{aligned} \hat{\beta}_{controls}^{HOMO} &= \frac{1}{n-1} \sum_{s=1}^{n-1} \hat{\beta}_s \\ \hat{\beta}_{patient}^{HOMO} &= \hat{\beta}_n, \end{aligned} \quad (8.9)$$

The associated sampling variances are:

$$\begin{aligned} \widehat{\text{Var}}(\hat{\beta}_{controls}^{HOMO}) &= \frac{\hat{\sigma}_{G_C}^2}{n-1} \\ \widehat{\text{Var}}(\hat{\beta}_{patient}^{HOMO}) &= \hat{\sigma}_{G_C}^2 \\ \text{where } \hat{\sigma}_{G_C}^2 &= \frac{1}{(n-1)-1} \sum_{s=1}^{n-1} (\hat{\beta}_s - \hat{\beta}_{controls}^{HOMO})^2. \end{aligned} \quad (8.10)$$

As a consequence, the patient versus control group contrast  $b$  is estimated by:

$$\hat{b}^{HOMO} = \hat{\beta}_{controls}^{HOMO} - \hat{\beta}_{patient}^{HOMO} = \frac{1}{n-1} \sum_{s=1}^{n-1} \hat{\beta}_s - \hat{\beta}_n, \quad (8.11)$$

and the sampling variance of this estimator is:

$$\widehat{\text{Var}}(\hat{b}^{HOMO}) = \hat{\sigma}_{G_C}^2 \left( \frac{1}{n-1} + 1 \right). \quad (8.12)$$

**Heteroscedastic model** In the heteroscedastic model, heterogeneous variances are accounted for. As described earlier, in the general case, each element of  $\gamma_{G_C}$  follows a normal distribution:

$$\gamma_{G_C}^s \sim \mathcal{N}\left(0, \sigma_G^2 + \frac{\sigma_s^2}{V}\right). \quad (8.13)$$

Due to the non-sphericity of the measurement errors (as a consequence of heteroscedasticity), the system (8.6) is solved by using a weighted least square to get:

$$\begin{aligned}\hat{\beta}_{controls}^{HETERO} &= \frac{1}{\sum_{j=1}^{n-1} w_j} \sum_{s=1}^{n-1} w_s \hat{\beta}_s \\ \text{where } w_s &= \frac{1}{\hat{\sigma}_G^2 + \hat{\sigma}_s^2} \\ \hat{\beta}_{patient}^{HETERO} &= \hat{\beta}_n.\end{aligned}\tag{8.14}$$

The associated sampling variances are:

$$\begin{aligned}\widehat{\text{Var}}(\hat{\beta}_{controls}^{HETERO}) &= \frac{1}{\sum_{s=1}^{n-1} w_s} \\ \widehat{\text{Var}}(\hat{\beta}_{patient}^{HETERO}) &= \hat{\sigma}_G^2 + \hat{\sigma}_n^2.\end{aligned}\tag{8.15}$$

The within-subject variance estimates,  $\hat{\sigma}_s^2$ , are computed at the subject level as described in eq. (8.3). In this paper, we compute  $\hat{\sigma}_G^2$  using a recent computationally efficient approach available in the MEMA function of the AFNI software package and described in [Chen 2012a].

The patient versus control group contrast  $b$  is estimated by:

$$\begin{aligned}\hat{b}^{HETERO} &= \hat{\beta}_{controls}^{HETERO} - \hat{\beta}_{patient}^{HETERO} \\ &= \frac{1}{\sum_{j=1}^{n-1} w_j} \sum_{s=1}^{n-1} w_s \hat{\beta}_s - \hat{\beta}_n\end{aligned}\tag{8.16}$$

and the sampling variance of this estimator is:

$$\widehat{\text{Var}}(\hat{b}^{HETERO}) = \frac{1}{\sum_{s=1}^{n-1} w_s} + \hat{\sigma}_G^2 + \hat{\sigma}_n^2.\tag{8.17}$$

If homoscedasticity is respected then it can easily be proven that this model reduces to the homoscedastic model. However, in the presence of heteroscedasticity, the heteroscedastic model has two main advantages by comparison to the homoscedastic model:

1. In the control group, observations with high within-subject variances are down-weighted in order to provide a more efficient estimate of the control group parameter,  $\hat{\beta}_{controls}^{HETERO}$ .
2. While with the homoscedastic model the sampling variance of the patient group estimate,  $\widehat{\text{Var}}(\hat{\beta}_{patient}^{HOMO})$ , depends solely on variance estimations performed in the control group, the heteroscedastic sampling variance,  $\widehat{\text{Var}}(\hat{\beta}_{patient}^{HETERO})$ , takes advantage of both the control group (estimation of  $\sigma_G^2$ ) and the patient (estimation of  $\sigma_n^2$ ) data.

The impact on the efficiency of the control group estimate (point 1) might be subtle and leads to no substantial improvement in one-sample t-tests if the homoscedasticity assumption is not overly altered [Mumford 2009]. However, the

impact on the variance estimate of the patient parameter (point 2) can be large if the within-subject variance of the patient of interest is very different from the control subjects. In ASL studies, the large influence of within-subject variance has been described in [Viviani 2009]. Furthermore, patients are known to be less cooperative than control subjects which could potentially lead to within-subject variance inflation due to increased movement. We therefore expect that adjusting for the proper patient within-subject variance, using the heteroscedastic instead of the homoscedastic model, will lead to more accurate results in the detection of patient-specific perfusion abnormalities.

### 2.1.3 Hypothesis testing

Under the null hypothesis:

$$H_0 : \beta_{controls} = \beta_{patient}, \quad (8.18)$$

the estimated patient versus control group contrast,  $\hat{b}$ , divided by its estimated sampling standard deviation  $(\widehat{\text{Var}}(\hat{b}))^{\frac{1}{2}}$  follows a t-distribution with  $n - 1$  degrees of freedom.

A probability under the null hypothesis can therefore be calculated for each voxel with:

$$P(X < x) \text{ where } X = \frac{\hat{b}}{\sqrt{\widehat{\text{Var}}(\hat{b})}} \sim \mathcal{T}_{n-1}, \quad (8.19)$$

and  $x$  is the value taken by  $X$  at the voxel of interest. Equation (8.19) gives the probability to have a hyper-perfusion, similarly, hypo-perfusions can be detected by substituting  $P(X < x)$  by  $P(X > x)$ . The thresholding of this probability map gives the detections.

## 2.2 Evaluation framework

In this subsection, we first present how the ground truth hypo- and hyper-perfusions were determined based on anatomical and perfusion information (2.2.1). Then, the experiments undertaken to verify the assumptions of the homoscedastic model (2.2.2) and to compare the homoscedastic and heteroscedastic models (2.2.3) are presented.

### 2.2.1 Ground truth

Quantitative assessment of the detections and comparison between detection methods are challenging tasks. This is mainly because, like in many other medical imaging problems, the ground truth is not clearly stated. We chose to evaluate this framework on patients diagnosed with tumour pathology because perfusion abnormalities have been widely studied in this clinical context [Chalela 2000, Sugahara 2000, Warmuth 2003, Weber 2006, Chawla 2007]. Also, as perfusion is a useful clinical information, DSC is part of the clinical routine for these patients. Based

on clinical knowledge, we used a semi-automatic procedure that took advantage of the complementary anatomical (T1w-Gd, T2w FLAIR) and perfusion (DSC) information to get an estimation of the ground truth.

**True positives** According to clinical knowledge, T1w-Gd hyper-signals are usually indicative of the presence of hyper-perfusions [Weber 2006]. However, in some tumour types they are related to a different phenomenon. For instance, in lymphomas the hyper-signal identified on the T1w-Gd map does not correspond to a hyper-perfusion [Weber 2006]. Moreover, neoangiogenesis, which is characterized by hyper-perfusions in ASL, can spread out of the T1w-Gd hyper-signal in particular in high grade tumours such as glioblastoma [Hakyemez 2005]. Ground truth hyper-perfused regions based on the T1w-Gd hyper-signal are therefore imprecise.

To overcome these limitations, we applied a two-step procedure taking advantage of both the anatomical and perfusional information (delivered by DSC CBF). In order to get an estimation of the ground truth, we implemented a method inspired by the hotspot technique [Noguchi 2008] commonly used in clinical practice. To this aim, the tumour was first segmented using a semi-automated method based on the T2w and T1w-Gd images and visually inspected by an expert neuro-radiologist. Then, we compared the tissue segmented as part of the tumour to its contralateral counterpart in the DSC CBF map. Voxels overtaking the lower and upper deciles were identified as potential hypo- and hyper-perfusions. Each potential perfusion abnormality was then visually inspected by an expert neuro-radiologist and manually corrected if needed. Special care was taken in order to avoid inclusion of hyper-perfusions related to the presence of arteries.

Due to its low SNR, ASL is not well suited to measure low levels of perfusion [Wintermark 2005]. That is why, we focused on hyper-perfusions for sensitivity estimation. Hypo-perfusions were nevertheless retained for specificity calculations.

Out of the 17 patients included in this study who underwent a DSC sequence, 9 presented hyper-perfusions, 16 hypo-perfusions and 8 both.

**False positives** According to clinical knowledge, in the absence of metastasis, the perfusion abnormalities should be confined to the affected tissue (tumour and oedema) identifiable on T1w-Gd and T2w. The proportion of non-affected tissue detected as a perfusion abnormality was used as a measure of the false positive rate.

In the control group, in which no detections were expected, an additional estimate of the specificity was calculated by leave-one-out cross-validation.

### 2.2.2 Validity of the assumptions of the homoscedastic model

The homoscedastic model makes the assumption that within-subject variance is either negligible by comparison to between-subject variance, or roughly constant across subjects [Mumford 2009]. In order to test each of these assumptions, we performed two experiments.

**Negligible within-subject variance** First, following [Chen 2012a], we measured the proportion of total variability that occurred within subjects with the following index, defined at each voxel for each subject  $s$ :

$$\lambda_s = \frac{\hat{\sigma}_s^2}{\hat{\sigma}_G^2 + \hat{\sigma}_s^2} \quad (8.20)$$

Values close to 1 mean that the within-subject variance ( $\sigma_s^2$ ) is preponderant compared to between-subject variance ( $\sigma_G^2$ ) and values close to 0 mean that between-subject variance holds the majority of the total variance. This measure is provided as an output of the MEMA function in AFNI [Chen 2012a].

**Constant within-subject variance across subjects** In a second experiment, we focused on the within-subject variance to verify whether it could be assumed roughly constant across subjects. To this aim, we calculated an average within-subject variance for each subject. This index was computed as suggested by [Mumford 2009] by averaging the within-subject variance for voxels within the interquartile range of the non-zero between-subject variance. Given this index of within-subject variance, we searched for outliers in the control group with Rosner's test [Rosner 1983], assuming that the distribution of the variance estimate was approximately normal. Then, we checked whether the control and patient groups had significantly different medians with a non-parametric Kruskal-Wallis test.

### 2.2.3 Comparison of the homoscedastic and heteroscedastic models

**Detections at a fixed threshold** In order to assess the difference that heteroscedasticity modelling would induce, we compared the sensitivity and specificity of the homoscedastic and heteroscedastic models. We compared both approaches in a usual setting, where the detections were identified at a threshold  $p < 0.05$  with FDR correction for multiple comparisons [Benjamini 1995]. As the data is usually pre-smoothed with a Gaussian kernel at the end of the pre-processing, we studied 6 kernel sizes defined by their FWHM: 0 mm<sup>3</sup> (i.e. no smoothing), 4 mm<sup>3</sup>, 6 mm<sup>3</sup>, 8 mm<sup>3</sup>, 10 mm<sup>3</sup>, 12 mm<sup>3</sup>.

**ROC analysis** In order to further assess the difference that heteroscedasticity modelling would induce, we compared the sensitivity and specificity of the homoscedastic and heteroscedastic models with ROC curves. In order to draw the ROC curves, we used 122 p-values (uncorrected), equally spaced in the logarithmic space. ROC curves provide a way to measure the performance of a classifier without focusing on a single threshold. We calculated the area under the curve as an indicator of the classification accuracy. Since the size of the expected detections is much smaller than the number of voxels that must not be detected, the area of interest in the ROC curve is the one of high specificity. That is why, as previously proposed in the literature, we focused on the area under the curve corresponding to

false positive rates ranging from 0% to 10% [Skudlarski 1999]. As in the previous experiment, different smoothing kernels were studied.

### 3 Materials

This section starts with a presentation of the datasets under study (3.1). The pre-processing steps applied to the data before the statistical analysis are then described (3.2).

#### 3.1 Data acquisition

25 patients diagnosed with brain tumours and 36 healthy volunteers were involved in this study. One control subject and four patients were excluded because of strong borderzone signs [Zaharchuk 2009]. The final dataset therefore included 21 patients (13 males, 8 females, age:  $55.2 \pm 14.1$  years) and 35 healthy volunteers (16 males, 19 females, age:  $27.7 \pm 6.4$  years).

Data acquisition was performed on a 3T Siemens Verio MR scanner with a 32-channel head-coil. Patients were scanned in the context of clinical practice. The imaging protocol included a 3D T1w (TR: 1900 ms, TE: 2.27 ms, FOV: 256 mm  $\times$  256 mm  $\times$  176 mm, flip angle: 9°, resolution: 1 mm  $\times$  1 mm  $\times$  1 mm), a PICORE Q2TIPS sequence [Wong 1998] with crusher gradients (TR: 3000 ms, TE: 18 ms, FOV: 192 mm  $\times$  192 mm, flip angle: 90°, resolution: 3 mm  $\times$  3 mm, slice thickness: 7 mm, inter-slice gap: 0.7 mm, TI: 1700 ms, bolus width:  $TI_{\text{wd}} = 700$  ms, 60 repetitions, mSENSE parallel imaging with accelerating factor of 2). In addition to these sequences, the patients also underwent a 3D T1w-Gd sequence (TR: 1900 ms, TE: 2.27 ms, flip angle: 9°, FOV: 250 mm  $\times$  250 mm  $\times$  176 mm, resolution: 1 mm  $\times$  1 mm  $\times$  1 mm) and a 2D T2w FLAIR sequence (TR: 9000 ms, TE: 90 ms, FOV: 220 mm  $\times$  199.4 mm, flip angle: 150°, resolution: 0.69 mm  $\times$  0.69 mm, slice thickness: 4 mm). Out of the 21 patients, 17 subjects also underwent a DSC sequence (GRE EPI, TR: 1500 ms, TE: 30 ms, FOV: 230 mm  $\times$  230 mm, flip angle: 90°, in plane resolution: 1.8 mm  $\times$  1.8 mm, slice thickness: 4 mm, inter-slice gap: 1.2 mm).

#### 3.2 Pre-processing

In this subsection, we present how raw ASL images were processed in order to compute CBF maps (3.2.1) and then normalized in intensity (3.2.2). Lastly, DSC pre-processing is briefly described (3.2.3)

##### 3.2.1 CBF estimation with ASL

Image pre-processing was performed using SPM8 (Statistical Parametric Mapping 8, Wellcome Department of Imaging Neuroscience, University College, London) in Matlab R2012a (Mathworks, Natick, MA). The anatomical image of each subject was segmented using the unified segmentation [Ashburner 2005]. A subject-specific anatomical brain mask was created, excluding voxels with less than 50% of brain



tissue in subsequent statistical analyses. A six-parameter rigid-body registration of the ASL volumes was carried out in order to reduce undesired effects due to subject motion. Rigid coregistration onto the whole-brain anatomical map was then performed based on mutual information. The average of unlabelled volumes was used to estimate the geometrical transformation to apply to each volume. Pair-wise subtraction of the control and labelled scans was then computed. A standard kinetic model [Buxton 1998] was applied in order to get ASL CBF, according to the following equation:

$$f = 6000 \times \frac{\lambda \Delta M}{2 M_0 \alpha TI_{wd} \exp^{-(TI + idx_{sl} * TI_{sl}) / T1_b}} \quad (8.21)$$

where  $f$  is the 4D CBF map in  $\text{mL} \cdot 100\text{g}^{-1} \cdot \text{min}^{-1}$ ,  $M_0$  the acquired  $M_0$  map (first volume of the ASL series),  $\lambda = 0.9 \text{ mL} \cdot \text{g}^{-1}$  the blood/tissue water partition coefficient,  $\alpha = 0.95$  the labelling efficiency,  $\Delta M$  the 4D perfusion-weighted map,  $TI = 1.7 \text{ s}$  the inversion time [Ferré 2012],  $idx_{sl}$  the slice index (0 for the first slice),  $TI_{sl} = 0.045 \text{ s}$  the duration of acquisition of one slice,  $TI_{wd} = 0.7 \text{ s}$  the temporal width of the bolus,  $T1_b = 1.5 \text{ s}$  the T1 of blood [Wang 2011]. The 6000 factor allows the conversion from  $\text{mL} \cdot \text{g}^{-1} \cdot \text{s}^{-1}$  to  $\text{mL} \cdot 100\text{g}^{-1} \cdot \text{min}^{-1}$  which is the standard unit for CBF. We bring the attention of the reader to the fact that, contrary to what is usually done in ASL pre-processing,  $\Delta M$  represents the set of perfusion-weighted maps (one volume per repetition) instead of a single perfusion-weighted map obtained by averaging across the repetitions. This is necessary in order to allow for the measurement of the within-subject variance.

Spatial normalization parameters estimated during the segmentation step were then applied to the T1 and ASL CBF maps in order to normalize the subjects into the ICBM-452 T1 template space [Mazziotta 2001]. This registration algorithm was selected since it gives good results even in the presence of large anatomical lesions [Crinion 2007a].

### 3.2.2 Intensity normalization

Intensity normalization is a common pre-processing step in PET or SPECT analysis [Arndt 1996] where the measured values are not quantitative. ASL produces quantitative CBF maps, but the large inter-subject variability in global CBF [Parkes 2003] induces a strong correlation across voxels. In voxelwise ASL detection studies, where the focus is on local variations across the brain, intensity normalization is therefore advised to increase the sensitivity [Aslan 2010]. Given  $CBF_{s,v}$ , the original CBF value of subject  $s$  at voxel  $v$ ,  $Perf_{s,v}$ , the normalized intensity value and  $\theta_s$  the cross-voxel normalization parameter, we have:

$$CBF_{s,v} = \theta_s \times Perf_{s,v}. \quad (8.22)$$

The most widespread approach to estimate  $\theta_s$  in ASL data processing is to compute the mean CBF signal found in a given ROI. This ROI is either limited to an anatomical region, which is known to be avoided by the pathology under study,

or covers the whole normal grey matter [Aslan 2010]. A threshold of 70% of grey matter is often chosen [Petr 2011]. The latter approach was preferred in this study since, in a general setting, no brain region is free of anomalies across all pathologies. Unphysiological negative perfusion estimates were excluded from the normalization mask.

Furthermore, as areas of abnormal perfusion should not be included in the normalization mask, we used an iterative scheme. At the first step, the normalization parameter was computed based on a mask covering the entire grey matter of the patient map under study. A first pass outlined the corresponding perfusion abnormalities as described in section 2.1.3. On the next step, the detected voxels were excluded from the intensity normalization estimation. These two steps were iterated until convergence.

### 3.2.3 CBF estimation with DSC

The DSC images were processed using MR manufacturer software by manually choosing an arterial input function to calculate: the CBF, cerebral blood volume, and mean transit time maps. The method is based on a deconvolution algorithm as described in [Ostergaard 1996]. Similarly to ASL, DSC CBF maps were coregistered on anatomical maps and spatially normalized.

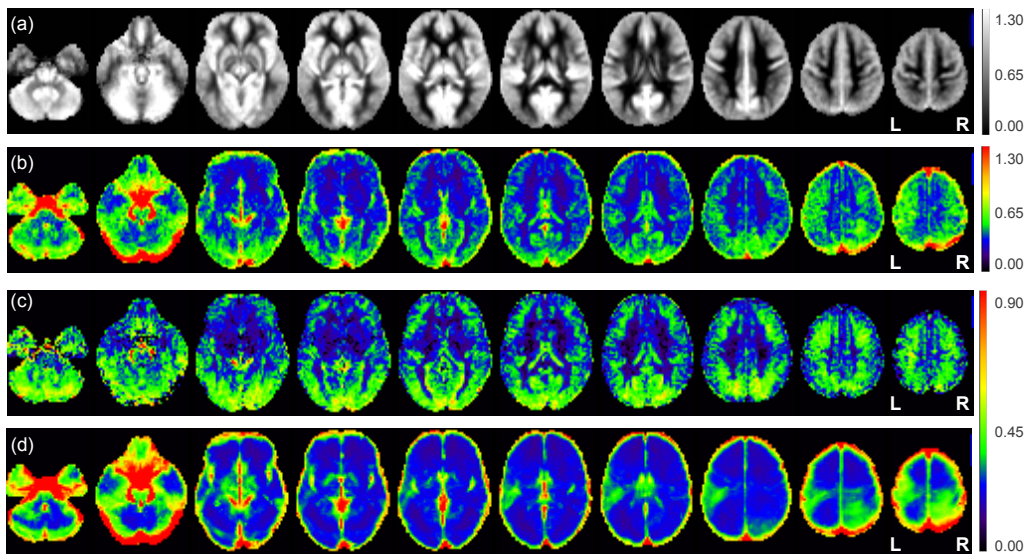


Figure 8.1: Parameter estimates of the homoscedastic and heteroscedastic model in the control group, computed from 35 healthy subjects. a) Mean perfusion estimate  $\hat{\beta}_{controls}^{HETERO}$ . b) Combined within- and between-subject standard deviation estimate from the homoscedastic model  $\hat{\sigma}_{G_C}$ . c) Between-subject standard deviation estimate from the heteroscedastic model  $\hat{\sigma}_G$ . d) Square root of the average within-subject variance in the control group. Perfusion is expressed in normalized units (ratio to mean grey matter perfusion). Axial slices are displayed in neurological convention.

## 4 Results

In this section we start with a graphical presentation of the parameter estimates of the homoscedastic and heteroscedastic models in the control group (4.1). Then, we present the results of the experiments regarding the validity of the assumptions of the homoscedastic model (4.2). In a last subsection, we display the results of the quantitative comparison between the homoscedastic and heteroscedastic models (4.3).

### 4.1 ASL template: a model of normal perfusion

In fig. 8.1, the parameter estimates of the homoscedastic and heteroscedastic models are displayed. Each parameter is defined as a voxelwise map. First, an estimate of the control group parameter is provided ( $\hat{\beta}_{controls}^{HETERO}$  is displayed, the homoscedastic estimate  $\hat{\beta}_{controls}^{HOMO}$  is visually nearly identical). As expected, the CBF is higher in the cortex and the basal ganglia than in white matter. Then, three standard deviation estimates are displayed:

- the combined within- and between-subject standard deviation estimate  $\hat{\sigma}_{G_C}$  from the homoscedastic model (8.1 b);
- the between-subject standard deviation estimate  $\hat{\sigma}_G$  from the heteroscedastic model (8.1 c);
- the root square of the average within-subject variance estimates in the control group  $(\frac{1}{n-1} \sum_{s=1}^{n-1} \hat{\sigma}_s^2)^{\frac{1}{2}}$ . This map is not part of the estimated standard deviations but is provided as a visual example of expected within-subject standard deviation in the control group (8.1 d).

The high variance values observed in the vascular structures, such as the transverse sinus, are in concordance with the findings of [Viviani 2009]. This pattern is clearly visible in the combined within- and between-subject variance estimate of the homoscedastic model and captured by the within-subject variance estimate in the heteroscedastic model. The main variations observed in the between-subject variance, as estimated in the heteroscedastic model, are related to inter-subject misregistrations in the cortex. An increased variance is also visible in the occipital lobe, probably related to increased arterial transit times in these regions [MacIntosh 2010].

### 4.2 Testing the assumptions of the homoscedastic model

The homoscedastic model stands on one of the following assumptions: either the within-subject variance is negligible by comparison to between-subject variance, or the within-subject variance is roughly constant across subjects.

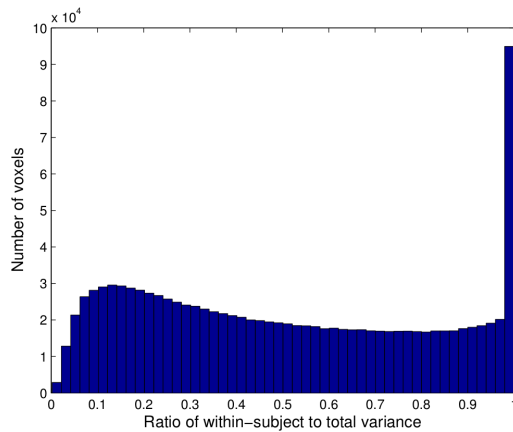


Figure 8.2: Histogram of the ratio of within-subject variance onto total variance in the control group. Both variance estimates have a significant impact depending on the voxels.

#### 4.2.1 Relative weights of between-subject and within-subject variances in the control group

In order to verify if the within-subject variance can be assumed negligible by comparison to between-subject variance, fig. 8.2 presents the histogram of the ratio of within-subject to total variance in the control group ( $\lambda_s$  from eq. (8.20)). Overall, a total of 1 094 790 voxels are considered. Values close to one indicate a preponderance of the within-subject variance, whereas values close to zero outline a dominating between-subject variance. Clearly, both components of variance have an important impact as the  $\lambda_s$  index spans the complete range of values between 0 and 1. The large peak indicating voxels with a zero between-subject variance was also observed in [Chen 2012a] and might be the consequence of calculation inaccuracies. While the true between-subject variance might not be exactly zero, these voxels nevertheless present a very small between-subject variance, negligible by comparison to the within-subject variance. The quartiles of the distribution are 0.23 and 0.78, so that the outermost 25% of the voxels (on each side) have either a dominant within-subject variance or a dominant between-subject variance. In the remaining 50% voxels, both variance components have a significant impact. In conclusion, it cannot be assumed that within-subject variance is negligible in comparison to between-subject variance.

#### 4.2.2 Cross-subject comparison of within-subject variances

In this experiment, we tested whether within-subject variance could be assumed roughly constant across subjects. As proposed by [Mumford 2009], we calculated an average within-subject variance across voxels. This led to an index of within-subject variance per subject. Fig. 8.3 displays the within-subject variance indexes in the control and patient groups.

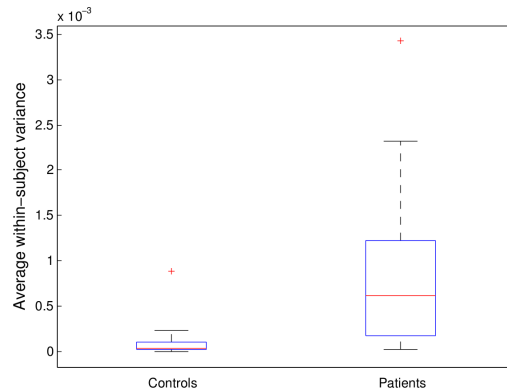


Figure 8.3: Average estimated within-subject variances in control and patient groups. For each box, the red line corresponds to the median and the top and bottom lines of the blue square are the upper and lower quartiles of the distribution.

In the control group, it is clear that one of the subjects presents an unexpected high variance by comparison to the other controls. Retrospectively, we found out that this subject moved substantially more than other controls and was identified as uncooperative by the MR physicist during the acquisition. As expected from fig. 8.3, we found one outlier in the control group with Rosner’s test.

The within-subject variance was significantly higher in the patient group than in the control group ( $p < 0.05$  with Kruskal-Wallis test). This might be explained by the fact that patients tend to have more difficulties to lie still during the acquisition due to their pathological condition.

The index proposed by [Mumford 2009] is an average across voxels and therefore focused on global variations of the within-subject variance. However, strong variations, sometimes caused by artefacts, that appear locally can also be a concern. As an example, fig. 8.4a displays the estimated within-subject standard deviation for a control subject presenting locally atypical patterns. While the high variance induced by the presence of large vessels is a pattern shared across subjects (as previously described in fig. 8.1, last panel), the high variance observed bilaterally in the frontal lobe is specific of this control subject. We hypothesize that these strong variations are the consequence of motion during the acquisition that was not correctly compensated during the pre-processing. Even if this subject was not previously outlined as an outlier, these atypical patterns of variance might have a detrimental impact on the analysis if not properly taken into account.

In conclusion, variations across subjects of the within-subject variance appear to be important in ASL. This might be a consequence of the low SNR of this technique, since small artefacts in the original control and labelled scans can lead to substantial variations in the perfusion-weighted map after subtraction. In the next sections we investigate whether modelling heteroscedasticity can improve the detections of perfusion abnormalities.

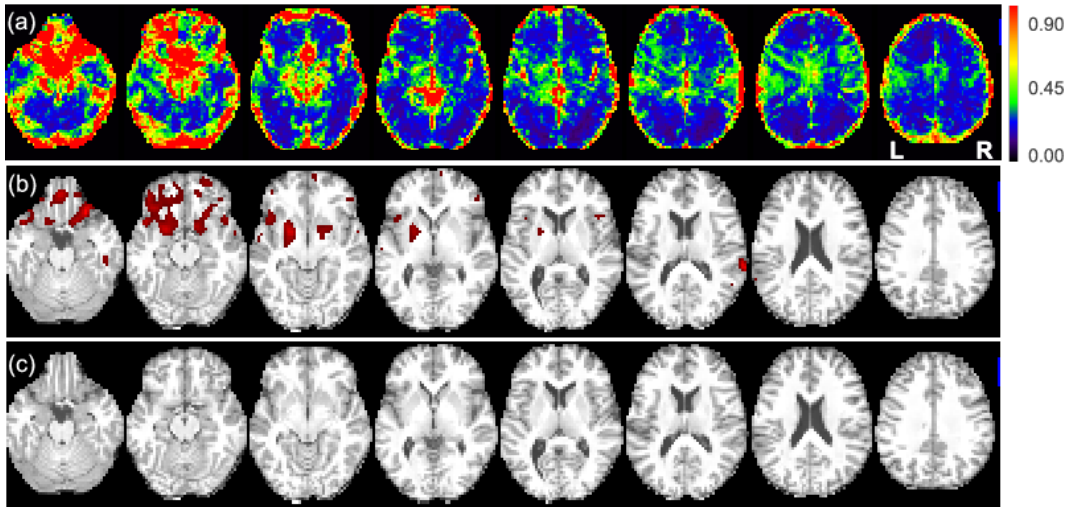


Figure 8.4: a) Estimated within-subject standard deviation in a control subject presenting locally atypical values. b) T1w map with false positive detections (in red) by leave-one-out cross-validation (smoothing  $\text{FWHM} = 8 \text{ mm}^3$ ) with the homoscedastic model (b) and the heteroscedastic model (c). No false positive detections are observed with the heteroscedastic model.

Table 8.1: Sensitivity and specificity in the control (first row) and patient (second and third rows) groups with the homoscedastic and heteroscedastic models for different smoothing kernels ( $\text{FWHM} = [0, 4, 6, 8, 10, 12] \text{ mm}^3$ ), at  $p < 0.05$  FDR corrected.

	Homoscedastic model						Heteroscedastic model					
	0	4	6	8	10	12	0	4	6	8	10	12
Specificity (controls)	1.00	0.99	0.99	<b>0.99</b>	0.99	0.99	1.00	1.00	<b>1.00</b>	1.00	1.00	1.00
Specificity (patients)	0.92	0.89	0.87	<b>0.86</b>	0.85	0.84	1.00	0.99	<b>0.99</b>	0.99	0.99	0.98
Sensitivity (hyper)	0.49	0.59	0.63	<b>0.66</b>	0.67	0.67	0.51	0.59	<b>0.61</b>	0.58	0.57	0.55

### 4.3 Comparison of homoscedastic and heteroscedastic models

In this subsection, the homoscedastic and heteroscedastic models are compared in terms of sensitivity and specificity to detect patient-specific perfusion abnormalities with ASL (as described in 2.1.3).

#### 4.3.1 Fixed threshold

**Quantitative analysis** Table 8.1 presents the sensitivity to detect hyper-perfusions and the specificity both in the patient group and by leave-one-out cross-validation in the control group with the homoscedastic and heteroscedastic models for different smoothing kernels ( $\text{FWHM} = [0, 4, 6, 8, 10, 12] \text{ mm}^3$ ), at  $p < 0.05$  FDR corrected.

Overall, the heteroscedastic model leads to a decrease in false positive rate. This is particularly noticeable in the patient group where the specificity is improved

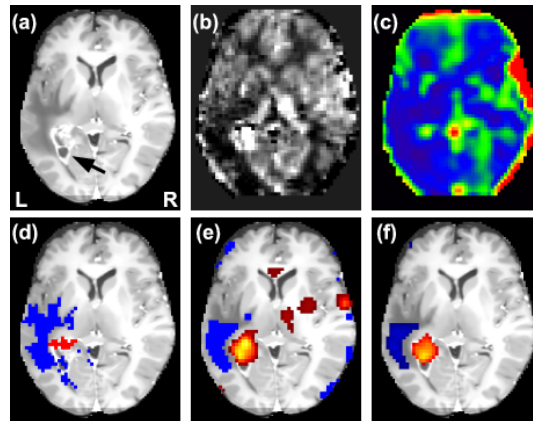


Figure 8.5: Detections of perfusion abnormalities with the homoscedastic and heteroscedastic models in a patient suffering from a high grade glioma. a) T1w-Gd map, the tumour site is pointed by a black arrow. b) Patient ASL CBF estimate  $\hat{\beta}_n$ . c) Within-subject standard deviation of ASL CBF  $\hat{\sigma}_n^2$ . d) T1w-Gd map with ground truth overlaid. T1w-Gd map with hypo- (blue colour-map) and hyper-perfusions (hot colour-map) overlaid for the homoscedastic (e) and heteroscedastic (f) models. Modelling heterogeneous variances (heteroscedastic model) reduces the false positive detections while preserving the true detections. Axial slices are displayed in neurological convention.

for each smoothing studied. As expected, in the control group, this effect is also observed but to a lesser extent. This is probably due to the fact that the hypothesis of homoscedasticity is better suited for the control subjects. The increase in false positive rate, with the homoscedastic model, is accompanied by a relative increase in sensitivity, which does not seem significant given the specificity loss.

**Qualitative analysis** At the individual level, as illustrated in fig. 8.4, local atypical patterns of variance can further lead to a substantial increase in false positives with the homoscedastic model. The unexpected high standard deviation in the frontal lobe for the control subject presented in fig. 8.4 leads to false positive hyper-perfusions with the homoscedastic model whereas the heteroscedastic model does not get any false positive (smoothing  $\text{FWHM} = 8 \text{ mm}^3$ ). Fig. 8.5 further illustrates the benefits of the heteroscedastic model compared to the homoscedastic model in a patient subject. The motion artefacts and hyper-signals induced by arteries indeed correspond to regions of high within-subject standard deviation. Modelling heterogeneous variances reduces the artefactual detections in these regions while preserving the quality of the true detections.

To investigate whether, at a same specificity rate, the sensitivity differs between the two approaches, we employed a ROC analysis as described in the next section.

## 4.3.2 ROC analysis

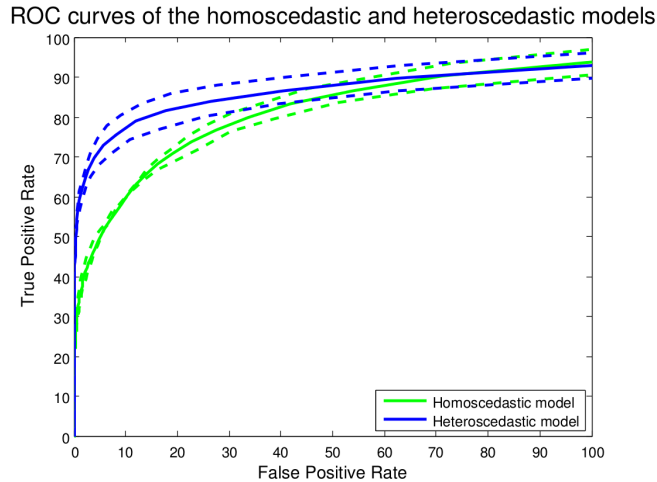


Figure 8.6: ROC curves for perfusion abnormality detections with the homoscedastic and heteroscedastic models. The average ROC curves across the studied smoothings are plotted in plain line. Dotted lines are plotted one standard deviation away from the average.

Table 8.2: Area under the ROC curve for false positive rates ranging between 0% and 10% with the homoscedastic and heteroscedastic models. The heteroscedastic model outperforms the homoscedastic model.

	Homoscedastic model						Heteroscedastic model					
	0	4	6	8	10	12	0	4	6	8	10	12
ROC area	0.46	<b>0.49</b>	<b>0.49</b>	<b>0.49</b>	0.48	0.48	0.63	0.70	<b>0.72</b>	<b>0.72</b>	0.69	0.65

**Quantitative comparison** Fig. 8.6 presents the ROC curves in the patient group for the homoscedastic and heteroscedastic models. The average over the studied smoothing kernels is plotted along with the spread provided by the standard deviation over the 6 measurements. The ROC curve of the heteroscedastic model is substantially closer to the ideal classifier than the homoscedastic curve.

Table 8.2 presents the area under the ROC curves for false positives rates ranging from 0 to 10%, for the homoscedastic and heteroscedastic models with different smoothing kernels. The heteroscedastic model outperforms the homoscedastic model with an increased area under the curve. The best values are 0.72 and 0.49 with the heteroscedastic and homoscedastic model respectively. For both approaches, a maximum area under the curve is reached for a smoothing kernel of  $8 \text{ mm}^3$ .



**Qualitative comparison** In order to illustrate the advantage of the heteroscedastic over the homoscedastic model, we chose 3 representative subjects and compared the methods at fixed false positive rate and true positive rate. To this aim, we selected the uncorrected p-values that would lead to a pre-specified false positive rate (respectively true positive rate) from the ROC analysis. We worked with data smoothed with a Gaussian kernel of  $8 \text{ mm}^3$  that led to the best area under the curve with both models. It is worth noting that in the previous section the sensitivity and specificity were estimated at the group level and that the 3 subjects presented in this part were chosen so that they would best illustrate the group findings. Fig. 8.7 presents the detections obtained with both methods on the 3 selected patients. The first subject, which presents a small hyper-perfusion, is studied at a true positive rate of 50%. In the two remaining patients, the methods are compared at a false positive rate of 0.1%.

Patient 15 suffers from a gliosarcoma in the left hemisphere close to the parahippocampal region. The lesion displays a small hyper-perfusion in its dorsal part, small hypo-perfusions are seen in the surrounding oedema. At a true positive rate of 50%, the homoscedastic model displays a larger number of false positives than the heteroscedastic model. Patient 16 was diagnosed with a high grade tumour in the left temporal lobe. The lesion is characterized by a large hyper-perfusion and a surrounding hypo-perfusion. With a false positive rate of 0.1%, the hyper-perfusion is correctly located with both methods. The extent of the hyper-perfusion is however better covered by the heteroscedastic model. Patient 6 suffers from a meningioma partly hyper-perfused. Similarly to patient 16, at a false positive rate of 0.1%, both methods detect the hyper-perfusion but the heteroscedastic model is clearly more sensitive.

These 3 cases illustrate the loss of sensitivity of the homoscedastic model by comparison to the heteroscedastic model at the same specificity level.

## 5 Discussion and conclusion

We have compared two approaches to quantitatively outline patient-specific pathological patterns of abnormal perfusion in ASL data based on the massively univariate GLM: the homoscedastic and heteroscedastic models.

We demonstrated that the assumptions underlying the homoscedastic model are not verified in ASL studies. More precisely, the within-subject variance cannot be considered as negligible by comparison to between-subject variance, nor constant across subjects. In fMRI data, small deviations from homoscedasticity have shown to not overly alter the results in one-sample mixed-effects GLM analysis [Mumford 2009]. Here, we showed that modelling heterogeneous within-subject variances is essential in order to reach a satisfactory level of specificity in a mixed-effect two-sample t-test comparing a patient to a group of controls in ASL. These results are in line with a recent study by [Chen 2012a] where heteroscedastic mixed-effects GLM were shown to provide more accurate results in fMRI.

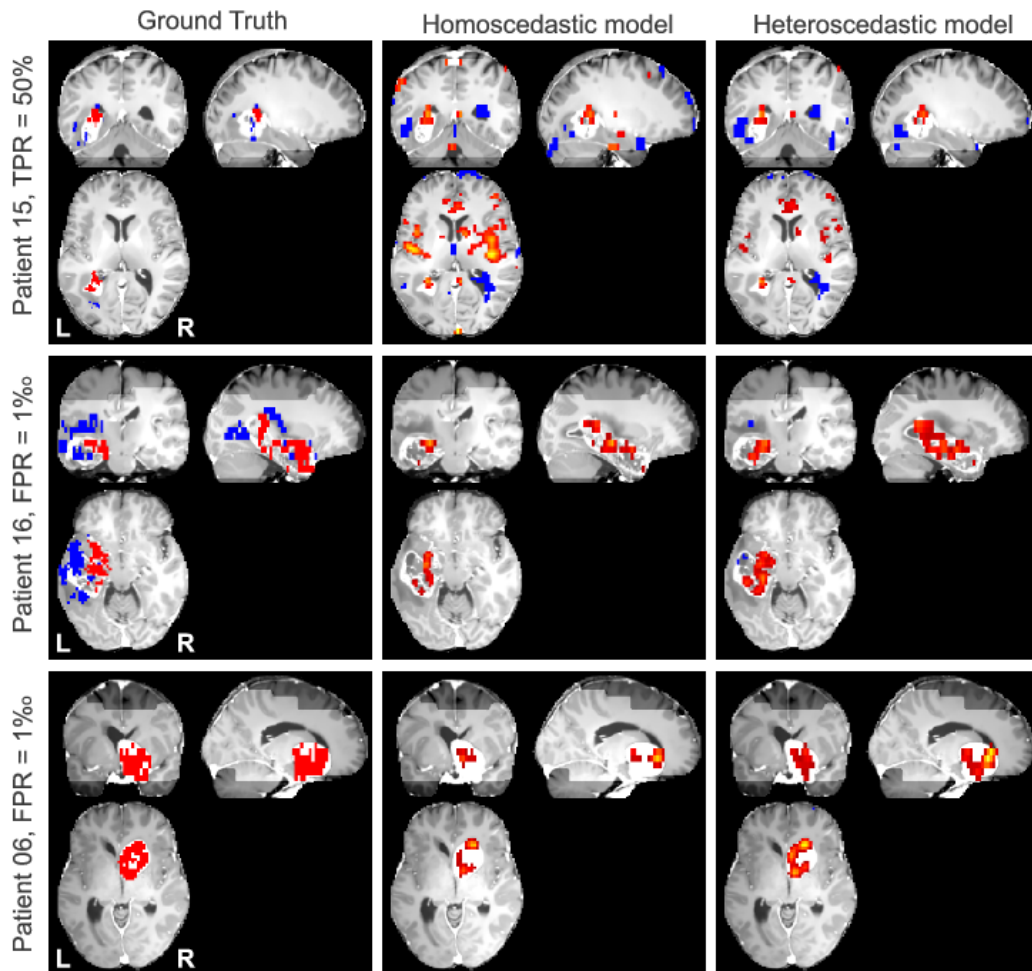


Figure 8.7: Perfusion abnormalities detections in 3 patients with homoscedastic and heteroscedastic models. From left to right: ground truth hyper-perfusions (red) and hypo-perfusions (blue) overlaid on the  $T1w$ -Gd map; detections with the homoscedastic model with a smoothing kernel of  $FWHM = 8 \text{ mm}^3$ , hyper-perfusions (hot colormap) and hypo-perfusions (blue); detections with the heteroscedastic model with a smoothing kernel of  $FWHM = 8 \text{ mm}^3$ , hyper-perfusions (hot colormap) and hypo-perfusions (blue).

In the context of this paper, we defined a single parameter in the subject-level design matrix and focused on perfusion-weighted images obtained after pair-wise subtraction of the control and labelled scans. Other authors have suggested that taking the complete ASL time-course into account (before subtraction) would lead to more efficient estimates of perfusion in fMRI [Mumford 2006]. Also, additional regressors can be introduced as nuisance covariates for denoising purposes in the subject-level design matrix as suggested in [Wang 2012]. Modifying the subject-level design matrix would change the estimated subject parameters and their sampling variances. However, an heteroscedastic model would still be appropriate in this setting so that the conclusions of this paper remain valid even with a different subject-level model.

We demonstrated that within-subject variance captures important information regarding the subject-specific spatial distribution of noise in ASL data. We also outlined that patient-specific brain perfusion abnormalities can be correctly detected using ASL if the heterogeneous within-subject variances are properly modelled. We therefore advise the use of heteroscedastic models in ASL studies.

## Chapter 9

# *A contrario* detection of perfusion abnormalities

---

This chapter is presented in the form of a paper. This work was submitted to NeuroImage. Preliminary versions of this work were published in the IEEE 9th International Symposium on Biomedical Imaging (ISBI) in 2012 [Maumet 2012a] and in the 15th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) in 2012 [Maumet 2012b]

### Abstract

In this paper, we introduce a new locally multivariate procedure to quantitatively extract individual patterns of abnormal perfusion in ASL. This method, called the *a contrario* approach comes from the computer vision community and allows to take into account the local information without smoothing of the data. To improve on the standard *a contrario* approach, which assumes white noise, we introduce a non-independent *a contrario* approach which takes into account the noise spatial coherency to provide the distribution of the statistic.

Validation is undertaken on a dataset of 60 healthy volunteers and 25 patients diagnosed with brain tumours. The *a contrario* approach outperforms the massively General Linear Model usually employed for this type of analysis. The non-independent *a contrario* approach provides a statistic that is more valid than the standard *a contrario* approach and than the GLM.

**Contents**

<b>1</b>	<b>Introduction</b>	<b>163</b>
<b>2</b>	<b>Methods</b>	<b>164</b>
2.1	The massively univariate General Linear Model	164
2.2	The <i>a contrario</i> approach: a locally multivariate procedure	166
2.2.1	Partition of the image into sub-regions	167
2.2.2	From voxel-based to region-based probabilities	167
<b>3</b>	<b>Material</b>	<b>169</b>
3.1	Data	170
3.1.1	Data acquisition	170
3.1.2	Pre-processing	170
3.2	Ground truth	171
3.2.1	Positives	172
3.2.2	Negatives	172
<b>4</b>	<b>Results</b>	<b>172</b>
4.1	Comparison of the <i>a contrario</i> approach with the General Linear Model	173
4.1.1	Individual ROC curves	173
4.1.2	Group ROC curves	175
4.1.3	Qualitative comparison	175
4.2	Distribution of the p-values under the null hypothesis	176
<b>5</b>	<b>Discussion and conclusion</b>	<b>179</b>

## 1 Introduction

Brain perfusion is the biological process that ensures the delivery of oxygen and nutrients to the cerebral tissues by means of microcirculation. **ASL**, a **MRI** technique introduced in the early 1990's, allows to measure the level of perfusion through a quantitative index: the **CBF**. Contrary to standard perfusion imaging, including **PET** and **SPECT** in nuclear medicine or **DSC** in **MRI**, **ASL** is completely non-invasive and does not require the injection of an exogenous contrast agent. In **ASL**, blood water, used as an endogenous tracer, is labelled with a radio-frequency pulse. After a delay called inversion time, a labelled image of the brain is acquired. A control image is also acquired without prior labelling. In **ASL** with a single readout time, the difference between the labelled and the control image leads to a perfusion-weighted map, proportional to the **CBF**. To increase the **SNR** of the measure, the acquisition is repeated several times.

The absence of radiation and the use of an endogenous tracer, which removes the need for injection of a contrast agent, are clear advantages of **ASL** over standard perfusion imaging techniques. The **ASL** sequence however suffers from a low **SNR**, which is still a serious obstacle for its use in clinical practice.

Perfusion is an indicator of the well-being of the tissues and is a useful tool for diagnosis in clinical practice. Identifying regions of abnormal perfusion, either hypo-perfusions or hyper-perfusions, can help understanding the mechanism of a disease and taking care of the patient. For instance, for patients diagnosed with tumours, the clinician is interested in hyper-perfusions that would reveal the grade of the tumour [Weber 2006], or help differentiating between post-radiation necrosis and tumour recurrence [Sugahara 2000]. In strokes, the extent of hypo-perfusion and mismatch with diffusion imaging provide an insight on the possible recovery of the tissue [Huck 2012], an indication of tremendous importance to help the clinician decide on the adequate treatment. Since the introduction of **ASL**, and despite its low **SNR**, a large number of studies have demonstrated its usefulness in identifying patterns of abnormal perfusion at the group level, for example [Pinkham 2011]. The large majority of these quantitative statistical analyses are based on a **GLM**, a massively univariate approach widespread in the brain imaging community. Very few **ASL** studies focused on voxelwise quantitative perfusion abnormality detections at the individual level. In [Petr 2013], a template-based analysis was presented in order to detect individual activation patterns in functional **ASL** data. Interestingly, they also applied their method to a pathological case in order to detect a hypo-perfusion co-located with a dysplasia in an epileptic patient. While their approach opens the field to the detection of patient-specific perfusion abnormalities with **ASL**, a single pathological case was presented and no quantitative validation performed.

With other perfusion methods such as **SPECT**, template based perfusion abnormality detections at the individual level in patients have demonstrated their interest [Vik 2005, Stamatakis 2002, Lee 2000]. Thanks to its non-invasiveness, **ASL** is more suitable than **DSC** or **SPECT** to acquire data in healthy controls. However, the identification of individual patterns of hypo- and hyper-perfusions in patients

with ASL mainly relies on visual analysis [Zaharchuk 2012, Chen 2012b].

In this paper, we propose a new statistical framework for the detection of patient-specific brain perfusion abnormalities with ASL. Our method is based on an *a contrario* analysis. This recent statistical approach, first introduced in [Desolneux 2003], has later been applied to medical image processing [Rousseau 2007, Aguerrebere 2009]. It allows to take into account neighbouring voxels without pre-smoothing the data,

The proposed method is compared to the GLM on quantitative grounds by means of ROC analyses. The model of normal perfusion is computed out of the data of 60 healthy volunteers. Quantitative validation is performed on a dataset of 25 patients diagnosed with brain tumours. As a matter of fact, perfusion abnormalities have been widely studied in this context [Chalela 2000, Sugahara 2000, War-muth 2003, Weber 2006, Chawla 2007].

We previously investigated the ability to detect hypo- and hyper-perfused regions using a template of normal perfusion in two conference papers [Maumet 2012a, Maumet 2012b]. On the methodological side, we present here a new variant of our model, which takes into account the spatial autocorrelation of the noise. Furthermore, a larger database is studied, a novel quantitative validation is provided, and a comprehensive quantitative comparison with the GLM is performed.

In section 2, the principles of *a contrario* analyses are presented, and the proposed approach for the detection of perfusion abnormalities is introduced. Section 3 outlines the experimental context: data acquisition and pre-processing are described, followed by the quantitative metrics and experiments used for validation. The results are presented in section 4. Section 5 gives a discussion and concludes.

## 2 Methods

This section presents the methods developed in order to detect areas of abnormal perfusion in a patient map by comparison to a group of control subjects. First, in 2.1, we review the standard massively univariate GLM [Beckmann 2003, Mumford 2009]. Then, in 2.2, we describe the proposed locally multivariate *a contrario* approach.

### 2.1 The massively univariate General Linear Model

The General Linear Model used to detect patterns of abnormal perfusion in a patient by comparison to a group of subjects can be defined as a hierarchical model with two levels: subject and group.

At the subject level, we observe a series of  $n$  CBF volumes (one per repetition of the ASL acquisition) and aim at obtaining an estimate of the true CBF for each voxel  $v$ . Thus, we have:

$$\mathbf{Y}_{(s,v)} = \mathbf{X}_{(s,v)}\beta_{(s,v)} + \boldsymbol{\varepsilon}_{(s,v)} \quad (9.1)$$

where  $\mathbf{Y}_{(s,v)}$  is a  $d$ -dimensional vector of replicated CBF values.  $\mathbf{X}_{(s,v)}$  is the subject-level design matrix that reduces to a  $d$ -dimensional vector of ones in the absence of covariate.  $\beta_{(s,v)}$  is the subject parameter, i.e. the actual (unknown) CBF value of the subject  $s$  in the voxel  $v$ . Finally, the  $d$ -dimensional vector  $\varepsilon_{(s,v)}$  is the error term and follows a central Gaussian distribution with standard deviation  $\sigma_{(s,v)}$  representing the intra-subject variability.

Equation (9.1) can be solved independently for each subject. If no nuisance regressors are defined in the model, the ordinary-least-square estimate of the subject parameter  $\beta_{(s,v)}$  is the sample average across the ASL repeats:

$$\hat{\beta}_{(s,v)} = \frac{1}{n} \sum_{i=1}^n y_{(s,v)}^i, \quad (9.2)$$

where  $y_{(s,v)}^i$  is the  $i^{\text{th}}$  element of the vector  $\mathbf{Y}_{(s,v)}$ . Similarly, the sampling variance of  $\hat{\beta}_{(s,v)}$  is estimated by:

$$\widehat{\text{Var}}(\hat{\beta}_{(s,v)}) = \frac{\hat{\sigma}_{(s,v)}^2}{n} \quad \text{where} \quad \hat{\sigma}_{(s,v)}^2 = \frac{1}{n-1} \sum_{i=1}^n (y_{(s,v)}^i - \hat{\beta}_{(s,v)})^2. \quad (9.3)$$

At the group level, we have an estimated CBF map for each subject,  $\hat{\beta}_{(s,v)}$ , and we want to compare the level of perfusion in the patient map to the level of perfusion in the control group. The model is thus:

$$\hat{\beta}_v = \mathbf{X}_{Gv} \beta_{Gv} + \varepsilon_{Gv}. \quad (9.4)$$

The  $n$ -dimensional vector  $\hat{\beta}_v$  contains the subject parameter estimates. The 2-dimensional vector of group parameters  $\beta_{Gv}$  contains the control group and patient parameters. The  $n \times 2$  group design matrix  $\mathbf{X}_{Gv}$  affects each subject to the group it belongs to. The  $n$ -dimensional vector  $\varepsilon_{Gv}$  is the error term and follows a central Gaussian distribution with standard deviation  $\sigma_{Gv}$  representing the inter-subject variability.

Without loss of generality, we assume that subjects 1 to  $n-1$  are part of the control group. Using an heteroscedastic model, the level of perfusion in the patient map at voxel  $v$  is compared to the level of perfusion in the control group with the following  $t$ -statistic [Maumet 2013a]:

$$t_v = \frac{\hat{\beta}_{(n,v)} - \frac{1}{\sum_{s=1}^{n-1} w_{(s,v)}} \sum_{s=1}^{n-1} w_{(s,v)} \hat{\beta}_{(s,v)}}{\hat{\sigma}_{Gv}^2 + \hat{\sigma}_{(n,v)}^2 + \frac{1}{\sum_{s=1}^{n-1} w_{(s,v)}}} \quad (9.5)$$

where  $\frac{1}{w_{(s,v)}} = \hat{\sigma}_{Gv}^2 + \hat{\sigma}_{(s,v)}^2$ .

$\hat{\sigma}_{Gv}^2$  is the estimated inter-subject variance and  $\hat{\sigma}_{(s,v)}^2$  the estimated intra-subject variance (over the ASL repetitions) for subject  $s$ .



Under the null hypothesis  $H_0$  that the perfusion value in the patient map is equal to the control group mean, the  $t$ -statistic  $t_v$  follows a Student distribution  $T_{n-1}$  with  $n - 1$  degrees of freedoms. The probability  $\pi_v^{GLM}$  of being in presence of a hyper-perfusion at voxel  $v$  of the patient map is therefore estimated by:

$$\pi_v^{GLM} = \Pr(T_{n-1} \geq t_v). \quad (9.6)$$

Hypo-perfusions can respectively be studied with  $\Pr(T_{n-1} \leq t_v)$ . Valid detections are then obtained by thresholding the probability map using a correction for multiple testing.

## 2.2 The *a contrario* approach: a locally multivariate procedure

The *a contrario* approach [Desolneux 2003] is a locally multivariate procedure that produces region-based probabilities (as opposed to voxel-wise probabilities generated by the massively univariate GLM). The *a contrario* probability estimation is usually based on a two-step procedure. First, a voxel-wise probability map is computed under a so-called “background model”. Second, voxel-wise probabilities are converted into region-based probabilities assuming a given type of local noise (usually white).

In the first step, a model of the background is defined.

This model can represent the background noise or be learned from a control population. Let  $J_v$  be the random variable representing the value observed at the voxel  $v$  of the image of interest. Under the background model  $\mathcal{M}$ , depending on the set of parameters  $\gamma$ , we have:

$$J_v \sim \mathcal{M}(\gamma_v). \quad (9.7)$$

Given  $x_v$ , the value observed at voxel  $v$ , each voxel of the image can be associated with a probability under the background model:

$$\Pr(J_v \geq x_v), \text{ where } J_v \sim \mathcal{M}(\gamma_v) \quad (9.8)$$

Rousseau et al. first noticed that the term “background model” commonly employed in the *a contrario* literature could be well considered as a null hypothesis in standard hypothesis testing [Rousseau 2007]. So that the standard massively univariate GLM (based on unsmoothed data) can in fact be used to produce the input voxel-wise probability map of the *a contrario* analysis leading to:

$$\Pr(J_v \geq x_v) = \pi_v^{GLM}. \quad (9.9)$$

In the second step, in order to extend the probabilities from voxels to regions of the image we then need:

1. A partition of the image into sub-regions;
2. A model that link the voxel-based probabilities observed in a region to its region-based probability.

### 2.2.1 Partition of the image into sub-regions

Given an *a priori* shape of the structures to be detected the regions can be selected accordingly. For instance, in the first *a contrario* application, which aimed at detecting alignments, segments of variable sizes were chosen as regions of analysis [Desolneux 2003].

In our case, the shape of the expected detections is unknown and, as it has been proposed [Rousseau 2007, Aguerrebere 2009], the sub-regions are defined as spheres centred at each voxel.

### 2.2.2 From voxel-based to region-based probabilities

The conversion from voxel-based to region-based probabilities involves the notion of “rare events”. By definition, a rare event occurs at voxel  $v$  if the probability (under the null hypothesis) to observe such value, or a more extreme, is smaller than a pre-defined threshold  $p_{PRE}$ . The initial voxel-wise probability map is thus thresholded to produce a binary map :

$$k_v = \begin{cases} 1, & \text{if } \pi_v^{GLM} \leq p_{PRE}, \\ 0, & \text{otherwise.} \end{cases} \quad (9.10)$$

Then, the number of rare events in a region  $R$ ,  $l_{(r)}$ , is determined by:

$$l_{(r)} = \sum_{v \in R} k_v. \quad (9.11)$$

Let  $L_{(r)}$  be the random variable representing the number of rare events observed in the region  $R$ . Then, we have:

$$L_{(r)} = \sum_{v \in R} K_v \quad \text{where } K_v \sim \text{Bern}(p_{PRE}) \quad (9.12)$$

where  $\text{Bern}(p_{PRE})$  is a Bernoulli distribution with probability  $p_{PRE}$ .

The t-statistic used in the massively univariate GLM, is therefore replaced by a locally multivariate statistic in the *a contrario* approach: the number of rare events  $l_{(r)}$ . In order to be able to make inference using the *a contrario* approach, we must determine the distribution of the statistic  $l_{(r)}$  under the null hypothesis.

**Independence assumption** The standard *a contrario* approach assumes spatial independence of the residuals, i.e. the probability (under the background model or equivalently under the null hypothesis) to observe a rare event at a given voxel is independent of the probability to observe a rare event at any other voxel in the image, and in particular in its neighbourhood. As a sum of independent Bernoulli random variables is a random variable following a Binomial distribution, the region-based probability  $\pi_{(r)}$  is then estimated using a Binomial distribution parametrized by the number of voxels in each region,  $e = \text{card}(R)$ , and the pre-defined p-value  $p_{PRE}$ :

$$\pi_{(r)} = \Pr(L_{(r)} \geq l_{(r)}), \quad \text{where } L_{(r)} \sim \text{B}(e, p_{PRE}). \quad (9.13)$$

In the context of this paper, the probability associated to each region  $R$  is affected to its centre voxel  $v$  to produce a voxel-wise probability map:

$$\pi_v^{CONT} = \pi_{(r)}. \quad (9.14)$$

**Non-independence of the residuals** The standard *a contrario* approach assumes spatially white noise in the residuals. However, the presence of spatial autocorrelation of the noise has been well described in the MRI literature, in particular when discussing the issue of cluster inference, see for example [Chumbley 2009], and cannot be ignored. We draw the attention of the reader on the fact that the dependence of the residuals is not related to the spatial coherency observed in the perfusion signal. Under the null hypothesis, it is the spatial autocorrelation of the noise (observable in the residuals) that impacts the locally multivariate probabilities. In order to estimate the region-based probability of observing a given number of rare events per sphere, we propose to rely on a multivariate normal distribution as developed hereafter. In the remainder of the paper, this alternative method will be referred as the non-independent *a contrario* approach.

Let's first notice, that the probability of observing  $l_{(r)}$  or more rare events can be calculated as a sum of probabilities:

$$\Pr(L_{(r)} \geq l_{(r)}) = \sum_{i=l_{(r)}}^e \Pr(L_{(r)} = i) \quad (9.15)$$

Then, given  $C_{(r)}^{(i)}$  the set of combinations of  $i$  elements among  $R$ , we have:

$$\Pr(L_{(r)} = i) = \sum_{S \in C_{(r)}^{(i)}} \Pr(K_{(v)} = 1, \forall v \in S \text{ and } K_{(v)} = 0, \forall v \notin S) \quad (9.16)$$

The problem now lies in the calculation of the joint probability to observe a given combination of “rare”/“not rare” voxels in the region of interest. Also, by definition of the term “rare event” we have:

$$\begin{aligned} & \Pr( K_{(v)} = 1, \forall v \in S \quad \text{and} \quad K_{(v)} = 0, \forall v \notin S ) \\ &= \Pr( \pi_v \leq p_{PRE}, \forall v \in S \quad \text{and} \quad \pi_v > p_{PRE}, \forall v \notin S ) \\ &= \Pr( t_v \geq \Phi_{t_\nu}^{-1}(p_{PRE}), \forall v \in S \quad \text{and} \quad t_v < \Phi_{t_\nu}^{-1}(p_{PRE}), \forall v \notin S ), \end{aligned} \quad (9.17)$$

where  $\Phi_{t_\nu}^{-1}$  is the inverse cumulative function of a Student distribution with  $\nu$  degrees of freedom, outlining that the joint probability can be estimated from a Student's  $t$  random field. As previously proposed in the neuroimaging literature regarding the use of random field theory for multiple testing correction [Kiebel 1999, Zhang 2009], we focus on the Gaussianised  $t$  random field:

$$Z_v = \Phi^{-1}(\Phi_{t_\nu}(t_v))$$

Then we have:

$$\begin{aligned} & \Pr ( t_v \geq \Phi_{t_v}^{-1}(p_{PRE}), \forall v \in S \quad \text{and} \quad t_v < \Phi_{t_v}^{-1}(p_{PRE}), \forall v \notin S ) \\ & = \Pr ( Z_v \geq \Phi^{-1}(p_{PRE}), \forall v \in S \quad \text{and} \quad Z_v < \Phi^{-1}(p_{PRE}), \forall v \notin S ), \end{aligned} \quad (9.18)$$

where  $\Phi^{-1}$  is the inverse cumulative function of a standard Normal distribution.

As a Gaussian random field,  $Z_v$  is also representable by a multivariate Gaussian distribution, we have:

$$\begin{bmatrix} Z_{(1)} \\ \vdots \\ Z_{(e)} \end{bmatrix} \sim \mathcal{N}_e \left( \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{(1,2)}^2 & \cdots & \sigma_{(1,e)}^2 \\ \sigma_{(1,2)}^2 & & & \\ \vdots & \ddots & & \\ \sigma_{(1,e)}^2 & \cdots & & 1 \end{bmatrix} \right). \quad (9.19)$$

where  $\mathcal{N}_e$  is a multivariate Gaussian distribution of dimension  $e$  and  $\sigma_{(i,j)}^2 = \text{cov}(Z_{(i)}, Z_{(j)})$ . The problem now lies in the estimation of the covariance between each pair of voxel  $\sigma_{(i,j)}^2, i \neq j$ . To this aim, we use the great amount of work developed in the context of random field theory. In particular, assuming that  $Z_v$  can be described as a stationary homogeneous Gaussian random field, [Worsley 1996] proposed an estimation of the spatial correlation. This estimate corresponds to the full-width-at-half-maximum of a Gaussian kernel that, applied to white noise, would lead to the same amount of autocorrelation. We computed this estimate using SPM8 (Statistical Parametric Mapping 8) on a subset of our subjects and found a full-width-at-half-maximum approximately equal to 1.5 voxels (in each direction). This value is in line with previous estimates obtained in ASL data [Wang 2003a]. Given this estimate of the autocorrelation,  $\sigma_{(i,j)}^2, i \neq j$  is easily computed.

The region-based probability taking into account the correlation between residuals  $\pi_{(r)}^*$  of the non-independent *a contrario* approach can therefore be calculated as the sum of joint probabilities calculated from a multivariate Gaussian distribution:

$$\pi_{(r)}^* = \sum_{i=l(r)}^e \sum_{S \in C_{(r)}^{(i)}} \Pr ( Z_v \geq \Phi^{-1}(p_{PRE}), \forall v \in S \quad \text{and} \quad Z_v < \Phi^{-1}(p_{PRE}), \forall v \notin S ) \quad (9.20)$$

In the context of this paper, the probability associated to each region  $R$  is affected to its centre voxel  $v$  to produce a voxel-wise probability map:

$$\pi_v^{*CONT} = \pi_{(r)}^*. \quad (9.21)$$

The standard *a contrario* approach assuming independence of the residuals can be seen as a special case where  $\sigma_{(i,j)} = 0 \quad \forall (i,j) \in (1 \dots e)^2, i \neq j$ .

### 3 Material

This section starts with a presentation of the datasets under study in 3.1. The metrics used for validation purposes are then described in 3.2.

### 3.1 Data

#### 3.1.1 Data acquisition

25 patients diagnosed with brain tumours and 61 healthy volunteers were involved in this study. One control subject and three patients were excluded because of strong borderzone sign [Zaharchuk 2009]. The final dataset therefore included 22 patients (13 males, 9 females, age:  $55.3 \pm 15.4$  years) and 60 healthy volunteers (28 males, 32 females, age:  $29.4 \pm 7.6$  years).

Data acquisition was performed on a 3T Siemens Verio MR scanner with a 32-channel head-coil. Patients were scanned in the context of clinical practice. The imaging protocol included a 3D T1-weighted anatomical sequence (T1w) (TR: 1900 ms, TE: 2.27 ms, FOV: 256 mm  $\times$  256 mm  $\times$  176 mm, flip angle: 9°, resolution: 1 mm  $\times$  1 mm  $\times$  1 mm), a PICORE Q2TIPS sequence [Wong 1998] with crusher gradients (TR: 3000 ms, TE: 18 ms, FOV: 192 mm  $\times$  192 mm, flip angle: 90°, resolution: 3 mm  $\times$  3 mm, slice thickness: 7 mm, inter-slice gap: 0.7 mm, TI: 1700 ms,  $TI_{wd}$ : 700 ms, 60 repetitions (35 subjects) or 30 repetitions (25 subjects), mSENSE parallel imaging with accelerating factor of 2). In addition to these sequences, the patients also underwent a 3D T1w post gadolinium (T1w-Gd) sequence (TR: 1900 ms, TE: 2.27 ms, flip angle: 9°, FOV: 250 mm  $\times$  250 mm  $\times$  176 mm, resolution: 1 mm  $\times$  1 mm  $\times$  1 mm) and a 2D T2w FLAIR sequence (TR: 9000 ms, TE: 90 ms, FOV: 220 mm  $\times$  199.4 mm, flip angle: 150°, resolution: 0.69 mm  $\times$  0.69 mm, slice thickness: 4 mm). Out of the 22 patients, 17 subjects also underwent a DSC sequence (GRE EPI, TR: 1500 ms, TE: 30 ms, FOV: 230 mm  $\times$  230 mm, flip angle: 90°, in plane resolution: 1.8 mm  $\times$  1.8 mm, slice thickness: 4 mm, inter-slice gap: 1.2 mm).

#### 3.1.2 Pre-processing

**Arterial Spin Labelling** Image pre-processing was performed using SPM8 (Statistical Parametric Mapping 8, Wellcome Department of Imaging Neuroscience, University College, London) in Matlab R2012a (Mathworks, Natick, MA). The anatomical image of each subject was segmented using the unified segmentation [Ashburner 2005]. A subject-specific anatomical brain mask was created, excluding voxels with less than 50% of brain tissue in subsequent statistical analyses. A six-parameter rigid-body registration of the ASL volumes was carried out in order to reduce undesired effects due to subject motion. Pair-wise subtraction of the control and labelled scans was then computed. Rigid coregistration onto the whole brain anatomical map was then performed based on mutual information. A standard kinetic model [Buxton 1998] was then applied in order to get ASL CBF, according to the following:

$$f = 6000 \times \frac{\lambda \Delta M}{2 M_0 \alpha TI_{wd} \exp^{-(TI + id_{x_{sl}} * TI_{sl}) / T1_b}} \quad (9.22)$$

where  $f$  is the CBF map in  $\text{mL} \cdot 100\text{g}^{-1} \cdot \text{min}^{-1}$ ,  $M_0$  the acquired  $M_0$  map,  $\lambda = 0.9 \text{ mL} \cdot \text{g}^{-1}$  the blood/tissue water partition coefficient,  $\alpha = 0.95$  measures

the labelling efficiency,  $\Delta M$  is the perfusion-weighted map,  $TI = 1.7$  s the inversion time [Ferré 2012],  $idx_{sl}$  the slice index (0 for the first slice),  $TI_{sl} = 0.045$  s the readout time for one slice,  $TI_{wd} = 0.7$  s the temporal width of the bolus,  $T1_b = 1.5$  s the T1 of blood [Wang 2011]. The 6000 factor allows the conversion from  $\text{mL.g}^{-1}.\text{s}^{-1}$  to  $\text{mL.100g}^{-1}.\text{min}^{-1}$  which is the standard unit for CBF. We bring the attention of the reader to the fact that, contrary to what is usually done in ASL pre-processing,  $\Delta M$  represents the set of perfusion-weighted maps (one volume per repetition) instead of a single perfusion-weighted map obtained by averaging across the repetitions. This is necessary in order to allow for the measurement of the within-subject variance.

Spatial normalization parameters estimated during the segmentation step were then applied to the T1 and ASL CBF map in order to normalize the subjects into the ICBM-452 T1 template space [Mazziotta 2001]. This registration algorithm was selected since it gives good results even in the presence of large anatomical lesions [Crinion 2007a]. Normalisation in intensity was then applied to each ASL CBF map in order to reduce the inter-subject variability [Aslan 2010]. The normalisation parameter was calculated as the average CBF in grey matter similarly to [Petr 2013].

**Dynamic Susceptibility weighted Contrast imaging** The DSC images were processed using MR manufacturer software by manually choosing an arterial input function to calculate: CBF, cerebral blood volume and mean transit time maps. The method is based on a deconvolution algorithm as described in [Ostergaard 1996]. Similarly to ASL, DSC CBF maps were co-registered on anatomical maps and spatially normalized.

### 3.2 Ground truth

Quantitative assessment of the detections and comparison between detection methods are challenging tasks. This is mainly because, like in many other medical imaging problems, the ground truth is not clearly stated. We chose to evaluate this framework on patients diagnosed with tumour pathology because perfusion abnormalities have been widely studied in this context. Also, as perfusion is a useful clinical information, DSC is part of the clinical routine for these patients. Based on clinical knowledge, we used a semi-automatic procedure that took advantage of the complementary anatomical (T1w-Gd, T2w FLAIR) and perfusion (DSC) information to get an estimation of the ground truth.

Due to its low SNR, ASL is not very well suited to measure low level of perfusion [Wintermark 2005]. That is why we focused on hyper-perfusions for sensitivity estimation. Hypo-perfusions were nevertheless retained for specificity calculation. Fig. 9.1 displays an example of ground truth for a representative subject.

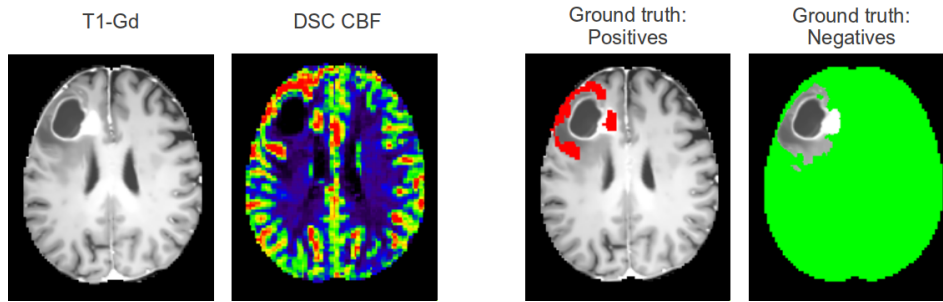


Figure 9.1: Ground truth for a representative subject: T1w-Gd and DSC CBF, Ground truth: hyper-perfusions (i.e. Ground truth positives) in red and iso-perfusions (i.e. Ground truth negatives) in green overlaid on the T1w-Gd.

### 3.2.1 Positives

In order to get an estimation of the positives (hyper-perfusions), we implemented a method inspired by the hotspot technique [Noguchi 2008] commonly used in clinical practice. First, the affected tissue (tumour and oedema) was segmented using a semi-automated method based on the T2w and T1w-Gd images and visually inspected by an expert neuro-radiologist. Second, we compared the segmented regions to their contralateral counterparts in the DSC CBF map. Voxels overtaking the upper decile were identified as potential hyper-perfusions. Each potential perfusion abnormality was then visually inspected by an expert neuro-radiologist and manually corrected if needed. Special care was taken in order to avoid inclusion of hyper-perfusions related to the presence of arteries. Out of the 17 patients included in this study who underwent a DSC sequence, 9 presented hyper-perfusions.

### 3.2.2 Negatives

According to clinical knowledge, in the absence of metastasis, the perfusion abnormalities should be confined to the affected tissue (tumour and oedema) identifiable on T1w-Gd and T2w. Therefore, the non-affected tissue (brain voxels not segmented as part of the tumour or oedema) was considered as iso-perfusion (negatives).

Here, all the 22 patients were included, even those without DSC images.

In the control group, in which no detections were expected, an additional estimate of the specificity was calculated by leave-one-out cross-validation.

## 4 Results

In 4.1, we start with a comparison of the quality of the detections obtained by our *a contrario* approach and by the General Linear model. We then focus on the validation of our *a contrario* statistical model by checking the distribution of the p-values under the null hypothesis in 4.2.

## 4.1 Comparison of the *a contrario* approach with the General Linear Model

In this section, we measure the performance in terms of sensitivity and specificity of the proposed *a contrario* approach and the massively univariate GLM. First, in 4.1.1, the ROC curves corresponding to each individual patient are computed and compared. Second, in 4.1.2, the average ROC curve is computed for each method. And finally a qualitative comparison is presented in 4.1.3, in order to illustrate the typical differences between both approaches.

### 4.1.1 Individual ROC curves

We first focused on the nine patients presenting hyper-perfusions.

To plot the ROC curve for a given subject  $s$ , the true positive rate,  $TPR_{(s,p_{unc})}$ , and the false positive rate,  $FPR_{(s,p_{unc})}$ , are estimated for a range of uncorrected p-values  $p_{unc}$  between 0 and 1, using the ground truth defined as described in section 3.2.

Here, it is interesting to note that, since we focus on the complete set of p-values, thresholding the probability map in *a contrario* is equivalent to thresholding the rare event count map (eq. (9.11)) and is therefore independent of the method selected to compute the region-based probabilities (standard or non-independent *a contrario*). For ease of calculation, the ROC curves were estimated using the standard *a contrario* approach.

We studied seven smoothing kernels for the GLM (with full-width-at-half-maximum  $w \in \{0, 2, 4, 6, 8, 10, 12\}$  mm<sup>3</sup>) and six parameter sets for the *a contrario* approach (the pre-defined p-value  $p_{PRE} \in \{0.01, 0.005, 0.001\}$  and the sphere radius  $rd \in \{1, 2\}$ ). The area under the curve was estimated for false positive rates ranging from 0 to 10%, as this is the area of interest in detection analysis [Skudlarski 1999].

**Best parameters** For each subject, the parameter set (pair  $(p_{PRE}, rd)$  in *a contrario* or  $w$  in GLM) leading to the highest area under the ROC was identified and used to plot the corresponding “best” ROC curve in fig. 9.2.

Visually, it is clear that the proposed *a contrario* approach outperforms the GLM for all subjects. A paired two-sample  $t$ -test comparing the *a contrario* and the GLM with the best parameter set for each subject outlines a significant difference ( $p = 0.004$ ).

**Best parameter set across all subjects** As statistical analyses are usually performed using a single set of parameters across all subjects (and not adjusting for the best parameter set independently for each subject as performed in the previous paragraph), we also compared the *a contrario* and GLM methods across all subjects given a fixed parameter set.

Table 9.1 presents the area under the ROC curves (average over the 9 patients) for false positives rates ranging from 0 to 10%, for the *a contrario* approach and the



GLM. The best area under the curve is obtained for the *a contrario* approach with  $p_{PRE} = 0.001$  and  $rd = 2$  and for the GLM with  $w = 8 \text{ mm}^3$ . A paired two-sample  $t$ -test between the best configurations outlines the superiority of the *a contrario* approach ( $p < 0.001$ ).

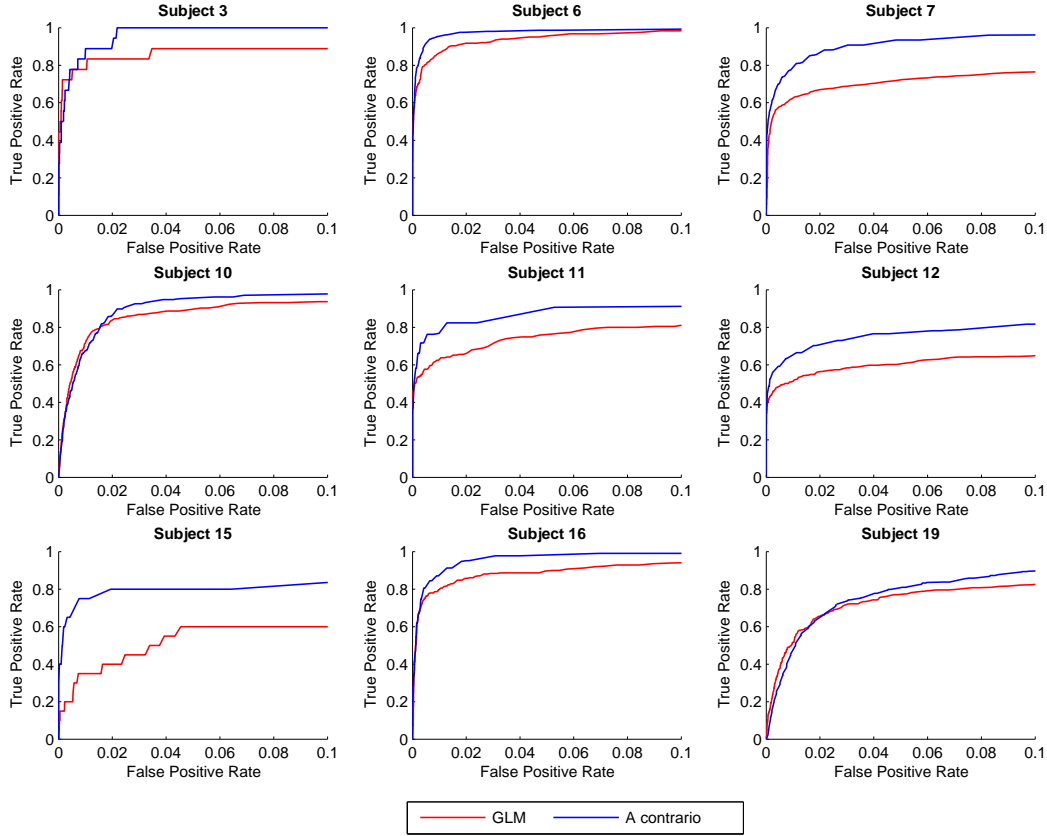


Figure 9.2: Best ROC curves for false positive rates ranging between 0% and 10%, for the massively univariate GLM (red) and the proposed *a contrario* approach (blue) in nine patients. Overall, the *a contrario* approach outperforms the GLM.

Table 9.1: Area under the ROC curve, averaged over the 9 subjects presenting hyper-perfusions, for false positive rates ranging between 0% and 10%, with the *a contrario* approach (with a sphere radius  $rd \in \{1, 2\}$  and a pre-defined p-value  $p_{PRE} \in \{0.01, 0.005, 0.001\}$ ) and the GLM (smoothed with a Gaussian kernel of full-width-at-half-maximum  $w \in \{0, 2, 4, 6, 8, 10, 12\} \text{ mm}^3$ ). The proposed *a contrario* approach outperforms all tested configurations of the GLM.

	<i>a contrario</i>						GLM						
	$rd = 1$			$rd = 2$			$w = 0$	$w = 2$	$w = 4$	$w = 6$	$w = 8$	$w = 10$	$w = 12$
$p_{PRE}$	0.01	0.005	0.001	0.01	0.005	0.001							
ROC area	0.68	0.73	0.78	0.76	0.81	<b>0.87</b>	0.63	0.68	0.71	0.72	<b>0.72</b>	0.69	0.64

### 4.1.2 Group ROC curves

In order to combine the data of all the patients included in this analysis, we computed an average ROC curve (across subjects). Without loss of generality, we assume that patient 1 to 9 present hyper-perfusions. The average true positive rate was then estimated by:

$$TPR_{G(p_{unc})} = \frac{1}{9} \sum_{s=1}^9 TPR_{(s,p_{unc})} \quad (9.23)$$

On the other hand, the group false positive rate was estimated across all subjects by:

$$FPR_{G(p_{unc})} = \frac{1}{22} \sum_{s=1}^{22} FPR_{(s,p_{unc})} \quad (9.24)$$

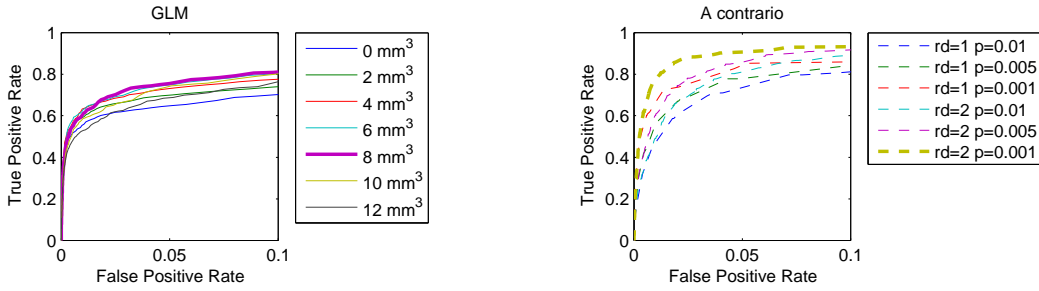


Figure 9.3: Average ROC curve (based on 22 subjects), for false positive rates ranging between 0% and 10%, with the GLM (smoothed with a Gaussian kernel of full-width-at-half-maximum  $w \in \{0, 2, 4, 6, 8, 10, 12\}$  mm<sup>3</sup>) and the a contrario approach (with a sphere radius  $rd \in \{1, 2\}$  and a pre-defined p-value  $p_{PRE} \in \{0.01, 0.005, 0.001\}$ )

Fig. 9.3 plots the resulting average ROC curve for each parameter set for the *a contrario* approach and the GLM. The corresponding area under the curves are summarized in table 9.2. The results are very close to the one obtained on the set of 9 subjects in the previous section, outlining the fact the the false positive rate at a given uncorrected p-value is not varying much across subjects. As before, the *a contrario* approach outperforms the GLM.

### 4.1.3 Qualitative comparison

In order to illustrate the advantage of the *a contrario* approach over the GLM, we chose 3 representative subjects and compared the methods at fixed false positive rate and true positive rate. To this aim, we selected the uncorrected p-values that would lead to a pre-specified false positive rate (respectively true positive rate) from the ROC analysis. We worked with the best set of parameters identified for each subject. Fig. 9.4 presents the detections obtained with both methods on the 3 selected patients. The first subject, which presents a small hyper-perfusion, is

Table 9.2: Area under the average ROC curve (based on 22 subjects), for false positive rates ranging between 0% and 10%, with the *a contrario* approach (with a sphere radius  $rd \in \{1, 2\}$  and a pre-defined p-value  $p_{PRE} \in \{0.01, 0.005, 0.001\}$ ) and the GLM (smoothed with a Gaussian kernel of full-width-at-half-maximum  $w \in \{0, 2, 4, 6, 8, 10, 12\}$  mm<sup>3</sup>). The proposed *a contrario* approach outperforms the GLM.

	<i>a contrario</i>						GLM						
	$rd = 1$			$rd = 2$			$w = 0$	$w = 2$	$w = 4$	$w = 6$	$w = 8$	$w = 10$	$w = 12$
$p_{PRE}$	0.01	0.005	0.001	0.01	0.005	0.001							
ROC area	0.68	0.73	0.79	0.75	0.80	<b>0.87</b>	0.63	0.67	0.70	0.72	<b>0.73</b>	0.70	0.65

studied at a true positive rate of 50%. In the two remaining patients, the methods are compared at a false positive rate of 5%.

Patient 15 suffers from a gliosarcoma in the left hemisphere near the parahippocampal region. The lesion displays a small hyper-perfusion in its dorsal part, small hypo-perfusions are seen in the surrounding oedema. At a true positive rate of 50%, the GLM displays a larger number of false positives than the *a contrario* approach.

Patient 7 presents a meningioma of both occipital lobes characterized by a large hyper-perfusion. At a false positive rate of 5%, most of the hyper-perfusion was properly detected by both methods. The *a contrario* approach is however more sensitive and detect a larger proportion of the hyper-perfusion.

Patient 11 suffers from a high grade tumour in the left temporal lobe. A hyper-perfused ring is observed, while the central necrosis of the tumour appears hypo-perfused. At a false positive rate of 5%, most of the hyper-perfused ring was properly detected by both the *a contrario* approach and the GLM. However, the GLM detection does not cover the most anterior part of the ring.

These 3 cases illustrate the reduced sensitivity of the GLM by comparison to the *a contrario* approach at the same specificity level.

## 4.2 Distribution of the p-values under the null hypothesis

We now focus on the validation of our *a contrario* statistical model. To this aim, we checked the distribution of the p-values under the null hypothesis. In the control group, no hypo- or hyper-perfusion is expected so that all the voxels are supposed to respect the null hypothesis.

We studied seven smoothing kernels for the GLM ( $w \in \{0, 2, 4, 6, 8, 10, 12\}$  mm<sup>3</sup>). The procedure to compute  $\pi_{(r)}^*$  in (9.20) is computationally expensive and we therefore restricted our study to the case  $rd = 1$ , for this section, focusing on three parameter sets for the *a contrario* approach ( $p_{PRE} \in [0.01; 0.005; 0.001]$ ,  $rd = 1$ ).

Fig. 9.5 plots the observed two-tailed probabilities against the expected probabilities for the standard and the non-independent *a contrario* approaches along

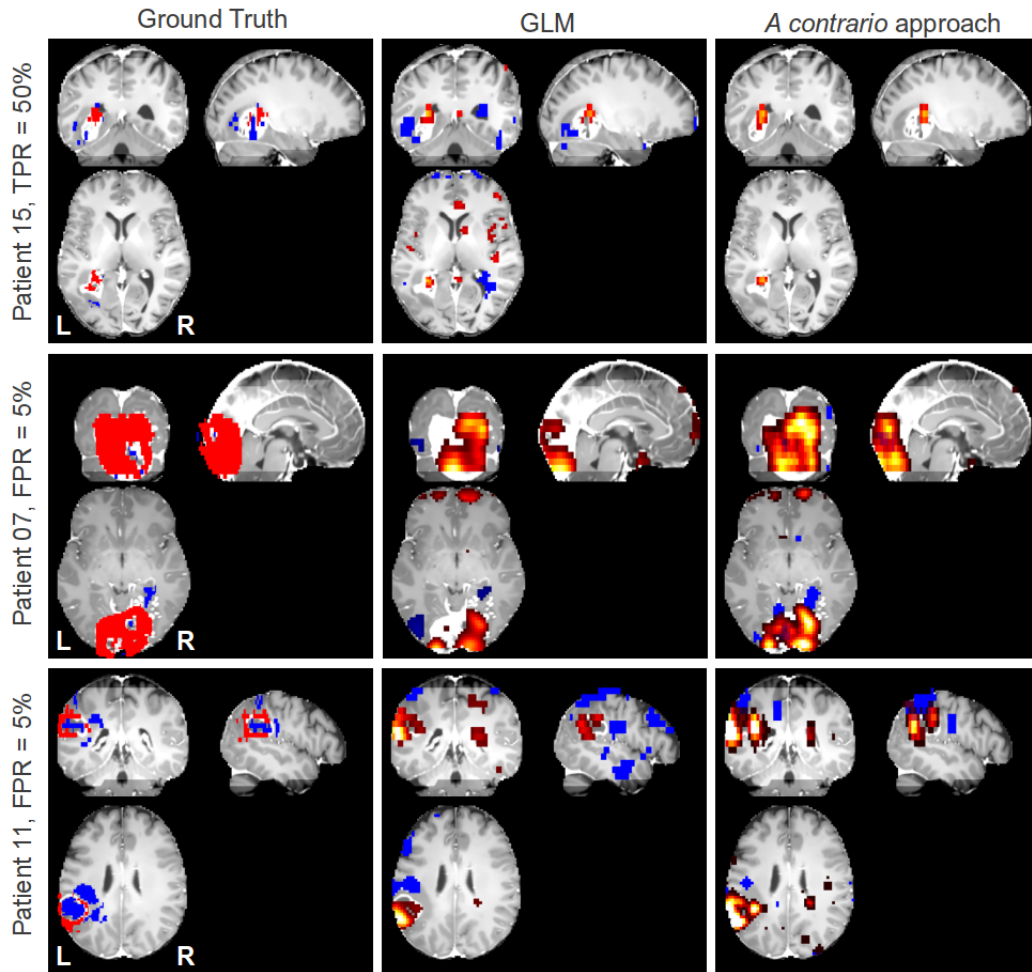


Figure 9.4: Perfusion abnormalities detections in 3 patients with the GLM and the *a contrario* approach using the best parameter set for each method. From left to right: ground truth hyper-perfusions (red) and hypo-perfusions (blue) overlaid on the T1w-Gd map; GLM, hyper-perfusions (red) and hypo-perfusions (blue); *a contrario* detections hyper-perfusions (red) and hypo-perfusions (blue).

with the GLM for comparison purposes. The expected distribution was estimated across subjects and voxels, leading to a total of about 15 000 x 60 samples. The 95% confidence interval was plotted as suggested in [Ge 2012].

The non-independent *a contrario* approach clearly outperforms the standard *a contrario* and leads to an observed distribution of the p-values under the null hypothesis very close to what is expected. The remaining difference might be accounted by potential mis-modelling due to mis-registrations.

The massively univariate GLM without smoothing is markedly anti-conservative and invalid for p-values smaller than  $10^{-3}$ . This behaviour is coherent with what was described in single-subject voxel-based morphometry [Scarpazza 2013, Viviani 2007a].

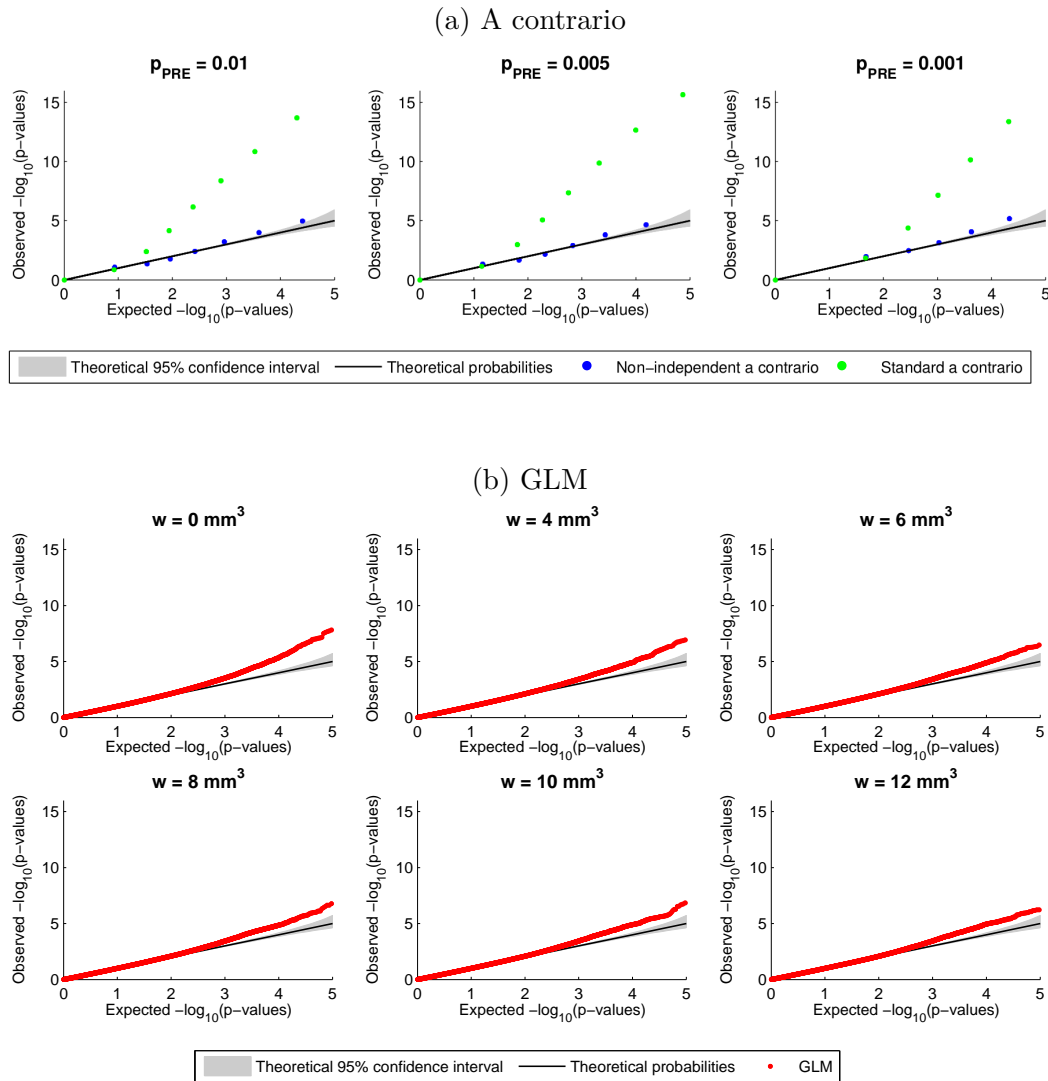


Figure 9.5: QQ-plot of the (a) region-based probabilities,  $\pi_{(r)}$  and  $\pi^*_{(r)}$ , in a *contrario* (for  $p_{PRE} \in [0.01; 0.005; 0.001]$ ), and (b) voxel-wise probabilities for the GLM  $\pi_v$  (for  $w \in \{0, 2, 4, 6, 8, 10, 12\} \text{ mm}^3$ ). The non-independent *a contrario* approach provides a distribution that is closer to the theory than the standard *a contrario* and the GLM.

Smoothing helps to provide a more valid estimate, the best models are obtained for smoothing of  $6 \text{ mm}^3$  and over.

While the standard *a contrario* approach clearly leads to invalid under-conservative statistics due to the spatial correlation of the residuals, the non-independent *a contrario* approach is better-behaved under the null hypothesis than the standard massively univariate GLM. As, in practice, statistical inference is performed using a single threshold (as opposed to ROC curves analyses), the validity of the approach is of tremendous importance. Indeed, validity ensures that, at a given threshold,

the false positive rate will not be arbitrarily inflated.

## 5 Discussion and conclusion

We have presented an *a contrario* approach for the detection of patient-specific perfusion abnormalities in ASL. Using ROC curves, we outlined that this locally multivariate procedure outperforms the massively univariate GLM, commonly employed in neuroimaging.

Furthermore, we introduced a non-independent version of this *a contrario* procedure in order to deal with the spatial autocorrelation of the noise observed in MRI. This non-independent *a contrario* approach corrected the inflation in false positive rate observed with the standard *a contrario*. Moreover, the distribution of the p-values under the null hypothesis of the non-independent *a contrario* is closer to the theoretical distribution than the massively univariate GLM.

We evaluated our approach in the context of brain tumours since perfusion patterns have been widely studied in this clinical context. This allowed us to get an estimate of the ground truth and perform a quantitative validation. Our approach is however suitable to a broader range of applications and we plan in the future to study pathologies presenting more subtle patterns of abnormal perfusion.



## Chapter 10

# *A contrario* detection of activation in single-subject BOLD fMRI

---

In this chapter, we propose to study the ability of the *a contrario* approach, previously adapted for basal perfusion abnormalities detection (cf. chapter 9), to detect areas of functional activity.

Part of this preliminary work was published in the 21st Annual Meeting & Exhibition of the International Society for Magnetic Resonance in Medicine (ISMRM) 2013 [Maumet 2013b].



Contents

<b>1</b>	<b>Introduction</b>	<b>183</b>
<b>2</b>	<b>Materials and methods</b>	<b>183</b>
2.1	Data	183
2.1.1	Acquisition	183
2.1.2	Pre-processing	183
2.2	Methods	184
2.2.1	An <i>a contrario</i> approach for activation detection	184
2.2.2	Validation	184
<b>3</b>	<b>Results</b>	<b>186</b>
3.1	Voxel-wise p-values under the null hypothesis	186
3.2	Comparison between the <i>a contrario</i> approach and the GLM	186
<b>4</b>	<b>Discussion and conclusion</b>	<b>190</b>

## 1 Introduction

BOLD fMRI is now a widespread imaging technique to study task-related activity in the brain. However, getting the areas of activation at the individual subject level is still an open issue. The standard massively univariate statistical analysis is usually performed after smoothing the data and makes use of a single p-value for final thresholding of the results [Gorgolewski 2012]. In group fMRI studies, the need for compensation of cross-subjects misregistrations clearly justifies the smoothing. However, at the individual level, where neat delineations of the activated areas are of interest, the use of Gaussian smoothing as a pre-processing step is more questionable.

## 2 Materials and methods

### 2.1 Data

#### 2.1.1 Acquisition

12 right-handed healthy volunteers were involved in this study (5 males, 7 females, age:  $28.6 \pm 2.7$  years) and scanned three times, with a one week period between two consecutive sessions. Data acquisition was performed on a 3T Siemens Verio MR scanner with a 32-channel head-coil. For each session, the imaging protocol included a 3D T1w MPRAGE (TR: 1900 ms, TE: 2.98 ms, FOV: 256 mm  $\times$  256 mm  $\times$  256 mm, flip angle: 9°, resolution: 1 mm  $\times$  1 mm  $\times$  1 mm) and a BOLD fMRI sequence (TR: 3000 ms, TE: 30 ms, FOV: 192 mm  $\times$  192 mm, flip angle: 90°, in plane resolution: 3 mm  $\times$  3 mm, 36 slices, slice thickness: 3.5 mm).

A block design paradigm made of 7 rest and 7 action blocks lasting 30 s each was implemented. During the action phases, the subject was asked to perform a maximum amplitude flexion-extension task of the right hand. For more details on the acquisition protocol and paradigm implementation, the interested reader is referred to the original paper presenting this imaging protocol [Raoult 2011].

#### 2.1.2 Pre-processing

MRI data were pre-processed with SPM8<sup>1</sup> (Statistical Parametric Mapping 8; Wellcome Department of Imaging Neuroscience, University College, London). Slice timing and motion correction were applied. Structural MRI were segmented using the unified segmentation [Ashburner 2005], and then normalized. fMRI data were registered on structural images and normalized. For the statistical analysis, both fMRI data smoothed using an isotropic 6 mm FWHM 3D Gaussian kernel and unsmoothed fMRI data was considered.

In view of the statistical analysis, at the subject level, a high-pass filter was applied to fMRI data to remove slow signal drifts due to undesired effects. In the

---

<sup>1</sup>[www.fil.ion.ucl.ac.uk](http://www.fil.ion.ucl.ac.uk)

GLM analyses, the haemodynamic response was modelled by the Informed Basis Set [Friston 1998] to account for possible delay and dispersion of the response from the canonical haemodynamic response function.

## 2.2 Methods

### 2.2.1 An *a contrario* approach for activation detection

In the *a contrario* approach, pre-smoothing of the data is avoided by taking into account the spatial neighbourhood into the statistical analysis. A brief overview of the technique is provided here, more details are available in the previous chapters (chapter 9, section 2.2 and chapter 3, section 3.4.2).

Starting from a voxel-wise probability map (obtained from a standard massively univariate GLM), a region-based probability map is built. To this aim, a grid of shapes is defined in the image of interest. In this study, we work with grids of spheres, as we have no a priori shape for the detections. The initial voxel-wise probability map is thresholded using a set of pre-defined p-values  $P = \{0.01, 0.005, 0.001\}$  and then, the number of over-threshold voxels is counted for each pre-defined p-value in each region (sphere). In the general *a contrario* formulation, the region-based probability is then estimated, assuming spatial independence of the residuals, using a binomial distribution parametrized by the number of voxels in each sphere and the pre-defined p-values.

To account for potential mismodelling, the null distribution used to define the voxel-wise probabilities was estimated with a non-parametric approach in the background of the image as illustrated in fig. 10.1. In the t-statistic map obtained from a standard GLM analysis, we estimated a non-parametric distribution using solely the out-of-brain voxels (as identified based on the segmentation of grey and white matter). Then, updated probabilities associated with the t-statistic of the in-brain voxel was computed using this empirical distribution. This can be seen as a recalibration of the p-values.

### 2.2.2 Validation

**Distribution of the voxel-wise p-values under the null hypothesis** For each subject, we identified the voxel that contained more than 99% of white matter. As we did not expect to observe activations in white matter, we considered that the null hypothesis (no activation) was verified in these voxels. For each subject, we were therefore able to compare the p-value distribution under the null hypothesis provided by the standard massively univariate GLM to the updated probabilities produced using the background information.

**Comparison between the *a contrario* approach and the GLM** In order to compare the performances of the standard massively univariate GLM with the proposed *a contrario* approach, we used ROC curves and estimated the area under

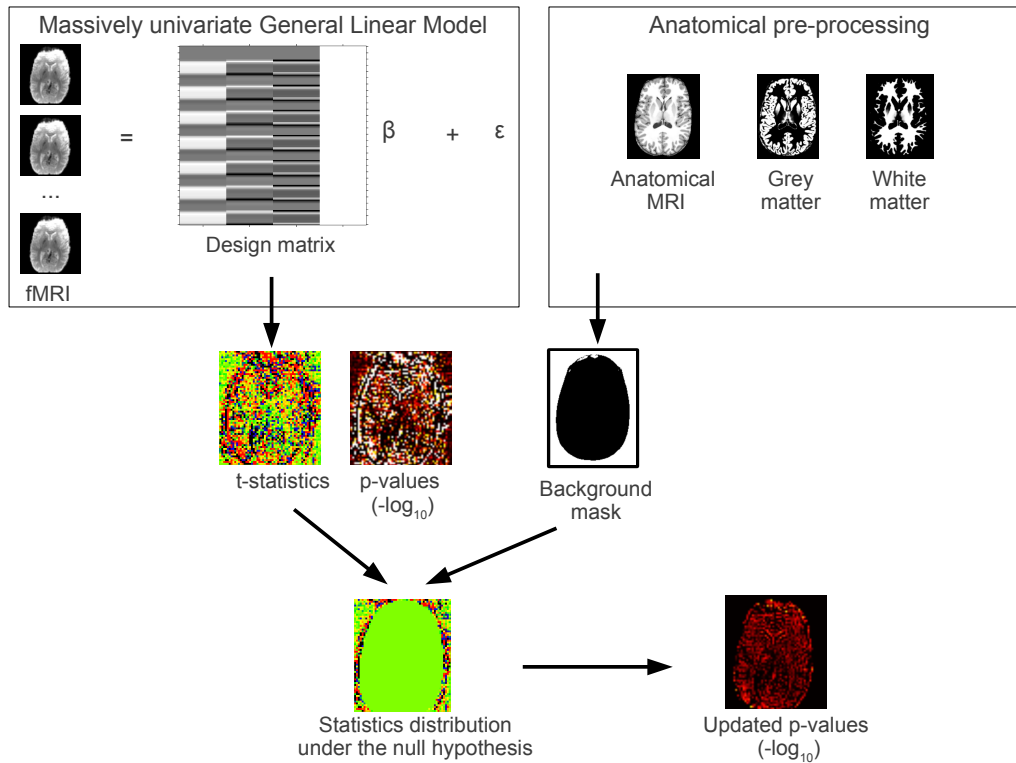


Figure 10.1: Subject-specific rescaling of the p-values using background information

the curve for false positive rates ranging from 0 to 0.1 (as only a small percentage of the voxels are expected to be active) [Skudlarski 1999].

Validation of fMRI analysis is still challenging due to the lack of ground truth. We focused on a well-studied right-hand motor paradigm in which the main activations are expected in the precentral and postcentral sulci (Left and Right) and in the Supplementary Motor Area (Left and Right) [Raoult 2011]. We studied two different “ground truths”.

First, as suggested in [Raoult 2011], we targeted the activation of the grey matter in the right hand motor area as manually delineated by an expert neuro-radiologist, while the surrounding white matter was expected to be inactive. This first ground truth benefited from a thorough definition, as it was precisely delineated by an expert. However it was only able to provide a measurement of the detection accuracy in a very local region.

Second, in order to capture more global information, we also tested a global criterion, in which activation in both the hand motor area and part of the Supplementary Motor Area (as defined by the AAL atlas) were activation targets, while whole-brain white matter (>70%) was expected to be inactive. As we did not expect an activation of the full Supplementary Motor Area, we considered that the detection was perfect in this region if at least 20% of the voxels were activated.

We compared the standard massively univariate GLM (no smoothing and smooth-

ing with a Gaussian FWHM kernel of  $6 \text{ mm}^3$ ) with the *a contrario* approach.

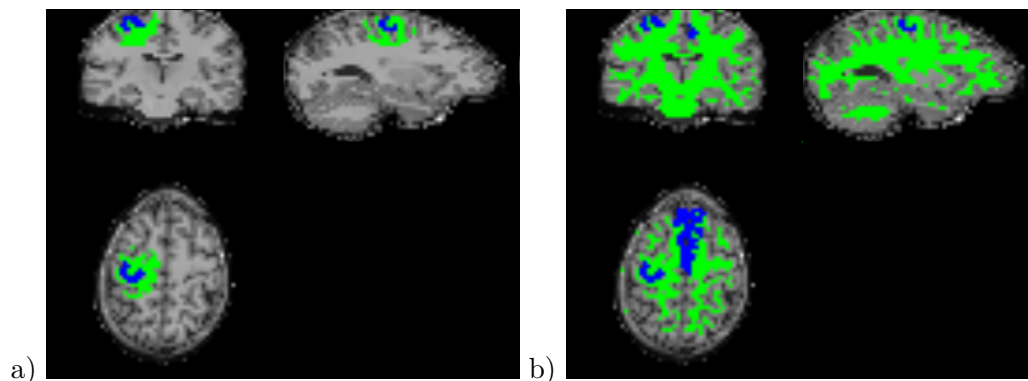


Figure 10.2: Local (a) and global (b) delineations of the Ground Truth for a representative subject: areas of iso-activity (no detection expected) are in green, areas of hyper-activity (i.e. activation) are in blue.

### 3 Results

#### 3.1 Voxel-wise p-values under the null hypothesis

Fig. 10.3 plots, for each subject, the observed probability against the expected probability. This kind of plot is sometimes referred as a Q-Q plot. In theory, the observed probability must be equal to the expected probabilities. A confidence interval can also be estimated (for more details on this procedure please refer to [Ge 2012]).

From fig. 10.3, it is very clear that the updated probabilities are always closer to the theoretical distribution than the probabilities extracted from the standard GLM. The anti-conservativeness of the GLM is striking. We guess that this might be related to a badly estimated temporal autocorrelation structure but further investigation would be needed in order to identify the precise cause of this invalidity. The proposed procedure, taking advantage of the background, while very simple, greatly improve the p-values and will therefore be used in the remainder of the chapter.

#### 3.2 Comparison between the *a contrario* approach and the GLM

Fig. 10.4 presents the ROC curves of each subject. The cross-session false positive rates (respectively true positive rates) were estimated by averaging the false positive rates (respectively true positive rates) of the three sessions (Day 1, Day 2 and Day 3). Table 10.1 displays the mean area under the ROC curve (average across subjects) for each session and each ground truth. First, we can notice that the performances in terms of area under the curve are lower with the local ground truth than with the global criterion. We hypothesize that, due to the small extent of its the activation target (less than 40 voxels per subject), the sensitivity calculated

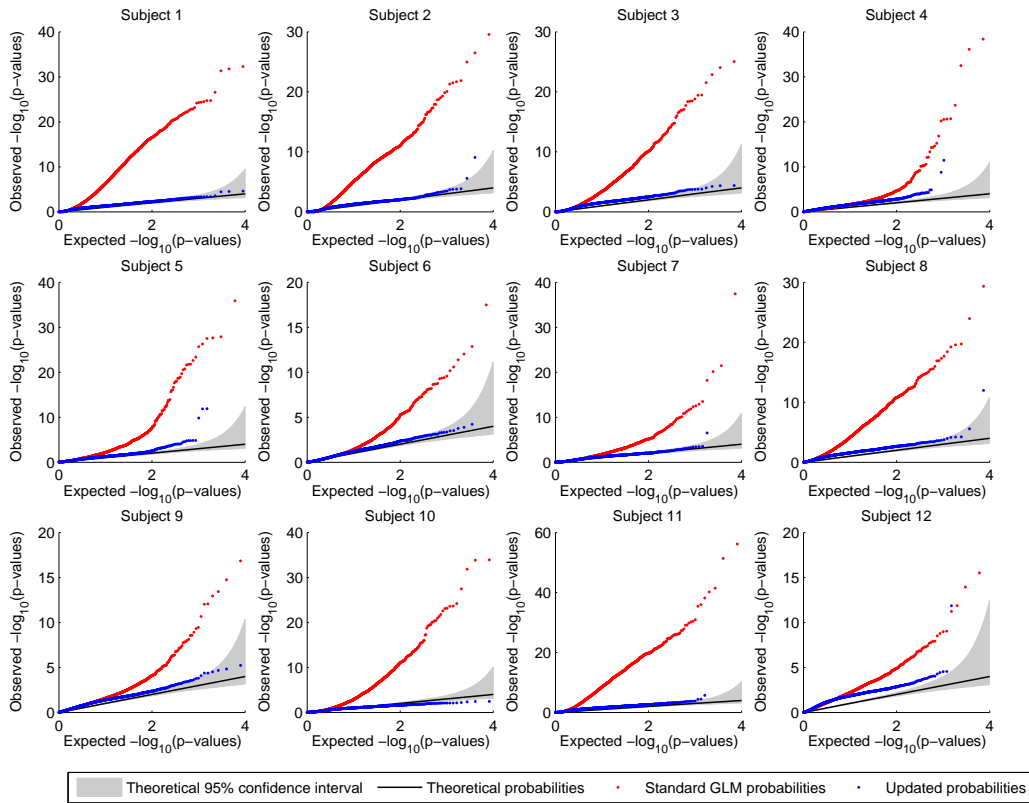


Figure 10.3: Subject Q-Q plots of the probabilities provided by the standard General Linear Model and of the updated probabilities (with background information). The theoretical distribution is displayed by a black line along with its 95% confidence interval in grey. The standard GLM probabilities are markedly invalid. The updated probabilities always provide more valid probabilities.

using the local ground truth might be more sensitive to mis-registrations between the anatomical image (on which the ROI was drawn) and the fMRI series.

Table 10.1: Average area under the ROC curves for Day 1, Day 2 and Day 3, using the local or the global ground truth.

	Local			Global		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
<i>a contrario</i>	0.54	0.50	0.50	0.75	0.68	0.73
GLM unsmoothed	0.43	0.46	0.44	0.59	0.62	0.63
GLM smoothed	0.45	0.41	0.37	0.66	0.60	0.60

Fig. 10.5 displays the local and global ROC curves averaged over sessions and subjects. Overall the *a contrario* approach outperformed both the unsmoothed and smoothed massively univariate GLM in terms of area under the curve. Paired

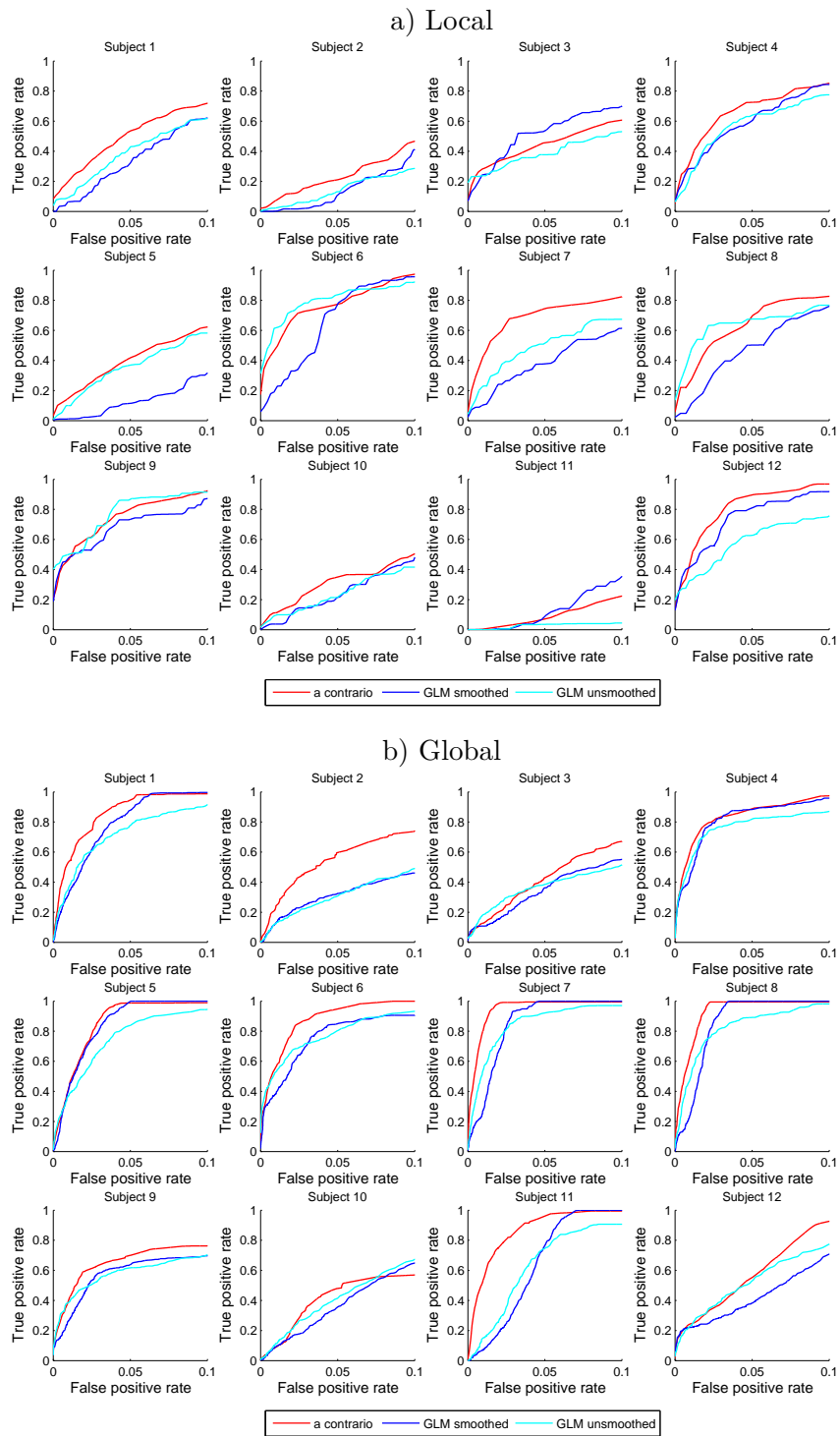


Figure 10.4: ROC curves for each subject averaged over the 3 sessions.

two sample t-tests were performed in order to detect significant improvement at

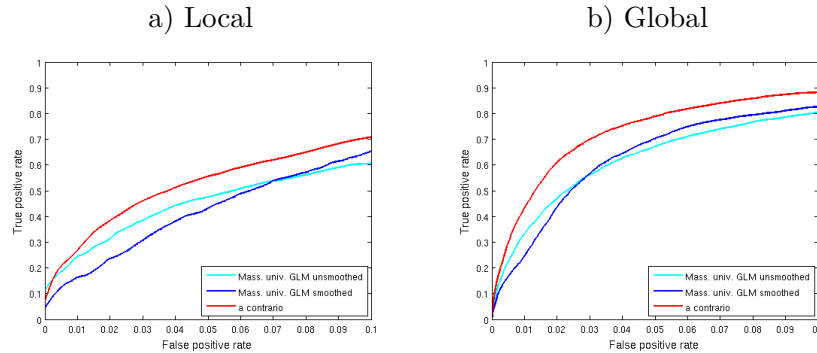


Figure 10.5: Group ROC curves averaged over the 3 sessions for the massively univariate GLM (unsmoothed and smoothed) and the *a contrario* approach based on the local (a) and global (b) ground truths.

the group level (false discovery rate  $q < 0.05$ ). The *a contrario* approach with local ground truth was significantly better than the smoothed GLM for all sessions ( $p = 0.023$ ,  $p = 0.030$ ,  $p = 0.012$ ) and than the unsmoothed GLM for 1 session out of 3 ( $p = 0.002$ ). Moreover, using the global ground truth, the *a contrario* approach performed significantly better than the smoothed GLM ( $p = 0.002$ ,  $p = 0.007$ ,  $p = 0.006$ ) and unsmoothed GLM ( $p = 0.001$ ,  $p = 0.026$ ,  $p = 0.003$ ) for all sessions.

In order to outline qualitatively the improvement that can be provided by the *a contrario* approach, we selected two subjects that would best illustrate the quantitative findings. In fig. 10.6, the detections in subject 2 with a false positive rate of 5% based on the global ground truth are provided. For the three methods (*a contrario*, smoothed GLM and unsmoothed GLM) the hand motor area is correctly detected. In the unsmoothed GLM there is a great amount of “salt and pepper” detections showing spatial incoherency. Regarding the hand motor area, the detections of the *a contrario* are very close of the smoothed GLM. Interesting differences are however observed when focusing on the Supplementary Motor Area. While, for each approach, a detection of at least a few voxels is located in the Supplementary Motor Area, only the *a contrario* approach provides a nicely delineated cluster in this area.

In fig. 10.7, the detections in subject 7 with a true positive rate of 75% based on the local ground truth are provided. For the three methods (*a contrario*, smoothed GLM and unsmoothed GLM) the hand motor area is correctly detected. With the unsmoothed GLM there are a couple of “salt and pepper” detections but the tendency is less pronounced than for subject 2, studied in fig. 10.6. Overall, the *a contrario* approach is more specifically located in grey matter than both the smoothed and the unsmoothed GLM. Though, for this display, the parameters were selected to get a true positive rate of 75% using the local ground truth, it is also worth noting that the supplementary motor area is correctly detected with the *a contrario* approach and the unsmoothed GLM but not with the smoothed GLM.



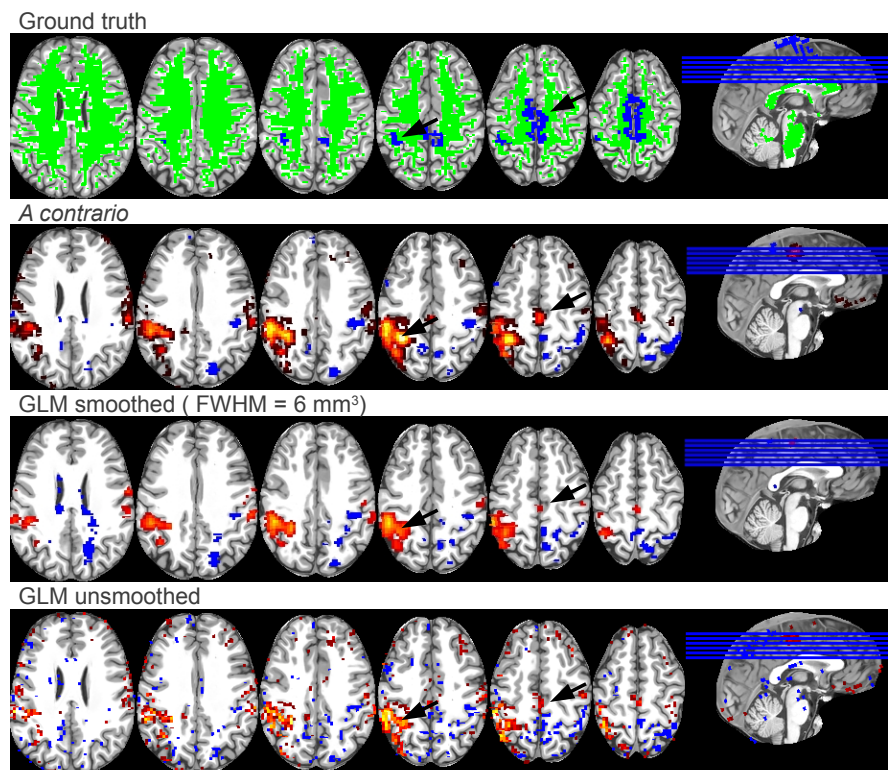


Figure 10.6: Global ground truth: example of detections in subject 2 with a false positive rate of 5%. The hand motor area (outlined by a black arrow on the fourth slice) is correctly detected by the three methods. The supplementary motor area (outlined by a black arrow on the fifth slice) is better detected with the *a contrario* approach.

## 4 Discussion and conclusion

In this preliminary work, we proposed a new *a contrario* approach to detect fMRI activations in single subjects. In order to “re-calibrate” the voxel-wise p-values obtained from the GLM, we used a very simple approach taking into account the background information. For validation purposes, and since no well defined ground truth is available in activation studies, we focused on a widely studied hand motor paradigm and investigated two different ground truths targeting local and global activation patterns.

The updated voxel-wise p-values were built based on an empirical distribution estimated in the image background (non-brain voxels). While this method is very simple, it yielded a clear improvement of the p-values distribution under the null-hypothesis. While, for consistency with previous studies, we decided to compare our *a contrario* approach to the standard GLM, an interesting extension of this work could investigate the effect of the p-values recalibration on the GLM.

The performances estimated with the local ground truth were lower than the

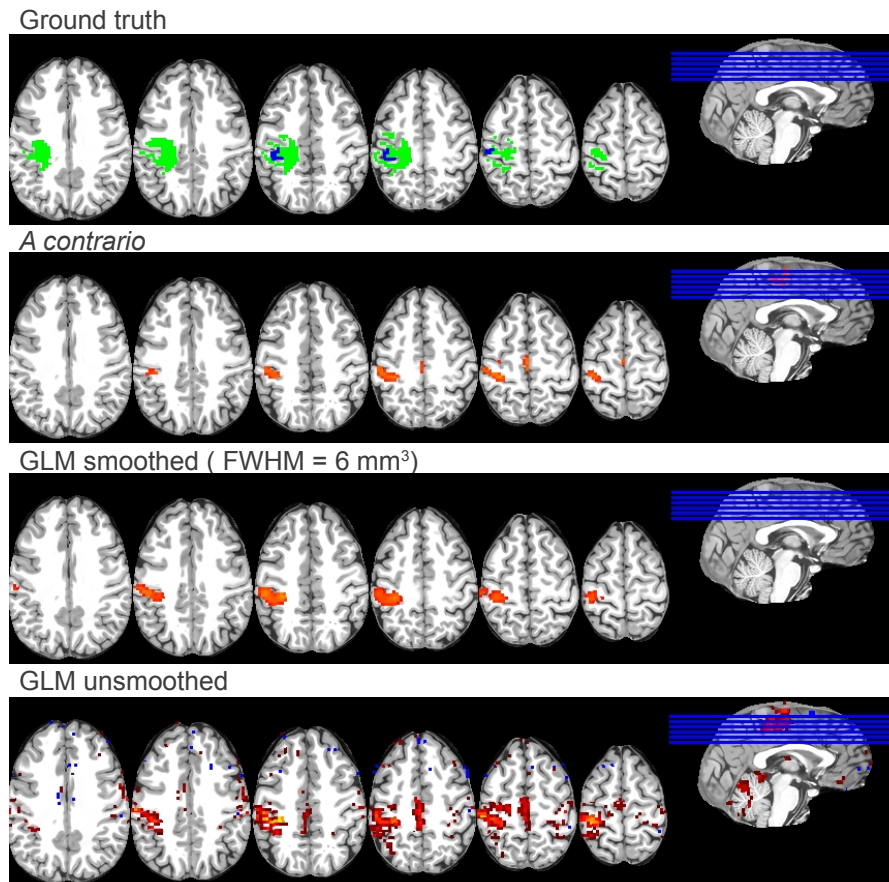


Figure 10.7: Local ground truth: example of detections in subject 7 with a true positive rate of 75%. The hand motor area is correctly detected by the three methods. There is a larger number of false positive detections with the *GLM* than the *a contrario* approach.

one obtained from the global ground truth. We hypothesize that this effect is a consequence of the small size of the local activation target, namely the grey matter located in the primary hand motor area. Indeed, small inaccuracies arising in the registration of the *fMRI* series onto the anatomical image can strongly impact the final estimate of the true positive rate.

Overall, the *a contrario* approach exhibited better spatially defined activations with a more interesting trade-off between sensitivity and specificity by comparison to the standard massively univariate *GLM*.



# Conclusion

---

In this thesis we have focused on the analysis of brain function in Magnetic Resonance Imaging. In conclusion, we will start with a review of our contributions and then discuss the limitations and perspective for future work.

## Contributions

Using standard group analyses based on a massively univariate general linear model, we outlined atypical patterns of activation and lateralisation in children diagnosed with typical specific language impairment by comparison to typically developing children (chapters 5 and 6). We demonstrated: a hypoactivation in the left hemisphere at the junction of the superior temporal gyrus and supramarginal gyrus; a hyperactivation in the right hemisphere in the anterior insula and the adjacent inferior frontal gyrus and the head of the caudate; and the absence of left lateralisation in the core language areas.

Then, three methodological contributions were provided in the context of patient-specific analyses.

First, we proposed to apply robust M-estimators in the computation of cerebral blood flow maps from a series of Arterial Spin Labelling acquisitions (chapter 7). On data corrupted by simulated perturbations, we illustrated the theoretical advantage of the proposed M-estimators by comparison to z-score thresholding or to the sample mean that is currently the standard approach. On the tested clinical cases, the superiority of M-estimates towards z-score thresholding was less pronounced but still present. Overall, both robust methods outperformed the sample mean on simulated and real data.

Second, we investigated the presence of heteroscedasticity in Arterial Spin Labelling data and its impact on patient-specific analyses comparing a single patient to a group of control subjects (chapter 8). Using metrics from the literature, we outlined the heteroscedasticity of Arterial Spin Labelling data. We then demonstrated an inflation in the false positive rate if the heteroscedasticity is not properly taken into account.

Third, we proposed an *a contrario* approach, which is a locally multivariate procedure focusing on the size of an excursion set in the neighbourhood of each voxel (chapter 9 and 10). In line with current development in the field of multivari-

ate neuroimaging, we outlined the superiority of this locally multivariate approach onto the massively univariate general linear model. This superiority was demonstrated in two different applications: the detection of patient-specific perfusion abnormalities and the detection of activation patterns in task-evoked analyses.

## Discussions and future work

While the first results provided on the atypical brain function in specific language impairment are promising, it would be of great interest to re-analyse these results in view of additional information provided by other sequences or modalities. A paper focusing on the anatomical data is currently in preparation. The diffusion data has also shown promising preliminary results. Also, it would be interesting to apply the *a contrario* analysis proposed for BOLD fMRI in the last chapter to this dataset.

Our use of robust statistics for the production of cerebral blood flow maps in arterial spin labelling, though validated quantitatively, was confined to qualitative analyses and not introduced in the next chapters focusing on quantitative detection of perfusion abnormalities. In our work on quantitative detections, we preferred to focus on the general linear model which is clearly the most widespread approach. Robust statistics could nevertheless be used in order to draw inference in patient-specific analyses.

Our experiments with Arterial Spin Labelling were performed with a PICORE Q2TIPS PASL sequence commercially available on our scanner. More recent ASL sequences, such as pseudo-continuous ASL, provide a better signal-to-noise ratio. It would be interesting to estimate and compare the extent of heteroscedasticity in a dataset acquired with a different sequence.

An interesting extension would be to perform a group *a contrario* approach and check whether the increase in sensitivity observed at the patient level is generalisable to the group.

# Glossary

---

**AAL atlas** Anatomic Automatic Labelling atlas. 81, 85, 95, 107, 108, 112, 113, 185, 202

**ANOVA** Analysis Of Variance. 85

**ASL** Arterial Spin Labelling. 1, 2, 7, 10, 16–26, 30, 31, 33, 35–37, 43–45, 47, 57, 58, 123, 125, 126, 129–137, 139–142, 146, 147, 149–152, 154–156, 158, 160, 161, 163–165, 169–171, 194, 199, 200, 203, 204, 208

**BAT** Bolus Arrival Time. 13, 19, 22–24

**BOLD** Blood-Oxygen-Level-Dependent. 1–3, 7, 29–31, 33, 35–37, 39, 41–44, 47, 183, 194, 199, 200

**CASL** Continuous Arterial Spin Labelling. 21, 23, 24, 199

**CBF** Cerebral Blood Flow. 13, 15–20, 22–25, 30, 31, 37, 44, 45, 47, 57, 123, 125, 126, 129–136, 138, 139, 142, 147, 149–152, 156, 163–165, 170–172, 198–200, 202–204, 208

**CBV** Cerebral Blood Volume. 13, 14, 16–19, 30, 57, 198, 199

**CMRO<sub>2</sub>** Cerebral Metabolic Rate of Oxygen. 30

**CT** Computed Tomography. 17, 18

**DSC** Dynamic Susceptibility weighted Contrast imaging. 1, 13, 14, 16–19, 123, 126, 128–131, 133–136, 138, 139, 146, 147, 149, 151, 163, 170–172, 198, 199, 203, 204, 208

**EEG** electro-encephalography. 29, 31, 199

**EPISTAR** EPI-Signal Targeting with Alternating Radio-frequency pulses. 21

**FAIR** Flow-sensitive Alternating Inversion Recovery. 21

**fASL** functional Arterial Spin Labelling. 30, 31, 36, 37, 41–43, 47, 199

- FDG** Fludeoxyglucose. 16, 17, 199
- FDR** False Discovery Rate. 53, 148, 155, 208
- FLAIR** Fluid Attenuated Inversion Recovery. 14, 76, 103, 147, 149, 198
- fMRI** Functional Magnetic Resonance Imaging. 1, 25, 30, 31, 35–37, 41–44, 47, 59, 60, 76, 77, 79, 80, 82, 84, 86, 87, 93, 94, 98, 102, 103, 107, 109, 110, 113–115, 119, 140, 142, 158, 160, 183, 185, 187, 190, 191, 194, 199–202, 207
- FOV** Field of View. 84, 106, 128, 149, 183
- FWE** Family Wise Error. 53, 84, 86, 87, 89, 107, 109, 110, 201, 202, 207
- FWHM** Full Width at Half Maximum. 38, 84, 107, 148, 155, 156, 159, 183, 186, 204, 208
- GLM** General Linear Model. 33, 34, 38–42, 47, 53, 58, 65, 84, 107, 137, 139–141, 158, 161, 163, 164, 166, 167, 173–179, 184–187, 189–191, 204, 205, 208
- HRF** Haemodynamic Response Function. 30, 42, 84, 200
- IFG** Inferior Frontal Gyrus. 74, 80, 81, 85–89, 91, 93–98, 102, 105, 107–117, 119, 201, 202, 207
- ITG** Inferior Temporal Gyrus. 85
- LI** Lateralization Index. 81, 85, 88, 89, 92–94, 103, 108–113, 119, 202
- MEG** magneto-encephalography. 29, 31, 199
- MNI** Montreal Neurological Institute. 86, 90, 109, 201, 202
- MR** Magnetic Resonance. 13, 18, 19, 21, 130, 149, 151, 154, 183
- MRI** Magnetic Resonance Imaging. 1–3, 7, 10, 18, 25, 29, 30, 33, 41, 46, 49, 53, 60, 69, 77, 81, 82, 84, 103, 104, 107, 123, 129, 139, 140, 163, 183, 199
- MTG** Middle Temporal Gyrus. 80, 85, 88, 89, 93, 96
- MTT** Mean Transit Time. 13, 18, 19, 198
- OLS** Ordinary Least Squares. 40, 44, 45, 48, 142, 144
- PASL** Pulsed Arterial Spin Labelling. 21–24, 194, 199
- PCASL** Pseudo-Continuous Arterial Spin Labelling. 21
- PCT** Perfusion Computed Tomography. 17–19

- PET** Positron Emission Tomography. 16–18, 29, 31, 139, 150, 163, 199
- PICORE** Proximal Inversion with a Control for Off-Resonance Effects. 21, 22, 128, 149, 194
- pSTG** posterior Superior Temporal Gyrus. 102, 108–110, 112, 113, 115, 117
- PT** Planum Temporale. 101, 102, 114, 115
- Q2TIPS** QUIPSS II with Thin-slice Seriodic Saturation. 22, 24, 128, 194
- QUIPSS II** Quantitative Imaging of Perfusion using a Single Subtraction. 24
- ROC** Receiver-Operating-Characteristics. 138, 148, 156–158, 164, 173–176, 178, 179, 184, 186, 187, 204, 205, 208
- ROI** Region of Interest. 76, 81, 85, 88, 89, 92–95, 103, 107, 108, 110–114, 150, 187, 201, 202
- SLI** Specific Language Impairment. 71, 75, 101, 102, 108, 113–117, 119
- SMG** Supramarginal Gyrus. 85, 88, 89, 91, 93–98, 107, 109, 110, 112–115, 117, 119, 202, 207
- SNR** Signal to Noise Ratio. 20, 21, 24, 26, 38, 62, 125, 139, 147, 154, 163, 171
- SPECT** Single Photon Emission Computed Tomography. 17, 18, 102, 114, 115, 139, 150, 163
- STG** Superior Temporal Gyrus. 74, 80, 85, 86, 88, 89, 91, 93–98, 105, 107, 108, 110, 112–114, 119, 201, 202
- STS** Superior Temporal Sulcus. 88, 89, 93, 95–98
- T-SLI** Typical Specific Language Impairment. 72, 76, 101, 103, 104, 108–119, 202
- T1w** T1 weighted. 76, 83, 103, 106, 128, 149, 155, 183, 204
- T1w-Gd** T1 weighted Gadolinium injected. 14–17, 147, 149, 156, 159, 172, 198, 199, 204
- T2\*w** T2\* weighted. 84, 106
- T2w** T2 weighted. 14, 76, 147, 149, 198
- TE** Echo time. 84, 106, 128, 149, 183
- TI** Inversion time. 23–25, 37, 128, 149
- TR** Repetition time. 37, 84, 106, 128, 149, 183



**TTP** Time To Peak. 13, 19, 198

**WLS** Weighted Least Squares. 40, 41, 44, 48, 51, 52

# List of Figures

---

1	Diagramme des troubles développementaux du langage. . . . .	xiii
2	Les quatre paradigmes langagiers inclus dans le protocole Dysphasie.	xiv
3	Hypo-activation du groupe dysphasique par rapport au groupe contrôle pour la tâche Définition ( $p < 0.05$ FWE). . . . .	xiv
4	Hyper-activation du groupe dysphasique par rapport au groupe contrôle pour la tâche Phon-diff ( $p < 0.05$ FWE). . . . .	xv
5	Moyenne : exemple d'estimation du débit sanguin cérébral en présence de valeurs atypiques. . . . .	xv
6	Seuillage par score $z$ : exemple d'estimation du débit sanguin cérébral en présence de valeurs atypiques. . . . .	xvi
7	M-estimateur de Huber : exemple d'estimation du débit sanguin cérébral en présence de valeurs atypiques. . . . .	xvi
8	Corrélation de la cartographie ASL CBF avec la carte DSC CBF estimée avec la moyenne, le M-estimateur de Huber ou la méthode des scores $z$ (a). Exemple de résultats sur données réelles (b), les flèches blanches soulignent des artefacts présents dans la carte obtenue avec moyennage qui sont correctement éliminé avec le score $z$ et le M-estimateur. . . . .	xvii
9	Exemple de détections individuelles d'anomalies de perfusion chez deux patients atteints de tumeurs cérébrales. De gauche à droite : T1-Gd, ASL CBF, détections. . . . .	xvii
10	La variance intra-sujet mesure la variabilité entre les répétitions. . .	xviii
11	La variance inter-sujets mesure la variabilité entre sujets d'un même groupe. . . . .	xviii
12	Vérité terrain chez un patient. De gauche à droite : T1-Gd, DSC CBF, Vérité terrain: hyper-perfusions (à détecter) en rouge et iso-perfusion (à ne pas détecter) en vert. . . . .	xix
13	Vue d'ensemble de l'approche <i>a contrario</i> : définition du modèle de fond, estimation d'une probabilité niveau région et correction des comparaisons multiples. . . . .	xxi

14	Anomalies de perfusion chez 2 patients avec le GLM et la méthode <i>a contrario</i> . De gauche à droite : vérité terrain hyper-perfusions (rouge) et hypo-perfusions (bleu) superposées à la carte T1w-Gd; GLM hyper-perfusions (couleurs chaudes) et hypo-perfusions (bleu); <i>a contrario</i> hyper-perfusions (couleurs chaudes) et hypo-perfusions (bleu). . . . .	xxiii
15	Exemple d'activations obtenues avec un taux de faux positifs de 5%. L'aire motrice primaire de la main (flèche noire sur la 4 <sup>ème</sup> coupe) est correctement détectée avec le GLM et l' <i>a contrario</i> . L'aire motrice supplémentaire (flèche noire sur la 5 <sup>ème</sup> coupe) est mieux détectée avec l'approche <i>a contrario</i> . . . . .	xxiv
1.1	Tag cloud computed from 8559 papers (title and abstract) referring to <i>brain function</i> indexed in Pubmed . . . . .	9
1.2	The three components of brain function according to [Vik 2005]. Illustration of connectivity reprinted from [Kaiser 2010] . . . . .	10
1.3	Main brain arteries, from [Gray 1918]. Lateral view displaying the carotid and vertebral arteries entering the skull on the right part of the brain (a). Ventral view of the artery ramifications at the basis of the brain (b). Close up on the circle of Willis: the carotid and vertebral arteries ramifies into the anterior, middle and posterior arteries (c). . . . .	11
1.4	Vascular territories of the middle (pink), posterior (yellow) and anterior (blue) brain arteries, from [Gray 1918]. Medial (a) and lateral (b) views. . . . .	11
1.5	Brain veins, from [Gray 1918]. Dorsal (a) and medial (b) views of the sinuses. The interior jugular vein brings the blood out of the skull (c). . . . .	12
1.6	Vascular tree and perfusion. Image courtesy of Isabelle Corouge. . . . .	12
1.7	Perfusion parameters estimated with DSC in a patient diagnosed with a brain tumour (not visible on the slice displayed): T1w, Cerebral Blood Flow (CBF), Cerebral Blood Volume (CBV), Mean Transit Time (MTT) and Time To Peak (TTP) maps. . . . .	13
1.8	Differentiation between 3 types of brain tumours: glioblastoma, lymphoma and low-grade glioma using the complementarity of T1w-Gd and perfusion imaging. T1w-Gd (A,D) or T2w FLAIR (G) anatomical map (first column) and DSC CBV map (second column). The malignant glioblastoma is associated with a contrast enhancement on the T1w-Gd map and a hyper-perfusion on the CBV map. The lymphoma is characterized by a contrast enhancement on the T1w-Gd but a low CBV ratio (tumour/normal). The low-grade glioma has also a low CBV ratio and has no enhancement on the T1w-Gd (T2w FLAIR displayed). Reprinted from [Ferré 2012], Copyright (2013), with permission from Elsevier. . . . .	14

- 
- 1.9 Ischaemic core and penumbra in stroke from [www.strokeforum.com/stroke-background/pathophysiology.html](http://www.strokeforum.com/stroke-background/pathophysiology.html) . . . . . 15
- 1.10 Differentiation between post-radiation necrosis (pseudo-progression) and tumour recurrence (true progression) using perfusion imaging. T1w-Gd anatomical map (first line) and DSC CBV map (second line). In pseudo-progression (first column) the presence of radionecrotic lesions is characterised by a low CBV ratio. In true progression (second column) neoangiogenesis is characterised by hyperperfusions (high CBV ratio). Reprinted from [Ferré 2012], Copyright (2013), with permission from Elsevier. . . . . 16
- 1.11 Reduced glucose metabolism and perfusion (CBF) of the temporal lobes in a patient suffering from semantic dementia as imaged by FDG-PET (Left) and ASL (Right). Reprinted from [Corouge 2012], with permission. . . . . 16
- 1.12 Hypoperfusions (in blue) and hyperperfusions (in red) overlaid on T1w-Gd in a patient suffering from drug-resistant epilepsy. A hypoperfusion is co-located with the dysplasia observed in the right temporal lobe of this patient. Reprinted from [Petr 2013], Copyright (2013), with permission from Wiley. . . . . 17
- 1.13 Schematic view of an ASL acquisition. . . . . 19
- 1.14 Building blocks of an ASL sequence: preparation of tissue magnetization, labelling and readout. . . . . 20
- 1.15 Pulsed and Continuous ASL. . . . . 21
- 1.16 Schematic view of the difference signal  $\Delta M$  observed in an ASL experiment. . . . . 22
- 1.17 Buxton's model for  $\Delta M$ , examples of  $c(t)$  and  $m(t)$  functions for PASL and CASL. . . . . 23
- 1.18 PhonoDif paradigm, as described in [de Guibert 2010]. This block paradigm is made of 3 alternating periods of rest and action. . . . . 26
- 1.19 Neurovascular coupling, Reprinted from [Arthurs 2002], Copyright (2013), with permission from Elsevier. . . . . 27
- 1.20 Three levels of measurements for brain activity and associated imaging techniques. Studying task-evoked activity can be performed with a measurement of the neuronal electrical activity (with EEG or MEG) or with a measurement of change in metabolism (with PET imaging) or haemodynamics (with MRI or PET). . . . . 29
- 1.21 The balloon model linking the stimulus time-course with the physiological parameters and the observed BOLD (for BOLD fMRI, circled in blue) and CBF (for fASL, circled in green) effects. Reprinted from [Buxton 2004], Copyright (2013), with permission from Elsevier. 30

2.1	An example of pre-processing pipeline for BOLD fMRI or ASL data including: slice timing, spatial realignment, registration to the anatomical space, an optional registration into a population space (for group studies) and finally smoothing. . . . .	35
2.2	Example of slice delays in a sequential acquisition. . . . .	36
2.3	Example of use of the GLM to model the time-course of a BOLD acquisition related to a block paradigm, reprinted from “The General Linear Model for fMRI analyses” by Klaas Enno Stephan, SPM 2013 Course. . . . .	39
2.4	Matrix representation of the GLM, reprinted from “The General Linear Model for fMRI analyses” by Klaas Enno Stephan, SPM 2013 Course . . . . .	40
2.5	Example of subject activations for a language paradigm. . . . .	42
2.6	Creation of a task-related regressor (one column of the design matrix $\mathbf{X}_{s(v)}$ ) by convolution of the paradigm time-course with the HRF. In this example, a block paradigm with 3 action and rest blocks is presented. . . . .	42
2.7	Example of design matrix, $\mathbf{X}_{s(v)}$ , for a subject analysis of task-evoked activation (computed with SPM8). The matrix presents a total of 8 regressors, from left to right: 1 task-related regressor modelling a block paradigm (from fig. 2.6), 6 nuisance motion covariates (3 translations and 3 rotations parameters estimated during the registration step in the pre-processing) and 1 constant regressor. . . . .	43
2.8	Example of CBF map for a control subject. . . . .	44
2.9	Example of group activations for a language paradigm. . . . .	46
2.10	Example of group perfusion map. . . . .	46
2.11	Example of task-evoked hyper-activations (hot colormap) and hypo-activations (winter colormap) in a group of children diagnosed with typical specific language impairment by comparison to typically developing children. . . . .	49
2.12	Example of design matrix, $\mathbf{X}_{G(v)}$ , for a between-group analysis (computed with SPM8). The matrix presents a total of 3 regressors, from left to right: 2 regressors modelling the group effects and 1 nuisance regressor (age of the subjects). . . . .	50
3.1	Concept of multivariate classification in two dimensions. Reprinted from [Klöppel 2012], Copyright (2013), with permission from Elsevier.	59
3.2	10-fold cross-validation. . . . .	60
3.3	Principle of a searchlight analysis (a, b) and examples of activations obtained with the searchlight approach (c) and the massively univariate GLM (d). Reprinted from [Kriegeskorte 2007a], Copyright (2013), with permission from Elsevier. . . . .	61

3.4	Overview of the <i>a contrario</i> approach: background model definition, region-based probability estimation and correction for multiple comparisons. . . . .	62
4.1	Developmental language disorders diagram. . . . .	71
4.2	Main brain areas involved in language, adapted from The Brain From Top To Bottom website . . . . .	72
4.3	Geschwind-Wernicke’s model of language production following an auditory stimulus, adapted from the The Brain From Top To Bottom website . . . . .	73
4.4	Variability of Broca’s (in the frontal lobe) and Wernicke’s (in the temporal and parietal lobes) areas from The Brain From Top To Bottom website . . . . .	73
4.5	Activation in one subject during a language paradigm (Phon-seg paradigm described hereafter). The language activation patterns observed in the IFG (Broca’s area) and STG (Wernicke’s area) are clearly left-lateralised. . . . .	74
4.6	The arcuate fasciculus: (a) historical view from The Brain From Top To Bottom website and (b) modern three-segment view, reprinted from [Catani 2005], Copyright (2013), with permission from Wiley. . . . .	74
4.7	The four language paradigms used in the “Neuroimaging in Dysphasia” project. . . . .	76
5.1	fMRI group effects for each language task. Results are displayed at $p < 0.05$ FWE for category and definition tasks, and at $p < 0.001$ uncorrected for phon-diff and phon-seg tasks. The functional maps are superimposed onto an individual brain normalized with respect to our customized paediatric template, with x-coordinates in MNI space. Left panels display the left hemisphere. . . . .	86
5.2	Task comparisons, conjunctions, and union ( $p < 0.001$ uncorrected). Functional maps are superimposed onto an individual brain normalized with respect to our customized paediatric template, with x-coordinates in MNI space. Left panels display the left hemisphere. . . . .	90
5.3	ROI analyses for each paradigm. Bar plots depict the effect size of activation (mean $\pm$ standard deviation) in each left and right ROI for each paradigm. Significant activation or deactivation ( $p < 0.05$ ) are indicated by a double star. In addition, significant interparadigm differences ( $p < 0.05$ ) are indicated by brackets. Oper = opercularis; Tri = triangularis; Rol = rolandic; Inf = inferior; Mid = middle; Sup = superior. . . . .	92

5.4	LIs and significant lateralisations within the ROIs for each paradigm. The box plots depict lateralisation for each ROI, with positive LIs (resp. negative) reflecting left (resp. right). Significant lateralisations ( $p < 0.05$ ) are indicated by bold lines. Significant interparadigm differences of LIs are indicated by a bracket. Oper = opercularis; Tri = triangularis; Rol = rolandic; Inf = inferior; Mid = middle; Sup = superior. . . . .	92
6.1	fMRI group effects for each language task ( $p < 0.05$ FWE). The functional maps are super-imposed onto an individual brain normalized with respect to our customized paediatric template, with x-coordinates in MNI space. Left slices are left hemisphere. . . . .	109
6.2	fMRI between-group comparisons for definition and phon-diff tasks ( $p < 0.05$ FWE). Blue and yellow colours indicate the hypo- and hyperactivations, respectively, for the T-SLI group compared with the control group. The 3D view focuses on the peak contrasts. Functional maps are superimposed on an individual brain normalized with respect to our customized paediatric template. Coordinates are in Montreal Neurological Institute space. . . . .	110
6.3	LIs, significant group lateralisations and between-group comparisons within the ROIs for each single language task. The box plots depict group lateralisation for each ROI, with positive and negative lateralisation indices reflecting left and right, respectively, and with significant left or right lateralisations ( $p < 0.05$ ) outlined by bold lines. The p-values of between-group comparisons are indicated, with significant between-group differences outlined by a square bracket with an asterisk. (a) P-value of the group factor from analysis of covariance with factors age and group. . . . .	111
6.4	LIs, significant group lateralisations and between-group comparisons using combined tasks analysis within single (right) and extended (left) ROIs. The box plots depict group lateralisation for each ROI, with positive and negative lateralisation indices reflecting left and right, respectively, and with significant left or right lateralisations ( $p < 0.05$ ) outlined by bold lines. The P-values of between-group comparisons are indicated, with significant between-group differences outlined by a square bracket with an asterisk. (a) P-value from two-sided two sample t-test. Temporoparietal language ROI = STG and SMG; frontal language ROI = IFG (opercularis and triangularis) and insula; language ROI = frontal language and temporoparietal language ROIs; non-language ROI = whole brain (i.e. all AAL atlas ROIs) except the language ROI. . . . .	112
7.1	Sensitivity of the sample mean to outliers. A single corrupted volume can induce strong artefact in the final CBF maps computed with the sample mean. . . . .	125
7.2	Volume exclusion via z-score thresholding as proposed in [Tan 2009].	126

7.3	Robust ASL CBF map via Huber's M-estimator. . . . .	126
7.4	Psy functions for the three estimators: average, z-score thresholding and Huber's M estimator. . . . .	128
7.5	Simulation: corruption of a given percentage of ASL volumes (from 0% to 50%) by adding samples drawn from a uniform distribution to a given percentage of the voxels (2%, 20% or 50%) per volume. . . . .	130
7.6	Healthy subject dataset with simulated outliers: sum of squared differences (SSD) of ASL CBF map, computed by M-estimator, z-score thresholding [Tan 2009] and sample average, with the estimated ground truth. Low, medium and high level of volume corruption, from 0% to 50% of corrupted volumes. In all configuration Huber's M-estimators is either better or as good as z-thresholding to estimate robust CBF maps. In the presence of outliers, Huber's M-estimator is always more accurate than the sample average. . . . .	132
7.7	Patient dataset with simulated outliers: correlation of ASL CBF map, computed by M-estimator, z-score thresholding [Tan 2009] and sample average, with the DSC CBF map. Low, medium and high level of volume corruption, from 0% to 50% of corrupted volumes. The white arrow points the tumour site. A similar behaviour as for healthy subject simulation (fig. 7.6) is observed outlining that correlation with DSC CBF is a valuable indicator to measure the quality of the ASL CBF estimates. . . . .	133
7.8	Example of robust CBF map in one patient: three contiguous axial slices are depicted. White arrows outlines large artefacts presents in the averaged perfusion-weighted map and correctly corrected by both z-score thresholding and M-estimator. . . . .	134
7.9	Real clinical dataset: correlation of ASL CBF map with DSC CBF for different number of repeats $r$ . . . . .	135
7.10	DSC CBF estimate (a) and ASL CBF estimates based on 40 repetitions: average (b), z-thresholding (c), Huber's M-estimator (d). . . . .	135
8.1	Parameter estimates of the homoscedastic and heteroscedastic model in the control group, computed from 35 healthy subjects. a) Mean perfusion estimate $\hat{\beta}_{controls}^{HETERO}$ . b) Combined within- and between-subject standard deviation estimate from the homoscedastic model $\hat{\sigma}_{GC}$ . c) Between-subject standard deviation estimate from the heteroscedastic model $\hat{\sigma}_G$ . d) Square root of the average within-subject variance in the control group. Perfusion is expressed in normalized units (ratio to mean grey matter perfusion). Axial slices are displayed in neurological convention. . . . .	151
8.2	Histogram of the ratio of within-subject variance onto total variance in the control group. Both variance estimates have a significant impact depending on the voxels. . . . .	153



- 8.3 Average estimated within-subject variances in control and patient groups. For each box, the red line corresponds to the median and the top and bottom lines of the blue square are the upper and lower quartiles of the distribution. . . . . 154
- 8.4 a) Estimated within-subject standard deviation in a control subject presenting locally atypical values. b) T1w map with false positive detections (in red) by leave-one-out cross-validation (smoothing  $\text{FWHM} = 8 \text{ mm}^3$ ) with the homoscedastic model (b) and the heteroscedastic model (c). No false positive detections are observed with the heteroscedastic model. . . . . 155
- 8.5 Detections of perfusion abnormalities with the homoscedastic and heteroscedastic models in a patient suffering from a high grade glioma. a) T1w-Gd map, the tumour site is pointed by a black arrow. b) Patient ASL CBF estimate  $\hat{\beta}_n$ . c) Within-subject standard deviation of ASL CBF  $\hat{\sigma}_n^2$ . d) T1w-Gd map with ground truth overlaid. T1w-Gd map with hypo- (blue colour-map) and hyper-perfusions (hot colour-map) overlaid for the homoscedastic (e) and heteroscedastic (f) models. Modelling heterogeneous variances (heteroscedastic model) reduces the false positive detections while preserving the true detections. Axial slices are displayed in neurological convention. . . . 156
- 8.6 ROC curves for perfusion abnormality detections with the homoscedastic and heteroscedastic models. The average ROC curves across the studied smoothings are plotted in plain line. Dotted lines are plotted one standard deviation away from the average. . . . . 157
- 8.7 Perfusion abnormalities detections in 3 patients with homoscedastic and heteroscedastic models. From left to right: ground truth hyper-perfusions (red) and hypo-perfusions (blue) overlaid on the T1w-Gd map; detections with the homoscedastic model with a smoothing kernel of  $\text{FWHM} = 8 \text{ mm}^3$ , hyper-perfusions (hot colormap) and hypo-perfusions (blue); detections with the heteroscedastic model with a smoothing kernel of  $\text{FWHM} = 8 \text{ mm}^3$ , hyper-perfusions (hot colormap) and hypo-perfusions (blue). . . . . 159
- 9.1 Ground truth for a representative subject: T1w-Gd and DSC CBF, Ground truth: hyper-perfusions (i.e. Ground truth positives) in red and iso-perfusions (i.e. Ground truth negatives) in green overlaid on the T1w-Gd. . . . . 172
- 9.2 Best ROC curves for false positive rates ranging between 0% and 10%, for the massively univariate GLM (red) and the proposed *a contrario* approach (blue) in nine patients. Overall, the *a contrario* approach outperforms the GLM. . . . . 174

9.3	Average ROC curve (based on 22 subjects), for false positive rates ranging between 0% and 10%, with the GLM (smoothed with a Gaussian kernel of full-width-at-half-maximum $w \in \{0, 2, 4, 6, 8, 10, 12\}$ mm <sup>3</sup> ) and the <i>a contrario</i> approach (with a sphere radius $rd \in \{1, 2\}$ and a pre-defined p-value $p_{PRE} \in \{0.01, 0.005, 0.001\}$ ) . . . . .	175
9.4	Perfusion abnormalities detections in 3 patients with the GLM and the <i>a contrario</i> approach using the best parameter set for each method. From left to right: ground truth hyper-perfusions (red) and hypo-perfusions (blue) overlaid on the T1w-Gd map; GLM, hyper-perfusions (red) and hypo-perfusions (blue); <i>a contrario</i> detections hyper-perfusions (red) and hypo-perfusions (blue). . . . .	177
9.5	QQ-plot of the (a) region-based probabilities, $\pi_{(r)}$ and $\pi_{*(r)}$ , in <i>a contrario</i> (for $p_{PRE} \in [0.01; 0.005; 0.001]$ ), and (b) voxel-wise probabilities for the GLM $\pi_v$ (for $w \in \{0, 2, 4, 6, 8, 10, 12\}$ mm <sup>3</sup> ). The non-independent <i>a contrario</i> approach provides a distribution that is closer to the theory than the standard <i>a contrario</i> and the GLM. . . . .	178
10.1	Subject-specific rescaling of the p-values using background information	185
10.2	Local (a) and global (b) delineations of the Ground Truth for a representative subject: areas of iso-activity (no detection expected) are in green, areas of hyper-activity (i.e. activation) are in blue. . . . .	186
10.3	Subject Q-Q plots of the probabilities provided by the standard General Linear Model and of the updated probabilities (with background information). The theoretical distribution is displayed by a black line along with its 95% confidence interval in grey. The standard GLM probabilities are markedly invalid. The updated probabilities always provide more valid probabilities. . . . .	187
10.4	ROC curves for each subject averaged over the 3 sessions. . . . .	188
10.5	Group ROC curves averaged over the 3 sessions for the massively univariate GLM (unsmoothed and smoothed) and the <i>a contrario</i> approach based on the local (a) and global (b) ground truths. . . . .	189
10.6	Global ground truth: example of detections in subject 2 with a false positive rate of 5%. The hand motor area (outlined by a black arrow on the fourth slice) is correctly detected by the three methods. The supplementary motor area (outlined by a black arrow on the fifth slice) is better detected with the <i>a contrario</i> approach. . . . .	190
10.7	Local ground truth: example of detections in subject 7 with a true positive rate of 75%. The hand motor area is correctly detected by the three methods. There is a larger number of false positive detections with the GLM than the <i>a contrario</i> approach. . . . .	191
10.8	AutoMRI components. . . . .	249
10.9	AutoMRI: workflow. . . . .	249
A total of <b>94 figures</b> .		



# List of Tables

---

1	Aire sous la courbe ROC pour des taux de faux positifs compris entre 0% et 10% avec les modèles homoscédastique et hétéroscédastique. Le modèle hétéroscédastique est plus performant que le modèle homoscédastique. . . . .	xx
2	Aire sous la courbe ROC pour des taux de faux positifs compris entre 0% et 10% avec le GLM hétéroscédastique et la méthode <i>a contrario</i> . La méthode <i>a contrario</i> est plus performante que le GLM. . . . .	xxii
5.1	fMRI results for each language task: peak location, cluster extent – Z-score ( $p < 0.05$ FWE; $k = 5$ ). For visual language tasks, additional results among language areas are reported at $p < 0.001$ uncorrected in brackets. Clusters in italics correspond to sub-clusters belonging to cluster of the same extent given in the column. <sup>(1)</sup> Posterior part; <sup>(2)</sup> Middle part; <sup>(3)</sup> Anterior part; <sup>(4)</sup> Dorsal part; <sup>(5)</sup> Ventral part; <sup>(6)</sup> Extending in the adjacent IFG. Oper = opercularis; Tri = triangularis; Inf = inferior; Mid = middle; Sup = superior; Med = medial; Ant = anterior. . . . .	87
5.2	Task comparisons (N) and conjunctions (C). Peak locations, cluster extent–Z-score ( $p < 0.001$ uncorrected; $k = 10$ ). Clusters in italics correspond to sub-clusters belonging to cluster of the same extent given in the column. <sup>(1)</sup> Posterior part; <sup>(2)</sup> Middle part; <sup>(3)</sup> Anterior part; <sup>(4)</sup> Dorsal part; <sup>(5)</sup> Ventral part; <sup>(6)</sup> Ventral part, extending into the SMG, angular, superior temporal gyri; <sup>(7)</sup> Ventral part, extending into the adjacent insula; <sup>(8)</sup> Extending into the adjacent IFG; <sup>(9)</sup> Overlapping most of the adjacent dorsal and ventral IFG-Opercularis. Ph-s = Phonol-Seg task; Ph-d = Phonol-Diff task; Oper = opercularis; Tri = triangularis; Rol = rolandic; Inf = inferior; Mid = middle; Sup = superior; Med = medial; Ant = anterior; SMG = SMG. . . . .	91

6.1	Groups characteristics. Ranges are reported in brackets. Bold font indicates language scores smaller than one standard deviation below the normal mean. (a) Including scores of the four youngest children who performed equivalent subtests from the Nouvelles Epreuves pour l'Evaluation du Langage (N-EEL; see text). L2MA = battery Langage oral, Langage écrit, Mémoire, Attention; WISC-IV= Wechsler Intelligence Scale for Children-Fourth version; VCI and PRI= Verbal Comprehension and Perceptual Reasoning Indexes. . . . .	105
7.1	Real clinical dataset: correlation coefficient with DSC CBF map of ASL CBF map computed by M-estimator, z-score thresholding [Tan 2009] and sample average in 14 patients diagnosed with brain tumours. Last column: mean and standard deviation across subjects.	134
8.1	Sensitivity and specificity in the control (first row) and patient (second and third rows) groups with the homoscedastic and heteroscedastic models for different smoothing kernels (FWHM = [0, 4, 6, 8, 10, 12] mm <sup>3</sup> ), at $p < 0.05$ FDR corrected. . . . .	155
8.2	Area under the ROC curve for false positive rates ranging between 0% and 10% with the homoscedastic and heteroscedastic models. The heteroscedastic model outperforms the homoscedastic model. . .	157
9.1	Area under the ROC curve, averaged over the 9 subjects presenting hyper-perfusions, for false positive rates ranging between 0% and 10%, with the a contrario approach (with a sphere radius $rd \in \{1, 2\}$ and a pre-defined p-value $p_{PRE} \in \{0.01, 0.005, 0.001\}$ ) and the GLM (smoothed with a Gaussian kernel of full-width-at-half-maximum $w \in \{0, 2, 4, 6, 8, 10, 12\}$ mm <sup>3</sup> ). The proposed a contrario approach outperforms all tested configurations of the GLM. . . . .	174
9.2	Area under the average ROC curve (based on 22 subjects), for false positive rates ranging between 0% and 10%, with the a contrario approach (with a sphere radius $rd \in \{1, 2\}$ and a pre-defined p-value $p_{PRE} \in \{0.01, 0.005, 0.001\}$ ) and the GLM (smoothed with a Gaussian kernel of full-width-at-half-maximum $w \in \{0, 2, 4, 6, 8, 10, 12\}$ mm <sup>3</sup> ). The proposed a contrario approach outperforms the GLM. . . . .	176
10.1	Average area under the ROC curves for Day 1, Day 2 and Day 3, using the local or the global ground truth. . . . .	187
A total of <b>10 tables</b> .		

# Bibliography

---

- [Abel 2009] Stefanie Abel, Katharina Dressel, Ruth Bitzer, Dorothee Kümmerer, Irina Mader, Cornelius Weiller and Walter Huber. *The separation of processing stages in A. lexical interference fMRI-paradigm*. NeuroImage, vol. 44, no. 3, pages 1113–1124, 2009. (Cited on pages 97 and 98.)
- [Abrams 2012] Daniel A. Abrams, Srikanth Ryali, Tianwen Chen, Evan Balaban, Daniel J. Levitin and Vinod Menon. *Multivariate Activation and Connectivity Patterns Discriminate Speech Intelligibility in Wernicke’s, Broca’s, and Geschwind’s Areas*. Cerebral Cortex, 2012. (Cited on page 73.)
- [Ackermann 2010] Hermann Ackermann and Axel Riecker. *The contribution(s) of the insula to speech production: A review of the clinical and functional imaging literature*. Brain Structure and Function, vol. 214, no. 5-6, pages 419–433, 2010. (Cited on pages 115 and 116.)
- [Aguerreberre 2009] Cecilia Aguerreberre, Pablo Sprechmann, Pablo Musé and Rodolfo Ferrando. *A-contrario localization of epileptogenic zones in SPECT images*. In Proceedings of the IEEE 6th International Symposium on Biomedical Imaging (ISBI), pages 570–573, 2009. (Cited on pages xxi, 63, 64, 164 and 167.)
- [Aguirre 1997] Geoffrey K. Aguirre, Eric Zarahn and Mark D’Esposito. *Empirical analyses of BOLD fMRI statistics. II. Spatially smoothed data collected under null-hypothesis and experimental conditions*. NeuroImage, vol. 5, no. 3, pages 199–212, 1997. (Cited on pages 44 and 142.)
- [Aguirre 1998] Geoffrey K. Aguirre, Eric Zarahn and Mark D’Esposito. *The variability of human, BOLD hemodynamic responses*. NeuroImage, vol. 8, no. 4, pages 360–409, 1998. (Cited on page 31.)
- [Aguirre 2002] Geoffrey K. Aguirre, John A. Detre, Eric Zarahn and David C. Alsop. *Experimental design and the relative sensitivity of BOLD and perfusion fMRI*. NeuroImage, vol. 15, no. 3, pages 488–500, 2002. (Cited on pages 25, 31, 44, 45 and 142.)

- [Aguirre 2006] Geoffrey K. Aguirre. *Experimental Design and Data Analysis for fMRI*. In *Functional MRI: Basic principles and clinical applications*, chapitre 3, pages 58–74, 2006. (Cited on page 28.)
- [Ahmad 2003] Z. Ahmad, L. M. Balsamo, B. C. Sachs, B. Xu and W. D. Gaillard. *Auditory comprehension of language in young children: neural networks identified with fMRI*. *Neurology*, vol. 60, no. 10, pages 1598–1605, 2003. (Cited on page 96.)
- [Alm 2004] Per A. Alm. *Stuttering and the basal ganglia circuits: A critical review of possible relations*. *Journal of Communication Disorders*, vol. 37, no. 4, pages 325–369, 2004. (Cited on page 118.)
- [Alsop 2010] David C. Alsop, Weiyang Dai, Murray Grossman and John A. Detre. *Arterial spin labeling blood flow MRI: its role in the early characterization of Alzheimer’s disease*. *Journal of Alzheimer’s Disease*, vol. 20, no. 3, pages 871–880, 2010. (Cited on page 15.)
- [Alsop 2012] David C. Alsop. *Arterial spin labeling: its time is now*. *Magma*, vol. 25, no. 2, pages 75–77, 2012. (Cited on page 19.)
- [Ambroise 2002] Christophe Ambroise and Geoffrey J. McLachlan. *Selection bias in gene extraction on the basis of microarray gene-expression data*. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 10, pages 6562–6566, 2002. (Cited on page 60.)
- [Anderson 2006] Dianne P. Anderson, A. Simon Harvey, Michael M. Saling, Vicki Anderson, Michael Kean, David F. Abbott, R. Mark Wellard and Graeme D. Jackson. *fMRI lateralization of expressive language in children with cerebral lesions*. *Epilepsia*, vol. 47, no. 6, pages 998–1008, 2006. (Cited on pages 79 and 93.)
- [Annett 1985] M. Annett. *Left, right, hand and brain: the Right Shift Theory*. Numeéro 02. Hillsdale, NJ, Erlbaum, 1985. (Cited on page 101.)
- [Arndt 1996] S. Arndt, T. Cizadlo, D. O’Leary, S. Gold and N. C. Andreasen. *Normalizing counts and cerebral blood flow intensity in functional imaging studies of the human brain*. *NeuroImage*, vol. 3, no. 3 Pt 1, pages 175–184, 1996. (Cited on page 150.)
- [Arthurs 2002] Owen J. Arthurs and Simon Boniface. *How well do we understand the neural origins of the fMRI BOLD signal?* *Trends in Neurosciences*, vol. 25, no. 1, pages 27–31, 2002. (Cited on pages 27 and 201.)
- [Ashburner 2005] John Ashburner and Karl J. Friston. *Unified segmentation*. *NeuroImage*, vol. 26, no. 3, pages 839–851, 2005. (Cited on pages 38, 84, 107, 149, 170 and 183.)

- [Ashburner 2007] John Ashburner. *A fast diffeomorphic image registration algorithm*. NeuroImage, vol. 38, no. 1, pages 95–113, 2007. (Cited on page 38.)
- [Aslan 2010] Sina Aslan and Hanzhang Lu. *On the sensitivity of ASL MRI in detecting regional differences in cerebral blood flow*. Magnetic Resonance Imaging, vol. 28, no. 7, pages 928–935, 2010. (Cited on pages 150, 151 and 171.)
- [Balsamo 2002] Lyn M. Balsamo, Benjamin Xu, Cecile B. Grandin, Jeffrey R. Petrella, Suzanne H. Braniecki, Teresa K. Elliott and William D. Gaillard. *A functional magnetic resonance imaging study of left hemisphere language dominance in children*. Archives of Neurology, vol. 59, no. 7, pages 1168–1174, 2002. (Cited on pages 80, 82, 96 and 105.)
- [Balsamo 2006] L. M. Balsamo, B. Xu and W. D. Gaillard. *Language lateralization and the role of the fusiform gyrus in semantic processing in young children*. NeuroImage, vol. 31, no. 3, pages 1306–1314, 2006. (Cited on page 98.)
- [Bannier 2012] Elise Bannier, Camille Maumet, Anca Pasnicu, Jean-Christophe Ferré, Eduardo Pasqualini, Arnaud Biraben, Jean-Yves Gauvrit and Christian Barillot. *Voxel Based Analysis of 3D Double Inversion Recovery for the detection of cortical abnormalities in drug resistant epilepsy*. In Proceedings of the ISMRM 20th Annual Meeting & Exhibition, page 3244, Melbourne, Australia, 2012. (Cited on pages 58 and 250.)
- [Barlow 2002] Jessica A. Barlow and Judith A. Gierut. *Minimal pair approaches to phonological remediation*. Seminars in speech and language, vol. 23, no. 1, pages 57–68, 2002. (Cited on page 81.)
- [Bates 1997] Elizabeth Bates. *Origins of language disorders: A comparative approach*. Developmental Neuropsychology, vol. 13, no. 3, pages 447–476, 1997. (Cited on page 101.)
- [Beckmann 2003] Christian F. Beckmann, Mark Jenkinson and Stephen M. Smith. *General multilevel linear modeling for group analysis in FMRI*. NeuroImage, vol. 20, no. 2, pages 1052–1063, 2003. (Cited on pages 38, 140 and 164.)
- [Belliveau 1992] J. W. Belliveau, K. K. Kwong, D. N. Kennedy, J. R. Baker, C. E. Stern, R. Benson, D. A. Chesler, R. M. Weisskoff, M. S. Cohen and R. B. Tootell. *Magnetic resonance imaging mapping of brain function. Human visual cortex*. Investigative Radiology, vol. 27 Suppl 2, pages S59–65, 1992. (Cited on page 29.)
- [Belton 2003] Emma Belton, Claire H. Salmond, Kate E. Watkins, Faraneh Vargha-Khadem and David G. Gadian. *Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia*. Human Brain Mapping, vol. 18, no. 3, pages 194–200, 2003. (Cited on page 118.)



- [Benjamini 1995] Yoav Benjamini and Yosef Hochberg. *Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing*. Journal of the Royal Statistical Society. Series B. (Methodological), vol. 57, no. 1, pages 289–300, 1995. (Cited on pages [xxi](#), [53](#) and [148](#).)
- [Berl 2006] Madison M. Berl, Chandan J. Vaidya and William Davis Gaillard. *Functional imaging of developmental and adaptive changes in neurocognition*. NeuroImage, vol. 30, no. 3, pages 679–691, 2006. (Cited on page [79](#).)
- [Bernal 2003] Byron Bernal and Nolan R. Altman. *Speech delay in children: A functional MR imaging study*. Radiology, vol. 229, no. 3, pages 651–658, 2003. (Cited on page [114](#).)
- [Binder 1996] J. R. Binder, S. J. Swanson, T. A. Hammeke, G. L. Morris, W. M. Mueller, M. Fischer, S. Benbadis, J. A. Frost, S. M. Rao and V. M. Haughton. *Determination of language dominance using functional MRI: A comparison with the Wada test*. Neurology, vol. 46, no. 4, pages 978–84, 1996. (Cited on page [79](#).)
- [Bishop 1997] D. V. M. Bishop. *Uncommon understanding: Development and disorders of language comprehension in children*. Psychology Press, 1997. (Cited on page [101](#).)
- [Bishop 2000] D. V. M. Bishop. *How does the brain learn language? Insights from the study of children with and without language impairment*. Developmental Medicine and Child Neurology, vol. 42, no. 2, pages 133–142, 2000. (Cited on page [101](#).)
- [Bishop 2004a] Dorothy V. M. Bishop. *Diagnostic dilemmas in specific language impairment*. In Van Balkom J. Verhoeven L., editeur, *Classification of developmental language disorders*, pages 309–326. Mahwah, erlbaum édition, 2004. (Cited on page [101](#).)
- [Bishop 2004b] Dorothy V. M. Bishop and Margaret J. Snowling. *Developmental dyslexia and specific language impairment: same or different?* Psychological bulletin, vol. 130, no. 6, pages 858–886, 2004. (Cited on pages [101](#) and [117](#).)
- [Bishop 2006] Dorothy V. M. Bishop. *What Causes Specific Language Impairment in Children?* Current Directions in Psychological Science, vol. 15, no. 5, pages 217–221, 2006. (Cited on pages [xiii](#), [71](#) and [101](#).)
- [Bishop 2008] D. V. M. Bishop and M. E. Hayiou-Thomas. *Heritability of specific language impairment depends on diagnostic criteria*. Genes, Brain, and Behavior, vol. 7, no. 3, pages 365–372, 2008. (Cited on page [101](#).)
- [Blank 2002] S. Catrin Blank, Sophie K. Scott, Kevin Murphy, Elizabeth Warburton and Richard J. S. Wise. *Speech production: Wernicke, Broca and beyond*.

- Brain, vol. 125, no. Pt 8, pages 1829–1838, 2002. (Cited on pages 102, 114 and 118.)
- [Bohland 2006] Jason W. Bohland and Frank H. Guenther. *An fMRI investigation of syllable sequence production*. NeuroImage, vol. 32, no. 2, pages 821–841, 2006. (Cited on page 116.)
- [Bokkers 2011] Reinoud Bokkers. *Arterial spin labeling perfusion MRI in cerebral ischemia*. thesis, Utrecht University, The Netherlands, 2011. (Cited on pages 15 and 57.)
- [Booth 2007] James R. Booth, Lydia Wood, Dong Lu, James C. Houk and Tali Bitan. *The role of the basal ganglia and cerebellum in language processing*. Brain research, vol. 1133, no. 1, pages 136–144, 2007. (Cited on page 116.)
- [Brauer 2007] Jens Brauer and Angela D. Friederici. *Functional neural networks of semantic and syntactic processes in the developing brain*. Journal of Cognitive Neuroscience, vol. 19, no. 10, pages 1609–1623, 2007. (Cited on page 96.)
- [Brett 2002] M. Brett, J. L. Anton, R. Valabregue and J. Poline. *Region of interest analysis using an SPM toolbox*. In Conference on Functional Mapping of the Human Brain, Sendai, Japan, 2002. (Cited on page 85.)
- [Broca 1861] Paul Broca. *Bulletin de la société française d’anthropologie. séance du 18 avril 1861, tome 2*. pages 235–238, 1861. (Cited on page 72.)
- [Brown 2005] Steven Brown, Roger J. Ingham, Janis C. Ingham, Angela R. Laird and Peter T Fox. *Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies*. Human Brain Mapping, vol. 25, no. 1, pages 105–117, 2005. (Cited on pages 114, 115, 116, 117 and 118.)
- [Brown 2009] Steven Brown, Angela R. Laird, Peter Q. Pfordresher, Sarah M. Thelen, Peter Turkeltaub and Mario Liotti. *The somatotopy of speech: phonation and articulation in the human motor cortex*. Brain and cognition, vol. 70, no. 1, pages 31–41, 2009. (Cited on page 115.)
- [Brun 2003] V. Brun and C. L. Gérard. *Les dysphasies*. Collection Rencontre en rééducation. Masson, 2003. (Cited on pages xiii and 71.)
- [Buchsbaum 2001] Bradley R. Buchsbaum, Gregory Hickok and Colin Humphries. *Role of left posterior superior temporal gyrus in phonological processing for speech perception and production*. Cognitive Science, vol. 25, no. 5, pages 663–678, 2001. (Cited on pages 80 and 97.)
- [Burton 2000] M. W. Burton, S. L. Small and S. E. Blumstein. *The role of segmentation in phonological processing: an fMRI investigation*. Journal of Cognitive Neuroscience, vol. 12, no. 4, pages 679–690, 2000. (Cited on pages 80 and 81.)

- [Burton 2001] Martha W. Burton. *The role of inferior frontal cortex in phonological processing*. *Cognitive Science*, vol. 25, no. 5, pages 695–709, 2001. (Cited on pages 80, 81 and 97.)
- [Burton 2006] Martha W. Burton and Steven L. Small. *Functional neuroanatomy of segmenting speech and nonspeech*. *Cortex*, vol. 42, no. 4, pages 644–651, 2006. (Cited on page 80.)
- [Buxton 1998] Richard B. Buxton, Lawrence R. Frank, Eric C. Wong, B. Siewert, S. Warach and R. R. Edelman. *A general kinetic model for quantitative perfusion imaging with arterial spin labeling*. *Magnetic Resonance in Medicine*, vol. 40, no. 3, pages 383–396, 1998. (Cited on pages 22, 129, 150 and 170.)
- [Buxton 2004] Richard B. Buxton, Kâmil Uludağ, David J. Dubowitz and Thomas T. Liu. *Modeling the hemodynamic response to brain activation*. *NeuroImage*, vol. 23 Suppl 1, pages S220–33, 2004. (Cited on pages 30, 31 and 201.)
- [Catani 2005] Marco Catani, Derek K. Jones and Dominic H. Ffytche. *Perisylvian language networks of the human brain*. *Annals of Neurology*, vol. 57, no. 1, pages 8–16, 2005. (Cited on pages 74, 75, 114 and 203.)
- [Chalela 2000] J.A. Chalela, David C. Alsop, J.B. Gonzalez-Atavales, J.A. Maldjian, S.E. Kasner and John A. Detre. *Magnetic resonance perfusion imaging in acute ischemic stroke using continuous arterial spin labeling*. *Stroke*, vol. 31, no. 3, pages 680–687, 2000. (Cited on pages 146 and 164.)
- [Chappell 2010] Michael A. Chappell, Bradley J. MacIntosh, Manus J. Donahue, Matthias Günther, Peter Jezzard and Mark W. Woolrich. *Separation of macrovascular signal in multi-inversion time arterial spin labelling MRI*. *Magnetic Resonance in Medicine*, vol. 63, no. 5, pages 1357–1365, 2010. (Cited on page 25.)
- [Chappell 2012] Michael A. Chappell, Mark W. Woolrich, Esben T. Petersen, Xavier Golay and Stephen J. Payne. *Comparing model-based and model-free analysis methods for QUASAR arterial spin labeling perfusion quantification*. *Magnetic Resonance in Medicine*, vol. 69, no. 5, pages 1–10, 2012. (Cited on page 25.)
- [Chawla 2007] S. Chawla, Sumei Wang, Robert Christian Wolf, J. H. Woo, Jiongjiong Wang, D. M. O’Rourke, K. D. Judy, M. S. Grady, E. R. Melhem and H. Poptani. *Arterial spin-labeling and MR spectroscopy in the differentiation of gliomas*. *American Journal of Neuroradiology*, vol. 28, no. 9, pages 1683–1689, 2007. (Cited on pages 57, 146 and 164.)
- [Chen 2012a] Gang Chen, Ziad S. Saad, Audrey R. Nath, Michael S. Beauchamp and Robert W. Cox. *fMRI group analysis combining effect estimates and*

- their variances*. NeuroImage, vol. 60, no. 1, pages 747–765, 2012. (Cited on pages 140, 145, 148, 153 and 158.)
- [Chen 2012b] Tai-Yuan Chen, Lee Chiu, Tai-Ching Wu, Te-Chang Wu, Chien-Jen Lin, Shih-Chuan Wu and Yu-Kun Tsui. *Arterial spin-labeling in routine clinical practice: a preliminary experience of 200 cases and correlation with MRI and clinical findings*. Clinical Imaging, vol. 36, no. 4, pages 345–352, 2012. (Cited on pages 57, 139 and 164.)
- [Chevrie-Muller 1985] C. Chevrie-Muller, A.-M. Simon, M.-T. Le Normand and S. Fournier. Batterie d'évaluation psycholinguistique. Paris, 1985. (Cited on page 81.)
- [Chevrie-Muller 1997] C. Chevrie-Muller, A. M. Simon and S. Fournier. Batterie langage oral-langage écrit, mémoire-attention: L2ma. Paris, 1997. (Cited on page 104.)
- [Chevrie-Muller 2001] C. Chevrie-Muller and M. Plaza. Nouvelles épreuves pour l'examen du langage: bilan complet des capacités langagières et cognitives de l'enfant: N-eel. Paris, 2001. (Cited on page 104.)
- [Chiron 1999] C. Chiron, F. Pinton, M. C. Masure, C. Duvelleroy-Hommet, F. Leon and C. Billard. *Hemispheric specialization using SPECT and stimulation tasks in children with dysphasia and dystrophia*. Developmental Medicine and Child Neurology, vol. 41, no. 8, pages 512–520, 1999. (Cited on pages 102 and 114.)
- [Chomsky 1968] N. Chomsky and M. Halle. The Sound Pattern of English. New York, 1968. (Cited on page 80.)
- [Chumbley 2009] Justin R. Chumbley and Karl J. Friston. *False discovery rate revisited: FDR and topological inference using Gaussian random fields*. NeuroImage, vol. 44, no. 1, pages 62–70, 2009. (Cited on pages 53 and 168.)
- [Cohen 2004] Laurent Cohen and Stanislas Dehaene. *Specialization within the ventral stream: the case for the visual word form area*. NeuroImage, vol. 22, no. 1, pages 466–476, 2004. (Cited on page 97.)
- [Colliot 2006] Olivier Colliot, Neda Bernasconi, Samson B Antel, N Khalili, Andrea Bernasconi and V Naessens. *Individual voxel-based analysis of gray matter in focal cortical dysplasia*. NeuroImage, vol. 29, no. 1, pages 162–171, 2006. (Cited on pages 57 and 58.)
- [Commowick 2008] Olivier Commowick, Pierre Fillard, Olivier Clatz and Simon K. Warfield. *Detection of DTI white matter abnormalities in multiple sclerosis patients*. In Dimitris Metaxas, Leon Axel, Gabor Fichtinger and Gábor Székely, éditeurs, Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), volume

- 5241 of *Lecture Notes in Computer Science*, pages 975–82, New York City, United States, 2008. Springer-Verlag. (Cited on page 57.)
- [Commowick 2009] Olivier Commowick, Simon Warfield K. and Grégoire Malandain. *Using Frankenstein’s Creature Paradigm to Build a Patient Specific Atlas*. In Guang-Zhong Yang, David Hawkes, Daniel Rueckert, Alison Noble and Chris Taylor, editeurs, Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Part II, volume 5762 of *Lecture Notes in Computer Science*, pages 993–1000. Springer, 2009. (Cited on page 57.)
- [Conti-Ramsden 1997] G. Conti-Ramsden, A. Crutchley and N. Botting. *The extent to which psychometric tests differentiate subgroups of children with SLI*. *Journal of Speech, Language, and Hearing Research*, vol. 40, no. 4, pages 765–777, 1997. (Cited on page 101.)
- [Conti-Ramsden 2003] Gina Conti-Ramsden. *Processing and linguistic markers in young children with specific language impairment (SLI)*. *Journal of Speech, Language, and Hearing Research*, vol. 46, no. 5, pages 1029–1037, 2003. (Cited on page 101.)
- [Corouge 2012] Isabelle Corouge, Aurore Esquevin, Florence Lejeune, Jean-Christophe Ferré, Elise Bannier, Catherine Merck, Serge Belliard, Christian Barillot and Jean-Yves Gauvrit. *Arterial Spin Labeling at 3T in semantic dementia: perfusion abnormalities detection and comparison with FDG-PET*. MICCAI 2012 Workshop on Novel Biomarkers for Alzheimer’s Disease and Related Disorders, 2012. (Cited on pages 15, 16 and 201.)
- [Costafreda 2006] Sergi G. Costafreda, Cynthia H. Y. Fu, Lucy Lee, Brian Everitt, Michael J. Brammer and Anthony S. David. *A systematic review and quantitative appraisal of fMRI studies of verbal fluency: role of the left inferior frontal gyrus*. *Human Brain Mapping*, vol. 27, no. 10, pages 799–810, 2006. (Cited on page 95.)
- [Costafreda 2009] Sergi G. Costafreda, Carlton Chu, John Ashburner and Cynthia H. Y. Fu. *Prognostic and diagnostic potential of the structural neuroanatomy of depression*. *PloS one*, vol. 4, no. 7, page e6353, 2009. (Cited on page 60.)
- [Cox 1996] R. W. Cox. *AFNI: software for analysis and visualization of functional magnetic resonance neuroimages*. *Computers and Biomedical Research*, vol. 29, no. 3, pages 162–173, 1996. (Cited on pages 38 and 140.)
- [Crinion 2003] J. T. Crinion, M. A. Lambon-Ralph, E. A. Warburton, D. Howard and R. J. S. Wise. *Temporal lobe regions engaged during normal speech comprehension*. *Brain*, vol. 126, no. 5, pages 1193–1201, 2003. (Cited on pages 80, 102 and 118.)

- [Crinion 2007a] Jenny Crinion, John Ashburner, Alex Leff, Matthew Brett, Cathy Price and Karl Friston. *Spatial normalization of lesioned brains: performance evaluation and impact on fMRI analyses*. *NeuroImage*, vol. 37, no. 3, pages 866–875, 2007. (Cited on pages 38, 150 and 171.)
- [Crinion 2007b] Jenny T. Crinion and Alexander P. Leff. *Recovery and treatment of aphasia after stroke: functional imaging studies*. *Current Opinion in Neurology*, vol. 20, no. 6, pages 667–673, 2007. (Cited on page 116.)
- [Dai 2008] Weiyang Dai, Dairon Garcia, Cedric de Bazelaire and David C. Alsop. *Continuous flow-driven inversion for arterial spin labeling using pulsed radio frequency and gradient fields*. *Magnetic Resonance in Medicine*, vol. 60, no. 6, pages 1488–1497, 2008. (Cited on page 21.)
- [De Fossé 2004] Lies De Fossé, Steven M. Hodge, Nikos Makris, David N. Kennedy, Verne S. Caviness, Lauren McGrath, Shelley Steele, David A. Ziegler, Martha R. Herbert, Jean A. Frazier, Helen Tager-Flusberg and Gordon J. Harris. *Language-association cortex asymmetry in autism and specific language impairment*. *Annals of Neurology*, vol. 56, no. 6, pages 757–766, 2004. (Cited on pages 102 and 115.)
- [de Guibert 2010] Clément de Guibert, Camille Maumet, Jean-Christophe Ferré, Pierre Jannin, Arnaud Biraben, Catherine Allaire, Christian Barillot and Elisabeth Le Rumeur. *fMRI language mapping in children: a panel of language tasks using visual and auditory stimulation without reading or metalinguistic requirements*. *NeuroImage*, vol. 51, no. 2, pages 897–909, 2010. (Cited on pages 9, 26, 77, 102, 104, 118, 119, 201 and 250.)
- [de Guibert 2011] Clément de Guibert, Camille Maumet, Pierre Jannin, Jean-Christophe Ferré, Catherine Tréguier, Christian Barillot, Elisabeth Le Rumeur, Catherine Allaire and Arnaud Biraben. *Abnormal functional lateralization and activity of language brain areas in typical specific language impairment (developmental dysphasia)*. *Brain*, vol. 134, no. Pt 10, pages 3044–58, 2011. (Cited on pages 9, 99 and 250.)
- [Deblaere 2002] K. Deblaere, W. H. Backes, P. Hofman, P. Vandemaele, P. A. Boon, K. Vonck, P. Boon, J. Troost, J. Vermeulen, J. Wilmink, E. Achten and A. Aldenkamp. *Developing a comprehensive presurgical functional MRI protocol for patients with intractable temporal lobe epilepsy: A pilot study*. *Neuroradiology*, vol. 44, no. 8, pages 667–673, 2002. (Cited on pages 79 and 102.)
- [DeLeon 2007] Jessica DeLeon, Rebecca F. Gottesman, Jonathan T. Kleinman, Melissa Newhart, Cameron Davis, Jennifer Heidler-Gary, Andrew Lee and Argye E. Hillis. *Neural regions essential for distinct cognitive processes underlying picture naming*. *Brain*, vol. 130, no. Pt 5, pages 1408–1422, 2007. (Cited on page 81.)

- [Démonet 2005] Jean-François Démonet, Guillaume Thierry and Dominique Cardebat. *Renewal of the neurophysiology of language: functional neuroimaging*. *Physiological reviews*, vol. 85, no. 1, pages 49–95, 2005. (Cited on pages 80 and 81.)
- [Desolneux 2003] Agnès Desolneux, L. Moisan and J. Morel. *A grouping principle and four applications*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 4, pages 508–513, 2003. (Cited on pages xx, 61, 62, 64, 164, 166 and 167.)
- [Detre 1992] J. A. Detre, J. S. Leigh, D. S. Williams and A. P. Koretsky. *Perfusion imaging*. *Magnetic Resonance in Medicine*, vol. 23, pages 37–45, 1992. (Cited on pages xi, 19, 30 and 125.)
- [Detre 2012] John A. Detre, Hengyi Rao, Danny J. J. Wang, Yu Fen Chen and Ze Wang. *Applications of arterial spin labeled MRI in the brain*. *Journal of Magnetic Resonance Imaging*, vol. 35, no. 5, pages 1026–1037, 2012. (Cited on page 19.)
- [Dibbets 2006] Pauline Dibbets, Katja Bakker and Jellemer Jolles. *Functional MRI of task switching in children with Specific Language Impairment (SLI)*. *Neurocase*, vol. 12, no. 1, pages 71–79, 2006. (Cited on pages 102, 114, 116 and 117.)
- [Dick 2008] F. Dick, F. Richardson and M. C. Saccuman. *Using functional magnetic resonance imaging to investigate developmental language disorders*. In D.V.M, editeur, *Understanding developmental language disorders*, pages 53–66. Norbury, C.F., Tomblin, J.B., Bishop, Hove, psychology édition, 2008. (Cited on pages 79 and 93.)
- [Dosenbach 2010] Nico U. F. Dosenbach, Binyam Nardos, Alexander L. Cohen, Damien a Fair, Jonathan D. Power, Jessica a Church, Steven M. Nelson, Gagan S. Wig, Alecia C. Vogel, Christina N. Lessov-Schlaggar, Kelly Anne Barnes, Joseph W. Dubis, Eric Feczko, Rebecca S. Coalson, John R. Pruett, Deanna M. Barch, Steven E. Petersen and Bradley L. Schlaggar. *Prediction of individual brain maturity using fMRI*. *Science*, vol. 329, no. 5997, pages 1358–1361, 2010. (Cited on page 9.)
- [Dronkers 1996] N. F. Dronkers. *A new brain region for coordinating speech articulation*. *Nature*, vol. 384, no. 6605, pages 159–161, 1996. (Cited on page 115.)
- [Du 2006] A. T. Du, G. H. Jahng, S. Hayasaka, J. H. Kramer, H. J. Rosen, M. L. Gorno-Tempini, K. P. Rankin, B. L. Miller, M. W. Weiner and N. Schuff. *Hypoperfusion in frontotemporal dementia and Alzheimer disease by arterial spin labeling MRI*. *Neurology*, vol. 67, no. 7, pages 1215–1220, 2006. (Cited on page 15.)

- [Duffau 2001] H. Duffau, L. Bauchet, S. Lehericy and L. Capelle. *Functional compensation of the left dominant insula for language*. *Neuroreport*, vol. 12, no. 10, pages 2159–2163, 2001. (Cited on page 116.)
- [Dunn 1996] M. Dunn, J. Flax, M. Sliwinski and D. Aram. *The use of spontaneous language measures as criteria for identifying children with specific language impairment: an attempt to reconcile clinical and research incongruence*. *Journal of Speech, Language, and Hearing Research*, vol. 39, no. 3, pages 643–654, 1996. (Cited on page 101.)
- [Duong 2001] T. Q. Duong, D. S. Kim, K. Ugurbil and S. G. Kim. *Localized cerebral blood flow response at submillimeter columnar resolution*. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 19, pages 10904–10909, 2001. (Cited on page 31.)
- [Durstun 2006] Sarah Durstun and B. J. Casey. *What have we learned about cognitive development from neuroimaging?* *Neuropsychologia*, vol. 44, no. 11, pages 2149–2157, 2006. (Cited on page 79.)
- [Edelman 1994] R. R. Edelman, B. Siewert, D. G. Darby, V. Thangaraj, A. C. Nobre, M. M. Mesulam and S. Warach. *Qualitative mapping of cerebral blood flow and functional localization with echo-planar MR imaging and signal targeting with alternating radio frequency*. *Radiology*, vol. 192, no. 2, pages 513–520, 1994. (Cited on page 21.)
- [Feng 2013] Lu Feng, Jiangang Liu, Ling Li, Wenjuan Wei, Jie Tian and Kang Lee. *Functional Brain Mapping with Locally Smoothed Regression*. In *Proceedings of the 10th IEEE International Symposium on Biomedical Imaging (ISBI)*, 2013. (Cited on page 62.)
- [Ferré 2012] Jean-Christophe Ferré, Jan Petr, Elise Bannier, Christian Barillot and Jean-Yves Gauvrit. *Improving quality of arterial spin labeling MR imaging at 3 tesla with a 32-channel coil and parallel imaging*. *Journal of Magnetic Resonance Imaging*, vol. 1239, pages 1233–1239, 2012. (Cited on pages 14, 15, 16, 57, 150, 171, 200 and 201.)
- [Filippini 2012] N. Filippini, L. D. Nickerson, C. F. Beckmann, K. P. Ebmeier, G. B. Frisoni, P. M. Matthews, S. M. Smith and C. E. Mackay. *Age-related adaptations of brain function during a memory task are also present at rest*. *NeuroImage*, vol. 59, no. 4, pages 3821–3828, 2012. (Cited on page 10.)
- [Fisher 2003] Simon E. Fisher, Cecilia S. L. Lai and Anthony P. Monaco. *Deciphering the genetic basis of speech and language disorders*. *Annual Review of Neuroscience*, vol. 26, pages 57–80, 2003. (Cited on page 101.)
- [Focke 2008] Niels K. Focke, Mark R. Symms, John S. Duncan and Jane L. Burdett. *Voxel-based analysis of whole brain FLAIR at 3T detects focal cortical*



- dysplasia*. *Epilepsia*, vol. 49, no. 5, pages 786–793, 2008. (Cited on pages 57 and 58.)
- [Fox 1984] P. T. Fox and M. E. Raichle. *Stimulus rate dependence of regional cerebral blood flow in human striate cortex, demonstrated by positron emission tomography*. *Journal of Neurophysiology*, vol. 51, no. 5, pages 1109–1120, 1984. (Cited on page 29.)
- [Fox 1986] P. T. Fox and M. E. Raichle. *Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects*. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 83, no. 4, pages 1140–1144, 1986. (Cited on page 17.)
- [Friederici 2006a] Angela D. Friederici. *The neural basis of language development and its impairment*. *Neuron*, vol. 52, no. 6, pages 941–952, 2006. (Cited on pages 79, 93, 101 and 102.)
- [Friederici 2006b] Angela D. Friederici. *What's in control of language?* *Nature Neuroscience*, vol. 9, no. 8, pages 991–992, 2006. (Cited on page 116.)
- [Friston 1991] K. J. Friston, C. D. Frith, P. F. Liddle and R. S. Frackowiak. *Comparing functional (PET) images: the assessment of significant change*. *Journal of Cerebral Blood Flow and Metabolism*, vol. 11, no. 4, pages 690–699, 1991. (Cited on page 53.)
- [Friston 1994a] K. J. Friston, A. P. Holmes, K. J. Worsley, J.-P. Poline, C. D. Frith and R. S. J. Frackowiak. *Statistical parametric maps in functional imaging: A. general linear approach*. *Human Brain Mapping*, vol. 2, no. 4, pages 189–210, 1994. (Cited on pages 84 and 107.)
- [Friston 1994b] K. J. Friston, P. Jezzard and R. Turner. *Analysis of functional MRI time-series*. *Human Brain Mapping*, vol. 1, no. 2, pages 153–171, 1994. (Cited on page 25.)
- [Friston 1994c] K. J. Friston and K. J. Worsley. *Assessing the significance of focal activations using their spatial extent*. *Human Brain Mapping*, vol. 220, pages 210–220, 1994. (Cited on page 53.)
- [Friston 1998] Karl J. Friston, P. Thomas Fletcher, O. Josephs, A. Holmes, M. D. Rugg and R. Turner. *Event-Related fMRI: Characterizing Differential Responses*. *NeuroImage*, vol. 40, no. 7, pages 30–40, 1998. (Cited on pages 42, 84, 107 and 184.)
- [Friston 2005a] Karl J. Friston, William D. Penny and Daniel E. Glaser. *Conjunction revisited*. *NeuroImage*, vol. 25, no. 3, pages 661–7, 2005. (Cited on pages 79 and 85.)

- [Friston 2005b] Karl J. Friston, K. E. Stephan, T. E. Lund, A. Morcom and S. Kiebel. *Mixed-effects and fMRI studies*. *NeuroImage*, vol. 24, no. 1, pages 244–52, 2005. (Cited on page 140.)
- [Frith 2006] Chris D. Frith. *The value of brain imaging in the study of development and its disorders*. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 47, no. 10, pages 979–982, 2006. (Cited on page 79.)
- [Fu 2008] Cynthia H. Y. Fu, Janaina Mourao-Miranda, Sergi G. Costafreda, Akash Khanna, Andre F. Marquand, Steve C. R. Williams and Michael J. Brammer. *Pattern classification of sad facial processing: toward the development of neurobiological markers in depression*. *Biological Psychiatry*, vol. 63, no. 7, pages 656–662, 2008. (Cited on page 60.)
- [Gaa 1996] J. Gaa, S. Warach, P. Wen, V. Thangaraj, P. Wielopolski and R.R. Edelman. *Noninvasive perfusion imaging of human brain tumors with EPISTAR*. *European Radiology*, vol. 6, no. 4, 1996. (Cited on page 57.)
- [Gaillard 2000] W. D. Gaillard, L. Hertz-Pannier, S. H. Mott, A. S. Barnett, D. LeBihan and W. H. Theodore. *Functional anatomy of cognitive development: fMRI of verbal fluency in children and adults*. *Neurology*, vol. 54, no. 1, pages 180–185, 2000. (Cited on page 79.)
- [Gaillard 2001a] W. D. Gaillard, C. B. Grandin and B. Xu. *Developmental aspects of pediatric fMRI: considerations for image acquisition, analysis, and interpretation*. *NeuroImage*, vol. 13, no. 2, pages 239–249, 2001. (Cited on pages 79 and 94.)
- [Gaillard 2001b] W. D. Gaillard, M. Pugliese, C. B. Grandin, S. H. Braniecki, P. Kondapaneni, K. Hunter, B. Xu, J. R. Petrella, L. Balsamo and G. Basso. *Cortical localization of reading in normal children: an fMRI language study*. *Neurology*, vol. 57, no. 1, pages 47–54, 2001. (Cited on pages 79, 80 and 96.)
- [Gaillard 2002] W. D. Gaillard, L. Balsamo, B. Xu, C. B. Grandin, S. H. Braniecki, P. H. Papero, S. Weinstein, J. Conry, P. L. Pearl, B. Sachs, S. Sato, B. Jabbari, L. G. Vezina, C. Frattali and W. H. Theodore. *Language dominance in partial epilepsy patients identified with an fMRI reading task*. *Neurology*, vol. 59, no. 2, pages 256–265, 2002. (Cited on page 79.)
- [Gaillard 2003] William D. Gaillard, Bonnie C. Sachs, Joseph R. Whitnah, Zaira Ahmad, Lyn M. Balsamo, Jeffrey R. Petrella, Suzanne H. Braniecki, Christopher M. McKinney, Kevin Hunter, Benjamin Xu and Cecile B. Grandin. *Developmental aspects of language processing: fMRI of verbal fluency in children and adults*. *Human Brain Mapping*, vol. 18, no. 3, pages 176–185, 2003. (Cited on pages 80, 82, 95 and 105.)

- [Gaillard 2004] W. D. Gaillard, L. Balsamo, B. Xu, C. McKinney, P. H. Papero, S. Weinstein, J. Conry, P. L. Pearl, B. Sachs, S. Sato, L. G. Vezina, C. Frat-tali and W. H. Theodore. *fMRI language task panel improves determination of language dominance*. *Neurology*, vol. 63, no. 8, pages 1403–1408, 2004. (Cited on pages 79, 80, 98 and 102.)
- [Gaillard 2006] W. D. Gaillard, E. N. Moore, D. A. Weber, E. K. Ritzl and M. M. Berl. *Functional magnetic resonance imaging in normal and pathological language development*. In G. Riva, D., Rapin, I., Zardini, editeur, *Language: normal and pathological development*, pages 105–120. John libbe édition, 2006. (Cited on pages 79, 81 and 102.)
- [Gandour 2003] Jack Gandour, Yisheng Xu, Donald Wong, Mario Dzemidzic, Mark Lowe, Xiaojian Li and Yunxia Tong. *Neural correlates of segmental and tonal information in speech perception*. *Human Brain Mapping*, vol. 20, no. 4, pages 185–200, 2003. (Cited on pages 80 and 97.)
- [Gauger 1997] L. M. Gauger, L. J. Lombardino and C. M. Leonard. *Brain morphology in children with specific language impairment*. *Journal of Speech, Language, and Hearing Research*, vol. 40, no. 6, pages 1272–1284, 1997. (Cited on pages 102 and 115.)
- [Ge 2012] Tian Ge, Jianfeng Feng, Derrek P. Hibar, Paul M. Thompson and Thomas E. Nichols. *Increasing power for voxel-wise genome-wide association studies: the random field theory, least square kernel machines and fast permutation procedures*. *NeuroImage*, vol. 63, no. 2, pages 858–873, 2012. (Cited on pages 177 and 186.)
- [Geschwind 1985] N. Geschwind and A. M. Galaburda. *Cerebral lateralization. Biological mechanisms, associations, and pathology: I. A hypothesis and A. program for research*. *Archives of Neurology*, vol. 42, no. 5, pages 428–459, 1985. (Cited on page 101.)
- [Giraud 2008] Anne-Lise Giraud, Katrin Neumann, Anne-Catherine Bachoud-Levi, Alexander W. von Gudenberg, Harald A. Euler, Heinrich Lanfermann and Christine Preibisch. *Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering*. *Brain and Language*, vol. 104, no. 2, pages 190–199, 2008. (Cited on page 118.)
- [González 1999] R. G. González, P. W. Schaefer, F. S. Buonanno, L. H. Schwamm, R. F. Budzik, G. Rordorf, B. Wang, A. G. Sorensen and W. J. Koroshetz. *Diffusion-weighted MR imaging: diagnostic accuracy in patients imaged within 6 hours of stroke symptom onset*. *Radiology*, vol. 210, no. 1, pages 155–162, 1999. (Cited on page 15.)
- [Gorgolewski 2012] Krzysztof J. Gorgolewski, Amos J. Storkey, Mark E. Bastin and Cyril R. Pernet. *Adaptive thresholding for reliable topological inference*

- in single subject fMRI analysis*. *Frontiers in Human Neuroscience*, vol. 6, no. August, page 245, 2012. (Cited on page 183.)
- [Grahn 2008] Jessica A. Grahn, John A. Parkinson and Adrian M. Owen. *The cognitive functions of the caudate nucleus*. *Progress in neurobiology*, vol. 86, no. 3, pages 141–155, 2008. (Cited on page 116.)
- [Gray 1918] Henry Gray. *Anatomy of the human body*. Lea & Febiger, 1918. (Cited on pages 11, 12 and 200.)
- [Grosjean 2008] Bénédicte Grosjean and Lionel Moisan. *A-contrario Detectability of Spots in Textured Backgrounds*. *Journal of Mathematical Imaging and Vision*, vol. 33, no. 3, pages 313–337, 2008. (Cited on page 64.)
- [Guyon 2002] I. Guyon, J. Weston, Stephen Barnhill and Vladimir Vapnik. *Gene selection for cancer classification using support vector machines*. *Machine learning*, pages 1–39, 2002. (Cited on page 60.)
- [Hakyemez 2005] B. Hakyemez, C. Erdogan, I. Ercan, N. Ergin, S. Uysal and S. Atahan. *High-grade and low-grade gliomas: differentiation by using perfusion MR imaging*. *Clinical radiology*, vol. 60, no. 4, pages 493–502, 2005. (Cited on page 147.)
- [Hardyck 1977] C. Hardyck and L. F. Petrinovich. *Left-handedness*. *Psychological Bulletin*, vol. 84, no. 3, pages 385–404, 1977. (Cited on pages 82 and 103.)
- [Harris 2006] Gordon J. Harris, Christopher F. Chabris, Jill Clark, Trinity Urban, Itzhak Aharon, Shelley Steele, Lauren McGrath, Karen Condouris and Helen Tager-Flusberg. *Brain activation during semantic processing in autism spectrum disorders via functional magnetic resonance imaging*. *Brain and cognition*, vol. 61, no. 1, pages 54–68, 2006. (Cited on pages 117 and 118.)
- [Heim 2003] St. Heim, B. Opitz, K. Müller and A. D. Friederici. *Phonological processing during language production: fMRI evidence for A. shared production-comprehension network*. *Cognitive Brain Research*, vol. 16, no. 2, pages 285–296, 2003. (Cited on pages 80 and 97.)
- [Heim 2009] Stefan Heim, Simon B. Eickhoff, Angela D. Friederici and Katrin Amunts. *Left cytoarchitectonic area 44 supports selection in the mental lexicon during language production*. *Brain Structure and Function*, vol. 213, no. 4-5, pages 441–56, 2009. (Cited on page 97.)
- [Heim 2010] Stefan Heim, Marion Grande, Elisabeth Meffert, Simon B. Eickhoff, Helen Schreiber, Juraj Kukulja, Nadim Jon Shah, Walter Huber and Katrin Amunts. *Cognitive levels of performance account for hemispheric lateralisation effects in dyslexic and normally reading children*. *NeuroImage*, vol. 53, no. 4, pages 1346–1358, 2010. (Cited on page 114.)

- [Henson 1999] R.N.A. Henson, C. Buchel, O. Josephs and K.J. Friston. *The slice-timing problem in event-related fMRI*. In *NeuroImage*, volume 9, page 125, 1999. (Cited on page 36.)
- [Herbert 2005] M. R. Herbert, D. A. Ziegler, C. K. Deutsch, L. M. O'Brien, D. N. Kennedy, P. A. Filipek, A. I. Bakardjiev, J. Hodgson, M. Takeoka, N. Makris and V. S. Caviness. *Brain asymmetries in autism and developmental language disorder: A nested whole-brain analysis*. *Brain*, vol. 128, no. Pt 1, pages 213–226, 2005. (Cited on pages 102 and 115.)
- [Herbert 2007] Martha R. Herbert and Tal Kenet. *Brain abnormalities in language disorders and in autism*. *Pediatric Clinics of North America*, vol. 54, no. 3, pages 563–583, vii, 2007. (Cited on pages 101 and 102.)
- [Hertz-Pannier 1997] L. Hertz-Pannier, W. D. Gaillard, S. H. Mott, C. A. Cuenod, S. Y. Bookheimer, S. Weinstein, J. Conry, P. H. Papero, S. J. Schiff, D. Le Bihan and W. H. Theodore. *Noninvasive assessment of language dominance in children and adolescents with functional MRI: A preliminary study*. *Neurology*, vol. 48, no. 4, pages 1003–1012, 1997. (Cited on page 79.)
- [Hickok 2003] Gregory Hickok, Bradley Buchsbaum, Colin Humphries and Tugan Muftuler. *Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt*. *Journal of Cognitive Neuroscience*, vol. 15, no. 5, pages 673–682, 2003. (Cited on page 114.)
- [Hickok 2007] Gregory Hickok and David Poeppel. *The cortical organization of speech processing*. *Nature Reviews. Neuroscience*, vol. 8, no. 5, pages 393–402, 2007. (Cited on pages 80, 96, 114 and 115.)
- [Hillis 2004] Argye E. Hillis, Melissa Work, Peter B. Barker, Michael A. Jacobs, Elisabeth L. Breese and Kristin Maurer. *Re-examining the brain regions crucial for orchestrating speech articulation*. *Brain*, vol. 127, no. Pt 7, pages 1479–1487, 2004. (Cited on page 116.)
- [Hillis 2005] Argye E. Hillis, Melissa Newhart, Jennifer Heidler, Peter Barker, Edward Herskovits and Mahaveer Degaonkar. *The roles of the "visual word form area" in reading*. *NeuroImage*, vol. 24, no. 2, pages 548–559, 2005. (Cited on page 97.)
- [Hillis 2007] Argye E. Hillis. *Aphasia: progress in the last quarter of A. century*. *Neurology*, vol. 69, no. 2, pages 200–213, 2007. (Cited on page 114.)
- [Hirsch 2006] Joy Hirsch. *Brain Mapping for Neurosurgery and Cognitive Neuroscience*. In *Functional MRI: Basic principles and clinical applications*, chapitre 7. 2006. (Cited on page 28.)
- [Hoeffner 2004] Ellen G. Hoeffner, Ian Case, Rajan Jain, Sachin K. Gujar, Gaurang V. Shah, John P. Deveikis, Ruth C. Carlos, B. Gregory Thompson,

- Mark R. Harrigan and Suresh K. Mukherji. *Cerebral perfusion CT: technique and clinical applications*. Radiology, vol. 231, no. 3, pages 632–644, 2004. (Cited on page 18.)
- [Holland 2001] Scott K. Holland, Elena Plante, E. Plante, Anna Weber Byars, Richard H. Strawsburg, R. H. Strawsburg, Vincent J. Schmithorst and Williams S. Ball. *Normal fMRI brain activation patterns in children performing A. verb generation task*. NeuroImage, vol. 14, no. 4, pages 837–843, 2001. (Cited on page 79.)
- [Holland 2007] Scott K. Holland, Vannest Jennifer, Marc Mecoli, Lisa M. Jacola, Jan-Mendelt Tillema, Prasanna R. Karunanayaka, Vincent J. Schmithorst, Weihong Yuan, Elena Plante and Anna W. Byars. *Functional MRI of Language Lateralization During Development in Children*. Int J. Audiol., vol. 46, no. 9, pages 533–551, 2007. (Cited on pages 79, 80, 94, 102 and 119.)
- [Holmes 1998] A. P. Holmes and K. J. Friston. *Generalisability, Random Effects & Population Inference*. Proceedings of Fourth International Conference on Functional Mapping of the Human Brain, NeuroImage 7, S754, 1998. (Cited on pages 38 and 140.)
- [Hrabe 2004] J. Hrabe and D. P. Lewis. *Two analytical solutions for a model of pulsed arterial spin labeling with randomized blood arrival times*. Journal of Magnetic Resonance, vol. 167, no. 1, pages 49–55, 2004. (Cited on page 25.)
- [Huber 1964] P. J. Huber. *Robust estimation of a location parameter*. The Annals of Mathematical Statistics, 1964. (Cited on pages 126 and 127.)
- [Huck 2012] S. Huck, H. U Kerl, M. Al-Zghloul, C. Groden and I Nölte. *Arterial spin labeling at 3.0 Tesla in subacute ischemia: comparison to dynamic susceptibility perfusion*. Clinical Neuroradiology, vol. 22, no. 1, pages 29–37, 2012. (Cited on pages 15, 57, 139 and 163.)
- [Hugdahl 2004] Kenneth Hugdahl, Hilde Gundersen, Cecilie Brekke, Tormod Thomsen, Lars Morten Rimol, Lars Ersland and Jussi Niemi. *fMRI brain activation in a finnish family with specific language impairment compared with a normal control group*. Journal of Speech, Language, and Hearing Research, vol. 47, no. 1, pages 162–172, 2004. (Cited on pages 102, 114 and 116.)
- [Huppertz 2005] Hans-Jürgen Huppertz, C. Grimm, S. Fauser, Jan Kassubek, I. Mader, A. Hochmuth, J. Spreer and Andreas Schulze-Bonhage. *Enhanced visualization of blurred gray-white matter junctions in focal cortical dysplasia by voxel-based 3D MRI analysis*. Epilepsy Research, vol. 67, no. 1-2, pages 35–50, 2005. (Cited on pages 57 and 58.)

- [Indefrey 2004] P. Indefrey and W. J. M. Levelt. *The spatial and temporal signatures of word production components*. *Cognition*, vol. 92, no. 1-2, pages 101–144, 2004. (Cited on page 80.)
- [Jakobson 1951] R. Jakobson, C.G.M. Fant and M. Halle. *Preliminaries to speech analysis: the distinctive features and their correlates*. Cambridge, mit press édition, 1951. (Cited on page 80.)
- [Jäncke 2007] L. Jäncke, Th Siegenthaler, S. Preis and H. Steinmetz. *Decreased white-matter density in A. left-sided fronto-temporal network in children with developmental language disorder: evidence for anatomical anomalies in A. motor-language network*. *Brain and Language*, vol. 102, no. 1, pages 91–98, 2007. (Cited on pages 102, 115 and 116.)
- [Jernigan 1991] T. L. Jernigan, J. R. Hesselink, E. Sowell and P. A. Tallal. *Cerebral structure on magnetic resonance imaging in language- and learning-impaired children*. *Archives of Neurology*, vol. 48, no. 5, pages 539–545, 1991. (Cited on pages 101, 115, 116 and 117.)
- [Johnstone 2006] Tom Johnstone, Kathleen S. Ores Walsh, Larry L. Greischar, Andrew L. Alexander, Andrew S. Fox, Richard J. Davidson and Terrence R. Oakes. *Motion correction and the use of motion covariates in multiple-subject fMRI analysis*. *Human Brain Mapping*, vol. 27, no. 10, pages 779–788, 2006. (Cited on page 42.)
- [Just 2004] Marcel Adam Just, Vladimir L. Cherkassky, Timothy A. Keller and Nancy J. Minshew. *Cortical activation and synchronization during sentence comprehension in high-functioning autism: evidence of underconnectivity*. *Brain*, vol. 127, no. Pt 8, pages 1811–1821, 2004. (Cited on page 117.)
- [Kaiser 2010] Marcus Kaiser, Claus C. Hilgetag and Rolf Kötter. *Hierarchy and dynamics of neural networks*. *Frontiers in Neuroinformatics*, vol. 4, no. August, pages 10–12, 2010. (Cited on pages 10 and 200.)
- [Karnath 2009] Hans-Otto Karnath, Johannes Rüter, André Mandler and Marc Himmelbach. *The anatomy of object recognition–visual form agnosia caused by medial occipitotemporal stroke*. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 29, no. 18, pages 5854–5862, 2009. (Cited on page 97.)
- [Kell 2009] Christian A. Kell, Katrin Neumann, Katharina von Kriegstein, Claudia Posenenske, Alexander W. von Gudenberg, Harald Euler and Anne-Lise Giraud. *How the brain repairs stuttering*. *Brain*, vol. 132, no. Pt 10, pages 2747–2760, 2009. (Cited on page 118.)
- [Keller 2009] Simon S. Keller, Timothy Crow, Anne Foundas, Katrin Amunts and Neil Roberts. *Broca’s area: nomenclature, anatomy, typology and asymme-*

- try*. *Brain and Language*, vol. 109, no. 1, pages 29–48, 2009. (Cited on page 115.)
- [Kiebel 1999] S. J. Kiebel, J. B. Poline, K. J. Friston, A. P. Holmes and K. J. Worsley. *Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model*. *NeuroImage*, vol. 10, no. 6, pages 756–766, 1999. (Cited on page 168.)
- [Kim 1995] Seong-Gi Kim. *Quantification of relative cerebral blood flow change by flow-sensitive alternating inversion recovery (FAIR) technique: Application to functional mapping*. *Magnetic Resonance in Medicine*, vol. 34, no. 3, pages 293–301, 1995. (Cited on page 21.)
- [Kim 2006] Seong-gi Kim and Peter A. Bandettini. *Principles of Functional MRI*. In *Functional MRI: Basic principles and clinical applications*, chapitre 1, pages 3–23. Springer, 2006. (Cited on page 26.)
- [Kleinhans 2008] Natalia M. Kleinhans, Ralph-Axel Müller, David N. Cohen and Eric Courchesne. *Atypical functional lateralization of language in autism spectrum disorders*. *Brain research*, vol. 1221, pages 115–125, 2008. (Cited on page 114.)
- [Klöppel 2012] Stefan Klöppel, Ahmed Abdulkadir, Clifford R. Jack, Nikolaos Koutsouleris, Janaina Mourão Miranda and Prashanthi Vemuri. *Diagnostic neuroimaging across diseases*. *NeuroImage*, vol. 61, no. 2, pages 457–463, 2012. (Cited on pages 16, 58, 59, 60 and 202.)
- [Knaus 2008] Tracey A. Knaus, Andrew M. Silver, Kristen A. Lindgren, Nouchine Hadjikhani and Helen Tager-Flusberg. *fMRI activation during A. language task in adolescents with ASD*. *Journal of the International Neuropsychological Society*, vol. 14, no. 6, pages 967–979, 2008. (Cited on pages 114, 117 and 118.)
- [Krasker 1982] William S. Krasker and Roy E. Welsch. *Efficient bounded-influence regression estimation*. *Journal of the American Statistical Association*, vol. 77, no. 379, pages 595–604, 1982. (Cited on page 127.)
- [Kriegeskorte 2006] Nikolaus Kriegeskorte, Rainer Goebel and Peter Bandettini. *Information-based functional brain mapping*. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 10, pages 3863–3868, 2006. (Cited on page 61.)
- [Kriegeskorte 2007a] Nikolaus Kriegeskorte and Peter Bandettini. *Analyzing for information, not activation, to exploit high-resolution fMRI*. *NeuroImage*, vol. 38, no. 4, pages 649–662, 2007. (Cited on pages 61, 62 and 202.)



- [Kriegeskorte 2007b] Nikolaus Kriegeskorte and Peter Bandettini. *Combining the tools: activation- and information-based fMRI analysis*. *NeuroImage*, vol. 38, no. 4, pages 666–668, 2007. (Cited on pages 61 and 62.)
- [Kriegeskorte 2009] Nikolaus Kriegeskorte, W. Kyle Simmons, Patrick S. F. Bellgowan and Chris I. Baker. *Circular analysis in systems neuroscience: the dangers of double dipping*. *Nature Neuroscience*, vol. 12, no. 5, pages 535–540, 2009. (Cited on pages 59 and 60.)
- [Lambert 2001] E. Lambert and D. Chesnet. *Novlex: une base de données lexicales pour les élèves de primaire*. *L'année psychologique*, vol. 101, pages 277–288, 2001. (Cited on page 82.)
- [Law 2003] Meng Law, Stanley Yang, Hao Wang, James S. Babb, Glyn Johnson, Soonmee Cha, Edmond A. Knopp and David Zagzag. *Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging*. *American Journal of Neuroradiology*, vol. 24, no. 10, pages 1989–1998, 2003. (Cited on page 14.)
- [Lawrence 2009] Natalia S. Lawrence, Fabrice Jollant, Owen O'Daly, Fernando Zelaya and Mary L. Phillips. *Distinct roles of prefrontal cortical subregions in the Iowa Gambling Task*. *Cerebral Cortex*, vol. 19, no. 5, pages 1134–1143, 2009. (Cited on page 43.)
- [Leach 2010] James L. Leach and Scott K. Holland. *Functional MRI in children: clinical and research applications*. *Pediatric Radiology*, vol. 40, no. 1, pages 31–49, 2010. (Cited on page 102.)
- [Lee 2000] Jong Doo Lee, Hee-Joung Kim, Byung In Lee, Ok Joon Kim, Tae Joo Jeon and Min Jung Kim. *Evaluation of ictal brain SPET using statistical parametric mapping in temporal lobe epilepsy*. *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 27, no. 11, pages 1658–1665, 2000. (Cited on page 163.)
- [Leonard 2002] Christiana M. Leonard, Linda J. Lombardino, Katherine Walsh, Mark A. Eckert, Jennifer L. Mockler, Lisa A. Rowe, Sharyl Williams and Cheryl B. DeBose. *Anatomical risk factors that distinguish dyslexia from SLI predict reading skill in normal children*. *Journal of Communication Disorders*, vol. 35, no. 6, pages 501–531, 2002. (Cited on pages 101 and 115.)
- [Li 2009] Xiangrui Li, Zhong-Lin Lu, Arnaud D'Argembeau, Marie Ng and Antoine Bechara. *The Iowa Gambling Task in fMRI images*. *Human Brain Mapping*, vol. 31, no. 3, pages 1–14, 2009. (Cited on pages 9 and 28.)
- [Liégeois 2003] Frédérique Liégeois, Torsten Baldeweg, Alan Connelly, David G. Gadian, Mortimer Mishkin and Faraneh Vargha-Khadem. *Language fMRI*

- abnormalities associated with FOXP2 gene mutation*. Nature Neuroscience, vol. 6, no. 11, pages 1230–1237, 2003. (Cited on pages 114, 117 and 118.)
- [Lindquist 2012] Martin A. Lindquist, Julie Spicer, Iris Asllani and Tor D. Wager. *Estimating and testing variance components in a multi-level GLM*. NeuroImage, vol. 59, no. 1, pages 490–501, 2012. (Cited on page 140.)
- [Liu 2005] Thomas T. Liu and Eric C. Wong. *A signal processing model for arterial spin labeling functional MRI*. NeuroImage, vol. 24, no. 1, pages 207–215, 2005. (Cited on page 37.)
- [Logan 2004] Brent R. Logan and Daniel B. Rowe. *An evaluation of thresholding techniques in fMRI analysis*. NeuroImage, vol. 22, no. 1, pages 95–108, 2004. (Cited on page 53.)
- [Lou 1984] H. C. Lou, L. Henriksen and P. Bruhn. *Focal cerebral hypoperfusion in children with dysphasia and/or attention deficit disorder*. Archives of Neurology, vol. 41, no. 8, pages 825–829, 1984. (Cited on pages 102 and 115.)
- [Luh 1999] Wen Ming Luh, Eric C. Wong, P. A. Bandettini and J. S. Hyde. *QUIPSS II with thin-slice T1 periodic saturation: a method for improving accuracy of quantitative perfusion imaging using pulsed arterial spin labeling*. Magnetic Resonance in Medicine, vol. 41, no. 6, pages 1246–1254, 1999. (Cited on page 24.)
- [Luh 2000] Wen Ming Luh, Eric C. Wong, P. A. Bandettini, B. D. Ward and J. S. Hyde. *Comparison of simultaneously measured perfusion and BOLD signal increases during brain activation with T(1)-based tissue identification*. Magnetic Resonance in Medicine, vol. 44, no. 1, pages 137–143, 2000. (Cited on page 31.)
- [MacIntosh 2010] Bradley J. MacIntosh, Nicola Filippini, Michael A. Chappell, Mark W. Woolrich, Clare E. Mackay and Peter Jezzard. *Assessment of arterial arrival times derived from multiple inversion time pulsed arterial spin labeling MRI*. Magnetic Resonance in Medicine, vol. 63, no. 3, pages 641–647, 2010. (Cited on page 152.)
- [Maisog 2008] José M. Maisog, Erin R. Einbinder, D. Lynn Flowers, Peter E. Turkeltaub and Guinevere F. Eden. *A meta-analysis of functional neuroimaging studies of dyslexia*. Annals of the New York Academy of Sciences, vol. 1145, pages 237–259, 2008. (Cited on pages 114, 117 and 118.)
- [Mandonnet 2007] Emmanuel Mandonnet, Aurélien Nouet, Peggy Gatignol, Laurent Capelle and Hugues Duffau. *Does the left inferior longitudinal fasciculus play a role in language? A brain stimulation study*. Brain, vol. 130, no. Pt 3, pages 623–629, 2007. (Cited on page 75.)

- [Maumet 2012a] Camille Maumet, Pierre Maurel, Jean-Christophe Ferré and Christian Barillot. *A contrario detection of focal brain perfusion abnormalities based on an ASL template*. In Proceedings of the IEEE 9th International Symposium on Biomedical Imaging (ISBI 2012), pages 1176–1179, 2012. (Cited on pages 140, 161 and 164.)
- [Maumet 2012b] Camille Maumet, Pierre Maurel, Jean-Christophe Ferré and Christian Barillot. *A comprehensive framework for the detection of individual brain perfusion abnormalities using Arterial Spin Labeling*. In Proceedings of the 15th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2012), pages 542–549, 2012. (Cited on pages 57, 140, 161 and 164.)
- [Maumet 2012c] Camille Maumet, Pierre Maurel, Jean-Christophe Ferré and Christian Barillot. *Robust Cerebral Blood Flow Map Estimation in Arterial Spin Labeling*. In Proceedings of the Second International Workshop on Multimodal Brain Image Analysis (MBIA), held in conjunction with MICCAI 2012, pages 215–224, Nice, France, 2012. (Cited on page 123.)
- [Maumet 2013a] Camille Maumet, Pierre Maurel, Béatrice Carsin, Jean-Christophe Ferré and Christian Barillot. *Patient-Specific Detection of Perfusion Abnormalities Combining Within- and Between-Subject Variances in ASL*. Rapport de recherche RR-8216, INRIA, 2013. (Cited on page 165.)
- [Maumet 2013b] Camille Maumet, Pierre Maurel, Jean-Christophe Ferré and Christian Barillot. *An a contrario approach for the detection of activated brain areas in fMRI*. In Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM) 21st Annual Meeting & Exhibition, page 3260, Salt Lake City, États-Unis, 2013. (Cited on page 181.)
- [Maumet 2013c] Camille Maumet, Pierre Maurel, Jean-Christophe Ferré and Christian Barillot. *Robust perfusion maps in Arterial Spin Labeling by means of M-estimators*. In Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM) 21st Annual Meeting & Exhibition, page 3037, Salt Lake City, États-Unis, 2013. (Cited on page 123.)
- [Maumet 2013d] Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Béatrice Carsin-Nicol and Christian Barillot. *Patient-specific detection of perfusion abnormalities combining within-subject and between-subject variances in Arterial Spin Labeling*. NeuroImage, page in press, 2013. (Cited on pages 137 and 250.)
- [Mazziotta 2001] J. Mazziotta, A. Toga and A. Evans. *A four-dimensional probabilistic atlas of the human brain*. Journal of the American Medical Informatics Association, vol. 8, no. 5, pages 401–430, 2001. (Cited on pages 150 and 171.)

- [Medina 2007] L. Santiago Medina, Byron Bernal and Jeniffer Ruiz. *Role of functional MR in determining language dominance in epilepsy and nonepilepsy populations: A. Bayesian analysis*. *Radiology*, vol. 242, no. 1, pages 94–100, 2007. (Cited on pages 79 and 102.)
- [Melzer 2011] Tracy R. Melzer, Richard Watts, Michael R. MacAskill, John F. Pearson, Sina Rüeger, Toni L. Pitcher, Leslie Livingston, Charlotte Graham, Ross Keenan, Ajit Shankaranarayanan, David C. Alsop, John C. Dalrymple-Alford and Tim J. Anderson. *Arterial spin labelling reveals an abnormal cerebral perfusion pattern in Parkinson's disease*. *Brain*, vol. 134, no. Pt 3, pages 845–855, 2011. (Cited on page 10.)
- [Mériaux 2006] Sébastien Mériaux, Alexis Roche, Ghislaine Dehaene-Lambertz, Bertrand Thirion and Jean-Baptiste Poline. *Combined permutation test and mixed-effect model for group average analysis in fMRI*. *Human Brain Mapping*, vol. 27, no. 5, pages 402–410, 2006. (Cited on page 140.)
- [Moore 2005] David R. Moore, Joy F. Rosenberg and John S. Coleman. *Discrimination training of phonemic contrasts enhances phonological processing in mainstream school children*. *Brain and Language*, vol. 94, no. 1, pages 72–85, 2005. (Cited on page 81.)
- [Mosso 1881] Angelo Mosso. *Ueber den Kreislauf des Blutes im menschlichen Gehirn*. Leipzig : Vert, 1881. (Cited on page 27.)
- [Mourão Miranda 2011] Janaina Mourão Miranda, David R. Hardoon, Tim Hahn, Andre F. Marquand, Steve C. R. Williams, John Shawe-Taylor and Michael Brammer. *Patient classification as an outlier detection problem: an application of the One-Class Support Vector Machine*. *NeuroImage*, vol. 58, no. 3, pages 793–804, 2011. (Cited on pages 59 and 60.)
- [Mumford 2006] Jeanette A. Mumford, Luis Hernandez-Garcia, Gregory R. Lee and Thomas E. Nichols. *Estimation efficiency and statistical power in arterial spin labeling fMRI*. *NeuroImage*, vol. 33, no. 1, pages 103–114, 2006. (Cited on pages 38, 43, 45, 140, 142 and 160.)
- [Mumford 2009] Jeanette A. Mumford and Thomas E. Nichols. *Simple group fMRI modeling and inference*. *NeuroImage*, vol. 47, no. 4, pages 1469–1475, 2009. (Cited on pages 38, 140, 145, 147, 148, 153, 154, 158 and 164.)
- [Mumford 2012] Jeanette A. Mumford, Benjamin O. Turner, F. Gregory Ashby and Russell A. Poldrack. *Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses*. *NeuroImage*, vol. 59, no. 3, pages 2636–2643, 2012. (Cited on page 60.)
- [Musso 2011] Francesco Musso, Jürgen Brinkmeyer, Daniel Ecker, Markus K. London, Giesela Thieme, Tracy Warbrick, Hans-Jörg Wittsack, Andreas Saleh,

- Wolfgang Greb, Peter de Boer and Georg Winterer. *Ketamine effects on brain function—simultaneous fMRI/EEG during a visual oddball task*. *NeuroImage*, vol. 58, no. 2, pages 508–525, 2011. (Cited on page 9.)
- [Myaskovskey 2013] Artiom Myaskovskey, Yann Gousseau and Michael Lindenbaum. *Beyond Independence: An Extension of the a Contrario Decision Procedure*. *International Journal of Computer Vision*, vol. 101, no. 1, pages 22–44, 2013. (Cited on page 64.)
- [Nichols 2003] Thomas Nichols and Satoru Hayasaka. *Controlling the familywise error rate in functional neuroimaging: a comparative review*. *Statistical Methods in Medical Research*, vol. 12, no. 5, pages 419–446, 2003. (Cited on pages 38 and 53.)
- [Nichols 2005] Thomas Nichols, Matthew Brett, Jesper Andersson, Tor Wager and Jean-Baptiste Poline. *Valid conjunction inference with the minimum statistic*. *NeuroImage*, vol. 25, no. 3, pages 653–660, 2005. (Cited on pages 79 and 85.)
- [Noguchi 2008] T. Noguchi, T. Yoshiura, A. Hiwatashi, O. Togao, K. Yamashita, E. Nagao, T. Shono, M. Mizoguchi, S. Nagata, T. Sasaki, S. O. Suzuki, T. Iwaki, K. Kobayashi, F. Mihara and H. Honda. *Perfusion imaging of brain tumors using arterial spin-labeling: correlation with histopathologic vascular density*. *American Journal of Neuroradiology*, vol. 29, no. 4, pages 688–693, 2008. (Cited on pages 14, 147 and 172.)
- [Obata 2004] T. Obata. *Discrepancies between BOLD and flow dynamics in primary and supplementary motor areas: application of the balloon model to the interpretation of BOLD transients*. *NeuroImage*, vol. 21, no. 1, pages 144–153, 2004. (Cited on page 27.)
- [Ogawa 1990] Seiji Ogawa, Tso-Ming Lee, Asha S. Nayak and Paul Glynn. *Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields*. *Magnetic Resonance in Medicine*, vol. 14, no. 1, pages 68–78, 1990. (Cited on page 29.)
- [Oldfield 1971] R. C. Oldfield. *The assessment and analysis of handedness: the Edinburgh inventory*. *Neuropsychologia*, vol. 9, no. 1, pages 97–113, 1971. (Cited on pages 81 and 103.)
- [Ors 2005] Marianne Ors, Erik Ryding, Magnus Lindgren, Peik Gustafsson, Gösta Blennow and Ingmar Rosén. *SPECT findings in children with specific language impairment*. *Cortex*, vol. 41, no. 3, pages 316–326, 2005. (Cited on pages 102, 114 and 115.)
- [O’Shaughnessy 2008] Elizabeth Stief O’Shaughnessy, Madison M. Berl, Erin N. Moore and William D. Gaillard. *Pediatric functional magnetic resonance*

- imaging (fMRI): issues and applications*. Journal of child neurology, vol. 23, no. 7, pages 791–801, 2008. (Cited on pages 79, 93, 94, 102 and 119.)
- [Ostergaard 1996] L. Ostergaard, R. M. Weisskoff, D. A. Chesler, C. Gyldensted and B. R. Rosen. *High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis*. Magnetic Resonance in Medicine, vol. 36, no. 5, pages 715–725, 1996. (Cited on pages 151 and 171.)
- [Parisse 2009] Christophe Parisse and Christelle Maillart. *Specific language impairment as systemic developmental disorders*. Journal of Neurolinguistics, vol. 22, no. 2, pages 109–122, 2009. (Cited on page 101.)
- [Parkes 2000] Laura M. Parkes, E. A. Moore and Paul S. Tofts. *An Improvement to Perfusion Quantification - Accounting for Blood Water Exchange Time*. International Society of Magnetic Resonance Imaging (ISMRM), vol. 8, 2000. (Cited on page 25.)
- [Parkes 2002] Laura M. Parkes and Paul S. Tofts. *Improved accuracy of human cerebral blood perfusion measurements using arterial spin labeling: accounting for capillary water permeability*. Magnetic Resonance in Medicine, vol. 48, no. 1, pages 27–41, 2002. (Cited on page 25.)
- [Parkes 2003] Laura M. Parkes and John A. Detre. *ASL: Blood Perfusion Measurements Using Arterial Spin Labelling*. In Paul Tofts, editeur, Quantitative MRI of the Brain: Measuring Changes Caused by Disease., chapitre 13. John Wiley & Sons, Ltd, 2003. (Cited on page 150.)
- [Paulesu 1993] E. Paulesu, C. D. Frith and R. S. Frackowiak. *The neural correlates of the verbal component of working memory*. Nature, vol. 362, no. 6418, pages 342–345, 1993. (Cited on page 80.)
- [Paulesu 1996] Eraldo Paulesu, Uta Frith, Margaret Snowling, Alison Gallagher, John Morton, Richard S. J. Frackowiak and Christopher D. Frith. *Is developmental dyslexia a disconnection syndrome?* Brain, vol. 119, no. 1, pages 143–157, 1996. (Cited on page 117.)
- [Penny 2003] William D. Penny and A. J. Holmes. *Random-Effects Analysis*. In Elsevier, editeur, Human Brain Function, chapitre 12, pages 843–850. San Diego, 2003. (Cited on pages 53 and 140.)
- [Peskind 2011] Elaine R. Peskind, Eric C. Petrie, Donna J. Cross, Kathleen Pagulayan, Kathleen McCraw, David Hoff, Kim Hart, Chang-En Yu, Murray a Raskind, David G. Cook and Satoshi Minoshima. *Cerebrocerebellar hypometabolism associated with repetitive blast exposure mild traumatic brain injury in 12 Iraq war Veterans with persistent post-concussive symptoms*. NeuroImage, vol. 54 Suppl 1, pages S76–82, 2011. (Cited on page 9.)

- [Petr 2011] Jan Petr, Jean-Christophe Ferré, Elise Bannier, H el ene Raoult, Jean-Yves Gauvrit and Christian Barillot. *Construction and evaluation of a quantitative arterial spin labeling brain perfusion template at 3T*. In Proceedings of the 8th IEEE International Symposium on Biomedical Imaging (ISBI), pages 1035–1038, 2011. (Cited on page 151.)
- [Petr 2013] Jan Petr, Jean-Christophe Ferr e, H el ene Raoult, Elise Bannier, Jean-Yves Gauvrit and Christian Barillot. *Template-based approach for detecting motor task activation-related hyperperfusion in pulsed ASL data*. Human Brain Mapping, no. in Press, 2013. (Cited on pages 16, 17, 58, 139, 163, 171 and 201.)
- [Pi erart 2005] B. Pi erart, A. Comblain, J. Gr egoire and P. Mousty. *Isadyle : Instruments pour le screening et l’approfondissement des dysfonctionnements du langage chez l’enfant*. Bruxelles, tema  dition, 2005. (Cited on page 81.)
- [Pinkham 2011] Amy Pinkham, James Loughhead, Kosha Ruparel, Wen-Chau Wu, Eve Overton, Raquel Gur and Ruben Gur. *Resting quantitative cerebral blood flow in schizophrenia measured by pulsed arterial spin labeling perfusion MRI*. Psychiatry Research, vol. 194, no. 1, pages 64–72, 2011. (Cited on pages 139 and 163.)
- [Plante 1991] E. Plante, L. Swisher, R. Vance and S. Rapcsak. *MRI findings in boys with specific language impairment*. Brain and Language, vol. 41, no. 1, pages 52–66, 1991. (Cited on page 101.)
- [Poldrack 2008] Russell A. Poldrack. *The role of fMRI in cognitive neuroscience: where do we stand?* Current Opinion in Neurobiology, vol. 18, no. 2, pages 223–227, 2008. (Cited on pages 58 and 59.)
- [Poldrack 2011] Russell A. Poldrack, Jeanette Mumford and Thomas E. Nichols. *Handbook of functional MRI data analysis*. Cambridge University Press, Cambridge, 2011. (Cited on pages 35, 38, 59, 60 and 140.)
- [Preibisch 2003] Christine Preibisch, Katrin Neumann, Peter Raab, Harald A. Euler, Alexander W. von Gudenberg, Heinrich Lanfermann and Anne-Lise Giraud. *Evidence for compensation for stuttering by the right frontal operculum*. NeuroImage, vol. 20, no. 2, pages 1356–1364, 2003. (Cited on page 118.)
- [Preis 1998] S. Preis, L. J ancke, P. Schittler, Y. Huang and H. Steinmetz. *Normal intrasyllabic anatomical asymmetry in children with developmental language disorder*. Neuropsychologia, vol. 36, no. 9, pages 849–855, 1998. (Cited on pages 102 and 115.)
- [Price 2003] Cathy J. Price and Joseph T Devlin. *The myth of the visual word form area*. NeuroImage, vol. 19, no. 3, pages 473–481, 2003. (Cited on page 97.)

- [Price 2010] Cathy J. Price. *The anatomy of language: A review of 100 fMRI studies published in 2009*. Annals of the New York Academy of Sciences, vol. 1191, pages 62–88, 2010. (Cited on pages 114, 115 and 116.)
- [Radanovic 2011] M. Radanovic and L.L. Mansur. Language disturbances in adulthood: New advances from the neurolinguistics perspective. Bentham Science Publishers, 2011. (Cited on pages 73 and 74.)
- [Raichle 2010] Marcus E. Raichle. *Two views of brain function*. Trends in Cognitive Sciences, vol. 14, no. 4, pages 180–190, 2010. (Cited on pages 10 and 27.)
- [Ramsey 2001] N. F. Ramsey, I. E. Sommer, G. J. Rutten and R. S. Kahn. *Combined analysis of language tasks in fMRI improves assessment of hemispheric dominance for language functions in individual subjects*. NeuroImage, vol. 13, no. 4, pages 719–733, 2001. (Cited on pages 79, 98, 102, 108 and 119.)
- [Ramus 2008] Franck Ramus and Gayaneh Szenkovits. *What phonological deficit?* Quarterly journal of experimental psychology (2006), vol. 61, no. 1, pages 129–141, 2008. (Cited on page 117.)
- [Raoult 2011] H el ene Raoult, Jan Petr, Elise Bannier, Aymeric Stamm, Jean-Yves Gauvrit, Christian Barillot and Jean-Christophe Ferr e. *Arterial spin labeling for motor activation mapping at 3T with a 32-channel coil: Reproducibility and spatial accuracy in comparison with BOLD fMRI*. NeuroImage, vol. 58, no. 1, pages 157–167, 2011. (Cited on pages 183 and 185.)
- [Rapin 1983] I. Rapin and DA. Allen. *Developmental language disorders: nosological considerations*. In Editor Kirk U, editeur, Neuropsychology of language, reading and spelling, pages 155–184. Academic Press, New York, 1983. (Cited on page 101.)
- [Rapin 2003] I. Rapin, M. Dunn and D. A. Allen. *Developmental language disorders*. In S.J. Segalowitz, editeur, Handbooks of neuropsychology, pages 593–630. Elsevier, Amsterdam, 2003. (Cited on pages 79, 93, 101 and 102.)
- [Richlan 2009] Fabio Richlan, Martin Kronbichler and Heinz Wimmer. *Functional abnormalities in the dyslexic brain: A quantitative meta-analysis of neuroimaging studies*. Human Brain Mapping, vol. 30, no. 10, pages 3299–3308, 2009. (Cited on pages 117 and 118.)
- [Riva 2000] D. Riva, F. Nichelli and M. Devoti. *Developmental aspects of verbal fluency and confrontation naming in children*. Brain and Language, vol. 71, no. 2, pages 267–284, 2000. (Cited on page 95.)
- [Rolheiser 2011] Tyler Rolheiser, Emmanuel A. Stamatakis and Lorraine K. Tyler. *Dynamic processing in the human language system: synergy between the arcuate fascicle and extreme capsule*. The Journal of neuroscience : the



- official journal of the Society for Neuroscience, vol. 31, no. 47, pages 16949–16957, 2011. (Cited on page 75.)
- [Rosner 1983] B. Rosner. *Percentage points for a generalized ESD many-outlier procedure*. Technometrics, vol. 25, no. 2, pages 165–172, 1983. (Cited on page 148.)
- [Rousseau 2007] F. Rousseau, S. Faisan, F. Heitz, J. P. Armspach, Y. Chevalier, F. Blanc, J. de Seze and L Rumbach. *An a contrario approach for change detection in 3D multimodal images: application to multiple sclerosis in MRI*. Proceedings of the 29th International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), pages 2069–2072, 2007. (Cited on pages xxi, 63, 65, 164, 166 and 167.)
- [Rousseau 2008] F. Rousseau, F. Blanc, J. de Seze, L. Rumbach and J.-P. Armspach. *An a contrario approach for outliers segmentation: Application to Multiple Sclerosis in MRI*. Proceedings of the IEEE 5th International Symposium on Biomedical Imaging (ISBI), vol. 2, no. 1, pages 9–12, 2008. (Cited on pages xxi and 63.)
- [Rousseeuw 2003] P. J. Rousseeuw and A. M. LeRoy. Robust regression and outlier detection. Wiley series in probability and mathematical statistics. Probability and mathematical statistics. Wiley-Interscience, 2003. (Cited on pages 126 and 128.)
- [Roux 2003] Franck-Emmanuel Roux, Kader Boulanouar, Jean-Albert Lotterie, Mehdi Mejdoubi, James P. LeSage and Isabelle Berry. *Language functional magnetic resonance imaging in preoperative assessment of language areas: correlation with direct cortical stimulation*. Neurosurgery, vol. 52, no. 6, pages 1335–1345; discussion 1345–7, 2003. (Cited on pages 79, 98 and 102.)
- [Roy 1890] C. S. Roy and C. S. Sherrington. *On the regulation of the blood-supply of the brain*. The Journal of Physiology, 1890. (Cited on page 27.)
- [Rutten 2002] G. J. M Rutten, N. F. Ramsey, P. C. Van Rijen, H. J. Noordmans and C. W. M. van Veelen. *Development of A. Functional Magnetic Resonance Imaging Protocol for Intraoperative Localization of Critical Temporoparietal Language Areas*. pages 350–360, 2002. (Cited on pages 79, 102 and 108.)
- [Sachs 2003] Bonnie C. Sachs and William D. Gaillard. *Organization of language networks in children: functional magnetic resonance imaging studies*. Current neurology and neuroscience reports, vol. 3, no. 2, pages 157–162, 2003. (Cited on page 79.)
- [Sachs 2008] Olga Sachs, Susanne Weis, Timo Krings, Walter Huber and Tilo Kircher. *Categorical and thematic knowledge representation in the brain: neural correlates of taxonomic and thematic conceptual relations*. Neuropsychologia, vol. 46, no. 2, pages 409–418, 2008. (Cited on pages 79 and 97.)

- [Scarpazza 2013] C. Scarpazza, G. Sartori, M. S. De Simone and A. Mechelli. *When the single matters more than the group: Very high false positive rates in single case Voxel Based Morphometry*. *NeuroImage*, vol. 70, pages 175–188, 2013. (Cited on pages 57, 58 and 177.)
- [Schuff 2011] Norbert Schuff, Yu Zhang, Wang Zhan, Maryann Lenoci, Christopher Ching, Lauren Boreta, Susanne G. Mueller, Zhen Wang, Charles R. Marzmar, Michael W. Weiner and Thomas C. Neylan. *Patterns of altered cortical perfusion and diminished subcortical integrity in posttraumatic stress disorder: an MRI study*. *NeuroImage*, vol. 54 Suppl 1, pages S62–8, 2011. (Cited on page 9.)
- [Schulz 2008] Enrico Schulz, Urs Maurer, Sanne van der Mark, Kerstin Bucher, Silvia Brem, Ernst Martin and Daniel Brandeis. *Impaired semantic processing during sentence reading in children with dyslexia: combined fMRI and ERP evidence*. *NeuroImage*, vol. 41, no. 1, pages 153–168, 2008. (Cited on page 117.)
- [Seizeur 2012] Romuald Seizeur, Nicolas Wiest-Daessle, Sylvain Prima, Camille Maumet, Jean-Christophe Ferre and Xavier Morandi. *Corticospinal tractography with morphological, functional and diffusion tensor MRI: a comparative study of four deterministic algorithms used in clinical routine*. *Surgical and Radiologic Anatomy*, vol. 34, no. 8, pages 709–19, 2012. (Cited on page 250.)
- [Shaywitz 2008] Sally E. Shaywitz and Bennett A. Shaywitz. *Paying attention to reading: the neurobiology of reading and dyslexia*. *Development and Psychopathology*, vol. 20, no. 4, pages 1329–1349, 2008. (Cited on page 117.)
- [Shiffler 1988] R. E. Shiffler. *Maximum Z scores and outliers*. *The American Statistician*, 1988. (Cited on page 126.)
- [Sidaros 2005] Karam Sidaros, Kern Olofsson, Maria J. Miranda and Olaf B. Paulson. *Arterial spin labeling in the presence of severe motion*. *Journal of Cerebral Blood Flow and Metabolism*, 2005. (Cited on page 125.)
- [Singh 2012] Krish D. Singh. *Which "neural activity" do you mean? fMRI, MEG, oscillations and neurotransmitters*. *NeuroImage*, vol. 62, no. 2, pages 1121–1130, 2012. (Cited on page 30.)
- [Skudlarski 1999] P. Skudlarski, R. T. Constable and J. C. Gore. *ROC analysis of statistical methods used in functional MRI: individual subjects*. *NeuroImage*, vol. 9, no. 3, pages 311–329, 1999. (Cited on pages 149, 173 and 185.)
- [Smith 1999] A. M. Smith, B. K. Lewis, U. E. Ruttimann, F. Q. Ye, T. M. Sinnwell, Y. Yang, J. H. Duyn and J. A. Frank. *Investigation of low frequency drift in fMRI signal*. *NeuroImage*, vol. 9, no. 5, pages 526–533, 1999. (Cited on page 44.)

- [Soriano-Mas 2009] Carles Soriano-Mas, Jesus Pujol, Hector Ortiz, Joan Deus, Anna López-Sala and Anna Sans. *Age-related brain structural alterations in children with specific language impairment*. *Human Brain Mapping*, vol. 30, no. 5, pages 1626–1636, 2009. (Cited on pages 102, 115, 116 and 117.)
- [Stamatakis 2002] Emmanuel A. Stamatakis, J. T. Lindsay Wilson, Donald M. Hadley and David J. Wyper. *SPECT imaging in head injury interpreted with statistical parametric mapping*. *Journal of Nuclear Medicine*, vol. 43, no. 4, pages 476–483, 2002. (Cited on page 163.)
- [Steinbrink 2009] C. Steinbrink, H. Ackermann, T. Lachmann and A. Riecker. *Contribution of the anterior insula to temporal auditory processing deficits in developmental dyslexia*. *Human Brain Mapping*, vol. 30, no. 8, pages 2401–2411, 2009. (Cited on page 117.)
- [Stothard 1998] S. E. Stothard, M. J. Snowling, D. V. Bishop, B. B. Chipchase and C. A. Kaplan. *Language-impaired preschoolers: A follow-up into adolescence*. *Journal of Speech, Language, and Hearing Research*, vol. 41, no. 2, pages 407–418, 1998. (Cited on page 101.)
- [Studholme 1999] C. Studholme, D.L.G. Hill and D.J. Hawkes. *An overlap invariant entropy measure of 3D medical image alignment*. *Pattern Recognition*, vol. 32, no. 1, pages 71–86, 1999. (Cited on page 37.)
- [Sugahara 2000] T. Sugahara, Y. Korogi, S. Tomiguchi, Y. Shigematsu, I. Ikushima, T. Kira, L. Liang, Y. Ushio and M. Takahashi. *Posttherapeutic intraaxial brain tumor: the value of perfusion-sensitive contrast-enhanced MR imaging for differentiating tumor recurrence from nonneoplastic contrast-enhancing tissue*. *American Journal of Neuroradiology*, vol. 21, no. 5, pages 901–909, 2000. (Cited on pages 57, 139, 146, 163 and 164.)
- [Szaflarski 2002] J. P. Szaflarski, J. R. Binder, E. T. Possing, K. A. McKiernan, B. D. Ward and T. A. Hammeke. *Language lateralization in left-handed and ambidextrous people: fMRI data*. *Neurology*, vol. 59, no. 2, pages 238–244, 2002. (Cited on page 114.)
- [Tan 2009] Huan Tan, Joseph A. Maldjian, Jeffrey M. Pollock, Jonathan H. Burdette, Lucie Y. Yang, Andrew R. Deibler and Robert A. Kraft. *A fast, effective filtering method for improving clinical pulsed arterial spin labeling MRI*. *Journal of Magnetic Resonance Imaging*, vol. 29, no. 5, pages 1134–1139, 2009. (Cited on pages xvi, 123, 125, 126, 127, 129, 131, 132, 133, 134, 135, 204, 205 and 210.)
- [Taylor 1990] I. Taylor and M. M. Taylor. *Psycholinguistics: learning and using language*. Prentice Hall, 1990. (Cited on page 74.)

- [Thirion 2007] Bertrand Thirion, Philippe Pinel, Sébastien Mériaux, Alexis Roche, Stanislas Dehaene and Jean-Baptiste Poline. *Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses*. NeuroImage, vol. 35, no. 1, pages 105–120, 2007. (Cited on page 140.)
- [Tie 2008] Yanmei Tie, Stephen Whalen, Ralph O. Suarez and Alexandra J. Golby. *Group independent component analysis of language fMRI from word generation tasks*. NeuroImage, vol. 42, no. 3, pages 1214–25, 2008. (Cited on pages 79, 81, 93, 98, 102 and 119.)
- [Tie 2009] Yanmei Tie, Ralph O. Suarez, Stephen Whalen, Alireza Radmanesh, Isaiah H. Norton and Alexandra J. Golby. *Comparison of blocked and event-related fMRI designs for pre-surgical language mapping*. NeuroImage, vol. 47 Suppl 2, no. Suppl 2, pages T107–15, 2009. (Cited on page 79.)
- [Tomblin 1996] J. B. Tomblin, N. L. Records and X. Zhang. *A system for the diagnosis of specific language impairment in kindergarten children*. Journal of Speech, Language, and Hearing Research, vol. 39, no. 6, pages 1284–1294, 1996. (Cited on page 101.)
- [Tuunanen 2006] Pasi I. Tuunanen and Risto A. Kauppinen. *Effects of oxygen saturation on BOLD and arterial spin labelling perfusion fMRI signals studied in a motor activation task*. NeuroImage, vol. 30, no. 1, pages 102–109, 2006. (Cited on page 10.)
- [Tzourio-Mazoyer 2002] N. Tzourio-Mazoyer. *Automated Anatomical Labeling of Activations in SPM Using A. Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain*. NeuroImage, vol. 15, no. 1, pages 273–289, 2002. (Cited on pages 81, 85, 95 and 107.)
- [Tzourio 1994] N. Tzourio, A. Heim, M. Zilbovicius, C. Gerard and B. M. Mazoyer. *Abnormal regional CBF response in left hemisphere of dysphasic children during A. language task*. Pediatric Neurology, vol. 10, no. 1, pages 20–26, 1994. (Cited on pages 102 and 115.)
- [Uddin 2009] Lucina Q. Uddin and Vinod Menon. *The anterior insula in autism: under-connected and under-examined*. Neuroscience and Biobehavioral Reviews, vol. 33, no. 8, pages 1198–1203, 2009. (Cited on page 117.)
- [Uddin 2011] Lucina Q Uddin, Vinod Menon, Christina B Young, Srikanth Ryali, Tianwen Chen, Amirah Khouzam, Nancy J. Minshew and Antonio Y Hardan. *Multivariate searchlight classification of structural magnetic resonance imaging in children and adolescents with autism*. Biological Psychiatry, vol. 70, no. 9, pages 833–841, 2011. (Cited on page 62.)
- [van der Lely 2005] Heather K. J. van der Lely. *Domain-specific cognitive systems: insight from Grammatical-SLI*. Trends in Cognitive Sciences, vol. 9, no. 2, pages 53–59, 2005. (Cited on page 101.)

- [Van Hout 1997] A. Van Hout. *Acquired aphasia in children*. Seminars in Pediatric Neurology, vol. 4, no. 2, pages 102–108, 1997. (Cited on page 101.)
- [VarghaKhadem 1998] F. VarghaKhadem, K. E. Watkins, C. J. Price, John Ashburner, K. J. Alcock, Alan Connelly, Richard S. J. Frackowiak, Karl J. Friston, M. E. Pembrey, M. Mishkin, David G. Gadian and R. E. Passingham. *Neural basis of an inherited speech and language disorder*. Proceedings of the National Academy of Sciences of the United States of America, vol. 95, no. 21, pages 12695–12700, 1998. (Cited on pages 101 and 118.)
- [VarghaKhadem 2005] Faraneh VarghaKhadem, David G. Gadian, Andrew Copp and Mortimer Mishkin. *FOXP2 and the neuroanatomy of speech and language*. Nature Reviews. Neuroscience, vol. 6, no. 2, pages 131–138, 2005. (Cited on pages 115, 117 and 118.)
- [Veit 2006] Thomas Veit, Frédéric Cao and Patrick Bouthemy. *An a contrario Decision Framework for Region-Based Motion Detection*. International Journal of Computer Vision, vol. 68, no. 2, pages 163–178, 2006. (Cited on page 64.)
- [Vigneau 2006] M. Vigneau, V. Beaucousin, P. Y. Hervé, H. Duffau, F. Crivello, O. Houdé, B. Mazoyer and N. Tzourio-Mazoyer. *Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing*. NeuroImage, vol. 30, no. 4, pages 1414–1432, 2006. (Cited on pages 80, 95, 96 and 97.)
- [Vik 2005] Torbjorn Vik, Fabrice Heitz, Izzie Namer and Jean-Paul Arnsperger. *On the modeling, construction, and evaluation of a probabilistic atlas of brain perfusion*. NeuroImage, vol. 24, no. 4, pages 1088–1098, 2005. (Cited on pages 9, 10, 163 and 200.)
- [Vitali 2005] Paolo Vitali, Jubin Abutalebi, Marco Tettamanti, James Rowe, Paola Scifo, Ferruccio Fazio, Stefano F. Cappa and Daniela Perani. *Generating animal and tool names: an fMRI study of effective connectivity*. Brain and Language, vol. 93, no. 1, pages 32–45, 2005. (Cited on page 97.)
- [Viviani 2007a] Roberto Viviani, Petra Beschoner, Katja Ehrhard, Bernd Schmitz and Jan Thöne. *Non-normality and transformations of random fields, with an application to voxel-based morphometry*. NeuroImage, vol. 35, no. 1, pages 121–130, 2007. (Cited on pages 57, 58 and 177.)
- [Viviani 2007b] Roberto Viviani, Petra Beschoner, Tina Jaeckle, Peter Hipp, Jan Kassubek and Bernd Schmitz. *The bootstrap and cross-validation in neuroimaging applications: estimation of the distribution of extrema of random fields for single volume tests, with an application to ADC maps*. Human Brain Mapping, vol. 28, no. 10, pages 1075–1088, 2007. (Cited on pages 57 and 58.)

- [Viviani 2009] Roberto Viviani, Eun-Jin Sim, Hanna Lo, Sven Richter, Sebastian Haffer, Nadine Osterfeld, Jan Thöne and Petra Beschoner. *Components of variance in brain perfusion and the design of studies of individual differences: the baseline study*. NeuroImage, vol. 46, no. 1, pages 12–22, 2009. (Cited on pages xviii, 140, 146 and 152.)
- [Voets 2006] N. L. Voets, J. E. Adcock, D. E. Flitney, T. E. J. Behrens, Y. Hart, R. Stacey, K. Carpenter and P. M. Matthews. *Distinct right frontal lobe activation in language processing following left hemisphere injury*. Brain, vol. 129, no. Pt 3, pages 754–766, 2006. (Cited on page 116.)
- [Wagner 2001] A. D. Wagner, E. J. Paré-Blagoev, J. Clark and R. A. Poldrack. *Recovering meaning: left prefrontal cortex guides controlled semantic retrieval*. Neuron, vol. 31, no. 2, pages 329–338, 2001. (Cited on page 95.)
- [Wang 2003a] Jiongjiong Wang, Geoffrey K. Aguirre, Daniel Y. Kimberg and John A. Detre. *Empirical analyses of null-hypothesis perfusion FMRI data at 1.5 and 4 T*. NeuroImage, vol. 19, no. 4, pages 1449–1462, 2003. (Cited on pages 25 and 169.)
- [Wang 2003b] Jiongjiong Wang, Geoffrey K. Aguirre, Daniel Y. Kimberg, Anne C. Roc, Lin Li and John A. Detre. *Arterial spin labeling perfusion fMRI with very low task frequency*. Magnetic Resonance in Medicine, vol. 49, no. 5, pages 796–802, 2003. (Cited on page 31.)
- [Wang 2011] Yang Wang, Andrew J. Saykin, Josef Pfeuffer, Chen Lin, Kristine M. Mosier, Li Shen, Sungeun Kim and Gary D. Hutchins. *Regional reproducibility of pulsed arterial spin labeling perfusion imaging at 3T*. NeuroImage, vol. 54, no. 2, pages 1188–1195, 2011. (Cited on pages 26, 36, 150 and 171.)
- [Wang 2012] Ze Wang. *Improving cerebral blood flow quantification for arterial spin labeled perfusion MRI by removing residual motion artifacts and global signal fluctuations*. Magnetic Resonance Imaging, 2012. (Cited on pages 142 and 160.)
- [Warburton 1996] Elizabeth Warburton, Richard J. S. Wise, Cathy J. Price, Cornelius Weiller, Uri Hadar, Stuart Ramsay and Richard S. J. Frackowiak. *Noun and verb retrieval by normal subjects*. pages 159–179, 1996. (Cited on page 95.)
- [Warmuth 2003] Carsten Warmuth, Matthias Günther and Claus Zimmer. *Quantification of Blood Flow in Brain Tumors: Comparison of Arterial Spin Labeling and Dynamic Susceptibility-weighted Contrast-enhanced MR Imaging*. Radiology, vol. 228, no. 4, pages 523–532, 2003. (Cited on pages 57, 126, 130, 146 and 164.)
- [Watkins 2002] K. E. Watkins, F. Vargha-Khadem, John Ashburner, R. E. Passingham, Alan Connelly, Karl J. Friston, Richard S. J. Frackowiak, M. Mishkin

- and David G. Gadian. *MRI analysis of an inherited speech and language disorder: structural brain abnormalities*. *Brain*, vol. 125, no. Pt 3, pages 465–478, 2002. (Cited on page 118.)
- [Watkins 2008] K. E. Watkins, S. M. Smith, S. Davis and P. Howell. *Structural and functional abnormalities of the motor system in developmental stuttering*. *Brain*, vol. 131, no. Pt 1, pages 50–59, 2008. (Cited on pages 117 and 118.)
- [Weber 2006] M. A. Weber, S. Zoubaa, M. Schlieter, E. Jüttler, H. B. Huttner, K. Geletneky, C. Ittrich, M. P. Lichy, A. Kroll, J. Debus, F. L. Giesel, M. Hartmann and M. Essig. *Diagnostic performance of spectroscopic and perfusion MRI for distinction of brain tumors*. *Neurology*, vol. 66, no. 12, pages 1899–1906, 2006. (Cited on pages 57, 139, 146, 147, 163 and 164.)
- [Webster 2004] Richard I. Webster and Michael I. Shevell. *Neurobiology of specific language impairment*. *Journal of Child Neurology*, vol. 19, no. 7, pages 471–481, 2004. (Cited on page 101.)
- [Wechsler 1997] D. Wechsler. *Wais-iii: Administration and scoring manual : Wechsler adult intelligence scale—third edition*. Psychological Corporation, 1997. (Cited on page 104.)
- [Wechsler 2003] D. Wechsler. *Intelligence scale for children—fourth edition*. Psychological Corporation, 2003. (Cited on page 104.)
- [Weismer 2005] Susan Ellis Weismer, Elena Plante, Maura Jones and J. Bruce Tomblin. *A functional magnetic resonance imaging investigation of verbal working memory in adolescents with specific language impairment*. *Journal of Speech, Language, and Hearing Research*, vol. 48, no. 2, pages 405–425, 2005. (Cited on pages 102, 114, 115 and 116.)
- [Wernicke 1874] C. Wernicke. *Der aphasische Symptomencomplex. Eine psychologische Studie auf anatomischer Basis*. Breslau: Cohn und Weigert, 1874. (Cited on pages 72 and 74.)
- [Weygandt 2012] Martin Weygandt, Carlo R. Blecker, Axel Schäfer, Kerstin Hackmack, John-Dylan Haynes, Dieter Vaitl, Rudolf Stark and Anne Schienle. *fMRI pattern recognition in obsessive-compulsive disorder*. *NeuroImage*, vol. 60, no. 2, pages 1186–93, 2012. (Cited on page 62.)
- [Wheeler 2013] Hayley M. Wheeler, Michael Mlynash, Manabu Inoue, Aaryani Tipirneni, John Liggins, Greg Zaharchuk, Matus Straka, Stephanie Kemp, Roland Bammer, Maarten G. Lansberg and Gregory W. Albers. *Early Diffusion-Weighted Imaging and Perfusion-Weighted Imaging Lesion Volumes Forecast Final Infarct Size in DEFUSE 2*. *Stroke*, vol. 44, no. 3, pages 681–685, 2013. (Cited on pages 15 and 57.)

- [Whitehouse 2008] Andrew J. Whitehouse and Dorothy V. Bishop. *Cerebral dominance for language function in adults with specific language impairment or autism*. *Brain*, pages awn266+, 2008. (Cited on pages 102 and 114.)
- [Wilke 2002] Marko Wilke, Vincent J. Schmithorst and Scott K. Holland. *Assessment of spatial normalization of whole-brain magnetic resonance images in children*. *Human Brain Mapping*, vol. 17, no. 1, pages 48–60, 2002. (Cited on pages 81, 84 and 95.)
- [Wilke 2003a] M. Wilke, S. K. Holland, J. S. Myseros, V. J. Schmithorst and W. S. Ball. *Functional magnetic resonance imaging in pediatrics*. *Neuropediatrics*, vol. 34, no. 5, pages 225–233, 2003. (Cited on pages 79, 93 and 94.)
- [Wilke 2003b] Marko Wilke. *Automated detection of gray matter malformations using optimized voxel-based morphometry: a systematic approach*. *NeuroImage*, vol. 20, no. 1, pages 330–343, 2003. (Cited on pages 57 and 58.)
- [Wilke 2003c] Marko Wilke. *Bright spots: correlations of gray matter volume with IQ in A. normal pediatric population*. *NeuroImage*, vol. 20, no. 1, pages 202–215, 2003. (Cited on pages 85, 95 and 107.)
- [Wilke 2003d] Marko Wilke, Vincent J. Schmithorst and Scott K. Holland. *Normative pediatric brain data for spatial normalization and segmentation differs from standard adult data*. *Magnetic Resonance in Medicine*, vol. 50, no. 4, pages 749–757, 2003. (Cited on pages 81, 84, 95 and 107.)
- [Wilke 2005] Marko Wilke, Karen Lidzba, Martin Staudt, Karin Buchenau, Wolfgang Grodd and Ingeborg Krägeloh-Mann. *Comprehensive language mapping in children, using functional magnetic resonance imaging: what’s missing counts*. *Neuroreport*, vol. 16, no. 9, pages 915–919, 2005. (Cited on pages 79 and 80.)
- [Wilke 2006] Marko Wilke, K. Lidzba, M. Staudt, K. Buchenau, W. Grodd and I. Kragelohmann. *An fMRI task battery for assessing hemispheric language dominance in children*. *NeuroImage*, vol. 32, no. 1, pages 400–410, 2006. (Cited on pages 79, 80, 93, 94, 96, 98, 102 and 119.)
- [Wilke 2007] Marko Wilke and Karen Lidzba. *LI-tool: A new toolbox to assess lateralization in functional MR-data*. *Journal of Neuroscience Methods*, vol. 163, no. 1, pages 128–136, 2007. (Cited on pages 81, 85 and 108.)
- [Wilke 2008] Marko Wilke, Scott K. Holland, Mekibib Altaye and Christian Gaser. *Template-O-Matic: A toolbox for creating customized pediatric templates*. *NeuroImage*, vol. 41, no. 3, pages 903–913, 2008. (Cited on pages 81, 84, 95, 107 and 119.)



- [Willie 2011] C. K. Willie and K. J. Smith. *Fuelling the exercising brain: a regulatory quagmire for lactate metabolism*. The Journal of Physiology, vol. 589, no. Pt 4, pages 779–780, 2011. (Cited on page 12.)
- [Wintermark 2001] M. Wintermark, J. P. Thiran, P. Maeder, P. Schnyder and R. Meuli. *Simultaneous measurement of regional cerebral blood flow by perfusion CT and stable xenon CT: a validation study*. American Journal of Neuroradiology, vol. 22, no. 5, pages 905–914, 2001. (Cited on page 18.)
- [Wintermark 2005] Max Wintermark, Musa Sesay, Emmanuel Barbier, Katalin Borbély, William P. Dillon, James D. Eastwood, Thomas C. Glenn, Cécile B. Grandin, Salvador Pedraza, Jean-François Soustiel, Tadashi Nariai, Greg Zaharchuk, Jean-Marie Caillé, Vincent Dousset and Howard Yonas. *Comparative overview of brain perfusion imaging techniques*. Stroke, vol. 36, no. 9, pages e83–e99, 2005. (Cited on pages [xix](#), [17](#), [147](#) and [171](#).)
- [Wise 2000] R. J. Wise, D. Howard, C. J. Mummery, P. Fletcher, A. Leff, C. Büchel and S. K. Scott. *Noun imageability and the temporal lobes*. Neuropsychologia, vol. 38, no. 7, pages 985–994, 2000. (Cited on page [97](#).)
- [Wolk 2012] David A. Wolk and John A. Detre. *Arterial spin labeling MRI: an emerging biomarker for Alzheimer’s disease and other neurodegenerative conditions*. Current Opinion in Neurology, vol. 25, no. 4, pages 421–428, 2012. (Cited on page [15](#).)
- [Wong 1997] Eric C. Wong, Richard B. Buxton and Lawrence R. Frank. *Implementation of quantitative perfusion imaging techniques for functional brain mapping using pulsed arterial spin labeling*. NMR in biomedicine, vol. 10, no. 4-5, pages 237–249, 1997. (Cited on pages [21](#) and [43](#).)
- [Wong 1998] Eric C. Wong, Richard B. Buxton and Lawrence R. Frank. *Quantitative imaging of perfusion using a single subtraction (QUIPSS and QUIPSS II)*. Magnetic Resonance in Medicine, vol. 39, no. 5, pages 702–708, 1998. (Cited on pages [24](#), [149](#) and [170](#).)
- [Wong 2006] Eric C. Wong, Matthew Cronin, Wen-Chau Wu, Ben Inglis, Lawrence R. Frank and Thomas T. Liu. *Velocity-selective arterial spin labeling*. Magnetic Resonance in Medicine, vol. 55, no. 6, pages 1334–1341, 2006. (Cited on page [26](#).)
- [Woolrich 2004] Mark W. Woolrich, Timothy E. J. Behrens, Christian F. Beckmann, Mark Jenkinson and Stephen M. Smith. *Multilevel linear modelling for fMRI group analysis using Bayesian inference*. NeuroImage, vol. 21, no. 4, pages 1732–1747, 2004. (Cited on pages [38](#) and [140](#).)
- [Worbe 2012] Yulia Worbe, Caroline Malherbe, Andreas Hartmann, Melanie Péligrini-Issac, Arnaud Messé, Marie Vidailhet, Stéphane Lehericy and

- Habib Benali. *Functional immaturity of cortico-basal ganglia networks in Gilles de la Tourette syndrome*. *Brain*, vol. 135, no. Pt 6, pages 1937–1946, 2012. (Cited on pages 9 and 10.)
- [Worsley 1992] Keith J. Worsley, A. C. Evans, S. Marrett and P. Neelin. *A three-dimensional statistical analysis for CBF activation studies in human brain*. *Journal of Cerebral Blood Flow and Metabolism*, vol. 12, no. 6, pages 900–18, 1992. (Cited on page 53.)
- [Worsley 1996] Keith J. Worsley, S. Marrett, P. Neelin, A. C. Vandal, Karl J. Friston and A. C. Evans. *A unified statistical approach for determining significant signals in images of cerebral activation*. *Human Brain Mapping*, vol. 4, no. 1, pages 58–73, 1996. (Cited on page 169.)
- [Worsley 2002] Keith J. Worsley, C. H. Liao, J. Aston, V Petre, G. H. Duncan, F. Morales and A. C. Evans. *A general statistical analysis for fMRI data*. *NeuroImage*, vol. 15, no. 1, pages 1–15, 2002. (Cited on pages 38 and 140.)
- [Yang 2012] Zhi Yang, Fang Fang and Xuchu Weng. *Recent developments in multivariate pattern analysis for functional MRI*. *Neuroscience bulletin*, vol. 28, no. 4, pages 399–408, 2012. (Cited on pages 58 and 59.)
- [Zaharchuk 2009] Greg Zaharchuk, Roland Bammer, Matus Straka, Ajit Shankaranarayan, David C. Alsop, Nancy J. Fischbein, Scott W. Atlas and Michael E. Moseley. *Arterial spin-label imaging in patients with normal bolus perfusion-weighted MR imaging findings: pilot identification of the borderzone sign*. *Radiology*, vol. 252, no. 3, pages 797–807, 2009. (Cited on pages 149 and 170.)
- [Zaharchuk 2012] Greg Zaharchuk, Ibraheem S. El Mogy, Nancy J. Fischbein and Gregory W. Albers. *Comparison of arterial spin labeling and bolus perfusion-weighted imaging for detecting mismatch in acute stroke*. *Stroke*, vol. 43, no. 7, pages 1843–1848, 2012. (Cited on pages 57, 139 and 164.)
- [Zarahn 1997] Eric Zarahn, Geoffrey K. Aguirre and M. D’Esposito. *Empirical analyses of BOLD fMRI statistics. I. Spatially Unsmoothed Data Collected under Null-Hypothesis Conditions*. *NeuroImage*, vol. 5, no. 3, pages 199–212, 1997. (Cited on pages 25 and 44.)
- [Zatorre 1996] R. J. Zatorre, E. Meyer, A. Gjedde and A. C. Evans. *PET studies of phonetic processing of speech: review, replication, and reanalysis*. *Cerebral Cortex*, vol. 6, no. 1, pages 21–30, 1996. (Cited on page 80.)
- [Zhang 2009] Hui Zhang, Thomas E. Nichols and Timothy D. Johnson. *Cluster mass inference via random field theory*. *NeuroImage*, vol. 44, no. 1, pages 51–61, 2009. (Cited on page 168.)

A total of **352 references**.



# Software

---

The work presented in this thesis is available as a software called AutoMRI. This tool, based on SPM8<sup>2</sup>, defines a set of pipelines (cf. fig. 10.8) implemented as loosely coupled components.

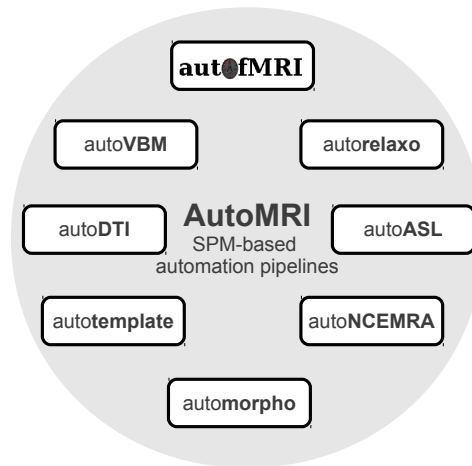


Figure 10.8: AutoMRI components.

Based on a predefined data organisation, each pipeline automatically access the data and update the data tree as needed as described in fig. 10.9. The result of each analysis is summarized in a PDF report.

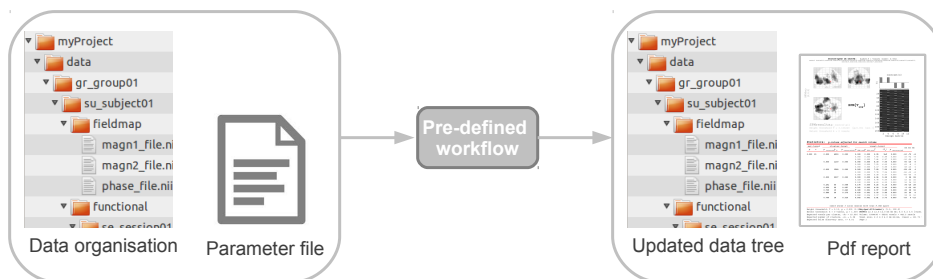


Figure 10.9: AutoMRI: workflow.

<sup>2</sup>[www.fil.ion.ucl.ac.uk/spm/software/spm8](http://www.fil.ion.ucl.ac.uk/spm/software/spm8)

This software, along with its user documentation, is currently distributed in the Visages team <sup>3</sup> and the Neurinfo platform<sup>4</sup> and is available on request. From its initial application to the study of specific language impairment [de Guibert 2010, de Guibert 2011], this tool was further expanded and used in multiple clinical (e.g. [Ban-  
nier 2012], [Seizeur 2012]) and methodological projects [Maumet 2013d].

---

<sup>3</sup>[www.irisa.fr/visages](http://www.irisa.fr/visages)

<sup>4</sup>[www.neurinfo.org](http://www.neurinfo.org)

# Publications

---

## Peer-reviewed journals

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot. *An a Contrario Approach for Patient-Specific Detection of Perfusion Abnormalities in Arterial Spin Labelling*. NeuroImage, submitted.

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot. *Robust Estimation of the Cerebral Blood Flow in Arterial Spin Labelling*. Magnetic Resonance Imaging, submitted.

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Béatrice Carsin, Christian Barillot. *Patient-Specific Detection of Perfusion Abnormalities Combining Within- and Between-Subject Variances in Arterial Spin Labelling*. NeuroImage, in press.

Clément de Guibert, Camille Maumet, Pierre Jannin, Jean-Christophe Ferré, Catherine Tréguier, Christian Barillot, Elisabeth Le Rumeur, Catherine Allaire, Arnaud Biraben. *Abnormal functional lateralization and activity of language brain areas in typical specific language impairment (developmental dysphasia)*. Brain, 2011, 134 (Pt 10), pp. 3044-3058.

Clément de Guibert, Camille Maumet, Jean-Christophe Ferré, Pierre Jannin, Arnaud Biraben, Catherine Allaire, Christian Barillot, Elisabeth Le Rumeur. *fMRI language mapping in children: a panel of language tasks using visual and auditory stimulation without reading or metalinguistic requirements*. NeuroImage, 2010, 51 (2), pp. 897-909.

Romuald Seizeur, Nicolas Wiest-Daessle, Sylvain Prima, Camille Maumet, Jean-Christophe Ferré, Xavier Morandi. *Corticospinal tractography with morphological, functional and diffusion tensor MRI: a comparative study of four deterministic algorithms used in clinical routine*. Surg Radiol Anat, 2012, 34 (8), pp. 709-19.

### International peer-reviewed conferences

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot. *A comprehensive framework for the detection of individual brain perfusion abnormalities using Arterial Spin Labelling*. International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2012, Nice, France. pp. 542-549.

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot. *Robust Cerebral Blood Flow Map Estimation in Arterial Spin Labelling*. International Workshop on Multimodal Brain Image Analysis held in conjunction with MICCAI 2012, 2012, Nice, France. pp. 215-224.

Camille Maumet, Pierre Maurel, Elise Bannier, Jean-Christophe Ferré, Christian Barillot. *A contrario detection of focal brain perfusion abnormalities based on an Arterial Spin Labelling template*. 9th IEEE International Symposium on Biomedical Imaging (ISBI), 2012, Barcelona, Spain. pp. 1176-1179.

### Abstract conferences

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot. *An a contrario approach for the detection of activated brain areas in fMRI*. International Society for Magnetic Resonance in Medicine 21st Annual Meeting & Exhibition (ISMRM), 2013, Salt Lake City, United States.

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot. *Robust perfusion maps in Arterial Spin Labelling by means of M-estimators*. International Society for Magnetic Resonance in Medicine 21st Annual Meeting & Exhibition (ISMRM), 2013, Salt Lake City, United States.

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot. *Automatic detection of perfusion abnormalities based on an Arterial Spin Labelling template*. ISMRM scientific workshop on Perfusion MRI, 2012, Netherlands. #8 – Oral.

Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Elise Bannier, Christian Barillot. *Using negative signal in mono-TI pulsed arterial spin labeling to outline pathological increases in arterial transit times*. ISMRM Scientific Workshop on Perfusion MRI, 2012, Amsterdam, Netherlands. #42.

### Software

Camille Maumet. *autoMRI 1.2.1* Dépôt à l'Agence pour la Protection des Programmes, number IDDN.FR.001.130017.000.S.A.2012.000.31230, June 2011.

Updated on September 8, 2013.

---

## Des études de groupe aux analyses individuelles dans l'exploration de la fonction cérébrale en imagerie de perfusion par marquage de spins et en IRM fonctionnelle BOLD

**Résumé:** Cette thèse aborde l'étude de la fonction cérébrale en Imagerie par Résonance Magnétique (IRM) à l'aide de deux séquences : l'IRM fonctionnelle (IRMf) BOLD et l'imagerie de perfusion par marquage de spins (ASL). Dans ce contexte, les analyses de groupe jouent un rôle important dans l'identification des dysfonctionnements globaux associés à une pathologie. D'autre part, les études individuelles, qui fournissent des conclusions au niveau d'un sujet unique, présentent un intérêt croissant. Dans ce travail, nous abordons à la fois les études de groupe et les analyses individuelles.

Dans un premier temps, nous réalisons une analyse de groupe en IRMf BOLD en vue d'étudier la dysphasie chez l'enfant, une pathologie peu explorée en neuroimagerie. Nous mettons ainsi en évidence un fonctionnement et une latéralisation atypiques des aires langagières.

Ensuite, nous nous concentrons sur les analyses individuelles. Nous proposons l'utilisation d'estimateurs robustes pour calculer les cartographies de débit sanguin cérébral en ASL. Puis, nous étudions la validité des hypothèses qui sous-tendent les analyses statistiques standard dans le contexte de l'ASL. Finalement, nous proposons une nouvelle méthode localement multivariée basée sur une approche *a contrario*. La validation de cette nouvelle approche est réalisée dans deux contextes applicatifs : la détection d'anomalies de perfusion en ASL et la détection de zones d'activation en IRMf BOLD.

**Mots clés:** Arterial Spin Labelling, IRM fonctionnelle BOLD, Analyses individuelles, Hétéroscélasticité, Modèle linéaire généralisé, Approches localement multivariées, Approche *a contrario*

### From Group to Patient-Specific Analysis of Brain Function in Arterial Spin Labelling and BOLD Functional MRI

**Abstract:** This thesis deals with the analysis of brain function in Magnetic Resonance Imaging (MRI) using two sequences: BOLD functional MRI (fMRI) and Arterial Spin Labelling (ASL). In this context, group statistical analyses are of great importance in order to understand the general mechanisms underlying a pathology, but there is also an increasing interest towards patient-specific analyses that draw conclusions at the patient level. Both group and patient-specific analyses are studied in this thesis.

We first introduce a group analysis in BOLD fMRI for the study of specific language impairment, a pathology that was very little investigated in neuroimaging. We outline atypical patterns of functional activity and lateralisation in language regions.

Then, we move forward to patient-specific analysis. We propose the use of robust estimators to compute cerebral blood flow maps in ASL. Then, we analyse the validity of the assumptions underlying standard statistical analyses in the context of ASL. Finally, we propose a new locally multivariate statistical method based on an *a contrario* approach and apply it to the detection of atypical patterns of perfusion in ASL and to activation detection in BOLD functional MRI.

**Keywords:** Arterial Spin Labelling, BOLD functional MRI, Patient-specific analysis, Heteroscedasticity, General Linear Model, Locally multivariate procedure, *a contrario* approach

---



