Au-delà des mots et des images, bases neurophysiologiques d'un système sémantique commun à la compréhension des phrases et des scènes visuelles

Anne-Lise Jouen

▶ **To cite this version:**

Anne-Lise Jouen. Au-delà des mots et des images, bases neurophysiologiques d'un système sémantique commun à la compréhension des phrases et des scènes visuelles. Neurosciences [q-bio.NC]. Université Claude Bernard - Lyon I, 2013. Français. NNT : 2013LYO10322 . tel-02613379

HAL Id: tel-02613379

https://theses.hal.science/tel-02613379

Submitted on 20 May 2020

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.

UNIVERSITÉ DE LYON

Université Claude Bernard Lyon 1

THÈSE

En vue de l'obtention du **DIPLOME DE DOCTORAT**

Délivré par *l'Université Claude Bernard – Lyon 1*

École Doctorale *Neurosciences et Cognition (ED 476 – NSCo)*

Discipline *Neurosciences*

Présentée et soutenue par

Anne-Lise Jouen

Le 20 Décembre 2013

Au-delà des mots et des images

Bases neurophysiologiques d'un système sémantique commun à la compréhension des phrases et des scènes visuelles

---

Jury

Directeurs de recherche :	Dr Peter Ford DOMINEY & Dr Jocelyne VENTRE-DOMINEY
Rapporteurs :	Dr Hugues DUFFAU
	Dr Emmanuel MELLET
Autres membres du jury :	Dr Chlöé FARRER-PUJOL
	Dr Pierre FOURNERET

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	**M. François-Noël GILLY**
Vice-président du Conseil d'Administration	M. le Professeur Hamda BEN HADID
Vice-président du Conseil des Etudes et de la Vie Universitaire	M. le Professeur Philippe LALLE
Vice-président du Conseil Scientifique	M. le Professeur Germain GILLET
Directeur Général des Services	M. Alain HELLEU

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur J. ETIENNE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. le Professeur Y. MATILLON
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : M. le Professeur P. FARGE

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. le Professeur F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. FLEURY
Département Chimie Biochimie	Directeur : Mme le Professeur H. PARROT
Département GEP	Directeur : M. N. SIAUVE
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur A. GOLDMAN
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : Mme S. FLECK
Département Sciences de la Terre	Directeur : Mme la Professeure I. DANIEL
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. C. COLLIGNON
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. P. FOURNIER
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. C. VITON
Institut Universitaire de Formation des Maîtres	Directeur : M. A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Administrateur provisoire : M. N. LEBOISNE

“But, said Alice, the world has absolutely no sense, who's stopping us from inventing one?”

— Lewis Carroll, Alice in Wonderland

“The creation of a single world comes from a huge number of fragments and chaos.”

— Hayao Miyazaki

Remerciements

Je tiens tout d'abord à remercier Jocelyne Ventre-Dominey et Peter Ford Dominey qui ont encadré mon travail de thèse. Je les remercie de m'avoir accueillie dans l'équipe dès mon master et d'avoir supporté ma boulimie de connaissances, en soutenant mon envie de suivre un double cursus mais aussi en modérant mes velléités parfois excessives de jeune chercheuse ! Un grand merci pour leurs suggestions durant l'élaboration de ce travail. Je leur sais tout particulièrement gré de m'avoir donné les moyens d'avoir l'ouverture interdisciplinaire que je souhaitais pour mon travail et de m'avoir financée pour le réaliser.

Je tiens ensuite à remercier vivement le Dr Hugues Duffau et le Dr Emmanuel Mellet (Rapporteurs) ainsi que le Dr Chlöé Farrer-Pujol et le Dr Pierre Fourneret qui ont montré leur intérêt pour ce travail en acceptant d'être membres du jury. Un grand merci en particulier au Dr Duffau, dont les travaux m'ont beaucoup inspirée, pour ses conseils lors de sa visite en mars dernier.

J'adresse mes remerciements aux Docteurs Henry Kennedy et Colette Dehay, directeurs successifs de l'Institut, qui ont accueilli notre équipe et m'ont permis de réaliser cette thèse dans les meilleures conditions. Je remercie également le professeur François Vital-Durand, la première personne de l'institut que j'ai rencontrée. Merci à lui pour sa jovialité et ses anecdotes, scientifiques ou non ; j'espère qu'il profitera de sa retraite bien méritée !

Je voudrais également exprimer toute ma reconnaissance à Monica Baciu et Emilie Cousin pour m'avoir encadrée durant mon master 2 : à la première pour avoir suivi mon parcours depuis et à la seconde, pour toutes les connaissances qu'elle m'a apportées sur l'IRMf.

Merci à Christophe Pallier pour avoir prêté son très asymétrique cerveau lors des pilotes de notre expérience IRMf, mais surtout pour ses connaissances en neuroimagerie et sa participation à ce projet par l'intermédiaire de Neurospin. Je dois aussi beaucoup à Timothy Ellmore qui a traversé l'Océan Atlantique pour nous apporter son savoir sur le DTI. Je tiens également à remercier Jean-Philippe Lachaux et Marcela Perrone pour avoir permis l'adaptation de notre protocole EEG en intra-cortical.

Un grand merci également au Dr Carol Madden Lombardi, qui a partagé son bureau avec moi pendant un certain temps : je la remercie pour sa collaboration, ses encouragements et ses conseils avisés. Merci à Michel Hoen, toujours disponible pour me donner des conseils.

Je remercie tous mes collègues et amis de l'Institut Cellules Souches et Cerveau que j'ai été heureuse de côtoyer. C'est en travaillant dans cette unité que j'ai découvert réellement l'esprit d'équipe et l'ambiance d'un laboratoire de recherche. Je garderai un souvenir ému des Noëls du labo, grands moments de convivialité.

Je veux remercier les doctorants -passés et présents- de mon équipe pour les bons moments partagés ensemble (Go Radical Dudes !!) : Stéphane Lallée (my BFF lab, pour les moments de déprime partagés mais surtout pour les franches rigolades et autres confidences), Xavier Hinaut (mon jumeau de thèse avec qui j'ai écumé l'Autriche, la Belgique et la Slovénie…), Pierre Enel (dernière ligne droite pour toi aussi, courage!), Maxime Petit (tu ne m'auras jamais fait autant rêver que le jour où tu as mangé une fourmi… et je suis ravie que tout roule pour toi), Grégoire Pointeau (à force de faire des blagues douteuses, il y en a forcément une drôle !)… Je n'oublie pas tous mes amis des autres équipes en particulier Pauline Aubel et Sophie Voisin (fournisseuses officielles de recettes gourmandes !) ainsi que les membres de la « team Cellules Souches », Pierre Osteil et Vincent Giudice, pour tous les super moments passés ensemble, depuis les débats scientifiques lors des pauses café jusqu'aux soirées jeux (même si j'en sais plus sur la cuniculiculture que je ne le devrais/voudrais !)

Plein de pensées pour mes chers amis de Sciences Co INPG Grenoble, la meilleure année de mes longues études ! Personne ne me comprend mieux que vous : nous avons vogué dans la même galère ! Un grand merci à Guillaume Garreau (GG pour les intimes !) qui a toujours une pensée pour nous depuis son île lointaine, un énorme beso pour ma « chérite » Chloe Gonseth alias Docteur Chlox et surtout un grand merci à Guillaume Lio (GL cette fois !) pour m'avoir aidée à maintes reprises (que ce soit à 3h du matin par téléphone pour terminer des cartes de Kohonen ou, plus récemment, pour m'initier aux mystères du traitement du signal EEG…).

Merci à tous ceux que j'ai sollicités pour la passation de mes expérimentations. Un grand merci à tous ceux que j'aurai oublié de citer mais qui se seront reconnus. Et merci aux lecteurs qui auront le courage de lire cette thèse jusqu'au bout !

Enfin, je terminerai par des remerciements aux personnes qui sont le plus chères à mon cœur…

Merci à Yann pour m'avoir (sup)portée ces trois dernières années (c'est-à-dire pendant une grande partie de ma thèse ! J'espère te rendre la pareille, même à distance). Merci pour tout ce que tu as apporté à ma vie : pour tout ce que tu m'as fait découvrir, pour toutes les belles choses que nous partageons et celles qui sont encore à venir… 私たちの人生を夢見ましょう、そして一緒に夢をかなえましょう。

Je pense que tu reconnaîtras ces quelques paroles, clin d'œil à une source d' « inspiration » que tu m'as fait découvrir récemment, en pleine rédaction de ma thèse…

《 きみ　が　いま　ぼく　を　ささえて , ぼく　が　いま　きみ　を　ささえる

だから　まよい　ながら　も　とも　に　いきていこうよ　みらい　へ　と。

きずついて　こわれ　そうな　ひ　も, なみだ　して　こまらせる　ひ　も

あるけれど　ぼくら　は　それ　を　こえて　いくんだ ,　だれ　より　うえ　を　めざして。

だいじ　な　もの　が　ある　ならば,　まもりぬいて　みせる　から

なく　したくない　もの　に, じぶん　の　すべて　を　かける　よ。

きよう　じゃ　ない　から　とき　に　きずつけ　きずついてゆく,

だけど　いま　なら　すこし　じしん　を　もって　あるける　よ。

きみ　が　いま　ぼく　を　ささえて , ぼく　が　いま　きみ　を　ささえる

だから　まよい　ながら　も　とも　に　いきていこうよ　みらい　へ　と。 》

Enfin (last but not least), je voudrais remercier ma mère, Annie, et lui dédier cette thèse.

Elle est la première lectrice, relectrice (et même parfois co-auteur !) de mes nombreux écrits : depuis ma première rédaction sur mon chat à l'école primaire jusqu'à ce travail de thèse. Sans son aide, son soutien sur tous les plans pendant toutes ces années, je n'en serais pas là aujourd'hui. Je ne te le dis pas assez mais merci pour tout…

Au-delà des mots et des images,

Bases neurophysiologiques d'un système sémantique commun à la compréhension des phrases et des scènes visuelles

Certaines théories du fonctionnement cognitif postulent l'existence d'un système cérébral impliqué dans la compréhension sémantique indépendamment de la modalité d'entrée des stimuli. L'objectif de ce travail de thèse était d'étudier le fonctionnement d'un tel réseau, impliqué à la fois dans la compréhension de phrases et de scènes visuelles, en lien avec la théorie de la cognition incarnée. Dans la littérature, un ensemble d'aires fronto-temporo-pariétales sensorimotrices et associatives sont décrites comme intervenant dans ces processus sémantiques, mais il existe un manque de consensus concernant la nature amodale de ce système et la plupart des travaux existants se sont concentrés sur l'identification de réseaux corticaux impliqués dans les représentations sémantiques, séparément pour l'une ou l'autre des modalités. De plus, les stimuli utilisés dans les protocoles expérimentaux sont généralement moins complexes que les situations interactives auxquelles nous sommes confrontés dans la vie de tous les jours. Une part importante de l'activité mentale humaine réside dans notre capacité à construire des représentations internes riches : ces modèles mentaux, impliqués dans une grande variété de processus cognitifs, nous permettent d'explorer certains souvenirs du passé, de planifier le futur ou encore de comprendre et de s'adapter à une situation en temps réel. Bien que les progrès des techniques d'Imagerie du Tenseur de Diffusion aient rendu possible la visualisation in vivo de fibres de matière blanche dans le cerveau humain, la connectivité du système sémantique amodal a très peu été étudiée jusque-là. Dans ce travail, nous avons utilisé différentes techniques (principalement de neuro-imagerie IRMf, DTI, EEG) pour mettre en évidence les bases neurophysiologiques d'un système sémantique commun impliqué dans la représentation et la compréhension de stimuli complexes verbaux et non-verbaux. Avec notre premier protocole combinant IRMf et DTI, nous nous sommes intéressés aux activations et à la connectivité cérébrales chez 19 sujets sains en train de lire des phrases ou d'observer des images représentant des événements quotidiens. Une analyse de l'activité cérébrale conjointe associée à la compréhension de ces deux types de stimuli a révélé un réseau fronto-temporo-pariétal commun, impliquant le gyrus frontal inférieur, le gyrus précentral, le cortex rétrosplénial, le gyrus temporal moyen avec une activité s'étendant jusqu'à la jonction temporo-pariétale (TPJ) et au lobe pariétal inférieur. La tractographie DTI a révélé une architecture spécifique de fibres de matière blanche, soutenant ce réseau sémantique et qui fait appel principalement aux faisceaux décrits comme la voie ventrale sémantique (IFOF, UF, ILF, MdLF). Notre seconde expérience (protocole comportemental) nous a permis d'étudier les différences interindividuelles dans la capacité à se représenter des phrases présentées visuellement ou auditivement. Nous avons démontré que les individus ne sont pas égaux quant à cette capacité de représentation et que ces différences se reflètent dans des marqueurs comportementaux tels que la facilité de représentation (évaluée par le COR, coefficient de représentabilité) et la vitesse de réponse (TR) ; mais aussi que ces différences interindividuelles trouvent une correspondance avec le nombre de fibres qui composent le MdLF, laissant supposer une implication de ce faisceau dans ces capacités de représentation. Les résultats de ce protocole comportemental, ainsi que ceux de notre troisième protocole en EEG, ont permis de mettre en évidence un effet contextuel particulièrement important pour la création d'une représentation dans les deux modalités : le contexte induit par la présentation d'un premier stimulus (phrase ou image) influence la représentation d'un second stimulus selon que celui-ci est sémantiquement cohérent ou non avec le premier stimulus présenté. Les résultats de notre expérience en EEG (ERPs) ont aussi permis de révéler l'existence de composantes spécifiquement influencées par l'information sémantique contextuelle disponible : il existerait des effets attentionnels précoces modalité-spécifiques et des effets plus tardifs d'intégration sémantique qui pourraient être équivalents pour les stimuli verbaux et non verbaux. Ces différentes données ont permis de mieux caractériser le "réseau sémantique" et de déterminer que la simulation incarnée, à la base de la compréhension sémantique, va bien au-delà d'une simple activation des aires sensorimotrices et d'une implication de la mémoire sémantique proprement dite. Le réseau sémantique recouperait notamment celui du « réseau du mode par défaut », qui joue un rôle dans l'introspection, et il ferait aussi intervenir des zones de convergence associées à de multiples processus cognitifs de haut niveau tels que la mémoire autobiographique, l'analyse de scènes, la création de modèles mentaux mais aussi de processus à forte composante sociale comme la théorie de l'esprit.

Mots clés: *système sémantique, action, cognition incarnée, langage, vision, représentation, information sémantique contextuelle, DTI, IRMf, EEG, différences interindividuelles*

Beyond words and images, neurophysiological basis of a common semantic system for sentences and visual scenes understanding

Certain theories of cognitive function postulate a neural system for processing meaning, independent of the stimulus input modality. The objective of this thesis work, in line with the embodied cognition domain, was to study functionalities of such a network involved in both sentence and visual scene comprehension. In the literature, a wide network of fronto-temporo-parietal sensorimotor and associative areas are described as being involved in this process, and while there's a lack of consensus on the amodal nature of this system, extensive research has focused on identifying distributed cortical systems that participate in meaning representations separately in the visual and language modalities. Moreover, the stimuli used are generally less complex than everyday life situations we meet. However, a significant portion of human mental life is built upon the construction of perceptually and socially rich internal scene representations and these mental models are involved in a large variety of processes for exploring specific memories of the past, planning the future, or understanding current situations. Although diffusion-tensor imagery based techniques makes feasible the visualization of white matter tracts in the human brain, the connectivity of the semantic network has been little studied. Through different experimental protocols involving mainly neuroimaging techniques (fMRI, DTI, EEG), we were able to reveal the neurophysiological basis of this common semantic network involved in the building of representation and comprehension of rich verbal and non-verbal stimuli. With our first experiment, we examined brain activation and connectivity in 19 subjects who read sentences and viewed pictures corresponding to everyday events, in a combined fMRI and DTI study. Conjunction of activity in understanding sentences and pictures revealed a common fronto temporo-parietal network that included inferior frontal gyrus, precentral gyrus, the retrosplenial complex, and medial temporal gyrus extending into the temporo-parietal junction (TPJ) and inferior parietal lobe. DTI tractography revealed a specific architecture of white matter fibers supporting this network which involves principally the pathways described as the ventral semantic route (IFOF, UF, ILF, MdLF). Our second experiment, which is a behavioral protocol, explored interindividual differences in the ability to represent sentences presented in auditory or visual modality. We demonstrated that individuals are not equal in this capacity to represent sentences, these differences were reflected in the effects on behavioral markers including scores of ease of representation (COR) and speed of responses (TR); they are also related to the number of fibers of the MdLF which supposes a role for this fasciculus in capacities of representation. Both the results of this behavioral protocol and results from our third EEG experiment also showed that the contextual effect was significant: the context induced by the presentation of a first stimulus has the ability to influence the representation of a second stimulus when is the second is semantically consistent or not with the first presented stimulus. Our EEG results (ERPs) revealed components influenced by the available semantic information: early attentional effects which could be modality-specific and later semantic integration process common for verbal and non-verbal stimuli. These data help to define a "meaning" network that includes components of recently characterized systems for semantic memory, embodied simulation, and visuo-spatial scene representation. The network substantially overlaps with the "default mode" network implicated as part of a core network of semantic representation, along with brain systems related to the formation of mental models, and reasoning. These data argue that the embodiment of understanding extends well beyond the sensorimotor system to the highest levels of cognition, including autobiographical memory, scene analysis, mental model formation, reasoning and theory of mind.

Keywords: *semantic system, action, embodied cognition, language, vision, representation, semantic contextual information, DTI, fMRI, EEG, interindividual differences*

SOMMAIRE

CHAPITRE 0 : INTRODUCTION GENERALE 1

0.1) DE LA DIFFICULTE A SAISIR LA NATURE ET LA COMPLEXITE DES REPRESENTATIONS CONCEPTUELLES 1

0.2 APERÇU HISTORIQUE ET PHILOSOPHIQUE : D'UNE VISION SEPARATISTE A UNE APPROCHE INTERACTIVE DE LA RELATION CORPS/ESPRIT 4

0.3) CADRE THEORIQUE ET OBJECTIFS DU PRESENT TRAVAIL DE RECHERCHE 11

CHAPITRE I : CADRE THEORIQUE 15

I.1) SYSTEMES IMPLIQUES : QUELQUES RAPPELS DE NOTIONS GENERALES 18

I.1.*1*) RAPPEL DE NOTIONS GENERALES SUR L'ORGANISATION FONCTIONNELLE DU SYSTEME SENSORI-MOTEUR 18

I.1.*2*) RAPPEL DE NOTIONS GENERALES SUR LA PERCEPTION VISUELLE 20

I.1.*3*) RAPPEL DE NOTIONS GENERALES SUR L'ORGANISATION DU LANGAGE 22

- **Réception vs. Production du langage** 24
- **Traitement cérébral de la lecture des mots** 26

I.1.*4*) RAPPEL DE NOTIONS GENERALES SUR LA MEMOIRE SEMANTIQUE ET LA MEMOIRE EPISODIQUE 28

I.2) NATURE DES REPRESENTATIONS ET DU SYSTEME CONCEPTUEL 33

- **Représentations conceptuelles localistes et amodales** 33
- **Représentations conceptuelles amodales et distribuées** 34
- **Représentations conceptuelles flexibles, distribuées et modalités-spécifiques** 35
- **Nature catégorie-spécifique des représentations** 36

I.3) SEMANTIQUE DE L'ACTION : CARACTERISTIQUES D'UN SYSTEME CONCEPTUEL COMMUN A LA COMPREHENSION VERBALE ET NON VERBALE 42

I.3.*1*) EXISTE-T-IL UN RESEAU SEMANTIQUE COMMUN IMPLIQUE DANS LA COMPREHENSION DES PHRASES ET DES IMAGES ? 44

A) TRAVAUX IMPLIQUANT DES STIMULI NON VERBAUX 44

- **« Hypothèse visuelle » vs « Hypothèse d'appariement direct »** 44
- **Le système des neurones miroirs** 45
- **Actions humaines** 48
- **Actions et objets** 49
- **Le réseau de l'observation de l'action** 50
  - *Rôle de l'EBA (« extrastriate body area ») dans la perception du corps* 51
  - *Sillon temporal supérieur (STS) et intégration multisensorielle* 51
  - *L'aire MT et le traitement du mouvement biologique* 52
  - *Jonction temporo-pariétale et représentations corporelles* 53
  - *Cervelet et évaluation de la séquentialité* 53

B) TRAVAUX IMPLIQUANT DES STIMULI VERBAUX 55

- **Travaux comportementaux** 55
- **Travaux en neuroimagerie** 56
  - *MNS et compréhension du langage* 56

- *Un réseau plus large d'aires sensorimotrices* 57
- *Organisation somatotopique* 58
- *Actions impliquant des objets* 60
- **Données neuropsychologiques** 60
- **Eléments verbaux responsables de la simulation sensorimotrice** 61
- *Abstrait vs Concret* 61
- *Verbes vs Noms* 62
- *Verbes et simulation sensorimotrice* 64
- *Influence des facteurs grammaticaux sur les représentations sensorimotrices* 66
- *Autres facteurs langagiers susceptibles d'influencer la simulation* 67

C) TRAVAUX COMPARATIFS STIMULI VERBAUX/NON VERBAUX 68

- **Modèles théoriques** 68
- *Des systèmes distincts pour traiter mots et images : le modèle du double codage de Paivio* 68
- *Système sémantique commun aux deux modalités : les modèles de Caramazza et collègues* 69
- **Travaux comportementaux** 69
- **Travaux en neuroimagerie** 70
- *Différences de traitement sémantique entre images et mots* 70
- *Régions du système sémantique commun aux mots et aux images* 74
- *Stimuli complexes* 74

Objets et actions 74

Phrases et Scènes visuelles 75

I.3.*2*) « OU EST LE SYSTEME SEMANTIQUE ? » 79

- **Les simulations sensori-motrices sont-elles nécessaires à la compréhension sémantique ?** 80
- **un réseau anatomico-fonctionnel étendu : importance des zones de convergence et de l'architecture des fibres de matière blanche** 82
- *Régions pariéto-temporales* 84
- *Régions frontales* 86
- *Connectivité anatomique* 87

CHAPITRE II : PARTIE EXPERIMENTALE 93

PREAMBULE A LA PARTIE EXPERIMENTALE 93

- **Création de protocoles expérimentaux : de la difficulté de normer des stimuli complexes** 93

II.1) PARTIE EXPERIMENTALE 1 : UTILISATION DE L'IRMF ET DU DTI POUR METTRE EN EVIDENCE LE RESEAU SEMANTIQUE COMMUN 97

II.1.*1*) INTRODUCTION 97

- **La méthodologie d'Imagerie par Résonance Magnétique fonctionnelle (IRMf)** 98
- *Principe de l'IRMf* 98
- *Applications aux structures biologiques* 98
- *Avantages et inconvénients de la technique* 99
- **La méthodologie d'Imagerie du Tenseur de Diffusion (DTI)** 100

II.1.*2*) BUTS DE L'EXPERIENCE ET HYPOTHESES 101

II.1.3) MATERIELS ET METHODES 103

- **Sujets** 103

- **Stimuli** 103

a) Création des Stimuli 103

Stimuli Images 103

Stimuli Phrases 103

Stimuli Contrôles 104

b) Création de 2 jeux de stimuli 105

c) Validation et répartition des stimuli 106

Validation des Stimuli Images 106

Validation des Stimuli Phrases 106

Validation du caractère dynamique 107

Suppression des stimuli ambigus 107

Répartition en 2 jeux de données équivalents 107

- **Protocole expérimental** 107

a) Sessions expérimentales 107

b) Procédure expérimentale 108

c) *Dispositif expérimental* 110

d) *Acquisition des données IRMf-DTI* 110

e) *Post-test* 111

II.1.*4*) ANALYSES DES DONNEES 111

A) ANALYSES DES DONNEES IRMF 111

- **Etape de prétraitement spatial** 111
- **Etape de traitement statistique** 113
- **Contrastes** 114

Contrastes Dynamiques/Statiques 114

Contrastes Images/Phrases 115

Analyse par conjonction 115

B) ANALYSES DES DONNEES ICA 116

- **Identification des réseaux à l'échelle du groupe** 116

Analyse en composantes spatiales 116

Regroupement hiérarchique et détermination des classes d'intérêt (COIs) 117

Cartes d'activations des t-scores et sélection des régions d'intérêt (ROIs) 117

- **Mesure de la connectivité fonctionnelle** 118

C) ANALYSES DES DONNEES DTI 118

- **Prétraitements des données DTI** 118
- **Tractographie déterministe des fibres dans l'espace propre de chaque sujet** 119
- **Identification des faisceaux de fibres pour le groupe de sujets et analyses individuelles** 119

II.1.*5*) RESULTATS 120

A) RESULTATS PARTIE I – RESULTATS IRMF 120

- **Analyse de dynamicité** 120
- **Analyses des aspects sémantiques** 121
  - *Phrases vs Contrôles Phrases (Pd + Ps) – 2 Pc* 122
  - *Images vs Contrôles Images (Id + Is) – 2 Ic* 123
  - *Conjonction Phrases et Images : [(Pd + Ps) – 2 Pc] ∩ [(Id + Is) – 2 Ic]* 124
- **Synthèse des résultats IRMf** 127

B) RESULTATS PARTIE II – ICA 128
- **Sélection des ROIs et des NOIs** 128
- **Corrélations à partir des ROIs issues des « Réseau Images » et « Réseau Phrases »** 131
- **Synthèse des résultats ICA** 132

C) RESULTATS PARTIE III – DTI 133
- **Délimitation des régions d'intérêt permettant la tractographie** 133
- **Résultats principaux : Connectivité intrahémisphèrique des faisceaux de la voie « ventrale » sémantique** 135
  - *Connexions fronto-pariétales* 135
  - *Connexions fronto-temporales* 138
  - *Connexions pariéto-temporales* 140
- **Résultats complémentaires** 143
  - *Connectivité du TPJ* 143
  - *Caractérisation du MdLF à travers la comparaison des trajets des différents faisceaux* 145
  - *Connectivité inter-hémisphérique* 149
  - *Connectivité cortico-striatale* 150
- **Synthèse résultats DTI** 150

D) RESULTATS PARTIE IV : RESULTATS COMPORTEMENTAUX 151
- **Mise en évidence du coefficient de représentabilité** 151
- **COR et données de neuroimagerie** 152

II.1.*6*) CONCLUSION ET PERSPECTIVES 155
- **Compréhension et Modèles mentaux** 156

(a) Aires temporo-pariétales et perception des objets et des personnes 156
(b) *Représentation des actions au niveau pariétal et prémoteur* 157
(c) *Réseau du mode par défaut (Default Mode Network) et Introspection* 157
(d) Représentation de la perspective spatiale au niveau du RSC 158
(e) Raisonnement et structure temporelle sous la dominance des régions frontales 158
- **Connectivité et zones de convergence** 159
  - *Jonction temporo-pariétale* 160
  - *Région temporale antérieure* 161
  - *Cortex rétrosplénial* 161
  - *Projections entre le cortex préfrontal et le noyau caudé* 163
  - *Voie ventrale sémantique* 163

SYNTHESE DES RESULTATS PARTIE EXPERIMENTALE 1 – IRMF ET DTI **167**

II.2) PARTIE EXPERIMENTALE 2 : EXPERIENCE COMPORTEMENTALE POUR EVALUER LES DIFFERENCES INDIVIDUELLES DANS LA CAPACITE A SE REPRESENTER UNE PHRASE **169**

II.2.*1*) INTRODUCTION 169

II.2.*2*) BUTS DE L'EXPERIENCE ET HYPOTHESES 172

II.2.*3*) MATERIELS ET METHODES 173
- **Sujets** 173
- **Stimuli** 174

a) Sessions et Conditions 174
Sessions 174
Conditions 174

b) Création de deux versions du protocole et validation des stimuli 175
c) Procédure expérimentale 175
II.2.*4*) ANALYSES DES DONNEES 177
a) Paramètres comportementaux 177
b) Facteurs d'intérêt et statistiques réalisées sur les données issues du protocole expérimental 178
- **Méthode des k-moyennes** 178
- **Analyses ANOVA** 179

II.2.*5*) RESULTATS 180
A) MISE EN EVIDENCE DE PROFILS COR DIFFERENTS 180
B) RESULTATS DES ANALYSES ANOVA 182
- **Effets simples** 182
  - *Effet principal du facteur Groupe (G1, G2, G3)* 182
  - *Effet principal du facteur Modalité (Lecture vs Audition)* 183
  - *Effet principal du facteur Ordre (Phrase1 vs Phrase 2)* 183
  - *Effet principal du facteur Suite (Suite vs NonSuite)* 183
- **Effets d'interactions** 184
  - *Interaction Suite x Ordre* 184
  - *Interaction Suite x Groupe* 185
  - *Interaction Suite x Ordre x Groupe* 187
  - *Interaction Mot x Groupe (interaction spécifique aux TL)* 188
- **Synthèse des résultats** 189

C) ANALYSE DES REPONSES AU QUESTIONNAIRE FINAL 191
- **Difficulté liée à la tâche** 191
- **Nature des représentations** 192
- **Influence de l'expertise en lecture sur les capacités de représentation** 194

II.2.*6*) CONCLUSION ET PERSPECTIVES 194
- **Différences interindividuelles et capacités de représentation** 196
- **Perspectives** 199

SYNTHESE DES RESULTATS PARTIE EXPERIMENTALE 2 – EXPERIENCE COMPORTEMENTALE 202
II.3) PARTIE EXPERIMENTALE 3 : UTILISATION DE L'EEG POUR METTRE EN EVIDENCE LES PROCESSUS TEMPORELS COMMUNS AUX STIMULI VERBAUX ET NON-VERBAUX AINSI QUE LES EFFETS LIES AUX INFORMATIONS CONTEXTUELLES SUR LA CREATION D'UNE REPRESENTATION 203
II.3.*1*) INTRODUCTION 203
- **La méthodologie EEG** 204
  - *La technique des potentiels évoqués (PEs)* 205
  - *Analyses Temps-fréquence* 207
  - *Avantages et Inconvénients des techniques d'analyse EEG* 208
- **Potentiels évoqués et traitement sémantique des mots et des images** 209
  - *la N400* 209
  - *la N300* 212
- **Autres composantes impliquées dans le traitement des informations verbales et non verbales** 213
- **Traitement sémantique et information contextuelle** 214
  - Information contextuelle et stimuli verbaux 215

- Information contextuelle et stimuli non-verbaux 218

II.3.*2*) BUTS DE L'EXPERIENCE ET HYPOTHESES 222

Hypothèses 223

II.3.*3*) MATERIELS ET METHODES 224

- **Sujets** 224
- **Stimuli** 225

a) Sélection des stimuli 225

Stimuli Images 226

Stimuli Phrases 227

b) Validation et répartition des stimuli 228

Validation des Stimuli Images 228

Validation des Stimuli Phrases 229

- **Protocole expérimental** 230

a) Session expérimentale 230

b) Procédure expérimentale 231

c) Acquisition des données EEG 233

II.3.*4*) ANALYSES DES DONNEES EEG 233

- **Prétraitements** 233
- **Analyse en composante indépendante** 233

Analyses statistiques ERPs 234

II.3.*5*) RESULTATS 235

- **ERPs Images Suite vs NonSuite** 235
- **ERPs Phrases Suite vs NonSuite** 238
- **Synthèse et discussion des résultats ERPs** 241
- différences entre Stimulus1 et Stimulus 2 241
- Stimuli Images 241
- Stimuli Phrases 246
- Processus communs et processus modalité-spécifiques 249
- Cohérence contextuelle et traitement sémantique 251

II.3.*6*) CONCLUSION ET PERSPECTIVES 253

- **Analyses FRPs (fixation related potentials)** 253
- **Analyses temps fréquences : mise en évidence des processus temporels amodaux** 259
- **Traitements sémantiques vs non sémantiques** 262
- **Analyses EEG et profils de représentateurs** 263
- **EEG intra-cortical** 264

SYNTHESE DES RESULTATS PARTIE EXPERIMENTALE 3 – EEG **266**

CONCLUSION **267**

III.1) SYNTHESE DES PRINCIPAUX RESULTATS **267**

III.3.*1*) SYSTEME SEMANTIQUE COMMUN A LA COMPREHENSION DES PHRASES ET DES SCENES VISUELLES 267

III.3.*2*) IMPORTANCE DE L'INFORMATION CONTEXTUELLE POUR LA CREATION D'UNE REPRESENTATION 271

III.3.*3*) DIFFERENCES INTERINDIVIDUELLES DANS LES CAPACITES DE REPRESENTATION A PARTIR D'UNE PHRASE 272

III.2) DISCUSSION **274**

III.2.*1*) PERTINENCE DU SYSTEME FRONTO-TEMPORO-PARIETAL DANS L'EVOLUTION DE LA COGNITION SOCIALE 275
III.2.*2*) EVOLUTION ET LANGAGE 276
- **L'hypothèse d'exploitation neuronale** 277
- **La théorie gestuelle de l'évolution du langage** 277

III.2.*3*) MULTIPLES ROLES DU SYSTEME FRONTO-TEMPORO-PARIETAL 280
III.3) PERSPECTIVES : INTERET DE LA COGNITION INCARNEE POUR LA RECHERCHE EN ROBOTIQUE **282**
- **ROBOTS : DES AGENTS « EMBODIED », UN CORPS ADAPTE POUR AFFUTER L'INTELLIGENCE MOTRICE** 283
- **APPRENDRE DE FAÇON NATURELLE EN INTERACTION AVEC L'ENVIRONNEMENT** 285
- **DEVELOPPER LES COMPETENCES SOCIALES DES ROBOTS** 287
- **ROBOTS ET COGNITION INCARNEE** 291
  - *Robots et coopération* 292
  - *Robots et langage : représentations conceptuelles et simulation* 294
  - *Robots « Incarnés » : le début d'une réflexion philosophico-scientifique* 297

REFERENCES BIBLIOGRAPHIQUES **300**

ANNEXES **342**

ANNEXE 1 **342**
ANNEXE 2 **343**
ANNEXE 3 **344**
ANNEXE 4 **345**
ANNEXE 5 **346**
ANNEXE 6 **347**
ANNEXE 7 **348**
ANNEXE 8 **349**
ANNEXE 9 **350**
ANNEXE 10 **351**

Liste des abréviations

AF : faisceau arqué (arcuate fasciculus)

AG : gyrus angulaire (angular gyrus)

AMS : aire motrice supplémentaire

BA : aire de Brodmann (Brodmann area)

BOLD : Blood Oxygen Level Dependent

COI : classe d'intérêt (class of interest)

COR : Coefficient de Représentabilité (Coefficient of Representability)

CPP : cortex pariétal postérieur

DCM : dynamic causal modeling

DL-PFC : cortex préfrontal dorsolatéral

DSC : débit sanguin cérébral (ou rCBF = regional Cerebral Blood Flow)

DTI : Imagerie du Tenseur de Diffusion (diffusion tensor imaging)

EBA : Extrastriate Body Area

EEG : Electroencéphalographie

ERP : potentiel lié à un événement (event related potential)

FA : anisotropie fractionnelle (fractional anisotropy)

FFA : aire fusiforme des visages (Fusiform Face Area)

GB : ganglions de la base

GFi : gyrus frontal inférieur

GFm : gyrus frontal médian

GTm : Gyrus temporal moyen

GTs : gyrus temporal supérieur

GTs : gyrus temporal Supérieur

HD : hémisphère droit

HG : hémisphère gauche

ICA : analyse en composantes indépendantes (independent component analysis)

ICs : composantes indépendantes

IFOF : faisceau fronto-occipital inférieur (inferior fronto-occipital fasciculus)

ILF : faisceau longitudinal inférieur (inferior longitudinal fasciculus)

IRM : imagerie par résonance magnétique (Magnetic Resonance Imaging)

IRMf : Imagerie par résonance magnétique fonctionnelle (functional MRI)

LPI : lobule pariétal inférieur

MdLF : faisceau longitudinal moyen (middle longitudinal fasciculus)

MEG : Magnétoencéphalographie

MNI : Montreal neurological institute

MNS : système des neurones miroirs (Mirror Neurons System)

MT/ V5 : cortex extrastrié temporal moyen

MTG : gyrus temporal moyen (ou GTM)

NOI : réseau d'intérêt (network of interest)

PEs : potentiels évoqués

PMv : cortex prémoteur ventral

PPA : aire parahippocampique des lieux (Parahippocampal place area)

ROI : région d'intérêt (region of interest)

RSC : cortex rétrosplénial

SED : stimulation électrique directe

sICA : analyse en composante indépendante spatiale

SLF : faisceau longitudinal supérieur (superior longitudinal fasciculus)

SPM : statistical parametric mapping

STS : sillon temporal supérieur (superior temporal sulcus)

TEP : tomographie par émission de positons

TMS : stimulation magnétique transcranienne (Transcranial Magnetic Stimulation)

TOM : théorie de l'esprit (theory of Mind)

TPJ : Jonction temporo-pariétale (temporo-parietal junction)

UF : faisceau unciné (unicate fasciculus)

VOI : volume d'intérêt (volume of interest)

VWFA : aire de la forme visuelle des mots (visual word form area)

CHAPITRE 0 : Introduction Générale

0.1) De la difficulté à saisir la nature et la complexité des représentations conceptuelles

« Le cheval et son cavalier galopaient sur la plage. »

Voici une phrase tout à fait banale. Nous avons tous appris à analyser grammaticalement ce type de phrase à l'école : sujet, verbe, complément… Toute analyse syntaxique est complexe quant aux mécanismes cérébraux qu'elle met en jeu, mais l'analyse du contenu sémantique de cette phrase l'est bien plus encore. Certes nous ne nous rendons pas compte de cette complexité du traitement sémantique, pourtant impliqué dans les situations que nous rencontrons au quotidien : quand nous manipulons des objets, quand nous interagissons avec des personnes, lisons un livre ou regardons un film… La compréhension du sens est à la fois commune à tous les êtres humains et propre à chacun. En quelque sorte, elle reflète l'essence même de l'être humain car elle est la base de la communication avec autrui spécifique à notre espèce et elle rend possible les interactions avec les autres ainsi qu'avec notre environnement.

La compréhension sémantique ne comporte pas de règles semblables, à proprement parler, aux règles de grammaire. Si l'apprentissage des règles de syntaxe débute principalement avec l'apprentissage scolaire, en ce qui concerne la compréhension, cela commence dès que s'établit une interaction avec l'environnement, c'est-à-dire dès le plus jeune âge. L'enfant, qui ne parle pas encore, aura déjà à « mettre du sens » sur tout ce qui l'entoure (objets, personnes, actions, situations…). Bien sûr, cela résulte aussi d'un apprentissage : ce sont les régularités dans ces interactions avec l'environnement qui permettront à l'enfant de développer sa compréhension du monde et de construire sa mémoire sémantique grâce à l'élaboration de concepts. Plus tard, avec l'acquisition du langage, l'enfant pourra associer des mots à ces concepts pour exprimer ce qu'il perçoit ou imagine.

Mais revenons à nos moutons, ou plutôt à nos chevaux : *« Le cheval et son cavalier galopaient sur la plage. »* Selon la théorie de la cognition incarnée de Barsalou (Barsalou, 1999), les concepts résultent des expériences sensorielles, motrices, sociales et émotionnelles et traiter un concept correspond à réactiver des expériences vécues avec des exemplaires de l'objet dont on aurait une représentation multimodale. En effet, à la lecture de cette phrase, certains imagineront la couleur de la robe du cheval : noire ou blanche ? D'autres visualiseront plus particulièrement le cavalier : homme ou

femme ? D'autres encore iront jusqu'à se représenter précisément la couleur de la mer et du ciel… En une fraction de seconde et sans effort particulier, chacun aura une compréhension instantanée de cette phrase et se forgera sa propre représentation, différente d'une personne à une autre car élaborée à partir de nos expériences personnelles vécues précédemment. Ce système de compréhension est tellement précis qu'il nous permet non seulement de déchiffrer le sens de la phrase mais aussi d'accéder, éventuellement, à une simulation émotionnelle : depuis les sensations basiques ressenties par le cavalier (l'odeur iodée de la mer, le bruit des vagues, les embruns sur le visage…) jusqu'à des sentiments plus abstraits et complexes comme la joie, la sensation de liberté…

On connaît l'ambiguïté que comportent les stimuli langagiers. Certes, dans le cas d'une phrase, les aspects sémantiques sont étroitement liés aux aspects syntaxiques qui nous renseignent déjà énormément sur le sens de la phrase. Cependant, il faut tenir compte également des données qui proviennent de notre système d'évaluation interne quant à la plausibilité de l'action. *« Le cheval et son cavalier galopaient sur la plage. »* pourrait suggérer qu'ils courent tous les deux côte à côte. Cette représentation, qui pourra surgir dans l'esprit de certains, n'est toutefois pas la situation la plus plausible ; le système de compréhension sémantique en tiendra compte pour évaluer et interpréter la situation. Par ailleurs, avec le langage, il suffit de l'introduction d'une petite variante comme un changement de préposition ou de temps du verbe pour que la phrase prenne un sens tout à fait différent. C'est une des particularités du langage, syntaxe et sémantique étant interdépendantes dans la construction d'une représentation. *« Le cheval et son cavalier galopaient sur la plage. »* et « *Le cheval du cavalier galopait sur la plage* » sont deux phrases à la signification très différente et pourtant les concepts de base sont les mêmes (cheval, cavalier, galoper, plage). Avec cette seconde phrase, il s'élabore de nouvelles prédictions quant à la situation et un scénario, qui va au-delà de la simple compréhension de cette phrase sans contexte, se construit : pourquoi ce cheval galope-t-il tout seul ? Est-ce que le cavalier a fait une chute ? La richesse de notre univers mental fait qu'il ne se limite pas à la compréhension de mots ; il intègre une simulation qui est sensorielle, motrice, émotionnelle mais aussi contextuelle et temporelle.

Regardons maintenant ces quelques images…

Avec les images, pas besoin de se créer une représentation imagée car elle nous est « imposée » et de ce fait, certaines ambiguïtés peuvent être levées. Ainsi, dans la première image : le cheval est

tacheté, le cavalier est une femme… L'image offre assez d'indices pour une compréhension instantanée. Cependant, ce n'est pas suffisant et notre système d'évaluation interne et de simulation est aussi mis en jeu, nous permettant de nous créer un modèle mental précis et d'interpréter la situation. Précédemment, à la lecture de la phrase, nous imaginions peut-être la mer ; ici, dans la deuxième image, nous ne la voyons pas et pourtant nous n'avons aucun doute que l'action se déroule sur une plage. Dans la troisième image, certains éléments présentent une ambiguïté : le cavalier est-il un homme, une femme ? Quelle est la couleur du cheval ? Mais ces informations sont-elles nécessaires pour comprendre la situation ? Il est évident que dans tous les cas, nous sommes à même de saisir la situation.

Malgré certains mécanismes cérébraux spécifiquement liés à la nature verbale ou non des stimuli (temps de lecture, attention…), quel que soit le niveau d'information fourni par la phrase ou l'image et malgré des représentations sémantiques différentes d'un individu à l'autre, nous sommes capables de comprendre le sens de la phrase ou de l'image de façon spontanée. Si la compréhension sémantique peut intervenir avec ou sans langage, nous pouvons nous interroger sur la nature de cette compréhension. Existe-t-il des mécanismes et des zones cérébrales communs à la compréhension sémantique, indépendamment de la façon dont le stimulus est présenté ? Comment, à partir de mécanismes d'analyse et d'indices fournis différents, parvient-on à un système de compréhension unifié ? Enfin comment savoir que j'ai bien la même « sémantique » que mon voisin, étant donné qu'elle est fonction de nos expériences passées, aboutissant à un système conceptuel propre, qui en quelque sorte reflète notre propre vision du monde.

0.2 Aperçu historique et philosophique : d'une vision séparatiste à une approche interactive de la relation corps/esprit

En 1825, F.J Gall écrivait *« Le système nerveux a été connu avec une lenteur qui ne s'explique que par les nombreuses difficultés qui se sont toujours opposées à ce genre de recherche. »*[1] En effet, le cerveau a toujours été un objet d'interrogations, de recherches et de polémiques, les connaissances sur le cerveau ayant été étroitement liées aux croyances religieuses, philosophiques et culturelles ainsi qu'aux progrès scientifiques et technologiques. La science, telle qu'on la définit aujourd'hui, est une construction récente qui date du milieu du 19ème siècle. Toutes les réflexions qui ont précédé l'approche scientifique actuelle proviennent de la philosophie et des spéculations de libres penseurs, écrivains, philosophes, éthiciens, médecins, théologiens… Ils étaient pourtant des précurseurs de l'approche scientifique, puisqu'ils ont inventé, à leur façon et en partant de rien, des méthodes déductives pour mieux comprendre le monde qui les entouraient.

C'est d'autant plus vrai dans le cadre des sciences cognitives où les réflexions actuelles qui concernent la manière dont on perçoit le monde et dont on se le représente s'inscrivent tout à fait dans cette perspective. Elles furent d'abord objets de préoccupations philosophiques, avant d'être aujourd'hui des thèmes neuroscientifiques qui bénéficient des technologies les plus modernes et qui sont largement étudiés et débattus. Cependant des concepts clés en sciences cognitives aujourd'hui, tels que la cognition incarnée, restent encore sujets de controverse et d'interrogations. C'est le cas, en particulier, en ce qui concerne la nature de nos concepts, des processus cérébraux associés à leur création et récupération en mémoire ainsi que le rôle des interactions corps/cerveau dans la création de ces représentations. L'objectif de ce bref rappel n'est donc pas de faire un « historique des neurosciences cognitives » mais plutôt de mettre en évidence dans quel contexte se sont élaborées les conceptions actuelles et de souligner ce qu'elles lui doivent et pourquoi elles sont souvent si polémiques.

Dès la préhistoire, on avait compris que le cerveau était un organe primordial pour la survie de l'être humain (exemples de trépanations remontant à plus de 12000 ans ou descriptions de blessures à la tête avec leurs symptômes au niveau moteur dans un papyrus égyptien de l'Ancien Empire). Toutefois, on a longtemps considéré que ce n'était pas le cerveau mais le cœur qui était le siège de l'intelligence et des sentiments. En Occident, Alcmeon de Crotone (VIe siècle avant notre ère) fut le premier à affirmer que le cerveau gouverne le corps humain et que les sensations sont transmises au cerveau, siège de la pensée. Un peu plus tard, Démocrite parle d'atomes psychiques

[1] F.J Gall, Sur les fonctions du cerveau et sur celles de chacune de ses parties, 1825, p.83

répartis dans le corps et avec des facultés différentes, permettant des échanges « surveillés » par le cerveau entre le corps et le monde extérieur : une conception relativement moderne qui préfigure l'activité nerveuse. Démocrite, tout comme Hippocrate (460-379 av. JC), considérait que le cerveau était le centre de la pensée et de l'intelligence tandis que pour Aristote (384-322 av. JC), le cœur était centre de l'intellect et le cerveau n'avait qu'un rôle de régulateur thermique de l'organisme. Aristote développait cependant une réflexion sur l'âme qui « ne pense jamais sans images », images qui proviennent des organes des sens. De l'Antiquité au Moyen Age, médecins et savants ont essayé de comprendre le fonctionnement biologique du cerveau et ont élaboré des modèles anatomiques. Galien (130-200 apr. JC) qui, comme Platon, faisait du cerveau le siège de la pensée rationnelle a tiré des conclusions erronées sur l'anatomie humaine à partir de ses observations sur des animaux ; mais l'étude du parcours de l'influx nerveux depuis le cerveau et de l'influence des nerfs sur le mouvement musculaire lui a permis de développer la notion de « pneuma psychique » et de démontrer que le cerveau jouait un rôle essentiel dans la commande du corps et dans l'activité mentale ayant pour origine la substance cérébrale elle-même. Il distinguait différentes facultés : motrice, sensible (5 sens) et raisonnable. Cette dernière faculté était subdivisée en imagination, raison et mémoire.

Les interdits et tabous qui caractérisaient le Moyen Age ont freiné les avancées scientifiques et il a fallu attendre la Renaissance, avec Leonard de Vinci (1452-1519) et André Vésale (1515-1564), pour que des descriptions assez précises de l'anatomie humaine et de la morphologie cérébrale soient faites grâce à la pratique de dissections d'animaux et de cadavres humains. Cependant, on n'osait pas remettre en question la doctrine de l'immatérialité de l'âme. Pour René Descartes (1596-1650), il existait une dualité corps-esprit : selon lui, le corps était une machine et les fonctions physiques étaient régies par les « esprits animaux », humeurs envoyées par le cœur et distribuées dans les nerfs pour agir sur les muscles. Les fonctions mentales étaient gouvernées par Dieu mais la glande pinéale servait d'interface pour réaliser la communication entre ces deux types de fonctions. Thomas Willis (1621-1675) distinguait la substance grise qui engendre les esprits animaux et la substance blanche qui les distribue au reste de l'organisme auquel ils communiquent sensibilité et mouvements.
Dans la deuxième moitié du XVIIe siècle, la polémique sur l'âme animale a continué à faire rage parmi les théologiens, physiciens, métaphysiciens, médecins…

Au XVIIIe siècle, la science progresse, les mentalités évoluent. La thèse de l'immatérialité de l'âme a disparu des ouvrages consacrés au cerveau et on se rapproche d'une conception plus biologique :

Cabanis écrivait « le cerveau secrète la pensée comme le foie secrète la bile »[2] ; Julien Onfray de La Mettrie (1709-1751) substituait le terme de cerveau à celui d'âme. Des philosophes tels que Hume, Condillac ou Diderot cessent de prendre comme base le dualisme cartésien : le corps, en interaction avec le monde extérieur, est le point de départ de l'analyse et n'est pas simplement l'incarnation d'un esprit. C'est sur la perception, la sensation et l'émotion que se fonde la connaissance. Pour Locke (1632-1704), le cerveau agissait comme une « unité centrale » vers laquelle arriveraient nos impressions, les objets extérieurs ayant une action sur nos sens. Pour Condillac (1715–1780), des faits psychiques (y compris la réflexion) s'élaboraient dans les sensations et il émit l'idée que le langage n'était pas seulement une expression de la pensée mais qu'il la structurait et en était un élément déterminant.

Dès la fin du XVIIIe siècle, les progrès de la physique ont également fait évoluer les théories de la transmission nerveuse (Luigi Galvani 1737-1798) et les descriptions anatomiques du corps humain, et plus spécialement du cerveau, se sont affinées (Félix Vicq d'Azyr 1748-1794).

Au XIXe siècle, la phrénologie avec Franz Joseph Gall (1758-1828) a révolutionné l'approche biologique. A partir de la morphologie du crâne avec ses protubérances et ses dépressions reflétant les facultés mentales et intellectuelles des individus, il conçut une cartographie corticale des « forces fondamentales, penchants et sentiments » correspondant à 35 régions du crâne. La phrénologie a été fortement décriée mais elle a eu le mérite de poser la question de la fragmentation de l'activité cérébrale, avec différentes fonctions localisées dans des régions cérébrales distinctes. Le physiologiste Pierre Flourens (1794-1867) fut l'un des détracteurs de la théorie de F.J Gall : il défendait l'idée d'un cerveau global, non compartimenté en zones cérébrales distinctes, où l'ensemble du cerveau est impliqué dans les comportements, préfigurant les futures conceptions holistiques du fonctionnement cérébral. « *Il n'y a donc pas des sièges divers pour les diverses facultés, ni pour les diverses sensations. La faculté de sentir, de juger, de vouloir une chose, réside dans le même lieu que celle d'en vouloir une autre ; et conséquemment cette faculté, essentiellement une, réside essentiellement dans un seul organe.* » écrivait Flourens en 1842 (Gazzaniga et al., 2001).

A la fin de XIXème siècle, les travaux de John Hughlings Jackson, de Paul Broca et de Carl Wernicke sur des patients cérébrolésés favorisèrent le retour des conceptions localisationnistes, avec l'idée que le fonctionnement neurologique et cérébral est organisé sous une forme hiérarchique et qu'une lésion focalisée provoque des déficits spécifiques. Paul Broca, s'appuyant sur des observations réalisées sur un patient qui souffrait d'aphasie (alors que la compréhension de ce patient restait intacte) due à une lésion cérébrale superficielle située dans le lobe frontal gauche, concluait qu'une

[2] J. G. Cabanis, Traité du physique et du moral de l'homme, 1802

circonvolution située dans cette partie du cerveau était responsable du langage articulé. Carl Wernicke publia le cas d'un patient ayant une lésion dans la partie postérieure du lobe temporal et présentant un symptôme inverse à celui du patient de Broca : il pouvait parler sans difficulté mais son discours était dépourvu de sens et il n'était pas non plus capable de comprendre le langage écrit ou parlé.

Par ailleurs, un certain nombre de révolutions scientifiques et technologiques dans des domaines tels que la physique, la chimie, la biologie, l'électronique, la génétique ont permis des avancées notables dans la connaissance des systèmes nerveux central et périphérique. Citons les découvertes concernant certaines maladies (James Parkinson 1755-1828), les progrès en matière de biologie cellulaire (Theodor Schwann) avec la découverte de la cellule nerveuse et les travaux sur les réseaux de neurones (Franz Nissls, Camillo Golgi, Ramon y Cajal). Parallèlement, les travaux sur l'évolution des espèces (B. Lamarck, C. Darwin) d'une part et d'autre part les questionnements sur la conscience menées par Charcot, Freud, Janet contribuèrent à remettre en question certaines idées et à ouvrir de nouvelles perspectives philosophiques ou psychologiques. Sorte de visionnaire de la cognition incarnée, Pierre Janet disait déjà : *« Percevoir un fauteuil c'est percevoir le mouvement qu'il faut pour s'asseoir »*.

Tout au long du XXème les recherches scientifiques fondamentales et les innovations technologiques initiées au XIXème siècle se sont poursuivies, ouvrant de nouvelles pistes d'investigation de plus en plus spécialisées et en même temps interdisciplinaires. Citons : les travaux de Brodmann en anatomie clinique qui permirent une meilleure cartographie des aires du cerveau, selon des critères anatomiques et fonctionnels précis ou encore les recherches sur les pathologies (maladie d'Alzheimer et travaux de Penfield sur l'épilepsie) qui apportèrent des informations précieuses sur les mécanismes cérébraux et également la mise en évidence par la neurobiologie de l'existence d'agents chimiques , les neurotransmetteurs, qui jouent un rôle dans la transmission des signaux, ainsi que les travaux du neurophysiologiste R. W Sperry sur la plasticité du cerveau et sur la relation entre conscience et cerveau.

Au début du XXe siècle des courants philosophiques prennent en compte les avancées scientifiques récentes pour poursuivre une réflexion originale. Ainsi la théorie de la Gestalt en rejetant l'antinomie âme et corps a introduit la notion de « structure » qui implique à la fois sujet et objet étroitement liés l'un à l'autre. Elle s'est intéressée à la perception et à la façon dont, face à la complexité de l'environnement, le cerveau s'efforce de mettre en forme, de donner une structure signifiante à ce qu'il perçoit, afin de le simplifier et de l'organiser. La réflexion la plus approfondie appartient au

philosophe Maurice Merleau-Ponty[3] qui, à partir de la notion de schéma corporel, issue de la physiologie et des travaux sur les membres fantômes, a essayé d'apporter des réponses à divers questionnements en se situant entre la neurologie, la psychologie et la phénoménologie. S'inscrivant dans la continuité des réflexions de philosophes tels que Hume, Condillac ou Diderot, privilégiant la notion de corps en interaction avec le monde extérieur, il cherche à comprendre la perception en action et s'intéresse surtout au rapport dynamique entre l'individu qui perçoit et le monde qu'il perçoit. A la différence de Husserl qui postulait que « toute conscience est conscience de quelque chose », Merleau Ponty développe plutôt la thèse selon laquelle « toute conscience est conscience perceptive ». Celui qui perçoit peut faire jouer une multitude de solutions différentes pour appréhender un même objet et les comportements sont « des dialectiques incarnées ». Le champ mental n'apparaît que comme médiateur entre conscience et nature, la perception appréhende la signification de la réalité et en s'exprimant dans le langage ou dans l'action. Loin d'une conception qui réduit la perception à une fonction sensorielle et loin d'une théorie qui privilégie la conscience, il considère que le corps propre est une condition de l'expérience que le sujet a du monde qu'il explore et qu'il réapprend en permanence à voir pour en incorporer les informations de façon dynamique, que la perception permet une ouverture au monde. Dans « Phénoménologie de la perception », il insiste sur l'idée que la perception est responsable du fait que l'univers perçu a un sens. Il y a une corporalité de la conscience et une intentionnalité corporelle. *« Le monde est inséparable du sujet, mais d'un sujet qui n'est rien que projet du monde, et le sujet est inséparable du monde, mais d'un monde qu'il projette lui-même. »* Par ailleurs, il écrit dans « La structure du comportement » (1942) : « *qu'il s'agisse de la compréhension d'un mot ou de la perception des couleurs et des positions spatiales, on ne peut voir dans le fonctionnement nerveux la mise en œuvre de dispositifs préétablis, que les stimuli, en raison de leurs propriétés objectives, viendraient déclencher du dehors* ». Dans « Phénoménologie de la perception », Merleau-Ponty précise que le langage n'est pas l'expression de représentations mentales et qu'il s'appuie sur une activité intentionnelle qui passe par le corps propre. On ne peut dissocier la parole et la pensée : *« La pensée n'est rien d'intérieur, elle n'existe pas hors du monde et hors des mots »*. Dans le même ouvrage, il écrit : *" Le sens d'un cours d'eau, ce mot ne veut rien dire si je ne suppose pas un sujet qui regarde d'un certain lieu vers un autre." (...)*

[3] L'ouvrage collectif « Le corps en acte » publié en 2010 à la suite du Colloque International réunissant neuroscientifiques, psychologues et philosophes et organisé en 2008 au Collège de France, sous la direction d' A. Berthoz et de B. Andrieu à l'occasion du Centenaire de la naissance de Maurice Merleau -Ponty souligne bien l'approche intuitive, très moderne du philosophe qui a mis en évidence le caractère multimodal de la perception et le rôle essentiel joué par l'action et le mouvement dans la perception.

" De même le sens d'une phrase, c'est son propos ou son intention, ce qui suppose encore un point de vue."

Mais les XXe et XXIe siècles voient surtout l'explosion des neurosciences cognitives avec l'émergence des nouvelles techniques de traitement du signal et avec l'apparition de l'imagerie au service de la recherche fondamentale et des pathologies. Cela a permis de répondre à certaines interrogations, tout en en faisant naître de nouvelles, confirmant ou infirmant certaines hypothèses antérieures tant sur le plan anatomique que sur celui de l'activité cérébrale et du diagnostic de certaines pathologies. Si les connaissances de plus en plus précises sur l'anatomie du cerveau, sur la biologie cellulaire, sur le fonctionnement des substrats cérébraux permettent une meilleure compréhension des fonctions cérébrales elles ne suffisent pas à tout expliquer.

On ne peut que constater quel a été et quel est encore aujourd'hui l'impact des démarches et positionnements philosophiques ou scientifiques qui ont marqué les siècles antérieurs mais, si on voit ressurgir des préoccupations qui ne sont pas si éloignées de celles de nos ancêtres, scientifiques et philosophes, elles s'expriment en des termes différents. Selon JP Changeux, *« Le clivage entre activités mentales et neuronales ne se justifie pas. Désormais à quoi bon parler d' «Esprit » ? Il n'y a plus que deux « aspects » d'un seul et même événement que l'on pourra décrire avec des termes empruntés soit au langage du psychologue (ou de l'introspection), soit à celui du neurobiologiste. »* (L'homme neuronal, 1983). Dichotomies entre raisonnement et émotions, entre perception et intelligence, entre corps et esprit, anatomie et comportement n'ont cessé d'exister et d'interpeller les philosophes, les médecins, les chercheurs au cours des siècles ; elles n'ont pas fini de susciter des interrogations. La confrontation entre une conception « globaliste » du fonctionnement du cerveau et l'idée parfois un peu simpliste d'un cerveau « découpé » en différentes zones anatomiques correspondant à des fonctions spécifiques trouve encore des échos dans certains questionnements actuels sur le fonctionnement complexe du cerveau et ce, malgré ou à cause de l'utilisation de la neuroimagerie qui peut tendre à vouloir faire correspondre un comportement à une zone cérébrale.

Stephen Kosslyn, l'un des fondateurs des neurosciences cognitives résume bien le conflit entre localisationnistes et globalistes : *« L'erreur des premiers localisationnistes fut de chercher à faire correspondre comportements et perceptions à des emplacements uniques du cortex. Les comportements et les perceptions sont tous produits par de multiples aires situées dans différentes parties du cerveau. Aussi, la façon de mettre fin à cette controverse consiste-t-elle à tenir compte du fait que les fonctions complexes telles que la perception, la mémoire, le raisonnement et le mouvement sont le résultat d'une multitude de processus sous-jacents qui s'effectuent chacun dans une seule région cérébrale. Effectivement, les habiletés elles-mêmes peuvent typiquement s'exprimer de très nombreuses façons, mettant en jeu différentes combinaisons de processus (…) Toute habileté complexe n'est donc pas le produit d'une seule partie du cerveau. Et en ce sens, les globalistes avaient*

raison. Les fonctions du genre que postulaient les phrénologistes ne sont pas localisées dans une seule région cérébrale. Toutefois les processus élémentaires auxquels il est fait appel pour exprimer ces habiletés ont une localisation précise. Et en ce sens, les localisationnistes avaient raison… » (Gazzaniga et al., 2001).

Au sein même des neurosciences cognitives, bien que les chercheurs postulent pour une approche qui considère l'homme dans sa globalité, il n'est pas si simple d'y parvenir. Les travaux du philosophe Bernard Andrieu apportent un éclairage intéressant sur les liens qui se sont établis entre les différentes disciplines modernes et qui prennent en compte la relation corps-cerveau-esprit. Il résume ainsi l'histoire de l'étude du cerveau *« Le cerveau psychologique est un objet, inventé dans l'histoire de la psychophysiologie, qui trouve aujourd'hui dans les neurosciences de l'action et de l'émotion un renouvellement épistémologique : le corps en mouvement est désormais au centre de la recherche psychologique du sujet vivant. »* Il souligne que, si tout au long du XXème siècle l'épistémologie du corps humain a continué de se heurter à la dispersion des disciplines (psychologie, psychanalyse, biologie, neurologie, sciences cognitives...) et à leur volonté d'affirmer leur indépendance les unes par rapport aux autres en imposant des modèles spécifiques, on sent aujourd'hui l'émergence de propositions allant dans le sens d'une unification interdisciplinaire afin de réunir toutes les connaissances. Selon lui, *« Le temps est venu d'un dialogue constructif entre neurosciences, philosophie, psychologie sur la base d'un principe commun : une connaissance corporelle d'un monde cérébré »*.

Certes, les sciences cognitives ont replacé l'être humain au centre des recherches (l'homme dans sa relation à l'autre et dans son rapport à son environnement) mais elles gagneront encore beaucoup à éviter de considérer le cerveau de façon trop mathématique, physiologiste ou physico-chimique et par là même dans une approche trop déterministe. Il est également important de ne pas négliger la part de réflexion portant sur l'existence humaine, sur le sens de l'existence et sur la liberté des individus dans leur rapport au monde. En effet on ne peut envisager de réfléchir sur le fonctionnement du cerveau humain sans se rappeler que l'être humain est avant tout un être social, fait pour la communication, que notre cerveau a évolué à partir de celui des primates, apparu approximativement il y a 20 millions d'années, et qu'il a été modelé par l'évolution jusqu'à aboutir au cerveau humain tel qu'on le connaît avec toutes ses aptitudes, des plus étonnantes aux plus banales.

0.3) Cadre théorique et objectifs du présent travail de recherche

Aujourd'hui, on ne conçoit plus l'esprit comme un ensemble de fonctions logiques/abstraites mais comme un système biologique, intrinsèquement lié à l'expérience corporelle et interconnecté avec l'action et l'interaction avec d'autres personnes. JP Changeux parle d' « homme neuronal »[4] doté de cette extraordinaire « machine » qu'est le cerveau, aux multiples possibilités combinatoires qui permettent de construire des « représentations » dynamiques, spontanées et éphémères tout à la fois, soit à partir d'une interaction avec son environnement soit à partir d'une « mobilisation » interne. A.Berthoz, quant à lui, utilise l'expression « homme sensible ». Il précise qu'après un siècle où prédominait une certaine conception de la raison qui a privilégié des modèles formels, on redonne leur place aux sens et à l'acte, au vécu : les sens n'ont pas seulement un rôle de capteurs, ils sont aussi fondateurs de l'expérience de l'action. Il y a un lien étroit entre percevoir et agir et la perception est une action simulée, les neurones étant capables de coder le mouvement que nous allons faire ou que quelqu'un d'autre effectue via un mécanisme de simulation incarnée (Gallese, 2008). Pour le neurobiologiste A. Damasio le cerveau ne peut penser sans le corps : il refuse l'idée d'un corps se réduisant à des précablages mécaniques et il insiste sur la capacité du corps à donner du sens à travers son action sur l'environnement ainsi que sa capacité à recevoir des signaux générés par des perceptions, des émotions et qu'il transmet au système neural capable de reconstruire des représentations mentales.

Ce travail s'inscrit dans une démarche qui consiste à s'interroger sur les capacités du cerveau à percevoir, à ressentir, à créer et conserver des représentations. En effet, une part importante de la cognition humaine réside en la capacité à se construire des représentations internes riches car ces représentations sont impliquées dans de nombreuses fonctions cognitives : elles sont utiles au raisonnement, à la rétention et à la récupération de souvenirs, à la planification d'événements futurs ou encore à la compréhension en temps-réel des situations de la vie de tous les jours.
Afin de mieux cerner l'objet d'étude de ce travail, il semble tout d'abord important de repréciser quelques aspects de la cognition incarnée, domaine dans lequel s'inscrit ce travail.

La « cognition incarnée » (encore appelée cognition située) est un champ majeur de la recherche actuelle en sciences cognitives. Elle part du constat que notre cerveau est un organe vivant relié à un corps (lui aussi vivant) et plongé dans un environnement sur lequel il agit. Cette inscription corporelle, vivante et active du cerveau a une incidence capitale sur la pensée (pris dans le sens large

[4] L'homme neuronal, 1983

de « cognition »). Cela signifie que toutes les idées que nous avons, ont une composante corporelle liée à notre système perceptif, émotionnel ou moteur. La théorie de la cognition incarnée s'est développée dans les années 90, avec l'ouvrage fondateur de F.J. Varela, E. Thompson, & E. Rosch *« L'inscription corporelle de l'esprit »* (1993) et la formalisation du terme « cognition incarnée » liée aux travaux de Barsalou. Cependant, le rôle des représentations sensorielles et motrices dans la création des concepts n'est pas une idée nouvelle, comme nous l'avons vu précédemment. Kiefer et Pullvermüler (2012) rappellent que ce débat remonte aux philosophes grecs qui ont spéculé pour savoir, si oui ou non, les concepts sont essentiellement fondés sur nos sens et nos interactions avec l'environnement. Ce débat oppose le point de vue des rationalistes (de Platon à Descartes, en passant par Leibniz et Kant) à celui des empiristes (Aristote et des philosophes plus modernes tels que Locke et Hume). Selon le point de vue rationaliste, les concepts sont des entités mentales fondamentalement distinctes des impressions sensorielles, ils sont basés sur des catégories innées ou sont formés par le raisonnement. Les concepts ne peuvent donc pas refléter les sensations parce que la perception n'est pas assez structurée pour fournir des concepts qui ont du sens. En revanche, pour les philosophes empiristes, tous les concepts sont tirés de l'expérience sensorielle (impressions) et des idées, qui sont des copies des impressions. Ces réflexions philosophiques sur le rôle des représentations sensorielles et motrices dans la création des concepts sont à mettre en parallèle avec un débat similaire dans le domaine de la recherche en neurosciences cognitives contemporaines (débat que nous détaillerons dans le prochain chapitre). Dans ce domaine des sciences cognitives modernes, les concepts sont classiquement envisagés comme des entités mentales abstraites, différents de ce qui est traité grâce aux systèmes cérébraux perceptifs et moteurs (Quillian, 1969; Anderson, 1983; Tyler et Moss, 2001; Pylyshyn, 1984). Les caractéristiques sensorielles ou motrices des objets et des événements sont transformées en une représentation amodale commune et abstraite, où ce qui est lié à l'information sensorimotrice initiale est perdu. Bien que les théories traditionnelles ne nient pas l'implication des systèmes sensorimoteurs dans des tâches faisant appel aux concepts, ils supposent que les implications sensorimotrices, que l'on peut retrouver lors de la compréhension du langage ou de pensées conceptuelles, sont secondaires ou sont des processus concomitants au traitement de l'information conceptuelle qui, elle, serait traitée amodalement dans les cortex associatifs (Machery, 2007; Weiskopf, 2010, McClelland & Rogers, 2003; Rogers et al, 2004).

L'approche plus moderne de la cognition incarnée remet en question ce point de vue classique et suppose l'existence de liens entre les systèmes cérébraux sensorimoteurs d'une part et le système conceptuel d'autre part. Ils supposent que les concepts sont essentiellement fondés sur la perception et l'action (Barsalou et al, 2003 ; Kiefer et Spitzer, 2001 ; Pulvermüller, 2005 ; Martin et Chao, 2001 ;

Lakoff et Johnson, 1999 ; Gallese et Lakoff, 2005). Les représentations des éléments externes et des états internes (perception, proprioception, émotion et introspection), ainsi que des actions, sont les éléments essentiels à l'élaboration des concepts et ainsi jouent un rôle fondamental dans la compréhension et la pensée conceptuelle.

Les concepts sémantiques correspondraient donc à des traces multimodales et on pourrait y accéder, quelle que soit la modalité d'entrée (auditive, visuelle, tactile). Action et représentation ne sont plus interprétées selon une dichotomie classique physique/état mental, mais seraient étroitement liées et ces processus impliqueraient, au moins en partie, les mêmes réseaux cérébraux. Le réseau de la motricité serait engagé quand on effectue une action mais aussi lorsqu'on doit interpréter les actions des autres. Les concepts sémantiques résulteraient ainsi de toutes nos expériences sensorielles, motrices, sociales et émotionnelles et traiter un concept correspondrait à réactiver la trace multimodale (visuelle, auditive, tactile…) de ces expériences vécues auparavant. Agir dans le monde, y interagir avec des objets et des individus, se représenter le monde, le percevoir, le comprendre constitueraient différents niveaux du même lien relationnel entre les organismes et l'environnement où ils agissent et pensent.

L'objectif principal de ce travail s'inscrit dans cette perspective : étudier les corrélats neuroanatomiques d'un réseau sémantique, accessible quelle que soit la modalité d'entrée, et mieux comprendre la nature de nos représentations à travers différentes méthodologies (IRMf, DTI, EEG, expérience comportementale).

Dans un premier chapitre théorique, après avoir fait quelques rappels généraux sur les systèmes impliqués (système sensorimoteur, langage, mémoire…) et sur les différentes théories de la nature conceptuelle, nous détaillerons les nombreux travaux qui abondent dans le sens d'un réseau cérébral commun, impliqué à la fois dans la compréhension sémantique de stimuli verbaux et non verbaux.

Dans le chapitre expérimental, nous expliquerons les différents protocoles qui nous ont permis d'apporter quelques éléments de réponse aux thématiques abordées dans ce travail. Nous commencerons par décrire notre premier protocole, qui combine IRMf et DTI, pour à la fois mettre en évidence les substrats neuronaux de la compréhension sémantique (de stimuli verbaux et non-verbaux) mais aussi la connectivité au sein de ce réseau (connectivité structurale via l'étude de données DTI et connectivité fonctionnelle basée sur nos données IRMf). Dans la seconde partie du chapitre expérimental, nous décrirons une expérience comportementale où nous nous sommes intéressés à la nature des représentations langagières et aux différences interindividuelles. Enfin, dans la troisième partie du chapitre expérimental, nous présenterons les premiers résultats d'un protocole EEG dont l'objectif initial est de mieux comprendre le décours temporel associé à la compréhension de stimuli verbaux et non-verbaux, mais aussi les mécanismes qui peuvent intervenir

lorsque la compréhension nécessite l'intégration d'une information sémantique contextuelle (créée par un scénario de deux images ou deux phrases successives).

Pour terminer, nous présenterons, en conclusion, une synthèse de nos résultats en illustrant nos propos par les réflexions les plus actuelles sur l'origine des liens entre système sémantique et systèmes sensori-moteurs, dans une perspective liée à l'évolution. Enfin, puisque ce travail de thèse s'inscrit dans les thématiques plus larges d'une équipe où neurosciences et robotique interagissent étroitement, nous ouvrirons le débat en proposant une réflexion sur l'apport de la cognition incarnée pour la robotique. En espérant que ce large panorama, de la philosophie de l'Antiquité à la robotique de demain, n'effrayera pas trop nos lecteurs…

CHAPITRE I : Cadre Théorique

Nous avons vu, dans le chapitre précédent, comment au fil des siècles nous sommes passés d'une vision dichotomique de la relation corps/esprit à une approche où ces deux systèmes seraient étroitement liés. Grâce à des théories nouvelles, comme celle de la cognition incarnée, et à de nombreux travaux récents dans ce domaine, on considère maintenant que ces systèmes pourraient interagir pour permettre de créer et d'accéder à des représentations conceptuelles. En effet, la cognition incarnée permettrait de faire le lien entre des processus cognitifs de « bas » niveau (action/perception) et ceux de plus haut niveau (compréhension sémantique /langage/pensées). Relier les capacités motrices à des capacités cognitives est en contradiction avec la vision classique qui suppose une séparation claire entre les processus de haut niveau et de bas niveau et qui stipule que la connaissance dérive de processus computationnels intervenant dans des domaines séparés (Jirak et al. 2010).

Un travail de revue récent (Kiefer & Pulvermüller, 2012) reprend les différentes théories et les expérimentations actuelles à propos du système conceptuel. Selon les auteurs, il est maintenant largement reconnu que les concepts représentent la somme de toutes nos expériences sensorielles et motrices avec l'environnement, les concepts correspondant à une généralisation de nos expériences passées avec différentes catégories de stimuli (objets, événements ou idées). Ils citent comme exemple le concept « chat » ; d'après les expériences que nous avons eues avec différents exemplaires de cette catégorie, nous savons qu'un chat a quatre pattes, de la fourrure, qu'il miaule, qu'on peut le caresser… Cette représentation est abstraite dans le sens où elle ne fait pas référence à un chat en particulier mais plutôt à une représentation prototypique du concept « chat », mais on peut aussi se référer à une représentation particulière (le chat de ma grand-mère). Cependant, pour ces auteurs, la nature précise des traces mnésiques conceptuelles et en particulier de l'implication des systèmes sensorimoteurs dans la création des concepts, reste un sujet de controverse. Il est vrai que, malgré une profusion de travaux dans le domaine de la cognition incarnée, la variabilité des résultats, des méthodologies et des interprétations théoriques ne donnent pas toujours une idée claire de la façon dont les systèmes sensorimoteurs pourraient participer à la compréhension sémantique, ne permettant pas non plus d'aboutir à une définition unifiée de la cognition incarnée. C'est peut-être principalement parce que cette notion a été étudiée à l'aune de disciplines diverses (intelligence artificielle, robotique, psychologie cognitive, neurosciences, philosophie, linguistique, anthropologie cognitive…), reflétant en fait des conceptions assez différentes (Borghi et Cimatti, 2010). Au sein même de l'approche neuroscientifique de la cognition incarnée, deux approches se dissocient. Selon la version la plus radicale (Gallese, 2008 ; Glenberg, 1997), la cognition est

contrainte par le corps que nous possédons et donc par les états du corps, soulignant l'importance de l'action dans la simulation. Dans cette perspective, le corps est toujours considéré comme un organisme agissant. La conception plus classique de la cognition incarnée (Barsalou, 1999) reconnaît, quant à elle, le rôle du système sensorimoteur mais défend l'idée que la simulation s'effectue de façon multimodale et pas seulement par rapport à l'action et aux états du corps. Selon Borghi et Cimatti (2010), ces deux approches ne sont pas contradictoires, mais elles ne tiennent pas compte d'un fait important : considérer le corps humain comme une entité sociale.

Il existe un certain nombre d'autres lacunes dans le domaine de la cognition incarnée, avec notamment un clivage entre les travaux portant sur la compréhension langagière d'une part et ceux qui reposent sur l'étude de stimuli imagés d'autre part. S'il existe un système conceptuel commun, lié à une simulation sensorimotrice, il devrait être accessible quels que soient les stimuli (mots, phrases, objets isolés, scènes visuelles…) ou la modalité d'entrée (audition, vision, synesthésie…). Si plusieurs travaux ont cherché à comparer directement l'activation de concepts à la fois à travers des stimuli verbaux et non verbaux, les stimuli utilisés sont parfois très simples (mots, objets isolés…) et ne reflètent pas toute la complexité d'une compréhension plus « naturelle », nécessaire lorsqu'il s'agit d'interpréter des situations complexes, qu'elles soient décrites dans des phrases, observées via des images ou rencontrées lors d'une interaction directe (comprendre les intentions d'autrui, assigner des rôles thématiques…). Si un certain nombre d'études mettent en évidence des régions communes impliquées dans la compréhension des deux types de stimuli (Bajo, 1988 ; Vandenberghe et al., 1996 , Bright et al. 2004 ; Kircher et al., 2009 ; Van Doren et al., 2010 ; Binder & Desai, 2011 ; Binder et al., 2009, Graves et al., 2010), les résultats et les interprétations ne sont pas toujours consensuels. En effet, localiser précisément un système sémantico-conceptuel indépendant de la modalité d'entrée reste difficile, notamment en raison de l'utilisation de techniques, stimuli et tâches variés. Il faut noter toutefois que, grâce aux techniques de neuroimagerie modernes (Electroencéphalographie EEG, Magnétoencéphalographie MEG, Imagerie par résonance magnétique fonctionnelle IRMf, l'Imagerie par Tenseur de Diffusion DTI), on a évolué d'une vision plutôt localisationiste de ce système conceptuel vers la notion d'un système distribué avec un rôle particulièrement important de certaines zones de convergence ou « hubs » intégratifs, soutenu par une architecture spécifique de fibres de matière blanche. Cependant, si la notion de « réseau sémantique » a largement été étudiée pour la compréhension langagière (notamment grâce au DTI), la connectivité au sein d'un réseau sémantique commun aux stimuli verbaux et non-verbaux reste un sujet très peu abordé.

Un certain nombre de points restent obscurs également concernant la participation de notre propre mémoire (sémantique mais pas seulement) à l'élaboration de ces représentations sensorimotrices.

Comme le précise Jeffrey Binder (Binder et Desai, 2011), si la nature des concepts a été débattue depuis des siècles, la mémoire sémantique est un sujet d'étude relativement récent, ce qui est assez surprenant compte tenu de sa large implication dans la cognition humaine, la mémoire sémantique étant l'un des éléments les plus caractéristiques de ce qui constitue l'être humain dans la mesure où elle englobe toutes les connaissances déclaratives que nous acquérons sur le monde. La mémoire ne sert pas uniquement à se souvenir du passé, elle permet aussi d'anticiper l'action, de prédire le futur. Par ailleurs, un certain nombre de liens avec d'autres domaines d'étude cognitifs restent à effectuer, notamment en ce qui concerne les aspects de la cognition sociale. En effet, les recherches sur la théorie de l'esprit (TOM) grâce à laquelle nous sommes capables de saisir les intentions d'autrui permettent également de réinterpréter le rôle des réseaux cérébraux sensorimoteurs dans l'ensemble du schéma du système nerveux central et de son développement.

Comment appréhender, dès lors, le système sémantique face à cette diversité des théories, des terminologies, des expériences, des stimuli… ? En effet, de nombreuses questions restent encore en suspens quant à la nature du système conceptuel (est-il amodal ou modalité spécifique ?) et quant à la variabilité des représentations en fonction des stimuli. Les représentations sont-elles « catégories spécifiques » (animaux, outils, actions…) impliquant des régions cérébrales localisées distinctement et des processus différents ? Ou un même système distribué est-il impliqué dans le traitement de toutes ces catégories ? Si toute notre compréhension se fait par l'intermédiaire d'une simulation sensorimotrice, qu'en est-il des concepts abstraits ? A quel point les représentations que nous formons sont-elles flexibles ?

En prenant en compte ces différents aspects, nous nous efforcerons d'aboutir à un modèle le plus clair possible de la compréhension sémantique des stimuli verbaux et non verbaux. Nous commencerons donc par un bref rappel sur les systèmes impliqués, tels que les systèmes sensorimoteurs et langagiers, mais aussi mnésiques. Puis, nous envisagerons les différents aspects théoriques afin de mieux comprendre la nature du système conceptuel et des représentations. Enfin, nous ferons une revue la plus exhaustive possible des travaux réalisés dans le domaine de la cognition incarnée, travaux qui utilisent diverses méthodes pour des stimuli verbaux et non verbaux, afin de mieux cerner s'il existe un réseau sémantique commun et d'appréhender son rôle dans la compréhension de stimuli complexes tels que des actions.

I.1) Systèmes impliqués : quelques rappels de notions générales

La première partie de ce chapitre théorique sera donc consacrée à un rappel de notions générales sur les systèmes sensorimoteurs, perceptifs[5] et langagiers puisque, d'après les théories de la cognition incarnée, ces systèmes seraient impliqués conjointement dans la compréhension sémantique. Le fonctionnement de ces systèmes sera abordé séparément puisque nous consacrerons ultérieurement une partie de ce chapitre aux interactions de ces systèmes et à leur possible participation à un système sémantique commun. Enfin, une réflexion sur le système conceptuel ne peut s'exonérer d'une réflexion sur la mémoire puisque pour comprendre le monde qui nous entoure, nous devons stocker des concepts et pouvoir y accéder ensuite. La notion de mémoire sémantique est fondamentale pour la compréhension, mais nous verrons aussi que l'on peut difficilement dissocier cette réflexion des autres systèmes mnésiques et en particulier de la mémoire épisodique/autobiographique (la mémoire du « soi »), car c'est à partir de nos propres expériences passées que nous créons les concepts auxquels nous aurons besoin de nous référer pour toute situation où la compréhension est mise en jeu.

I.1.*1*) Rappel de notions générales sur l'organisation fonctionnelle du système sensori-moteur

Le système sensori-moteur est organisé hiérarchiquement, du cortex associatif jusqu'aux muscles effecteurs. Cette organisation hiérarchique permet aux hauts niveaux hiérarchiques de se consacrer spécifiquement à l'exécution des tâches complexes. Son organisation en canaux parallèles ainsi que la ségrégation fonctionnelle au niveau neuronal rendent possible un contrôle complet du cortex associatif sur l'ensemble des niveaux hiérarchiques. Un feed-back sensoriel régule l'activité motrice de façon inconsciente au niveau inférieur de la hiérarchie sensorimotrice et un apprentissage induit un changement dans la nature et le lieu du contrôle sensori-moteur, d'abord conscient puis automatique.

Différentes zones composant le système sensori-moteur semblent impliquées dans le contrôle sensori-moteur. Le cortex pariétal postérieur (CPP) a un rôle d'intégration multimodale puisqu'il gère des informations en provenance de plusieurs systèmes sensoriels (visuel, auditif et somatique). De ce fait il joue un rôle dans l'évaluation du contexte dans lequel s'effectue un mouvement en gérant différentes données comme la position des parties du corps mobilisées, la localisation du corps et des objets dans l'espace. Il participe à la planification des mouvements à effectuer en produisant des

[5] Nous nous contenterons de rappels sur le système perceptif visuel puisque notre travail porte principalement sur des stimuli en modalité visuelle.

modèles internes, en amont des cortex prémoteur et moteur et en dirigeant l'attention (Halsband et Lange ; 2006 ; Andersen et al., 1997 ; Xing & Andersen, 2000 ; Colby & Goldberg, 1999). Intervient également la partie dorso-latérale du cortex préfrontal qui reçoit des informations issues du CPP et qui projette en retour sur le CPP (Petrides & Pandya, 1984, 1999). Le cortex préfrontal correspond à un haut degré d'intégration dans la hiérarchie du contrôle moteur, il aurait un rôle dans l'évaluation des stimuli externes et l'initiation des réactions volontaires envers ces stimuli. Le cortex préfrontal interviendrait dans des fonctions exécutives telles que la planification et la prise de décision, dans la mémoire de travail visuelle afin de permettre un apprentissage d'une séquence prévisible d'actions (Robertson, Tormos, Maeda, & Pascual-Leone, 2001 ; Fuster, 2008 ; D'Esposito et al., 2000 ; Miller & Cohen, 2001 ; Gazzaniga, 2001).

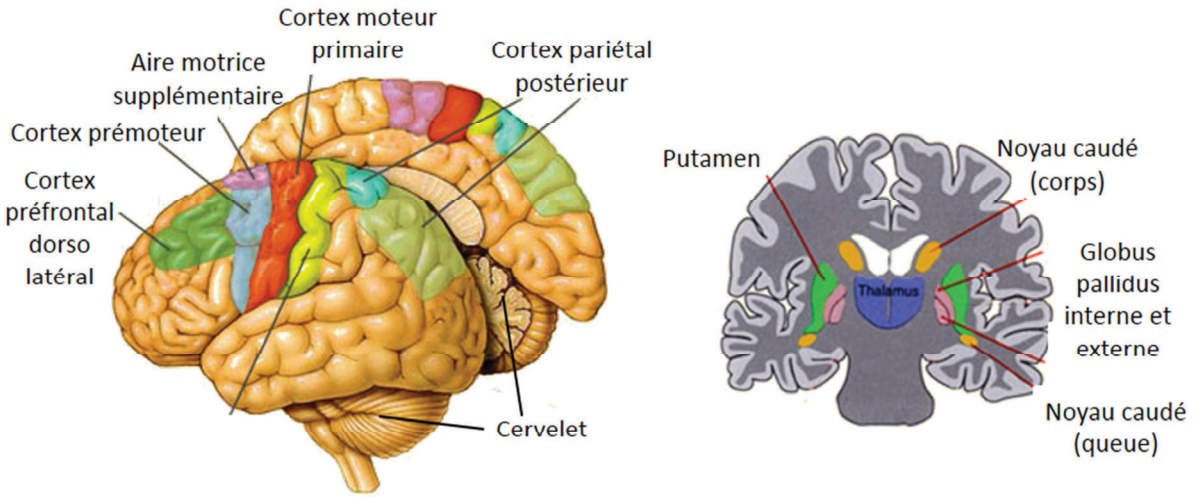


Fig 1) Organisation du système sensori-moteur
Les principales régions corticales qui composent le système sensorimoteur (à gauche) : cortex préfrontal-dorsolatéral, le cortex prémoteur, l'aire motrice supplémentaire, le cortex pariétal postérieur ; auxquelles s'ajoute le cervelet et, au niveau sous-cortical, les ganglions de la base (à droite).
Adapté de http://lecerveau.mcgill.ca/

Les aires du cortex moteur secondaire projettent en majeure partie sur le cortex moteur primaire. Ces aires (aire motrice supplémentaire AMS et cortex pré-moteur) seraient impliquées dans la production et le contrôle de mouvements complexes et dans la planification et la réalisation de mouvements volontaires, sous contrôle du cortex préfrontal dorso-latéral (Goldberg, 1985, Gazzaniga, 2001). L'AMS aurait un rôle critique dans la planification et l'initiation des mouvements en fonction des expériences motrices passées stockées en mémoire, en particulier lorsqu'elle sont intégrées dans un planning temporel précis (Halsband & al., 2006). Le simple fait d'anticiper un mouvement déclenche une activation de l'aire supplémentaire. Le cortex prémoteur jouerait un rôle critique dans la sélection d'un comportement moteur et dans l'apprentissage sensori-moteur dans la mesure où il permettrait la récupération de mouvements déjà acquis (Mitz, Godschalk et Wise ;

1991) et il interviendrait dans la régulation de la posture, en dictant au cortex moteur une position optimale pour un mouvement donné.

Le cortex moteur primaire (localisé au niveau du lobe frontal) est le point de convergence des signaux moteurs corticaux ainsi que le principal émetteur des signaux de commandes motrices, vers les motoneurones de la moelle épinière puis vers les muscles et ce, à travers quatre voies motrices descendantes : deux vers la région dorso-latérale et deux vers la région ventro-médiane. Le cortex moteur primaire est organisé de façon somatotopique (Penfield & Boldrey, 1937) et permettrait de générer des représentations des actions. Il serait un élément central du fonctionnement moteur, permettant l'anticipation de mouvements connus et l'actualisation de mouvements anticipés en fonction de la situation (Meier et al., 2008 ; Gazzaniga et al., 2001).

Le cervelet et les ganglions de la base (GB) sont deux importants systèmes sensori-moteurs qui coordonnent et modulent l'activité de la hiérarchie sensori-motrice en interagissant avec les différents niveaux de cette hiérarchie. En comparant des informations provenant de différentes sources afin de corriger éventuellement un mouvement en cours de réalisation, le cervelet aurait ainsi un rôle important dans les apprentissages de séquences de mouvements où l'organisation temporelle est cruciale. Les GB auraient aussi une fonction de modulation de par leur implication dans des boucles neuronales recevant des informations de différentes aires corticales ; ce sont eux qui les redirigent vers les aires motrices du lobe frontal par l'intermédiaire du thalamus. Ils seraient plus impliqués dans la phase de planification que dans celle d'exécution de séquences de mouvements générées de façon interne (Delong, 1990 ; Ivry, 1996 ; Boecker et al., 1998 ; Boecker, Jankowski, Ditter, & Scheef, 2008 ; Jouen et al., 2013).

I.1.*2*) Rappel de notions générales sur la perception visuelle

De même que pour les cortex sensoriels, moteurs ou auditifs, il existe pour la vision une organisation hiérarchique entre cortex primaire et secondaire. Les aires secondaires de ces différents systèmes convergent ensuite vers des zones associatives de haut niveau qui permettent un traitement plus global de l'information. Les aires associatives permettent de faire émerger une représentation intégrée et multisensorielle du monde, en faisant converger des signaux en provenance de différentes modalités sensorielles (Gazzaniga, 2001).

Une fois captée par l'œil, l'information visuelle est transmise au cortex visuel primaire (cortex strié ou V1) qui se situe dans la partie la plus postérieure du lobe occipital du cerveau, où l'image commencera à être reconstituée. Les informations seront ensuite dirigées vers d'autres aires

visuelles (de V2 à V6) qui codent pour différentes informations. Plus précisément, l'analyse des stimuli visuels, amorcée dans V1 et V2, se poursuit ensuite à travers deux grands systèmes corticaux de traitement de l'information visuelle : la voie inférieure ventrale « consciente » qui permet de reconnaître les objets et la voie supérieure dorsale « inconsciente » qui guide nos actes (Ungerleider et Mishkin, 1983). La voie ventrale (voie du « quoi ») s'étend vers le lobe temporal, après avoir traversé les aires de traitement visuel (chacune ajoutant un aspect spécifique à la perception : forme, couleur, profondeur…). La représentation ainsi obtenue et transmise à l'extrémité inférieure du lobe temporal, serait ensuite comparée à des souvenirs visuels afin d'être reconnue. Certaines informations rejoignent le lobe frontal où sont évaluées leur signification et leur importance, ce qui correspond à un stade où la perception consciente serait atteinte. La voie ventrale serait donc impliquée dans la reconnaissance des objets à travers une perception consciente et en traitant les propriétés visuelles intrinsèques des objets (forme, couleur…). La voie dorsale (voie du « où ») projette, quant à elle, vers le lobe pariétal et transporte les signaux déclenchés par les stimuli visuels depuis le cortex visuel en traversant des aires chargées de déterminer la position spatiale d'un objet par rapport à un observateur, de calculer sa taille, de recueillir des informations sur le mouvement et les séquences temporelles.

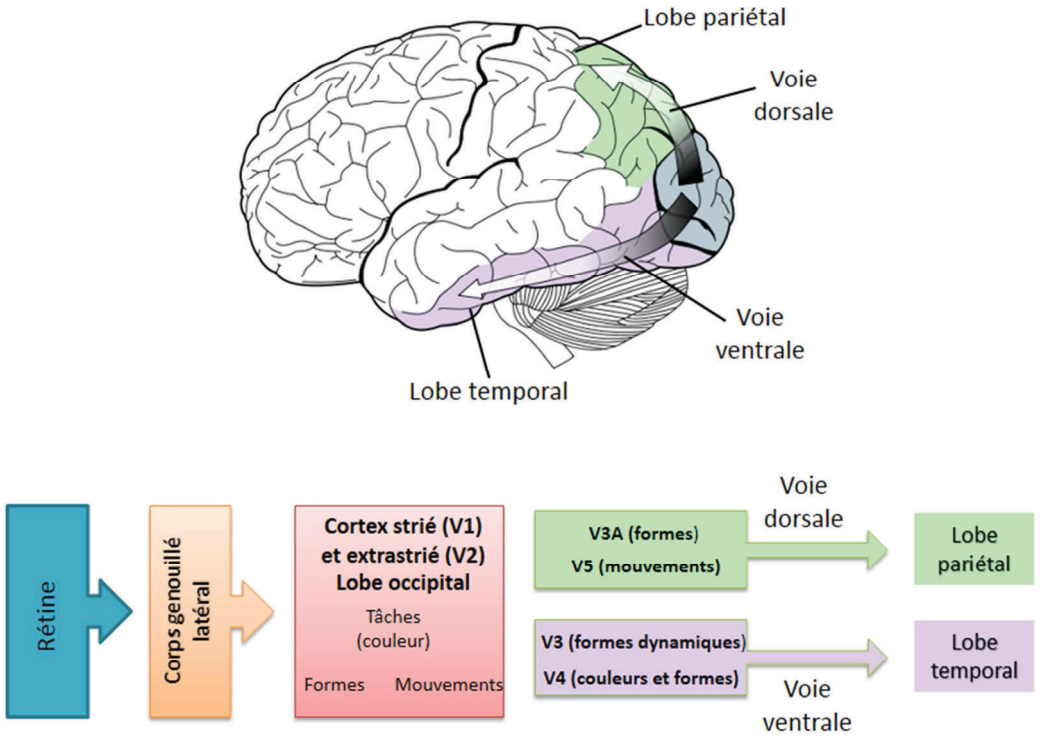


Fig 2) Organisation des voies ventrale et dorsale du système visuel
Après avoir été traitée au niveau de la rétine, l'information visuelle est envoyée au corps genouillé latéral puis au niveau du cortex extra-strié (V1 et V2) où des informations basiques sont analysées. Les informations sont ensuite plus précisément traitées par la voie dorsale (voie du où) qui se termine au niveau pariétal et la voie ventrale (voie du quoi) qui se termine au niveau temporal.
Adapté de http://lecerveau.mcgill.ca/

Au sein de cette voie dorsale, il faut noter le rôle particulièrement important de l'aire V5 (ou aire MT) qui semble contribuer à la perception du mouvement biologique. La voie dorsale serait donc essentielle dans le contrôle visuo-moteur et la localisation des objets en analysant les propriétés spécifiques des objets (taille, mouvement, localisation…), nécessaires pour pouvoir interagir avec eux, en proposant un plan d'action adapté à la situation (Carter et al, 2010). On désigne même, aujourd'hui, la voie dorsale sous le terme de « voie de l'action » : elle permettrait de guider en temps réels les actions que nous dirigeons vers des objets du monde visuel ; en intégrant les relations spatiales entre l'organisme et l'environnement, elle nous permet d'interagir efficacement avec lui.

Enfin, il ne faut pas oublier que la perception visuelle est extrêmement complexe, car dépendante de traitements à la fois descendants et ascendants. Nous ne voyons pas ce que nous pensons voir. Lorsque nous regardons une scène, nous avons l'impression de la voir entièrement d'un seul regard, mais en pratique nous ne percevons que quelques détails… La perception visuelle est temporaire, partielle et fragmentaire (Carter et al, 2010). Nous ne voyons pas les objets, les visages, les paysages directement en tant qu'objets, nous percevons des lignes, des surfaces, des contours et de ces perceptions mouvantes, nous devons extraire des éléments qui nous permettent de déduire la nature de ce qui nous entoure et de lui donner du sens.

I.1.*3*) Rappel de notions générales sur l'organisation du langage

L'analyse du langage requiert divers niveaux d'intégration complexes, c'est pourquoi nous avons choisi de consacrer un paragraphe au langage afin de prendre en compte ces différents aspects pour l'analyse ultérieure de nos résultats. Nous nous focaliserons essentiellement sur les aspects perceptifs/réceptifs du langage, ainsi que sur la lecture des mots. Comme nos protocoles expérimentaux comportent des phrases à lire, nous avons voulu intégrer un rappel assez synthétique des connaissances actuelles sur les zones cérébrales impliquées dans la lecture, sans rentrer dans les détails des différents modèles psycholinguistiques et neuroanatomiques. Par ailleurs, même si nous abordons les aspects sémantiques du langage, nous ne débattrons pas ici de la nature du système sémantique ; de même, nous aborderons plus en détail les aspects fonctionnels et neuroanatomiques du langage en lien avec notre thématique de recherche et les perspectives de l'approche de la cognition incarnée dans un prochain chapitre (cf. Chapitre I.3).

Longtemps, notre compréhension des bases cérébrales du langage était plutôt restreinte : avec l'aire de Wernicke et l'aire de Broca, la première intervenant dans la compréhension des mots et la seconde dans leur production, ces deux aires étant reliées par le faisceau arqué. Mais de nombreuses

expériences en neuroimagerie cérébrale ont mis en évidence une organisation en réseau beaucoup plus complexe avec l'intervention de nombreuses aires, pas forcément « langagières » et des rôles divers joués par les aires impliquées. La neuroimagerie a ainsi permis la découverte d'une troisième région indispensable au langage, le territoire de Geschwind (situé au niveau du lobule pariétal inférieur) qui est à la fois connecté à l'aire de Broca et à l'aire de Wernicke et qui, de ce fait, pourrait être un relais informatif indirect, alternative à la voie directe impliquant le faisceau arqué (Catani et al., 2005). Le territoire de Geschwind pourrait jouer un rôle clé pour appréhender les multiples propriétés d'un mot car cette région se situe à l'intersection des cortex auditif, visuel et somatosensoriel avec qui elle est massivement connectée ; de plus elle contiendrait des neurones mutlimodaux, permettant de combiner les différentes propriétés d'un mot (son, vue, sens). Son rôle serait donc d'aider le cerveau à classifier et à étiqueter les choses, une condition préalable pour former des concepts et une pensée abstraite. Le lobule pariétal inférieur serait d'ailleurs l'une des dernières structures du cerveau à s'être développée durant l'évolution, mais aussi pendant la maturation cérébrale individuelle.

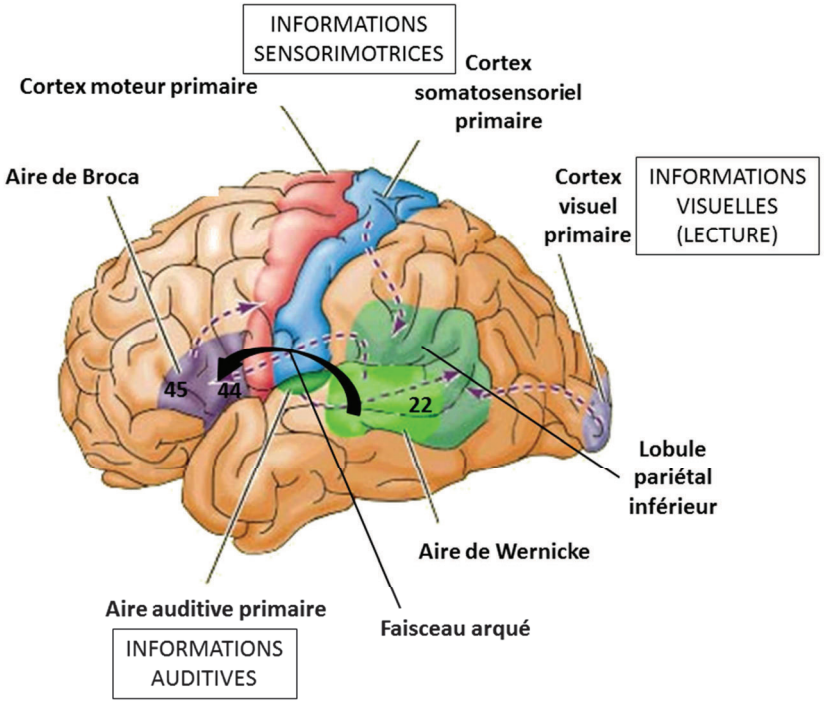


Fig 3) Organisation des principales aires du langage

L'aire de Broca chevauche les aires 44 et 45 de Brodmann, situées à l'avant de l'aire prémotrice dans la région inféro-postérieure du lobe frontal. Les aires de Broca et Wernicke sont reliées par le faisceau arqué. L'aire de Wernicke, dont une composante anatomique clé est le planum temporale, est située au niveau du gyrus temporal supérieur, dans la portion supérieure de l'aire 22 de Brodmann.Il s'agit d'une localisation stratégique compte tenu de ses fonctions dans le langage : l'aire de Wernicke se trouve en effet située entre le cortex auditif primaire (aires 41-42) et le lobule pariétal inférieur. Ce dernier est composé de deux régions distinctes : caudalement, le gyrus angulaire (aire 39), lui-même au voisinage des aires occipitales visuelles (aires 17, 18, 19), et dorsalement, le gyrus supramarginal (aire 40) qui chevauche l'extrémité du sillon latéral, adjacent à la partie inférieure du cortex somato-sensoriel. Le LPI est connecté par d'importants faisceaux de fibres nerveuses à la fois à l'aire de Broca et à l'aire de Wernicke.

Adapté de http://lecerveau.mcgill.ca/

Comme nous le verrons plus tard, de plus en plus de résultats obtenus en neuroimagerie remettent toutefois en question le modèle classique des fonctions langagières localisées (modèle Geschwind-Wernicke), plaidant plutôt en faveur de zones de convergence et d'une conception plus distribuée des aires du langage (Friederici & Gierhan, 2012).

S'il est indéniable que le langage est une fonction prise en charge de manière prédominante par l'hémisphère gauche (HG), cette fonction s'accompagnant d'une asymétrie anatomique (Marc et Gustave Dax en 1836 ; Broca, 1864 ; Wernicke, 1871 ; Geschwind & Levitsky, 1968 ; Galaburda et al. 1987 ; Josse et al. 2003 ; Dehaene-Lambertz & Dehaene, 2002 ; Vigneau et al., 2006), on a pu constater que les deux hémisphères traitant les informations de façon différente, l'hémisphère droit (HD) joue également un rôle important au niveau du langage (Lindell, 2006 ; Vigneau et al., 2011). Les travaux de Springer et al., (1999), Seghier et al., (2004), Gitelman et al., (2005) ont mis en évidence des activations de l'hémisphère droit lors de tâches langagières dans des études menées en IRMf ou TEP. Mais si les régions de l'HG sont pleinement impliquées dans les processus phonologiques, sémantiques et syntaxiques (Vigneau et al., 2006), les régions activées dans l'HD (au niveau frontal) refléteraient plutôt une composante liée à des processus attentionnels et de mémoire de travail (Vigneau et al., 2011). L'HD aurait un lexique limité, sans capacité phonologique, cependant il serait activement impliqué dans le traitement contextuel via des activations des zones temporales (Vigneau et al., 2011).

- **Réception vs. Production du langage**

Parmi les aires du langage, on distingue généralement le pôle postérieur ou réceptif et le pôle antérieur ou expressif. Le pôle dit réceptif du langage correspond à différentes régions du lobe temporal gauche (Gaillard, Balsamo & Xu, 2004 ; Lehéricy, Cohen & Bazin, 2000). Ce pôle recouvre essentiellement l'aire de Wernicke composée du gyrus temporal supérieur, avec le planum temporale (Vigneau et al. 2006) et les gyri de Heschl (Démonet, Thierry & Cardebat, 2005). Le pôle réceptif intervient dans des tâches de lecture et d'écoute de mots (Dehaene, 2007 ; Vigneau et al., 2006). Il serait impliqué à la fois dans les traitements phonologique et sémantique du langage (Seghier, Lazeyras & Pegna, 2004), notamment dans des tâches de compréhension du langage, dans l'accès au sens et dans la construction d'une représentation lexicale des mots perçus dans les deux modalités auditive et visuelle. Ces fonctions impliquent l'activation du cortex temporal inférieur, de la partie postérieure du cortex temporal moyen, incluant le gyrus angulaire, ainsi que de l'« aire temporale basale du langage » (Démonet et al. 2005).

Le pôle dit expressif du langage engage différentes régions du lobe frontal (Berl et al., 2005 ; Gaillard et al., 2004 ; Lehéricy et al., 2000), principalement le gyrus frontal inférieur gauche, où se situe l'aire

de Broca, ainsi que le gyrus précentral ou le cortex moteur (Seghier et al., 2004). Le gyrus frontal inférieur gauche est constitué anatomiquement de 3 parties : une partie postérieure (pars opercularis, BA44), une partie moyenne (pars triangularis, BA45) et une partie inférieure (pars orbitaris, BA47). Le pôle expressif est mis en jeu lors de tâches de production telles que des tâches de génération et de fluences verbales (Démonet et al., 2005). Un résultat surprenant apporté par la neuroimagerie est que le versant moteur de la production du langage ne serait pas pris en charge par l'aire de Broca (Houdé et al., 2002). En effet, dans des tâches de production simple du langage (récitation continue de mois de l'année, énumération de chiffres), qui nécessitent la planification des mouvements élémentaires de l'appareil bucco-phonatoire pour produire les sons mais pas des processus plus complexes tels qu'une analyse sémantique ou phonologique, l'aire de Broca n'est pas activée. Loin de se limiter à la production des mots, ce pôle interviendrait également dans les traitements phonologique et sémantique du langage, puisqu'on retrouve une activation de l'aire de Broca dans des tâches n'impliquant pas la production du langage et même dans des tâches purement sémantiques (Démonet et al., 1992 ; Paulesu et al., 1993 ; Mazoyer et al., 1993 ; Houdé et al., 2002). Les traitements phonologique et sémantique dépendent de l'interaction dynamique des deux pôles réceptif et expressif. Ils impliquent tous deux de vastes réseaux neuronaux, situés essentiellement dans les régions frontale et temporale. Le traitement phonologique impliquerait le gyrus précentral et la partie dorsale antérieure du gyrus frontal inférieur gauche (Seghier et al., 2004) ainsi que les gyri temporaux supérieur et moyen et le gyrus supramarginal (Vigneau et al., 2006) tandis que le traitement sémantique nécessiterait la participation de la partie antérieure ventrale du gyrus frontal inférieur d'une part (Seghier et al., 2004) et d'autre part des régions temporales (cortex temporal moyen et inférieur, sillon temporal supérieur postérieur, gyrus angulaire, gyrus fusiforme ; Vigneau et al., 2006 ; Démonet et al., 2005). Vigneau et al. (2006) ont également montré que deux aires temporales formaient une région commune pour les traitements phonologique et sémantique : le gyrus temporal supérieur antérieur et le gyrus temporal inférieur postérieur (Vigneau et al., 2006). La pars opercularis du GFi serait dédiée au traitement syntaxique et la partie postérieure du gyrus temporal supérieur serait particulièrement activée par les phrases et l'intégration de l'information textuelle (Vigneau et al., 2006).

Caramazza (1996) note que les progrès de l'imagerie ont permis des avancées considérables dans l'étude des processus cognitifs chez des sujets sains, mais l'étude des patients reste intéressante pour connaître les bases neurales du langage. Par rapport à la réflexion sur la nature du système sémantique, l'étude de l'aphasie optique est pertinente. Ce trouble se caractérise par une difficulté sélective de la dénomination d'objets présentés visuellement, trouble non observé lorsque les objets sont présentés dans une autre modalité (au toucher ou en réponse à une définition verbale). Certains

auteurs (Paivio, 1971, 1991; Shallice, 1988; Warrington & Shallice, 1984 ; Warrington & McCarthy, 1994) ont émis l'hypothèse de l'existence d'une disconnexion entre deux sous-systèmes sémantiques, dont l'un serait en charge de l'information visuelle (propriétés visuelles des objets et images) et l'autre traiterait l'information verbale (propriétés fonctionnelles des mots seulement). Une autre hypothèse à la base de ce trouble serait la représentation des informations sémantiques verbales seulement dans l'hémisphère gauche, tandis que les informations sémantiques visuelles relèveraient des deux hémisphères. Mais ces données neuropsychologiques n'empêchent pas non plus de considérer que le système sémantique est organisé de façon unitaire dans l'hémisphère gauche (HG) et que mots et images pourraient y accéder directement et indépendamment (Caramazza, Hillis, Rapp, et Romani, 1990 ; Ralph, Patterson, & Hodges, 1997; Lambon Ralph, Graham, Patterson, & Hodges, 1999; Hillis, Rapp, & Caramazza, 1995). La question de l'existence d'un système sémantique unique amodal ou de plusieurs systèmes spécialisés selon les modalités est bien un problème essentiel, toujours d'actualité, en ce qui concerne le traitement cérébral du langage.

- **Traitement cérébral de la lecture des mots**

Pour lire un mot, le cerveau doit exécuter trois opérations : il doit analyser ce qui est vu (l'écrit), établir une relation avec les sons de la langue parlée (phonèmes), puis il doit trouver la signification du mot. Ces trois opérations activent des zones distinctes du cortex cérébral gauche (Dehaene, 2007) impliquant les neurones de la lecture.

Les psychologues ont longtemps dissocié deux voies permettant toutes deux d'accéder au sens d'un mot (Coltheart, 1978, 2005 ; Marshall & Newcombe, 1973 ; Sprenger-Charolles, 1989 ; Shallice, 1983): d'une part la voie phonologique ou indirecte qui permettrait l'association des groupes de lettres (graphèmes) aux sons de la langue parlée (phonèmes), ce qui conduirait à la prononciation du mot à voix haute ou mentalement et donc d'accéder au sens ; d'autre part la voie directe, lexicale ou orthographique, où l'analyse visuelle du mot conduirait directement à une recherche en mémoire de ce mot dans une sorte de lexique, ce qui permettrait de retrouver le mot et sa signification sans avoir à le prononcer. Pendant de nombreuses années, seule l'étude des patients souffrant de pathologies pouvait renseigner les chercheurs sur les corrélats neuroanatomiques de la lecture. Ainsi, on a pu déterminer que des lésions majeures dans la région pariéto-occipitale gauche pouvaient se traduire par une incapacité à écrire et/ou à lire tout en laissant intactes les capacités orales de langage, alors que des lésions des aires associatives auditives comme celle de Wernicke empêchaient à la fois la compréhension orale et la capacité de lecture de l'individu. Grâce au progrès de la neuroimagerie, on peut suivre quasiment en temps réel le décours de la lecture et des expériences récentes ont montré que les deux voies de lecture postulées par les psychologues correspondaient bien à deux réseaux

distincts d'aires cérébrales (celui de la conversion graphèmes-phonèmes, le réseau du son et celui de la voie lexicale, le réseau du sens ; Jobard et al., 2003 ; Fisher et al., 2012). La voie grapho-phonologique est associée au système dorsal (temporo-pariétal) et la voie lexico-sémantique au système ventral (occipito-temporal). Le système dorsal active un réseau pariétal attentionnel (Cohen et al., 2008), comprenant le gyrus angulaire, le gyrus supra-marginal dans le lobule pariétal inférieur et la partie postérieure du gyrus temporal supérieur, c'est-à-dire l'aire de Wernicke (Pugh et al., 2001). Le système ventral conduit à un décodage visuel rapide des mots et implique l'activation des zones occipito-temporales ainsi que celle des zones associées à la fonction sémantique, telles que les zones latérales extrastriées, la région occipito-temporale inférieure gauche, les zones temporales supérieures, le gyrus supra-marginal et de la partie triangulaire/operculaire du gyrus frontal inférieur. Le système du codage phonologique, à proximité du gyrus frontal inférieur (GFi ; BA 44/45, l'aire de Broca) serait également impliqué (Pugh et al, 2001). Mais ces deux processus –voie indirecte et voie directe – seraient utilisés ensemble : chaque voie pouvant être davantage sollicitée pour certains mots et selon les lecteurs (Dehaene, 2007).

Globalement, qu'il s'agisse d'une lecture à voix haute ou mentale, le texte est d'abord traité au niveau du cortex visuel primaire, dans le pôle occipital : comme tout motif graphique, il active la voie de traitement de la reconnaissance qui transmet alors les informations aux aires du langage, avec une activation qui s'étend tout d'abord à la région occipito-temporale gauche, siège présumé de la reconnaissance des mots écrits (déchiffrage de la structure orthographique). En effet, dans les tâches expérimentales où les participants doivent lire mentalement des mots, on trouve toujours activée la région dite de la forme visuelle des mots (VWFA) dans le gyrus fusiforme, aux abords du sillon occipito-temporal latéral gauche (Dehaene et al., 2011 ; McCandliss et al., 2003). La lecture, c'est-à-dire la reconnaissance des mots écrits, activerait une aire occipito-temporale toujours située entre les réponses aux visages et les réponses aux objets (Ishai et al., 2000 ; Puce et al., 1996). Survient ensuite une explosion d'activités dans de multiples régions allant du cortex temporal jusqu'au frontal (Marinkovic et al, 2003), sachant qu'une partie du cortex temporal supérieur étant multimodale, elle est activée aussi bien par l'écrit que par l'oral. L'information serait donc acheminée vers l'aire de Wernicke où l'association entre la forme du mot et sa forme auditive serait effectuée. Plus précisément, ce serait la région temporale supérieure (planum temporale) qui interviendrait dans la conversion des lettres en sons, elle réagirait notamment à la compatibilité entre les lettres et les sons. L'information se retrouverait ensuite dans l'aire de Broca qui ajoute une structure syntaxique et un programme articulatoire, une fois le mot identifié, reliant ainsi le mot écrit au mot parlé et permettant de « prononcer » ce mot mentalement. En effet, la région pariétale inférieure située juste au-dessus du planum temporale forme vraisemblablement, avec une partie de l'aire de Broca

(région operculaire), un circuit qui s'active lorsque nous prononçons mentalement les mots. Les régions cérébrales activées lors de l'analyse du sens des mots (l'arrière de la circonvolution temporale moyenne, la face ventrale antérieure du lobe temporal et la partie triangulaire de la région frontale inférieure) interviennent également lorsque nous réfléchissons aux relations conceptuelles entre les mots parlés ou même entre les images (par exemple, association sémantique entre l'image d'un palmier et d'une pyramide, qui font partie du même champ lexical « Désert » et non sémantique, Dehaene, 2007). Une des particularités de la jonction temporo-pariétale, en tant qu'aire sémantique, c'est d'être active sans même que l'on ait demandé au participant d'effectuer un travail mental : on a montré notamment qu'elle était déjà activée dans l'état de repos, et qu'elle se "désactivait" en-dessous de son niveau initial, quand on présente des pseudo-mots dépourvus de sens (Dehaene, 2007). D'autres études confirment que la région temporale moyenne coderait aussi le sens des mots. Par exemple, chez les lecteurs japonais, la région postérieure du cortex inférieur temporal gauche est la seule à être activée lorsqu'un même mot est écrit dans deux formes écrites n'ayant rien à voir sur le plan visuel (en kanji et en kana) pour des tâches implicites et explicites (Ha Duy Thuy et al., 2004). Nous aborderons plus en détail le rôle de ces régions et la possibilité de l'existence d'un système conceptuel amodal dans les sous-chapitres à venir.

I.1.*4*) Rappel de notions générales sur la mémoire sémantique et la mémoire épisodique

Notre objectif n'est pas ici de décrire l'ensemble des systèmes mnésiques, nous nous concentrerons sur les aspects mnésiques qui peuvent être en lien avec la compréhension.

« La mémoire est la capacité à se souvenir d'un poème, à reconnaître un visage, à évoquer un événement très ancien, à savoir faire du vélo ou à se rappeler que les clés sont sur la table. Le point commun à tous ces éléments est d'impliquer l'apprentissage et la reconstruction totale ou partielle d'une expérience passée. » en synchronisant la décharge des neurones impliqués dans l'expérience d'origine (Carter et al. 2010).

Classiquement, on considère que la mémoire déclarative (l'un des deux grands systèmes de mémoire à long terme, avec la mémoire procédurale) renvoie à des souvenirs que l'on peut se rappeler consciemment, comme des faits et des connaissances, et qu'elle se divise en deux sous-systèmes : la mémoire épisodique qui stocke les expériences personnelles spécifiques à chacun et la mémoire sémantique qui contient les informations factuelles et les connaissances générales sur le monde, notamment la signification des objets et des mots (Tulving, 1972).

La mémoire épisodique permet à un sujet de se rappeler des événements vécus dans un lieu et à un instant donné (non seulement l'événement vécu, mais aussi tout le contexte particulier de cet événement) en reconstruisant ces expériences passées y compris des sensations et des émotions (Tulving & Szpunar, 2009). Elle a une forte composante émotionnelle. La mémoire autobiographique a longtemps été assimilée à la mémoire épisodique et il est vrai que ces deux systèmes partagent un certain nombre de caractéristiques même si on sait aujourd'hui que ce sont deux systèmes distincts (Tulving, 2005). La mémoire autobiographique est multiple ; elle contient des connaissances générales ainsi que des expériences uniques propres à un individu, qui ont été accumulées depuis l'enfance et qui permettent à cet individu de se construire un sentiment d'identité (Picard et al. 2009). Selon P.Piolino, « la mémoire autobiographique permet de « voir » des images, mais aussi de se promener dans ses souvenirs, en ayant « gardé en tête » les odeurs, les bruits… tout ce qui constitue une ambiance. Il s'agit d'une mise en scène de son propre corps » (Piolino, 2009).

La mémoire sémantique, moins affective, est le système par lequel l'individu stocke sa connaissance du monde. Une grande partie de ce stock de connaissances est accessible rapidement et sans effort. Elle nous permet la construction d'une représentation mentale du monde sans nécessité de sa perception immédiate et la mémoire sémantique est indépendante du contexte spatio-temporel d'acquisition. C'est la mémoire du sens des mots, celle qui permet de se souvenir du nom des objets, de leurs attributs (couleur, goût…), mais aussi d'éléments beaucoup plus complexes tels que la fonction des objets, les concepts, les croyances, les coutumes sociales... Ainsi elle interviendrait dans de nombreux processus cognitifs tels que le langage, le raisonnement, la résolution de problèmes et les interactions sociales (Binder et al., 2009 ; Binder & Desai, 2011).

On considère généralement que la mémoire épisodique implique des aires corticales variées en fonction du contenu de l'expérience d'origine (un souvenir plutôt visuel activera les aires visuelles, la voix d'une personne, le cortex auditif), mais elle impliquerait également le lobe frontal qui aurait pour rôle de vérifier que les souvenirs épisodiques ne sont pas erronés dans la réalité ; interviendrait également l'hippocampe, région où les événements seraient « convertis en souvenirs ». Les aires fronto-temporales joueraient un rôle important dans la mémoire autobiographique tandis que l'hippocampe serait responsable du tri des souvenirs récents et des souvenirs plus anciens (Piolino, 2009). La mémoire sémantique, quant à elle, impliquerait une large partie du lobe temporal antérieur (gyri temporaux inférieur et moyen, et portion antérieure du gyrus fusiforme) ainsi que le cortex pariétal inférieur qui est un cortex associatif ; ces deux régions auraient un rôle dans le stockage et la récupération des informations conceptuelles (Binder & Desai, 2011 ; Patterson et al., 2007; Binder et al., 2009). Le lobe frontal aurait aussi un rôle dans la mémoire sémantique, il permettrait de faire appel sélectivement aux connaissances stockées afin de guider et d'adapter

notre comportement actuel (Martin, 2007 ; Gold et al. 2005). Dans une expérience en TEP cherchant à comparer directement la récupération en mémoire épisodique et en mémoire sémantique, Wiggs et al. (1999) ont mis en évidence des activations bilatérales dans les cortex temporal et frontal pour la condition sémantique alors que la condition épisodique activait un réseau distinct localisé dans le cortex pariétal médian, le cortex rétrosplenial, le thalamus et le cervelet ainsi que dans les régions frontales gauches. De nombreuses autres études en neuroimagerie, chez des sujets sains et chez des patients (Maguire, 2001; Levine et al., 2004 ; Svoboda et al., 2006; Cabeza & St. Jacques, 2007; Schacter et al.,2007 ; Spreng et al., 2009), semblent montrer l'existence d'un réseau qui jouerait un rôle clé dans la mémoire épisodique. Ce réseau inclurait les régions temporales et frontales, l'hippocampe, les cortex associatifs sensoriels et des régions pariétales postérieures. Les données venant de la neuropsychologie ont aussi grandement participé à cette dissociation entre mémoire épisodique et mémoire sémantique ; on a par exemple observé des patients présentant une lésion au niveau du lobe temporal moyen et ayant une atteinte sévère de la mémoire épisodique alors que la mémoire sémantique était préservée dans la plupart des cas. Inversement, les patients atteints de démence sémantique qui est une variante de la démence frontotemporale (caractérisée par une dégénérescence au niveau du lobe temporal anterolatéral) présentent de lourds troubles sémantiques alors que la mémoire épisodique est généralement conservée (Greenberg & Verfaellie, 2010).

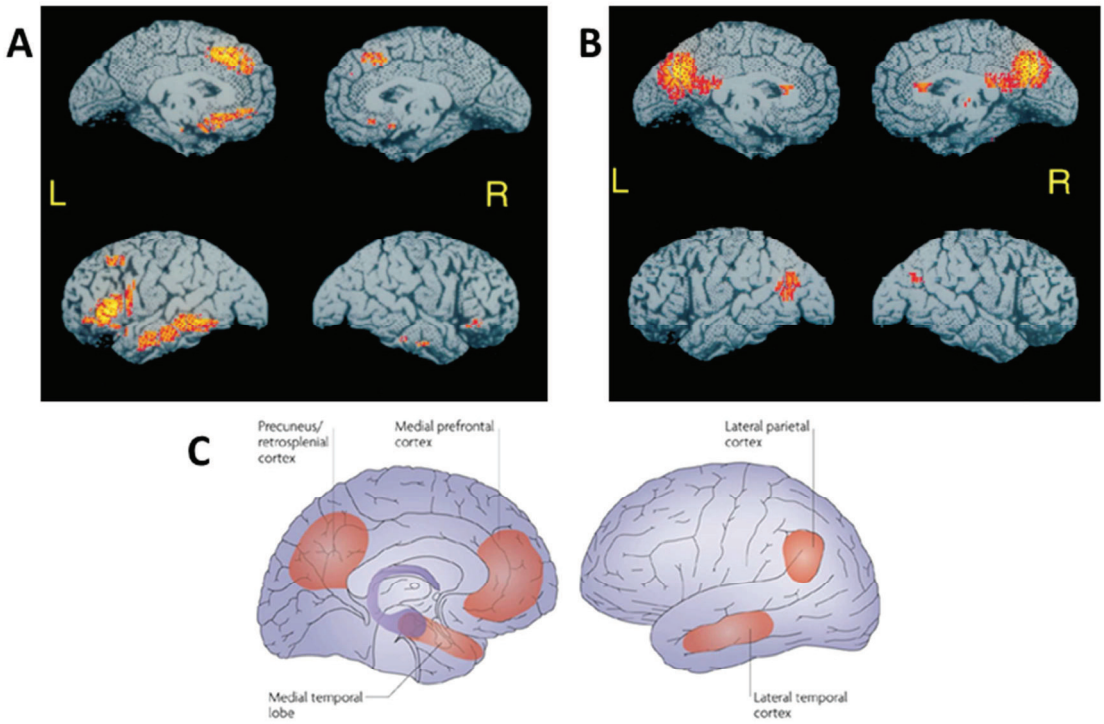


Fig 4) Principaux corrélats neuroanatomiques des différents systèmes mnésiques
La mémoire sémantique impliquerait plus spécifiquement le cortex temporal et le cortex frontal (A) tandis que la mémoire épisodique ferait appel au cortex pariétal médian, au cortex rétrosplenial, au thalamus, au cervelet ainsi qu'aux régions frontales gauches (B). La mémoire autobiographique, quant à elle, impliquerait le cortex rétrosplénial et le precuneus, le cortex temporal latéral, le cortex pariétal latéral et le cortex préfrontal (C). **Adapté de Wiggs et al. (1999) et Irish & Piguet (2013)**

Cependant, ces différents systèmes mnésiques que sont la mémoire épisodique, autobiographique et sémantique ne sont pas complètement dissociables : ils impliquent un certain nombre de régions cérébrales qui se recoupent mais surtout, ils seraient interdépendants et interagiraient constamment les uns avec les autres (Greenberg et al, 2010 ; Binder et al., 2009 ; Irish & Piguet, 2013). En effet, on peut considérer la mémoire sémantique comme résultant des expériences emmagasinées en mémoire épisodique/autobiographique. Il y aurait une transition progressive entre mémoire épisodique et mémoire sémantique, cette dernière permettant d'aboutir à une généralisation de l'information en retenant les traits communs à divers épisodes vécus et en les déconnectant de leurs contextes d'origine. Inversement la compréhension des expériences personnelles que nous vivons au quotidien fait sans cesse appel aux concepts et aux connaissances stockés en mémoire sémantique. La mémoire épisodique/autobiographique jouerait donc un rôle en facilitant l'ajout de nouvelles informations sémantiques et à l'inverse, la mémoire sémantique fournirait un cadre qui faciliterait l'acquisition de nouveaux souvenirs personnels. De même en ce qui concerne la récupération d'informations, les souvenirs épisodiques faciliteraient leur recherche en mémoire sémantique et la mémoire sémantique contiendrait les éléments de base à partir desquels des souvenirs épisodiques complexes et détaillés peuvent être construits et récupérés (voir Greenberg et al, 2010 pour une revue chez les patients). Dans une étude de fluence catégorielle (Greenberg et al., 2009), certaines des catégories sur lesquelles les sujets étaient testés semblaient générer des souvenirs purement épisodiques (e.g un participant qui cite une marque de voiture en pensant au concessionnaire automobile devant lequel il était passé en venant) mais dans d'autres cas, les sujets faisaient appel à des souvenirs plus difficilement classifiables (e. g un participant listant tous les ustensiles qu'il utilise quotidiennement dans sa cuisine). Il n'est alors pas évident de repérer à quel type de mémoire appartient ce souvenir : ce n'est pas quelque chose qui s'est déroulé à un instant temporel précis (un événement) et cela ne peut donc être considéré comme purement épisodique ; cependant il y a un contexte spatial spécifique, ce n'est donc pas tout à fait sémantique non plus. Pour Greenberg et Verfaellie (2010), il faut nuancer la distinction entre mémoires épisodique et sémantique, beaucoup de nos souvenirs ne rentrant pas vraiment dans l'une ou l'autre des catégories. Ces deux types de mémoire pourraient en fait constituer les deux extrêmes d'un même gradient mnésique préservant de nombreux souvenirs qui se situent, entre les deux, sur un continuum. Pour Binder & Desai (2011), le gyrus cingulaire postérieur (cortex rétrosplénial) ainsi que le précuneus pourraient former des zones d'interface entre mémoire sémantique et mémoire épisodique. Plusieurs autres travaux mettent également en évidence l'existence d'un recoupement anatomique entre ces deux systèmes mnésiques et, en particulier, le rôle important que pourrait jouer la mémoire sémantique à la fois pour se rappeler des événements passés mais aussi pour planifier le futur (Binder et al., 2009 ; Maguire, 2001; Svoboda et al., 2006 ; Burianova et al., 2010). Ces nouvelles conceptions de la

mémoire permettent d'ouvrir une réflexion sur l'interdépendance de ces deux systèmes mais aussi sur la nature même de nos souvenirs et la flexibilité de nos représentations d'un individu à un autre. La pensée humaine est tellement complexe et personnelle qu'il est difficile d'affirmer que nous avons le même système mnésique que notre voisin puisque mémoire sémantique, épisodique et autobiographique sont intimement liées et dépendantes de l'expérience de chacun.

I.2) Nature des représentations et du système conceptuel

La nature du système conceptuel et de ses représentations suscite, depuis longtemps, de nombreuses interrogations. Les représentations sont-elles amodales et symboliques ou, au contraire, correspondent-elles à des réactivations de traces sensorimotrices et donc dépendantes de la modalité ? Ces traces sensorimotrices ont-elles une organisation somatotopique et sont-elles alors dépendantes des effecteurs impliqués (différentes parties du corps) ? On s'interroge sur la façon dont sont conceptualisés les différents objets que nous traitons (animaux, outils, humains…) : dans un espace sémantique commun ou selon des processus spécifiques impliquant des zones cérébrales distinctes ? Se pose également la question de savoir si les représentations sont flexibles ou si elles correspondent à des invariants situationnels. Nous essaierons d'aborder ces différents points afin de préciser quelle est la nature du système conceptuel et de nos représentations. Dans cette partie, nous aborderons principalement les aspects théoriques de ces réflexions (car nous pourrons avoir besoin de nous y référer ultérieurement), nous étudierons plus en détails les aspects expérimentaux résultant des liens entre simulation sensorimotrice et compréhension conceptuelle dans la partie Chapitre I, 3.

Comme le précisent Kiefer et Pulvermüller (2012), les différents problèmes soulevés par ces interrogations (nature amodale du système conceptuel, localisé ou distribué, flexibilité des représentations) ne sont pas totalement indépendants les uns des autres, puisqu'on retrouve souvent certains aspects allant de pair. Par exemple, une théorie supposant une représentation conceptuelle flexible implique nécessairement une vision distribuée du système conceptuel car les théories prônant une vision localisationiste de ce système ne peuvent permettre d'envisager une flexibilité des représentations. Ces aspects étant indissociables, nous traiterons à la fois la nature amodale vs modalité-spécifique, localiste vs distribuée du système conceptuel, ainsi que la nature des représentations.

- **<u>Représentations conceptuelles localistes et amodales</u>**

En ce qui concerne la nature des concepts, les théories les plus anciennes et qui ont prédominé jusqu'à récemment présentaient une vision localiste et amodale du système conceptuel où les représentations sont fondamentalement distinctes des implications sensorimotrices au niveau cérébral. Selon ces théories, les informations sensorielles et motrices provenant de l'environnement sont transformées en des représentations symboliques et amodales, sans plus aucune relation avec l'information originelle et un concept est représenté comme un nœud au sein d'un réseau sémantique unitaire (Collins et Loftus, 1975 ; Collins et Quillian, 1969 ; Quillian, 1969 ; Fodor, 1975 ;

Smith & Medin, 1981 ; Landauer & Dumais, 1997 ; Bowers, 2009). Chaque nœud conceptuel est relié à d'autres nœuds et ensemble, ils regroupent toutes les informations qui sont contenues dans un concept (un chat a 4 pattes, il miaule, il a de la fourrure…). Comme chaque concept est associé à un nœud représentationnel spécifique, bien distinct des représentations sensori-motrices, ce type de modèles de réseau sémantique suppose une vision localisationiste et amodale. Selon ces modèles, l'acquisition des concepts est statistique : on retient un concept à partir des répétitions d'expériences interactionnelles avec notre environnement et les concepts sont alors très stables car ils correspondent à des invariants situationnels. Ces théories sont actuellement de plus en plus remises en question car la nature des concepts définie dans ces modèles manque de flexibilité.

- **Représentations conceptuelles amodales et distribuées**

D'autres modèles envisagent également la nature des concepts comme amodale mais ils ont une vision distribuée de l'organisation du système sémantique, dans la mesure où un concept est formé de différentes unités représentationnelles qui peuvent faire intervenir différentes zones cérébrales. C'est le cas de modèles assez anciens et basiques, comme le modèle de comparaison des fonctionnalités de la mémoire sémantique (Semantic feature-comparison model de Smith, Shoben et Rips, 1974), ainsi que de modèles PDP (parallel distributed processing, McClelland & Rumelhart, 1986). Dans les modèles distribués de la mémoire sémantique, on suppose que les concepts sont codés par des unités de représentation multiples ; en raison de cette structure, il y a une certaine flexibilité conceptuelle car la contribution d'une ou plusieurs unités peut varier en fonction du contexte dans lequel on accède à un concept. Contrairement aux approches localistes-amodales, dans ces modèles, la connaissance conceptuelle n'est pas explicitement représentée sous la forme d'éléments ou de noeuds simples, elle est récupérée par la propagation de l'activation entre les unités de traitement qui forment un réseau interconnecté. C'est la force des poids de connexion entre les unités de ces réseaux qui, modulée par l'expérience, permettra l'apprentissage conceptuel. Cependant, dans ces modèles comme dans les précédents, on considère que la connaissance conceptuelle est représentée dans un format amodal, distinct des systèmes de la perception et de l'action. Et malgré des composantes représentationnelles distribuées en réseau, ces théories rejoignent l'idée d'un système conceptuel unitaire, stockant toutes sortes d'informations indépendamment de la modalité (vision ou lié à l'action) ou de la catégorie (animaux, outils…) Il en découle l'hypothèse que le substrat neuronal de ce système unitaire conceptuel pourrait être situé au niveau du lobe temporal antérieur. Il est vrai qu'un certain nombre de travaux semblent montrer que les représentations pourraient avoir une composante amodale, notamment les travaux issus d'études neuropsychologiques de patients atteints de maladies neurodégénératives (voir Kiefer & Pulvermüller, 2012 pour plus de références). On a souvent décrit un trouble dans le traitement

conceptuel chez des patients atteints de démence sémantique (lésion localisée dans le lobe temporal), trouble caractérisé par une perte massive dans tous les domaines conceptuels (y compris les animaux, les objets..) et quelle que soit la modalité d'entrée (visuelle, auditive…). Bien que les lésions puissent être variables dans la démence sémantique (allant du cortex temporal postérieur jusqu'au cortex frontal), les lésions localisées dans le pôle temporal semblent suffisantes pour faire émerger des déficits sémantiques dès les stades précoces de la maladie (Patterson et al., 2007). Pobric et al. (2010) ont retrouvé des troubles proches de ceux observés chez des patients atteints de démence sémantique, en stimulant par TMS les aires temporales antérieures chez des sujets sains en train d'effectuer des tâches sémantiques, impliquant à la fois des images et des mots. Le cortex temporal antérieur pourrait donc être le substrat neuronal des représentations conceptuelles amodales (Kiefer & Pullvermuler, 2012). Certains modèles, eux aussi principalement basés sur l'étude de cas neurospychologiques, vont même plus loin et proposent également une interprétation selon laquelle les connaissances sémantiques seraient représentées de façon stable et amodale, distinctement de l'implication des systèmes sensorimoteurs, mais en fonction des différentes catégories (objets, animaux…). Ce sont les modèles dits « catégories spécifiques » (nous évoquerons plus amplement ces modèles ci-après, associés à la notion de représentations localistes lorsque ces dernières sont considérées comme des unités neuronales simples codant pour des stimuli très spécifiques).

- **<u>Représentations conceptuelles flexibles, distribuées et modalités-spécifiques</u>**

Actuellement, on tend de plus en plus vers des modèles où les représentations conceptuelles sont considérées comme flexibles, modalité-spécifiques et impliquant un réseau cérébral distribué. Dans ces modèles, l'idée essentielle est que les concepts sont « fondés » dans la perception et l'action. Ces modèles trouvent leur apogée dans la formalisation de la théorie de la cognition incarnée et grâce à de nombreux travaux actuels, mais cette idée n'est pas nouvelle et on la retrouve dans des modèles et des travaux plus anciens (Beauvois et al. 1978, 1982 ; Allport, 1985 ; Farah et McClelland, 1991 ; Warrington et Shallice, 1984 ; Gallese et al, 1996). Les théories basées sur une approche modalité spécifique contredisent le point de vue classique localisationiste et amodal et émettent l'idée que les représentations conceptuelles sont fonctionnellement et neuro-anatomiquement reliées aux représentations perceptives et motrices. Les représentations des états externes (perception) et internes (proprioception, émotion et introspection) ainsi que les actions sont des éléments essentiels à la création des concepts et jouent un rôle fonctionnel fondamental dans la compréhension linguistique et dans la pensée conceptuelle (Kiefer et Pulvermüller, 2012). Les concepts sont supposés être « incarnés » dans le sens où les interactions avec d'autres personnes et avec des objets conduisent à la formation de traces mnésiques conceptuelles dans des zones

cérébrales spécifiques à la modalité qui, généralement, traitent les informations liées aux aspects sensoriels et à l'action mais aussi aux aspects émotionnels. On trouve la mise en évidence de ces liens à travers des protocoles expérimentaux comportementaux (Barsalou, 2008 ; Fischer & Zwann, 2008 ; Helbig et al., 2010). Corroborant ces résultats comportementaux, de nombreux travaux issus de l'étude de patients semblent également contredire l'hypothèse d'un système conceptuel unique localisé dans le cortex temporal, en montrant que d'autres zones, en particulier les cortex sensorimoteurs, seraient importantes pour traiter les concepts. Des lésions des systèmes visuels, auditifs ou des cortex associatifs moteurs peuvent perturber le traitement conceptuel (Bak et al, 2001, Cotelli et al, 2006, Neininger & Pulvermüller, 2003). La neuroimagerie (EEG, IRMf, MEG…) chez des sujets sains permet aussi de mettre en évidence des activations des aires sensorielles et motrices dans un grand nombre de tâches conceptuelles dans la modalité auditive, visuelle et même pour les odeurs (Martin, 2007; Pulvermüller & Fadiga, 2010 ; Kiefer et Pullvermuler, 2012). Nous détaillerons plus amplement ce type d'expériences, à la fois pour des stimuli visuels et non-visuels dans le chapitre I, 3.

- **Nature catégorie-spécifique des représentations**

Une autre approche consiste à considérer la nature « catégorie-spécifique » des représentations, beaucoup d'études chez des patients cérébrolésés ayant montré des troubles spécifiques pour différentes catégories. Dans ces théories, la notion d'un système unitaire de la mémoire conceptuelle est totalement abandonnée et ces théories défendent l'existence de sous-systèmes conceptuels innés et distincts qui résultent en un « découpage » a priori de l'espace sémantique selon différentes catégories tels que les animaux, les fruits, les outils… (Caramazza et Shelton, 1998 ; Caramazza & Mahon 2003). A la suite d'un AVC, d'un traumatisme crânien ou d'une maladie neurodégénérative, on peut rencontrer chez les patients une déficience dans la récupération des informations concernant les éléments composant une catégorie d'objets par rapport à d'autres catégories (Martin, 2007). Les catégories touchées peuvent être variées et on recense plus de 100 profils différents avec des troubles portant notamment sur la reconnaissance des animaux, des fruits et légumes, des congénères, des objets inanimés, des personnes célèbres… (Mahon et Caramazza, 2008). En particulier, une double dissociation qui a souvent été constatée est celle des troubles concernant les connaissances relatives aux animaux ou autres objets animés versus une déficience sélective en ce qui concerne les objets inanimés tels que les outils (Capitani et al. 2003). Un aspect qui semble important dans les profils de performance de patients présentant des troubles catégorie-spécifique, c'est que la détérioration semble vraiment affecter les connaissances conceptuelles et pas seulement l'entrée ou la sortie des représentations en fonction d'aspects liés à la modalité ; cela signifie que le déficit catégorie-spécifique ne dépend pas de la présentation des stimuli ou de la

modalité de réponse des patients. On trouve par exemple des patients incapables de nommer (sur sollicitation) ou de reconnaître (à partir d'une image) des choses animées alors qu'ils n'ont pas de problèmes pour les catégories inanimées, mais ces patients sont aussi incapables de répondre à des questions de connaissances sur les catégories touchées alors qu'ils n'ont pas de problème pour les autres catégories (ex : est-ce qu'un chat a 4 pattes vs est-ce qu'une table a 4 pieds ?). Mais on observe aussi souvent des troubles pour reconnaître les différents items au sein d'une catégorie (par exemple, le patient reconnaît qu'un chien est un animal mais ne peut préciser que c'est un chien) (Mahon et al., 2009). La connaissance conceptuelle des objets serait organisée par les caractéristiques sensorielles (la forme, le mouvement, la couleur…) et les propriétés motrices associées à l'utilisation de l'objet et, dans certains modèles, par d'autres propriétés fonctionnelles ou verbales liées à l'objet, par exemple son importance sociale, etc. (pour une revue cf. Martin, 2007). C'est l'idée que défendent un certain nombre de modèles explicatifs classiques basés sur une conception liée aux propriétés inhérentes des objets (sensory-motor property models ; depuis des modèles anciens tels que ceux de Lissauer, 1890 ; Freud, 1891 à des modèles plus récents, voir Humphreys & Forde, 2001). Les troubles de connaissances propres à chaque catégorie auraient lieu quand une lésion perturbe la récupération de l'information sur une propriété ou plusieurs propriétés critiques pour définir une catégorie d'objets et pour distinguer un objet spécifique parmi les différents représentants de cette catégorie (Martin, 2007).

Une autre alternative aux modèles basés sur les propriétés, ce sont les modèles dits « domaines-spécifiques ». Deux grandes conceptions s'opposent quant à la nature des systèmes cérébraux qui pourraient expliquer les causes des déficits sémantiques liés aux catégories (Mahon et al., 2009). Les théories qui se situent dans le premier groupe (Caramazza et al. 1990, Devlin et al. 1998, Tyler & Moss, 2001) reposent sur la notion de « structure corrélée » et supposent que la connaissance conceptuelle des éléments venant de différentes catégories sémantiques n'est pas représentée dans des régions cérébrales fonctionnellement dissociables. Dans ces modèles, la mémoire sémantique est considérée comme un système qui représente les régularités statistiques de co-occurrence des propriétés des objets dans le monde réel. Les théories du second groupe, quant à elles, affirment que différents substrats neuronaux sont impliqués dans les représentations des différentes catégories sémantiques (sensory/functional theory et domain-specific hypothesis). Lorsque des déficits touchent une catégorie sémantique spécifique, c'est que cela correspond à une atteinte du substrat neuronal dont dépend la catégorie des stimuli touchés. La théorie sensori-fonctionnelle (Warrington & Shallice, 1984) postule que les connaissances sémantiques sur les différentes catégories d'objets sont modales (sensori-dépendantes) et qu'elles sont représentées dans différents sous-systèmes sémantiques catégories-spécifiques. Selon l'hypothèse domaine-

spécifique (Caramazza et Shelton, 1998), l'évolution serait responsable de cette organisation cérébrale des connaissances conceptuelles : la pression de la sélection aurait permis d'aboutir à des circuits cérébraux spécifiques, « machines neuronales » dédiées à la résolution rapide et efficace de problèmes liés à la survie (identification rapide des objets). Dans cette perspective, les représentations distinctes des animaux, des autres congénères, des plantes et des outils seraient des catégories pertinentes pour notre cerveau car ce seraient des connaissances conceptuelles nécessaires à la survie de l'espèce ; par conséquent des régions différentes contiendraient les infos conceptuelles relatives à ces catégories. Les concepts seraient représentés localement de façon stable et amodale, distinctement de l'implication des systèmes sensorimoteurs.

Bien sûr, différents travaux viennent étayer ces approches, mais les résultats et interprétations parfois contradictoires ne permettent pas vraiment de parvenir à un modèle unifié du fonctionnement des zones cérébrales impliquées dans le traitement catégoriel (voir Mahon et al., 2009 ; Martin, 2007). Cependant, une zone apparaît comme cruciale pour le traitement catégoriel : au sein de la voie ventrale occipito-temporale (voie de la reconnaissance des objets), le cortex occipito-temporal n'apparaît pas, en effet, comme un système homogène de traitement des objets, il aurait plutôt une organisation fine liée à la catégorie des objets. De nombreuses études en neuroimagerie ont permis de mettre en évidence des zones du cortex occipitotemporal recrutées spécifiquement par différentes catégories telles que les animaux, les visages, les outils, le corps humain, etc. (Kourtzi & Kanwisher et al. 2001; Martin, 2001; Chao et Martin, 1999; Chao et al., 1999 ; Grill-Spector, 2003 ; Martin, 2007). On trouverait ainsi des aires spécifiquement impliquées dans le traitement des visages (FFA Face Fusiform Area, Kanwisher et al. 1997) ou des lieux (PPA Parahippocampal place area, Epstein et Kanwisher, 1998). Le gyrus fusiforme (partie la plus latérale) et la partie postérieure du sillon temporal supérieur seraient activés bilatéralement par les animaux (Grill-Spector, 2003). Quant aux outils ils entraîneraient une activité bilatérale accrue de la partie médiale du gyrus fusiforme et du gyrus temporal moyen gauche (Chao et al., 1999 ; Martin, 2007). Le corps humain serait représenté dans une région adjacente à la FFA, dans la partie latérale du gyrus fusiforme (Peelen & Downing, 2005, Schwarzlose et al. 2005). Selon Martin (2007), les résultats des diverses études montrent une forte cohérence de ces localisations, malgré des différences dans le format des stimuli utilisés (verbal/non verbal) et dans les modalités (visuel/auditif), mais aussi dans la variabilité et la complexité des stimuli et des tâches (scènes, films, tâches verbales de traitement conceptuel, Bartels et Zeki, 2004 ; Hasson et al. 2004 ; Chao et al., 1999).

Toutefois, cette conception catégorielle reste objet de débats. Pour Huth et al. (2012), les humains peuvent reconnaître des milliers de catégories différentes d'objets et d'actions ; compte tenu de la taille limitée du cerveau humain, il est déraisonnable de penser que chaque catégorie

d'objets est représentée dans une zone distincte du cerveau. Pour ces auteurs, les études IRMf sur ce thème ne sont pas convaincantes car elles n'ont pas permis d'identifier de manière consensuelle les zones cérébrales associées à de nombreuses catégories de stimuli qui sont pourtant des objets communs de notre quotidien (objets de la maison, vêtements, nourriture…). Huth et al. (2012) proposent la notion d'un continuum sémantique, espace sémantique commun réparti à la surface corticale, même si cet espace est organisé autour de grandes dimensions sous-jacentes (vivants vs non vivants par exemple). Cela constituerait une façon plus efficace de représenter les différentes catégories sans avoir besoin qu'une zone spécifique joue ce rôle pour une catégorie spécifique de stimuli. Dans le but de chercher ce continuum sémantique, les auteurs ont recueilli les réponses BOLD (blood oxygen level dépendent) en IRMf de plusieurs sujets pendant qu'ils regardaient des films « naturels » pendant plusieurs heures ; ces films contenaient de nombreux objets avec lesquels nous interagissons et des actions que nous effectuons dans la vie quotidienne. Ils ont ainsi pu obtenir les représentations corticales associées à 1705 catégories d'objets et d'actions en calculant un modèle de régression linéaire pour chaque voxel enregistré dans le cerveau de chaque sujet. Le modèle de poids obtenu décrit comment divers objets et catégories d'action influencent les signaux BOLD enregistrés dans chaque voxel : les catégories de poids positifs ont tendance à augmenter BOLD, tandis que celles avec des poids négatifs ont tendance à le diminuer. La réponse d'un voxel à une scène particulière correspond à la somme des poids de toutes les catégories présentes dans cette scène. Les auteurs ont ensuite utilisé l'analyse en composantes principales (ACP) pour faire ressortir cet espace sémantique commun à partir des modèles de poids basés sur les catégories chez chaque sujet. Avec l'ACP, des catégories qui sont représentées par des voxels proches dans le cerveau projetteront des points proches dans l'espace sémantique estimé et inversement pour les catégories représentées de manières très différentes. Cette analyse en composantes principales leur a permis de récupérer les 4 dimensions sous-jacentes principales qui sous-tendraient l'organisation de ce continuum sémantique (dans le modèle d'ACP une dimension correspond à une composante expliquant le mieux la variance dans le signal). La première dimension opposerait les items mobiles (humains, animaux, véhicules) aux non-mobiles ; la seconde dimension regrouperait les catégories de stimuli sociaux (les personnes, les verbes de communication), la troisième opposant les catégories associées à la civilisation (personnes, objets artificiels, véhicules) aux catégories liées à la nature (animaux) ; la quatrième dimension correspondrait aux catégories biologiques (personnes, animaux, plantes, partie du corps…) en opposition aux stimuli non biologiques, mais elle pourrait aussi contenir les contrastes entre objets animés (humains et animaux) et objets inanimés. Les auteurs affirment que l'on retrouve cette organisation de l'espace sémantique de façon très stable d'un sujet à un autre ; ils ont ensuite projeté ces résultats sur des templates corticaux afin de visualiser la localisation de ce continuum sémantique. Ces cartes d'activation confirment que l'espace

sémantique commun aux différentes catégories serait situé principalement au niveau du cortex occipitotemporal (cortex temporal inférieur, sulcus temporal moyen), mais aussi au niveau du cortex pariétal et jusqu'au cortex frontal antérieur. Ces travaux remettent en question la vision qu'une catégorie spécifique de stimuli serait codée par une aire particulière et la notion de continuum sémantique est une perspective intéressante avec l'idée qu'un espace distribué commun aux différentes catégories serait plutôt organisé selon des grandes dimensions regroupant différentes catégories (certaines propriétés de ces différentes catégories pouvant être commune d'une catégorie à une autre).

Le principal problème des modèles catégoriels les plus radicaux (hypothèse domaine-spécifique) est qu'ils ont tendance à présenter une vision trop restreinte et localisée du système conceptuel. Comme nous l'avons déjà souligné précédemment, la vision amodale-localisée manque de flexibilité et est peu cohérente avec les découvertes récentes, en particulier dans le domaine de la cognition incarnée. En effet, dans ces modèles, les concepts ne seraient pas fondés sur des représentations sensorimotrices. Certes, les éléments d'information fournis par la perception issue des cortex sensoriels primaires seraient essentiels mais ils ne seraient que la première étape d'une cascade de traitements. Ces éléments de base seraient ensuite projetés à des cortex associatifs multimodaux et ce seraient de petites populations de neurones très spécifiques de ces cortex qui permettraient de générer des images mentales utiles à la reconnaissance et à la compréhension de la situation en cours. Cette vision, très localiste et hiérarchique du système conceptuel a été illustrée par les travaux réalisés dans le cadre de la « théorie du neurone grand-mère ». Cette approche a été très à la mode dans les années 2000 avec la découverte, grâce à des enregistrements unitaires chez des patients implantés pour raisons médicales, que certains neurones répondaient spécifiquement à la présentation du visage ou même simplement du nom de personnes célèbres, tel que Bill Clinton, Jennifer Aniston, Halle Barry (Quiroga et al., 2005) ou encore, à la présentation de certains objets, mots, personnes… (pour une revue, voir Bowers, 2009). L'idée a été poussée à l'extrême en affirmant qu'un neurone isolé (ou un petit groupe de neurones) pourrait porter à lui seul des concepts aussi complexes qu'une grand-mère ou que la tarte aux pommes qu'elle prépare tous les dimanches. La cascade de traitements est supposée aboutir dans les parties antérieures des lobes frontaux et temporaux, où les images mentales seraient générées soit pour les comparer à la situation en cours ou pour se remémorer un souvenir. Comme le rappellent Meyer et Damasio (2009), dans la logique de ce point de vue, des lésions dans les cortex temporal et frontal antérieur devraient alors empêcher à la fois la perception normale et le rappel normal. Toutefois, cette prévision n'est pas corroborée par les résultats neuropsychologiques : les patients atteints de ces lésions ont une perception normale et ne présentent pas de déficits sélectifs dans la reconnaissance et le rappel des

objets ou d'événements (Damasio, 1989). Des dommages des cortex censés être impliqués dans « le haut de la cascade » de traitements n'empêchent ni la perception sensorielle, ni la récupération des représentations mentales d'objets ou d'événements généraux («une fête d'anniversaire»). En revanche, ils peuvent perturber le rappel d'un objet ou d'un événement précis («la fête d'anniversaire de mes 40ans») (Meyer & Damasio, 2009). De même, les patterns de déficits que l'on observe en général dans la maladie d'Alzheimer ou dans la démence sémantique et qui sont dus à des lésions très étendues et non pas à des lésions focales, s'expliquent beaucoup mieux dans la perspective d'un fonctionnement distribué, que ce soit au niveau fonctionnel ou neuroanatomique. Par exemple, dans la démence sémantique, la connaissance à propos d'un concept n'est pas affectée dans sa totalité comme pourrait le prédire un modèle localiste. Il y a au contraire une détérioration progressive au cours de la maladie, touchant d'abord les propriétés spécifiques du concept de cet objet (un canari peut chanter) jusqu'à atteindre un niveau plus général des propriétés partagées par les différents exemplaires d'une catégorie conceptuelle plus large (comme c'est un oiseau, un canari peut voler). De plus, dans la démence sémantique, l'information catégorielle d'ordre supérieur (animal) peut être conservée alors que des différenciations conceptuelles plus fines (oiseau vs poulet) sont altérées (Kiefer et Pullvermuler, 2012 ; Rogers et al., 2004). Encore une fois, une vision trop localiste génère un manque de flexibilité de la notion de représentation conceptuelle et n'est pas en accord avec les résultats montrant qu'une large gamme d'aires cérébrales est impliquée dans la pensée conceptuelle. Il n'y aurait donc pas de neurone " grand-mère ", ni de neurone " tarte aux pommes ". Les neurones impliqués dans les phénomènes de reconnaissance et de rappel seraient organisés en réseaux fonctionnels interconnectés voire synchronisés, impliquant des aires cérébrales distribuées pour permettre d'aboutir à des niveaux complexes d'intégration sémantique. En 1949 déjà, Hebb proposait l'idée que les neurones forment des « assemblées » qui correspondent aux substrats des idées. La notion de réseau semble vraiment cruciale et au sein de ces réseaux, certaines zones clés dites « zones de convergence » (Damasio, 1989) pourraient correspondre à des processus associatifs de haut niveau, pouvant contenir des représentations peut-être en partie amodales (nous discuterons plus en détails la notion de réseau et de zones de convergences dans la partie 3 de ce chapitre). Pour Martin (Martin, 2007) l'idée qu'il y ait une zone spécifiquement impliquée dans le traitement catégorielle (le cortex occipito-temporal) n'est pas incompatible avec la notion de réseaux et l'implication des systèmes sensorimoteurs dans la représentation conceptuelle. Même si certaines régions du cortex occipito-temporal pourraient avoir une organisation catégorielle, les concepts d'objets ne seraient pas explicitement représentés mais émergeraient de l'activité conjointe des régions qui contiennent des informations sur les propriétés des objets. Lorsque le traitement conceptuel le nécessite, les propriétés sensorielles et motrices associées à un objet engageraient spécifiquement les aires sensorielles et motrices correspondantes.

I.3) Sémantique de l'action : caractéristiques d'un système conceptuel commun à la compréhension verbale et non verbale

L'action est le moyen par lequel « le soi » interagit ou réagit face au monde extérieur. Une action peut être décrite comme l'expression finale de plusieurs étapes de traitement de l'information : l'intention, la programmation, la préparation et l'exécution (Grèzes, J., & Decety, 2001). Les actions dirigées vers un but sont générées de façon interne et il en résulte une représentation motrice mentale (Jeannerod, 1997). La notion de représentation mentale correspondant à la fois au contenu mental lié au but et aux conséquences d'une action donnée mais aussi aux processus neuronaux supposés se produire avant l'action, un tel système de représentation mentale permet d'anticiper, de simuler et même de comprendre les actions et les intentions d'autrui, ce qui est crucial pour les interactions sociales (Grèzes, J., & Decety, 2001).

Pendant longtemps, les actions ont été considérées comme les appendices triviaux d'opérations mentales plus sophistiquées (identification d'objet, compréhension du langage, prise de décision) et les processus moteurs relégués au simple fait d'appuyer sur un bouton et isolés, autant que possible, des processus cognitifs d'intérêt ; en a résulté une « négligence du contrôle moteur dans la science de la vie mentale et du comportement » (Abrams & Balota 1991 ; Rosenbaum, 2005 ; Fisher & Zwann 2008). Cependant ces conceptions ont changé car, comme nous l'avons déjà précisé, un nombre important de travaux dans le domaine de la cognition incarnée ont montré qu'il existait une équivalence fonctionnelle cérébrale entre l'observation, la compréhension, la mentalisation et la réalisation des actions, caractérisée par une simulation sensorimotrice impliquant les mêmes zones cérébrales que lorsque nous sommes réellement engagés dans l'action. Cette activité des systèmes sensori-moteurs est décrite par certains comme une « résonance motrice » (Zwaan & Taylor, 2006). Même si les travaux dans le champ de la cognition incarnée ont vu leur essor dans les années 1990-2000, ils viennent conforter des idées déjà anciennes mais minoritaires, comme le souligne cette réflexion d'un défenseur précoce de la cognition incarnée, le psychologue et philosophe américain Williams James *«Toute représentation d'un mouvement réveille dans une certaine mesure le mouvement réel dont il est question »* (1890, The Principes of Psychology p. 526).

Le fait que la reconnaissance des actions d'autrui est fondamentale pour nos interactions sociales est largement reconnu aujourd'hui, mais on peut même aller plus loin, cette simulation sensorimotrice ayant été mise en évidence à la fois pour l'observation d'actions mais aussi dans le langage (Gallese, 2008 ; Glenberg, 2007 ; Pulvermüller, 2005 ; Rizzolatti et Craighero, 2007 ; Fischer & Zwaan, 2008 ; Binder et al., 2009). Pour certains (Arbib, 2005 ; Arbib, 2010 ; Glenberg & Gallese, 2012) l'émergence

même du langage chez l'être humain serait liée à ces implications des systèmes sensorimoteurs dans le traitement conceptuel.

Cette dernière partie théorique concerne plus directement les aspects bibliographiques spécifiquement reliés à ce travail visant à savoir si la compréhension sémantique de stimuli décrivant des actions repose sur une simulation sensorimotrice. Un nombre grandissant d'études tendent à montrer que le système conceptuel humain serait implémenté dans des réseaux cérébraux distribués, utilisés à la fois pour l'action, la perception mais aussi pour des processus cognitifs de plus haut niveau comme la compréhension et le langage. De telles représentations conceptuelles issues de ces systèmes devraient alors être accessibles quels que soient les stimuli (mots, phrases, scènes visuelles…) ou la modalité d'entrée (audition, vision, synesthésie…). Ce chapitre s'articulera principalement autour de 3 questions interdépendantes : existe-il un système sémantique commun aux images et aux phrases impliqué dans la compréhension des actions ? Quels sont les corrélats neuroanatomiques de ce système, s'il existe ? Et enfin, quels peuvent être les rôles et le fonctionnement de chaque région cérébrale impliquée dans ce système ?

Nous essaierons, dans un premier temps de recenser séparément les travaux qui ont mis en évidence des effets de simulations sensorimotrices pour des stimuli non-verbaux et pour les stimuli verbaux, tout en nous concentrant particulièrement sur les implications des systèmes sensori-moteurs dans la compréhension des actions, puisque les stimuli que nous avons utilisés dans nos protocoles sont des actions (il faut noter que, dans nos protocoles, nos stimuli se réfèrent au terme d'actions au sens large de « quelqu'un en train de faire quelque chose » et non pas à des actions forcément dynamiques). Puis, nous nous intéresserons aux travaux qui ont cherché à comparer directement l'implication des systèmes sensori-moteurs pour les deux types de stimuli. Malgré la profusion de travaux sur la cognition incarnée, peu de travaux ont utilisé les deux types de stimuli simultanément et les travaux abordant séparément le traitement de ces stimuli semblent parfois recouper des réalités théoriques différentes (avec une tendance à étudier le système des neurones miroirs pour l'observation de stimuli d'actions en modalité visuelle vs une approche plus « sémantique-conceptuelle » pour les travaux impliquant le langage). C'est pourquoi après nous être intéressés aux corrélats neuroanatomiques de ce réseau commun, nous nous interrogerons sur la pertinence de ces systèmes sensorimoteurs dans la compréhension sémantique et le rôle des régions impliquées, sans toutefois négliger celui des autres régions qui pourraient intervenir. Nous nous pencherons particulièrement sur la possibilité de l'existence de zones de convergence qui auraient un rôle clé au sein du système conceptuel et sur l'importance d'un fonctionnement en réseau pour la compréhension, soutenu par une architecture spécifique de fibres de matière blanche.

I.3.*1*) Existe-t-il un réseau sémantique commun impliqué dans la compréhension des phrases et des images ?

A) Travaux impliquant des stimuli non verbaux

- **« Hypothèse visuelle » vs « Hypothèse d'appariement direct »**

Comme nous l'avons mentionné dans l'introduction, la reconnaissance des actions effectuées par autrui est un processus fondamental à la base de nos comportements sociaux. Bien que de nombreuses théories aient tenté d'expliquer ces mécanismes, on oppose souvent deux hypothèses principales pour expliquer la reconnaissance des actions au niveau visuel (Buccino et al., 2004 ; Rizzolatti, Fogassi, & Gallese, 2001). La première, désignée comme ''l'hypothèse visuelle'', soutient qu'aucune implication motrice n'est nécessaire car la reconnaissance de l'action repose sur une analyse visuelle de tous les constituants d'une action particulière c'est-à-dire une analyse de l'effecteur impliqué, des objets sur lesquels l'action porte et du contexte dans lequel l'action se déroule. Les inférences faites sur les interactions entre tous ces éléments permettraient à l'observateur de reconnaître les actions effectuées par autrui et cette reconnaissance, basée sur une analyse visuelle de la scène, impliquerait le cortex visuel extrastrié, le lobe temporal inférieur et la région du sillon temporal supérieur (STS). En effet, un certain nombre de résultats ont mis en évidence que ces zones répondaient de manière sélective à des objets, des parties du corps, au mouvement biologique ou encore aux interactions mains-objets, chez le singe (Gross et al., 1972 ; Ungerleider et al., 1994) et chez l'humain (Tanaka et al., 1991 ; Carey et al., 1997; Logothetis, 1998; Allison et al., 2000; Kourtzi et Kanwisher, 2000).

La seconde hypothèse, dénommée « l'hypothèse d'appariement directe » soutient que l'on peut reconnaître les actions effectuées par les autres en associant l'observation de l'action à sa propre représentation motrice de l'action, ce qui supposerait que les mêmes structures nerveuses seraient impliquées à la fois dans l'observation de l'action et dans l'exécution réelle de l'action. Une action serait comprise quand son observation provoque une « résonance » du système moteur, lui-même impliqué dans la réalisation des actions. La seconde hypothèse a trouvé un écho important avec la découverte du système des neurones miroirs chez le singe (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996 ; di Pellegrino et al. 1992; Umilta et al. 2001; Kohler et al. 2002; Keysers et al. 2003), puis chez l'homme, notamment grâce aux nombreux travaux qui ont montré l'existence d'un lien fort entre la réalisation des actions et leur compréhension.

Pour autant, cela ne veut pas dire que ''l'hypothèse visuelle » et « l'hypothèse d'appariement directe » soient totalement antinomiques, le fait que la compréhension des actions soit basée sur une correspondance directe entre la réalisation des actions et leur représentation au niveau moteur n'excluant pas la possibilité que d'autres mécanismes, basés sur le traitement des objets et du mouvement, puissent participer également à cette fonction (Rizzolatti, Fogassi, & Gallese, 2001).

- **<u>Le système des neurones miroirs</u>**

Il est certain que lorsque l'on pense liens entre exécution et compréhension des actions, le système des neurones miroirs (ou MNS mirror neurons systems) apparaît d'emblée comme le candidat idéal pour expliquer la simulation sensorimotrice telle qu'elle est explicitée par les théories de la cognition incarnée. Les neurones miroirs ont été décrits initialement chez le singe comme une population de neurones situés dans la partie rostrale du cortex prémoteur ventral (aire F5) qui déchargent à la fois lorsque l'animal exécute une action manuelle orientée vers un but et lorsqu'il observe un autre individu (congénère ou expérimentateur humain) en train d'effectuer une action similaire.

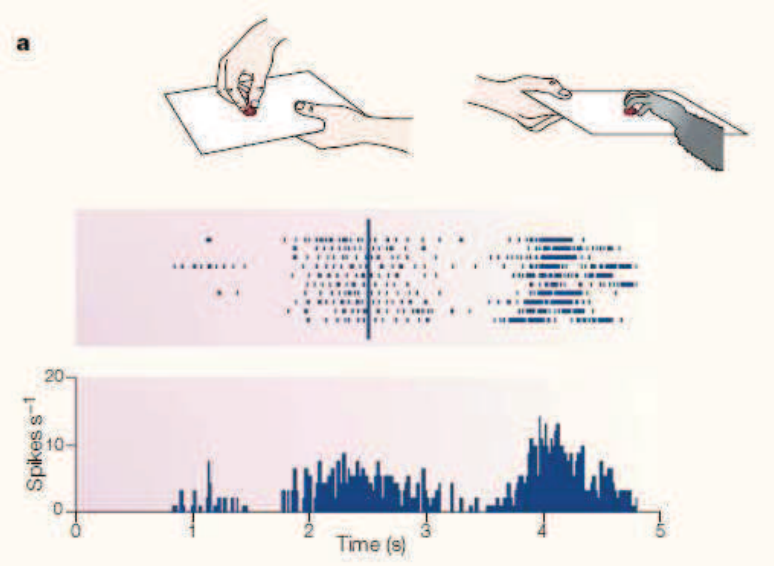


Fig 5) Réponses visuelles et motrices d'un neurone miroir de l'aire F5 chez le singe
L'expérimentateur saisit un morceau de nourriture placé devant le singe. On observe une forte activation neuronale en F5 lorsque le singe observe l'action, similaire à celle obtenue lorsque le singe réalise lui-même l'action. **Adapté de Rizzolatti, Fogassi, & Gallese, 2001**

Chez l'humain, il existerait un système miroir analogue à celui du singe qui ferait intervenir le cortex prémoteur ventral (PMv) ainsi que le cortex pariétal postérieur (CPP) somatosensoriel. Plus particulièrement, deux régions constitueraient le noyau clé du MNS humain (Rizzolatti & Craighero, 2004) : il s'agit de la partie rostrale du lobule pariétal inférieur (LPI) au niveau pariétal et, au niveau frontal, de la partie inférieure du gyrus précentral ainsi que de la partie postérieure du gyrus frontal inférieur (GFi). De nombreux travaux ont en effet montré que l'aire de Broca, essentiellement

connue pour ses fonctions langagières, recouperait partiellement le cortex prémoteur ventral, ayant ainsi également une composante motrice : la partie operculaire du gyrus frontal inférieur (BA44) correspondrait à l'homologue de la région F5 chez le singe (Buccino et al, 2004; Petrides & Pandya, 2002 ; Rizzolatti & Arbib, 1998, Binkofski et al. 1999, Ehrsson et al. 2000, Gerardin et al. 2000, Iacoboni et al. 1999, Krams et al. 1998 ; Skipper et al., 2007), la participation de l'aire de Broca dans la motricité ayant été soulignée à travers des expériences de « grasping » (Grafton, Arbib, Fadiga, & Rizzolatti, 1996), dans des tâches de manipulation d'objets (Binkofski et al., 1999) et d'imitation d'actions (Rizzolatti & Arbib, 1998). Il apparaît que la région de Broca serait plus qu'une aire impliquée uniquement dans le traitement du langage (Menz & Binkofski, 2008 ; Jirak et al., 2010). Fazio et al (2009) ont montré que des patients ayant des lésions frontales et dits « aphasiques de Broca » présentaient un trouble dissociatif spécifique de la compréhension des actions humaines : ils sont incapables de réordonner quatre images qui composent différentes étapes d'une action humaine (ex : saisir une bouteille) alors que cette capacité est épargnée lorsque la tâche est de réordonner les quatre images d'un événement physique ne faisant pas intervenir un humain (ex : un objet qui tombe).

Fig 2A

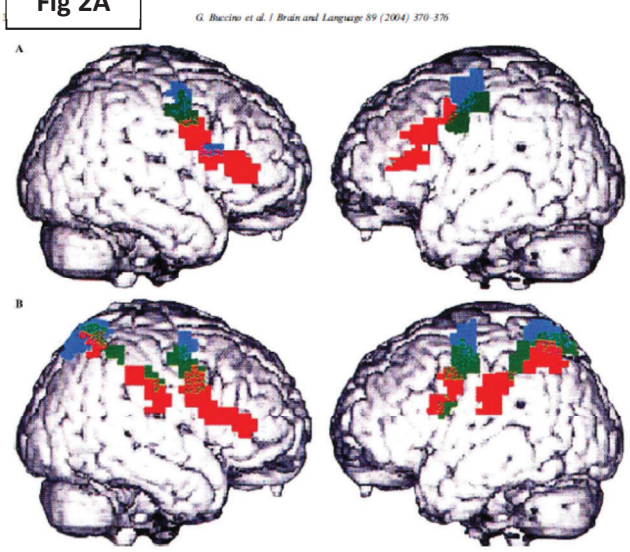


Fig 2B

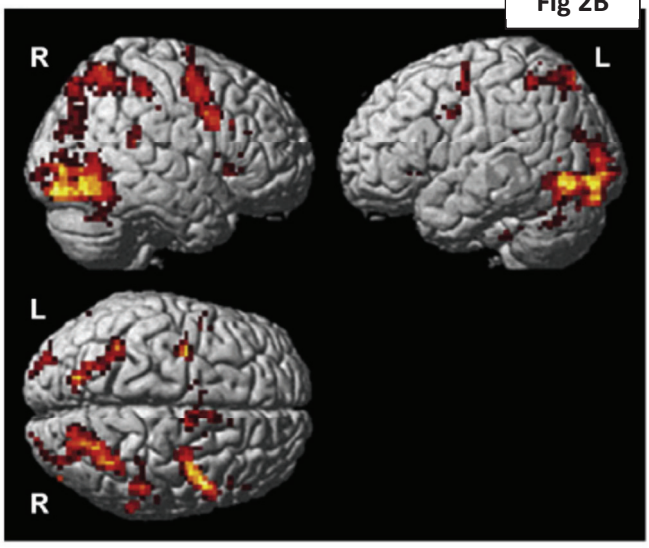


Fig 6) Activations pariétales et (pré)motrices durant l'observation d'actions

Fig 2A) Figure de gauche : Activations somatotopiquement organisées dans les cortex prémoteur et pariétal durant l'observation d'actions intransitives (en haut) et transitives (en bas). Les activations lors de l'observation d'actions effectuées avec la bouche sont représentées en rouge, celles avec la main en vert et celles avec le pied en bleu.
Adapté de De Buccino et al., 2004

Fig 2B) Figure de droite : Larges activations bilatérales situées dans le cortex parietal et s'étendant jusqu'au niveau occipito-temporal ; et dans le cortex prémoteur, jusqu'aux aires motrices supplémentaires et pré-supplémentaires.
Adapté de Ogawa & Inui, 2011

Ce système fronto-pariétal serait impliqué aussi bien dans l'apprentissage moteur, la programmation et l'exécution d'actions que dans la compréhension des actions effectuées par autrui (Jeannerod, 1994, Grafton et al., 1996, Buccino et al., 2001; Gazzola & Keysers, 2009; Grezes et al., 1998, 2001, 2003; Kilner et al., 2009; Iacoboni et al., 1999; Rizzolatti et al., 1996 ; Rizzolatti et al., 2001; Iacoboni & Dapretto, 2006 ; Fadiga et al., 1995). En effet, un nombre important de travaux en neuroimagerie chez l'humain font état d'activations des cortex pariétal et prémoteur lorsque les sujets observent ou imitent des actions, que ce soit via l'utilisation de la tomographie par émission de positons (TEP) (Grafton et al., 1996; Grezes et al., 1998), de la magnétoencéphalographie (MEG) (Nishitani & Hari, 2000) ou de l'IRMf (Iacoboni et al., 1999; Buccino et al., 2001; Tanaka & Inui, 2002; Johnson-Frey et al., 2003 ; Molnar-Szakacs et al. 2006 ; Ogawa & Inui, 2011) ou encore grâce à l'utilisation de la TMS (Fadiga et al., 1995 ; Strafella & Paus, 2000 ; Maeda et al., 2002 ; Gangitano et al., 2001; Watkins, Strafella, & Paus ; 2003, Urgesi et al., 2006). La première évidence directe de l'existence d'un système de neurones miroirs chez l'homme avait d'ailleurs été fournie par une expérience en TMS (Fadiga, Fogassi Pavesi, et Rizzolatti, 1995). Au cours de cette expérience, les potentiels évoqués moteurs (MEPs) étaient enregistrés à partir des muscles de la main et une impulsion TMS était appliquée alors que les sujets étaient en train d'observer un expérimentateur manipulant des objets ou dans différentes conditions contrôles (observation des objets seuls, observation de l'expérimentateur traçant des figures géométriques dans l'air avec son bras, détection d'un changement lumineux). Ils ont ainsi démontré que seule la condition « observation d'actions » augmentait considérablement les potentiels évoqués moteurs enregistrés au niveau des muscles des mains qui sont normalement recrutés quand l'action est réellement réalisée par l'observateur. Un travail récent (Mukamel et al., 2010) basé sur des enregistrements cellulaires a définitivement confirmé l'existence d'un système de neurones miroirs chez l'humain car ces enregistrements, chez des patients, ont permis d'obtenir une mesure directe de l'activité neuronale du cortex prémoteur frontal et du lobe temporal médial à la fois pendant l'observation et l'exécution d'actions. Ces auteurs ont montré qu'une proportion significative des neurones de l'aire motrice supplémentaire et de la région parahippocampique répondaient à la fois à l'observation et à l'exécution d'actions. Ils ont aussi relevé que des populations plus spécialisées de neurones étaient excitées pendant l'exécution de l'action mais inhibées pendant l'observation, cette propriété «miroir inverse » ayant déjà été mise en évidence chez le singe (Kraskov et al., 2009) et parfois chez l'homme via des diminutions d'activité du signal BOLD (Gazzola & Keysers, 2009). Elle est généralement interprétée comme une inhibition nécessaire de nos propres mouvements durant l'observation d'une action, permettant de faire la distinction entre nos propres mouvements et les mouvements des autres (Mukamel et al., 2010).

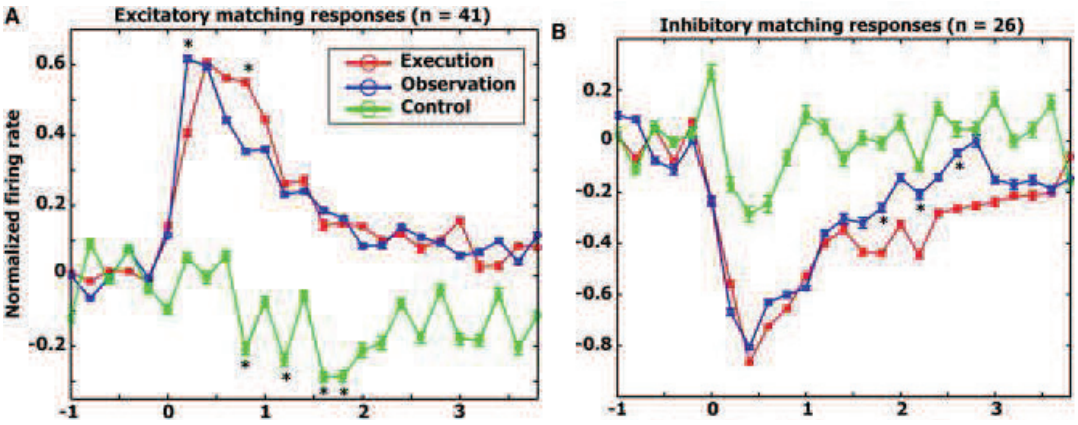


Fig 7) Représentation des réponses moyennées d'enregistrements effectués sur plusieurs neurones des cortex prémoteur et temporal

(A) Moyenne de 41 réponses excitratrices (provenant de 33 neurones différents) durant l'execution (rouge) ou l'observation (bleu) d'une action.
(B) Moyenne de 26 réponses inhibitrices (provenant de 21 neurones différents)

Dans les deux cas, le profil de réponse des neurones est le même pour la réalisation et l'observation de l'action : les neurones sont activés de façon similaire dans les deux conditions. Par ailleurs, ce profil de réponse est distinct de celui observé dans le cas de la condition contrôle (vert) où les patients n'avaient ni action à observer ou à effectuer. **Adapté de Mukamel et al., 2010**

- **<u>Actions humaines</u>**

Les activations du système fronto-pariétal seraient spécifiques du mouvement et des actions propres à l'humain. Dans une étude élécrophysiologique, Holz et al. (2008) ont montré des activations significativement plus importantes pour des mouvements effectués par des humains en comparaison avec des mouvements non-biologiques, au niveau des cortex moteur primaire et prémoteur, de l'aire motrice supplémentaire, ainsi que du cortex pariétal postérieur. Mason et al. (2004) ont comparé les activations cérébrales en IRMf de sujets qui devaient juger si des actions pouvaient être réalisées par des chiens ou des humains. Si les activations liées aux actions qui peuvent être effectuées uniquement par les chiens activent un réseau distribué avec de larges composantes occipitales et parahippocampiques, les activations liées aux actions humaines impliquent spécifiquement des aires du cortex préfrontal droit (notamment le gyrus frontal moyen et la partie médiale du cortex orbitofrontal) que les auteurs interprètent comme des activations liées aux processus de théorie de l'esprit (TOM). Sur un thème proche (comparaison d'actions buccales chez l'humain, le singe et le chien), les travaux en IRMf de Buccino et al (2004) suggèrent que la reconnaissance d'une action implique une connaissance « personnelle » de l'action (je peux reconnaitre cette action car je peux effectuer personnellement cette action). Nous sommes capables de reconnaître des actions effectuées par des individus d'une autre espèce, mais cette reconnaissance engagerait des régions cérébrales différentes selon que l'action observée génère une résonance motrice : les actions qui engagent le MNS fronto-pariétal sont celles qui appartiennent au

répertoire moteur de l'être humain (ex : mordre qui peut être effectuée par les différentes espèces) alors que les actions spécifiques d'autres espèces impliqueraient l'activation d'aires telles que le STS ou le gyrus fusiforme (ex : aboyer).

Cependant, la simulation sensorimotrice générée par l'observation d'actions ne serait pas dépendante du mouvement lui-même puisqu'on rencontre des activations similaires aussi bien dans des protocoles utilisant des vidéos (Buccino et al., 2001 ; Frey & Gerry, 2006) que dans des protocoles où une information dynamique est issue d'images statiques et donc où le mouvement est implicite (Urgesi et al, 2006 ; Johnson-Frey et al., 2003 ; Proverbio et al., 2009 ; Senior et al. 2000). Cette simulation serait très précise puisque les activations respecteraient une organisation somatotopique (Penfield & Boldrey, 1937) en fonction de l'effecteur employé (Buccino et al., 2001 Chong et al., 2008). En particulier, Buccino et al. (2001) ont montré que les réponses cérébrales générées par l'observation d'actions étaient organisées de façon somatotopique, à la fois dans les régions frontales et pariétales du MNS humain ; de plus l'organisation somatotopique du LPI correspondrait à celle trouvée chez le singe (Rizzolatti & Craighero, 2004). Les mouvements de la bouche activeraient bilatéralement le lobule pariétal inférieur, le gyrus précentral (BA6) et le GFi (au niveau de BA44), ainsi que le GFi droit (au niveau de BA45) ; les mouvements des mains activeraient bilatéralement l'aire interpariétale antérieure, ainsi que des portions plus dorsales de l'aire ventrale BA6 et de BA44, alors que les actions liées aux mouvements des pieds impliqueraient les lobules pariétaux supérieurs droit et gauche et une portion encore plus dorsale de l'aire 6 bilatéralement (Buccino et al., 2001). De même, la simulation sensorimotrice pourrait aussi refléter le degré d'effort que requiert une action, que cette action soit observée ou même imaginée (Decety et al., 1991 ; Decety et al., 1993 ; Proverbio et al., 2009).

- **<u>Actions et objets</u>**

Dans l'action, un rôle particulier est donné aux objets, qui peuvent être les médiateurs de l'action. Ainsi, Buccino et al., (2001) ont démontré que les mêmes systèmes sont impliqués lorsque l'action est transitive (impliquant un objet) ou intransitive (sans objet), l'observation des deux types de situations générant des activations similaires dans les régions prémotrices et au niveau de l'aire de Broca. Mais les actions avec objets nécessiteraient un traitement supplémentaire, caractérisé par des activations du lobe pariétal postérieur : les actions transitives activeraient à la fois les zones frontales et pariétales du MNS alors que les actions intransitives feraient intervenir uniquement le pôle frontal (ce manque d'activation du LPI pour les actions intransitives est cohérent avec les résultats de plusieurs autres travaux, Iacoboni et al. 1999, 2001 ; Koski et al. 2002, 2003). Cependant, pour Rizzolatti & Craighero (2004), si on considère que les zones prémotrices reçoivent des

informations visuelles en provenance du cortex pariétal postérieur (Rizzolatti, Luppino, & Matelli, 1998 ; Rizzolatti & Luppino, 2001), il est difficile de croire que ce dernier ne soit pas du tout activé lors de l'observation d'actions intransitives. Pour ces auteurs il serait plus probable que lorsqu'un objet est présent dans l'action, l'activation pariétale inférieure est beaucoup plus forte, ce qui masquerait statistiquement l'implication du LPI pour les actions sans objets. D'ailleurs, un certain nombre de données IRMf (Creem-Regehr & Lee, 2005 ; Boronat et al., 2005, Beauchamp & Martin, 2007; Martin, 2007 ; Chao & Martin, 2000) montrent en effet que l'observation passive d'objets/outils ou leur traitement sémantique dans des tâches de dénomination, activent également les cortex prémoteur ventral et pariétal inférieur. Ces activations sont généralement interprétées comme le résultat de l'action « affordée » par l'objet et par le fait qu'il existe des liens de compatibilité entre la reconnaissance de l'objet et les actions qu'il permet : identifier un objet manipulable impliquerait automatiquement la reconnaissance des actions qui lui sont associées. Tucker & Ellis (2003) ont relevé, dans une tâche d'identification d'objets, des temps de réaction plus courts lorsqu'il y a une compatibilité entre le geste de réponse et l'objet présenté visuellement. Non seulement, activer le réseau de la prise de l'objet faciliterait le traitement sémantique de l'objet mais un certain nombre de travaux montrent également que le réseau du codage de l'action recouperait celui de l'identification des objets (Vainio et al., 2008 ; Grèzes, Tucker & Ellis, 2003 ; Tucker & Ellis, 2004).

- **<u>Le réseau de l'observation de l'action</u>**

Par ailleurs, d'autres études en neuroimagerie suggèrent aussi que le MNS humain ferait partie d'un réseau plus large de régions cérébrales motrices et sensorielles, incluant notamment différentes régions pariétales et temporales et qu'il pourrait ainsi être associé à de nombreuses autres fonctions cognitives. On trouve ainsi l'implication de la jonction pariéto-temporale bilatéralement, du gyrus occipital supérieur, des lobules pariétaux inférieur et supérieur (Decety et al., 1997 ; Parsons et al., 1995; Buccino et al., 2001), du cortex prémoteur dorsal, l'aire motrice supplémentaire, de la portion moyenne du cortex cingulaire, (Gazzola & Keysers, 2009), du cortex temporal médian (aire MT) et du sillon temporal supérieur STS (Rizzolatti et al., 1996 ; Iacoboni et al., 2001, Beauchamp et al., 2002, Gazzola & Keysers, 2009) ainsi que des régions du système limbique (Wicker et al., 2003; Singer et al., 2004; Morrison et al., 2004), du cortex extra-strié (Buccino et al., 2004) et du cervelet (Gazzola & Keysers, 2009 ; Cattaneo et al., 2012). Différentes aires situées notamment au niveau du cortex occipito-temporal ne sont pas décrites initialement comme faisant partie du système des MNS mais seraient tout aussi importantes pour comprendre et effectuer les actions : certaines de ces régions seraient impliquées dans des traitements spécifiques des différents éléments d'une action (corps, effecteurs, objets…), alimentant alors « l'hypothèse visuelle ».

Cependant, cela ne signifie pas nécessairement que cela constituerait un traitement alternatif à celui effectué par le MNS. Pour des auteurs comme Astafiev (Astafiev et al., 2004), certaines de ces aires pourraient même appartenir au MNS humain ou du moins, agir de concert avec lui dans la compréhension des actions.

- *Rôle de l'EBA (« extrastriate body area ») dans la perception du corps*

Une aire semble particulièrement importante pour la perception du corps humain ou des parties du corps, il s'agit de l'EBA (pour « extrastriate body area »), une aire qui réagit spécifiquement à la présentation d'images représentant le corps humain par rapport à d'autres catégories de stimuli tels que des objets (Downing et al., 2001, 2006; Urgesi et al., 2006; Peelen & Downing, 2005; Saxe et al., 2006; Peelen et al., 2006; Spiridon et al., 2006). Plusieurs études (Astafiev et al., 2004 ; Peelen & Dowing, 2005, David et al., 2007) suggèrent que la fonction de cette aire ne serait pas purement perceptive et qu'elle aurait aussi des propriétés motrices puisque cette région serait responsable de l'identification du corps d'autres personnes, de la capacité d'inférer les actions des autres mais aussi de la perception de la position de son propre corps afin de guider une action. En plus d'être activée par des images représentant les différentes parties du corps d'autres personnes, l'activité de cette région serait modulée par la planification, l'exécution et la capacité d'imaginer ses propres mouvements (Astafiev et al., 2004 ; David et al., 2007). C'est ce qui donnerait à l'EBA sa capacité à distinguer si les actions sont réalisées par soi-même ou par une autre personne, cette capacité étant un prérequis préalable à l'agentivité (Gallagher, 2000 ; David et al., 2007). Selon Astavief et al., (2004), si l'EBA est impliquée à la fois dans la propre représentation du corps de l'observateur et dans la perception du corps des autres individus, une explication possible serait que l'EBA fait partie du système mis en oeuvre à la fois pour la perception et l'action, c'est-à-dire le MNS.

- *Sillon temporal supérieur (STS) et intégration multisensorielle*

De la même façon que pour l'EBA, certaines aires antérieures du sillon temporal supérieur (STS), qui contiennent également des neurones répondant à l'observation d'actions, sont parfois considérées comme des régions secondairement associées au MNS chez le singe (Perrett et al., 1989 ; Perrett et al., 1990 ; Iacoboni et al., 2001, Jellema et al., 2000 ; Jellema et al. 2002 ; Rizzolati & Craighero, 2004). Chez le singe, le STS coderait pour un nombre plus large de mouvements que l'aire F5, mais le STS n'aurait pas des propriétés motrices aussi fines que les autres régions du MNS et de ce fait ne pourrait être considéré comme appartenant pleinement au MNS selon Rizzolati & Craighero (2004). Chez l'humain, le STS a été associé à de nombreuses fonctions cognitives (Hein & Knight, 2008 ; Carter et al., 2011), de la théorie de l'esprit à l'intégration audiovisuelle, en passant par le traitement du langage et des visages… Mais cette région aurait également un rôle particulièrement

important dans la perception du mouvement biologique (Puce & Perrett, 2003; Allison, Puce, &McCarthy, 2000, Bonda et al., 1996; Puce et al., 1998), ainsi que dans la compréhension du but des actions et des intentions (Hoffman & Haxby, 2000 ; Carter et al., 2011; Pelphrey et al., 2004 ; Pelphrey et al., 2005 ; Vander Wyk et al., 2009). Le rôle du STS serait crucial car il permettrait une intégration multisensorielle audio-visuelle mais aussi forme-mouvement (Beauchamp, 2005 ; Beauchamp et al., 2002 ; Iacoboni et al., 2001).

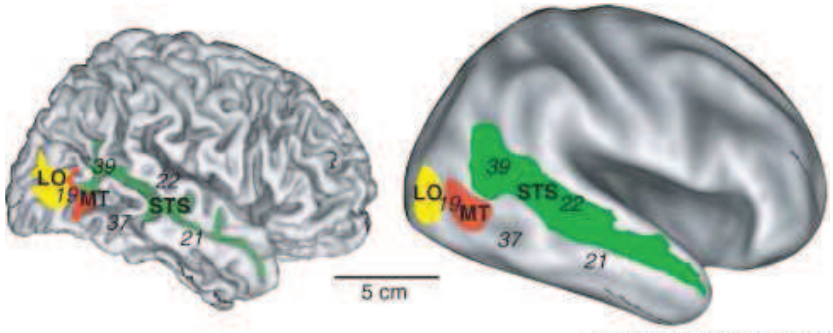


Fig 8) Localisation de trois régions multisensorielles du cortex occipito-temporal sur une vue latérale : aire latéro-occipital LO, aire MT (V5) et sillon temporal supérieur (STS)

Adapté de Beauchamp, 2005

- *L'aire MT et le traitement du mouvement biologique*

De la même façon que de nombreux travaux montrent que l'EBA et le STS n'ont pas un rôle qui se limite aux traitements perceptifs, des suggestions similaires ont été faites concernant l'aire MT connue pour avoir un rôle spécifique dans le traitement du mouvement et dont la localisation recoupe parfois ou du moins est adjacente à celle de l'EBA (Oreja-Guevara et al., 2004; Hamilton et Grafton, 2006; Jackson et al.,2006). En effet, le rôle du cortex extrastrié temporal moyen (plus souvent appelé aire MT/ V5 ou hMT) dans le traitement visuel du mouvement biologique a largement été démontré chez l'humain et le singe (Dupont et al., 1994; Tootell et al., 1995 ; Born & Bradley, 2005 ; Zeki et al., 1991 ; Maunsell & Van Essen, 1983), aussi bien lorsqu'il s'agit d'un mouvement réellement observé que lorsque le mouvement est illusoire, imaginé ou implicitement représenté dans des stimuli statiques (Williams & Wright, 2009 ; Kourtzi & Kanwisher, 2000, Goldenberg, 2001, Senior et al., 2000, Urgesi et al., 2006 ; Dumoulin et al., 2001). Pour certains auteurs, le rôle de l'aire MT ne se limiterait pas au traitement d'aspects perceptifs car cette aire participerait à des étapes précoces mais essentielles à la fois de la perception et de la réalisation des actions (Schenk et al., 2000; Oreja-Guevara et al., 2004; Hamilton et al., 2005). Cependant, l'aire MT resterait majoritairement sous l'influence descendante de processus attentionnels de plus hauts niveaux (Kourtzi & Kanwisher, 2000 ; Williams & Wright, 2009).

- *Jonction temporo-pariétale et représentations corporelles*

Il est également intéressant de constater que d'autres aires du cortex postérieur latéral, situées à proximité de l'EBA et du STS, sont aussi impliquées dans les représentations corporelles (body schema), notamment la jonction temporo-pariétale (Peigneux et al., 2000) et, en particulier, le gyrus angulaire. Une étude de Blanke et al., (2002) démontre que l'on peut induire une perception illusoire de mouvement corporel en stimulant le gyrus angulaire. Des lésions du gyrus angulaire et du cortex occipital latéral, particulièrement si elles sont latéralisées à gauche, peuvent induire des perturbations, une perception faussée du schéma corporel (Halligan et al., 1995 ; Coslett et al., 1998). Mais le gyrus angulaire a aussi un rôle plus général dans la capacité à créer mentalement la représentation d'un événement ou à retrouver en mémoire un événement passé (Huijbers et al., 2011 ; Vilberg & Rugg, 2008; Wagner et al.,2005). Plus globalement, le cortex pariétal postérieur (CPP) semble également avoir un rôle essentiel dans la compréhension des actions de par son rôle dans la détection de l'incongruence sensorimotrice, la représentation des agents effectuant une action et la prise de perspective (Chaminade & Decety, 2002; David et al., 2006; Farrer & Frith, 2002; Jeannerod, 2004, MacDonald & Paus, 2003).

- *Cervelet et évaluation de la séquentialité*

Le cervelet participerait également à la compréhension des actions puisqu'on retrouve des activations de cette région dans plusieurs travaux (Calvo-Merino et al., 2006 ; Gazzola & Keysers, 2008 ; Hamilton et al., 2006 ; Cattaneo et al., 2012). Le rôle du cervelet serait lié à l'évaluation des aspects temporels de l'action, il permettrait d'ordonner les différents éléments séquentiels qui composent une action ou un événement et aurait ainsi un rôle essentiel et spécifique dans leur compréhension. Cattaneo et al. (2012) ont comparé les performances de patients présentant des lésions du cervelet pendant qu'ils effectuaient une tâche consistant à identifier et à ordonner une séquence de 4 images représentant soit des actions effectuées par un acteur (action biologique), soit un événement modifiant l'état d'un objet inanimé mais sans intervention humaine. Même si les patients présentent des difficultés à effectuer la tâche dans les deux conditions, il s'avère qu'ils sont plus fortement perturbés dans la condition impliquant une action biologique. Pour Cattaneo et al (2012), le cervelet n'aurait pas seulement un rôle dans le contrôle et l'organisation temporelle des séquences de mouvements produites ou générées de façon interne ; il aurait aussi un rôle important dans les processus de cognition incarnée ainsi que dans l'activité liée au MNS et pourrait même jouer un rôle dans les domaines affectif et de la cognition sociale.

On voit donc que l'observation d'actions fait appel aux aires traditionnellement désignées comme le MNS, à savoir les cortex prémoteur ventral et dorsal, le lobule pariétal inférieur, le cortex moteur primaire (Rizzolatti & Craighero, 2004; Morin & Grezes, 2008; Gazzola & Keysers, 2009; Kilner et al., 2009; Rizzolatti & Sinigaglia, 2010). Toutefois, l'idée que le MNS humain constitue l'exact homologue du système miroir observé chez le singe reste encore controversée, les études en neuroimagerie ne permettant pas pleinement de répondre aux interrogations quant à la nature des neurones et de leur profil de décharge chez l'humain et certains auteurs préfèrent se référer au terme de réseau de l'observation de l'action (AON, action observation network, Press et al., 2011). Par ailleurs, le MNS différerait chez l'humain et le singe pour plusieurs raisons. Chez le singe, ce serait essentiellement les mouvements dirigés vers un but qui activeraient le système des neurones miroirs. Chez l'humain, même si l'analyse des buts de l'action reste une fonction primordiale du MNS humain, ce système pourrait aussi être activé par des actions sans signification particulière, par des mouvements formant les différentes parties constitutives d'une action ou encore quand le but d'une action est anticipé sans avoir besoin d'observer l'action complète (Decety et al., 1997 ; Grezes, Costes & Decety, 1998 ; Fadiga et al. 1995, Maeda et al. 2002, Patuzzo et al. 2003, Umilta et al., 2001 ; et al, 2003). C'est ce qui donnerait au MNS humain une forte composante sociale et un rôle déterminant dans la capacité de compréhension des intentions mais aussi d'imitation (Rizzolatti & al., 2001 ; Rizzolatti, & Craighero, 2004 ; Frey & Gerry, 2006). Enfin, comme nous l'avons déjà mentionné précédemment, le MNS serait intégré au sein d'un réseau plus large d'aires cérébrales (incluant une grande variété d'aires pariéto-temporales postérieures aux fonctions plus ou moins spécifiques) dont le rôle ne se limiterait pas à la compréhension des actions. Ce réseau plus large et sa capacité à faire correspondre les perceptions issues de l'environnement à des représentations sensori-moteurs internes, joueraient un rôle essentiel dans de multiples aspects de la cognition sociale allant de la perception des actions à l'empathie en passant par l'imitation, la compréhension des intentions d'autrui, la communication gestuelle et via les expressions faciales, le langage… (Oberman, Pineda & Ramachandran, 2007 ; Rizzolati & Craighero, 2004, Iacoboni et al, 2005 ; Buccino et al., 2004).

B) Travaux impliquant des stimuli verbaux

Appliquées au langage, les théories de la cognition incarnée prétendent que lorsque nous nous représentons et comprenons les mots, les mêmes aires sensorimotrices sont recrutées, comme lorsque nous interagissons, dans le monde réel, avec les objets et autres entités auxquelles les mots se réfèrent. Comprendre des phrases fait intervenir une simulation interne de l'état du monde, décrit par les phrases (Zwaan, 2003), ce que Glenberg et Robertson ont intitulé « The Indexical Hypothesis » (Glenberg et Robertson, 1999 ; Glenberg & Kaschak, 2002). Ces dernières années, de nombreux travaux comportementaux et en neuroimagerie ont accumulé les preuves que la compréhension langagière faisait intervenir une simulation sensorimotrice et impliquait le système moteur (au moins partiellement) car traiter les stimuli langagiers ferait appel aux mêmes unités de neurones que les actions désignées par les mots (Willems & Hagoort, 2007 ; Gallese, 2008 ; Barsalou, 2008; Fischer & Zwaan, 2008; Pulvermüller, 2005 ; Jirak et al., 2010 ; Hamzei et al., 2003 ; Tettamanti et al., 2005). De même que l'implication de l'aire de Broca dans l'observation et la compréhension des actions a été largement démontrée, des aires autrefois considérées comme purement motrices (telles que le cortex primaire moteur et le cortex prémoteur) sont aujourd'hui considérées comme essentielles pour la compréhension du langage (Jirak et al. 2010). La compréhension du langage générerait, elle aussi, une « résonance motrice » comparable à celle étudiée dans l'observation d'actions effectuées par autrui (Zwaan & Taylor, 2006). On ne peut plus, dès lors, considérer les aires cérébrales de la motricité et du langage comme des aires fonctionnellement spécifiques et indépendantes les unes des autres, mais plutôt comme des aires dont l'activation conjointe fournit le lien entre action et langage (Jirak et al. 2010 ; Pulvermüller, 2005).

- **Travaux comportementaux**

Au niveau comportemental, Glenberg et Kaschak (2002) ont mis en évidence un effet de compatibilité Action-phrase (Action-sentence Compatibility Effect, ACE) : les participants devaient produire un jugement à propos de phrases décrivant le transfert d'un objet concret ou d'un concept abstrait, le transfert pouvant être dirigé vers soi ou vers une autre personne ("Andy delivered the pizza to you/You delivered the pizza to Andy" vs "Liz told you the story/You told Liz the story") en répondant par l'intermédiaire de deux boutons (un proche et un éloigné). En fonction de la réponse, le mouvement à effectuer pour sélectionner l'un de ces boutons de réponse pouvait être compatible ou non avec la direction de l'action décrite dans la phrase (donner à quelqu'un = déplacement vers le bouton le plus éloigné vs recevoir de quelqu'un = déplacement vers le bouton le plus proche). Les auteurs ont trouvé une interaction entre le temps pour juger les phrases et la direction du mouvement : le jugement était plus rapide quand l'action décrite par la phrase correspondait à

l'action à produire pour répondre et ce, pour les phrases abstraites et concrètes. D'autres études comportementales ont montré des effets de compatibilité actions-phrases assez similaires (De Vega, 2008 ; Zwaan & Taylor, 2006 ; Tucker et Ellis, 2004) ou des effets de congruence actions-mots pour des items désignant des objets manipulables (Rueschemeyer et al., 2010) ou encore une modulation des réponses motrices en fonction du contenu du matériel verbal (Buccino et al., 2005; Scorolli et al., 2009). Les études comportementales ont également démontré que ces effets de compatibilité dépendaient des effecteurs impliqués : la simulation activée lors de la compréhension d'éléments langagiers (noms et verbes) serait sensible à la congruence entre l'effecteur impliqué (bouche/pied/jambe) et l'effecteur utilisé pour produire la réponse motrice (Scorolli et Borghi, 2007 ; Borghi et Scorolli, 2009). Par exemple, dans l'étude de Scorolli et Borghi (2007), lorsque les participants utilisaient leur bouche pour répondre, ils étaient plus rapides pour répondre à des paires de mots décrivant des actions effectuées avec la bouche qu'à des paires de mots décrivant des actions de la main. De même, Borghi et Scorolli (2009) ont montré un effet de facilitation lorsque il y avait une compatibilité entre l'effecteur dominant (la main droite pour les droitiers) utilisé pour produire une réponse motrice et celui évoqué par la phrase décrivant une action.

- **Travaux en neuroimagerie**

Les résultats d'un certain nombre d'études réalisées en IRMf suggèrent l'implication d'un large circuit fronto-temporo-pariétal, principalement latéralisé à gauche, dans la compréhension des stimuli langagiers (Jirak et al., 2010). L'aire de Broca, considérée comme l'homologue de l'aire F5 chez le singe et dans laquelle on a démontré l'existence de neurones avec des propriétés miroirs chez l'homme (Buccino et al., 2005 ; Geyer et al., 2000) pourrait fournir le lien entre langage et action puisque cette aire était considérée principalement langagière jusque-là ; mais l'ensemble du MNS pourrait également avoir un impact sur les traitements langagiers et sur la compréhension (Jirak et al. 2010 ; Glenberg et al., 2008). Plusieurs études en IRMf présentent effectivement des activations des zones caractérisées comme le MNS humain en réponse à la présentation de phrases d'actions ou de verbes, en particulier la partie operculaire du gyrus frontal inférieur et le lobule pariétal inférieur (Tranel et al., 2001 ; Pulvermuller, Lutzenberger, & Preissl, 1999; Martin et al., 1995 ; Tettamanti et al., 2005 ; Jirak et al., 2010).

  - *MNS et compréhension du langage*

Mais si les régions sensorimotrices décrites semblent proches de celles qui appartiennent au MNS humain, l'implication de ce système dans la compréhension du langage reste débattue. Les travaux qui ont cherché à comparer directement l'observation d'actions et la lecture de phrases d'actions rapportent des résultats parfois contradictoires et décrivent des activations proches, mais

ne se recoupant pas totalement, au niveau des aires prémotrices et préfrontales (Aziz-Zadeh et al., 2006; Aziz-Zadeh & Damasio, 2008 ; Hamzei et al., 2003 ; nous traiterons ce point plus en détail ultérieurement). Pour certains, le MNS ne serait pas responsable de la compréhension du langage per se, mais il pourrait jouer un rôle important en tant que précurseur de son développement (Rizzolatti & Arbib, 1998 ; Aziz-Zadeh & Damasio, 2008 ; Aziz-Zadeh et al. 2006). Cependant, plusieurs travaux font également état des implications du MNS, au niveau de l'aire de Broca et du lobule pariétal inférieur, lorsqu'il s'agit de traiter les sons spécifiques d'une action (Kohler et al., 2002 ; D'Ausilio et al. 2009 ; McNamara et al. 2008). Pour Jirak et al. (2010), l'implication des régions motrices dans le traitement du langage resterait tout de même étroitement liée aux régions du MNS, avec une possible utilisation de ce système pour intégrer ce qui est lié aux sons dans la compréhension des actions mais aussi éventuellement pour permettre la simulation nécessaire à la compréhension des mots d'actions.

- *Un réseau plus large d'aires sensorimotrices*

Indépendamment d'une implication spécifique du MNS, un large réseau d'aires sensorimotrices semble toutefois impliqué dans la compréhension du langage. La méta-analyse ALE (Activation likelihood estimation) de Jirak et al. (2010) qui est basée sur l'étude d'une vingtaine d'expériences IRMf comportant des stimuli langagiers relatifs à des actions met en évidence des aires à la fois langagières et motrices majoritairement latéralisées à gauche et situées principalement au niveau frontal dans la région des gyri frontal inférieur (BA44/45/46) et précentral (BA6/4/9), au niveau pariétal dans le gyrus supramarginal (BA40) et le précuneus (BA19), au niveau temporal dans le gyrus temporal moyen (BA22/39) et le gyrus fusiforme (BA37), ainsi qu'au niveau de l'insula (BA13). Les auteurs notent aussi des activations bilatérales dans le cervelet et au niveau du gyrus cingulaire postérieur (BA30) ainsi que des activations de l'hémisphère droit au niveau du lobule pariétal supérieur (BA5/7) et du gyrus précentral (BA4). Pour Jirak et al. (2010), les résultats de cette méta-analyse confirment bien l'implication de régions sensori-motrices dans la compréhension du langage puisque des aires motrices (telles que la SMA, le gyrus précentral, le cortex prémoteur mais aussi le cervelet) sont activées alors même que les tâches utilisées sont purement langagières, ne requérant aucune action motrice pour répondre de la part des sujets. Les activations frontales et pariétales de l'hémisphère droit confirment l'existence d'un réseau bilatéral, la participation de l'hémisphère droit dans des aspects sémantiques ayant déjà été mise en évidence par d'autres études (Canessa et al., 2008; Gronau, Neta, & Bar, 2008; Kemmerer et al., 2008). Les activations cérébelleuses sont tout aussi intéressantes car elles suggèrent que la compréhension du langage serait « incarnée », non pas uniquement en raison des activations sensorimotrices mais aussi parce qu'elle ferait intervenir un réseau fronto-pariéto-temporal très étendu avec une activation sous-

corticale (Jirak et al., 2010); plusieurs études confirment en effet l'implication du cervelet dans la compréhension sémantique langagière (Liljeström et al., 2008 ; Saccuman et al., 2006).

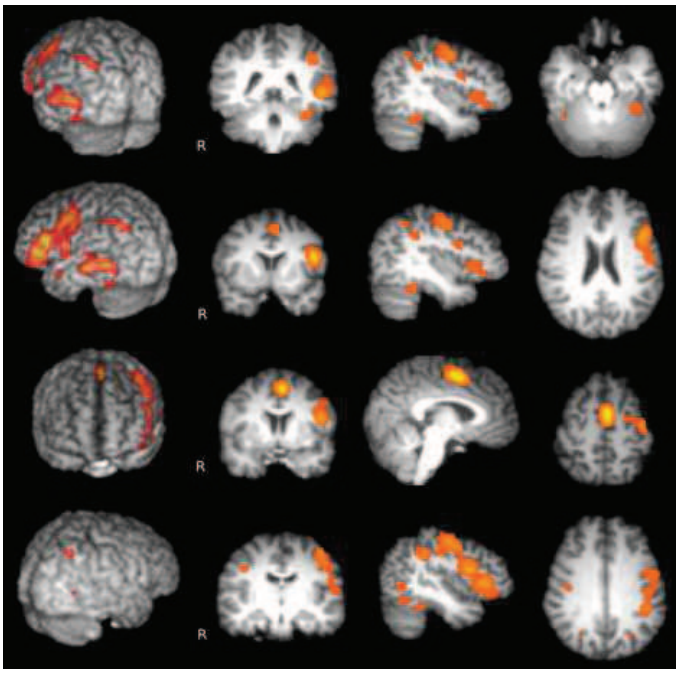


Fig 9) Résultats de la méta-analyse ALE réalisée par Jirak et al. (2010)
Basée sur une vingtaine d'expériences IRMf comportant des stimuli langagiers relatifs à des actions, cette étude met en évidence un large réseau d'aires à la fois langagières et motrices.
Adapté de Jirak et al., 2010

- *Organisation somatotopique*

Par ailleurs, de nombreuses travaux, notamment en neuroimagerie, ont également souligné que les activations provoquées par des stimuli verbaux engageaient le système sensorimoteur selon une organisation somatotopique (Jirak et al., 2010 ; Kemmerer et al, 2008 ; Rueschemeyer et al., 2010). La compréhension de verbes ou de mots d'action (Hauk et al, 2004; Pulvermüller et al, 2005 ; Kemmerer et al, 2008 ; Kable et al., 2002 ; Hauk & Pulvermüller, 2011), de phrases d'action (Tettamanti et al, 2005; Aziz-Zadeh et al, 2006) mais aussi des mots désignant des objets manipulables (Chao et Martin, 2000; Saccuman et al, 2006) activent le système moteur cérébral de façon différente selon l'effecteur employé. Ainsi les actions des bras provoqueraient des activations au niveau du gyrus frontal moyen et du gyrus précentral, tandis que les actions liées aux jambes impliqueraient des activations dans des régions plus dorsales au niveau des gyri precentral et postcentral ainsi qu'au niveau du cortex prémoteur dorsal (Hauk et al., 2004).

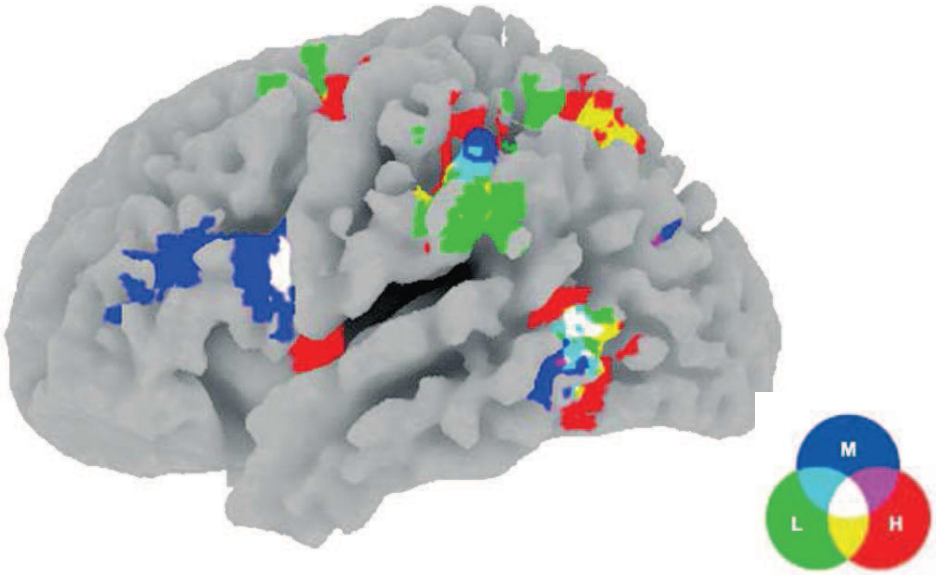


Fig 10) Réseau sensori-moteur activé durant l'écoute de phrases d'action
Les activations sont principalement localisées dans le cortex pariétal inférieur, les gyri frontal inférieur et précentral, les régions temporales moyennes gauches. Elles sont organisées de façon somatotopique (en rouge : les activations liées aux phrases impliquant la main ; en vert : celles concernant le pied ; en bleu : celles concernant les actions effectuées avec la bouche.
Adapté de Tettamanti et al., 2005

Dans un travail en IRMf de Tettamanti et al (2005), les sujets devaient écouter des phrases décrivant des actions effectuées avec la bouche, la main ou la jambe/pied. On note des activations variées en fonction des zones du corps mentionnées dans les phrases, principalement localisées dans l'hémisphère gauche au niveau du sillon intrapariétal postérieur gauche, du gyrus précentral et de la partie postérieure du cortex inféro-temporal pour les phrases impliquant la main, tandis que les phrases liées à la motricité des jambes impliquent des activations au niveau du sillon intrapariétal, mais aussi au niveau de la partie postérieure du cortex prémoteur et du lobule pariétal inférieur gauche ; les actions effectuées avec la bouche, quant à elles, activent principalement le gyrus frontal inférieur (pars opercularis + pars triangularis) et le lobule pariétal inférieur. Pour ces différentes actions, le gyrus temporal moyen (MTG), à proximité du TPJ, semble être une zone de convergence essentielle car il est activé par les différents types d'actions (principalement à gauche mais aussi bilatéralement) avec en particulier une activation de la zone MT/V5 qui, comme nous l'avons déjà mentionné, est connue pour être impliquée dans le traitement visuel du mouvement. Kable et al. (2002, 2005) ont aussi trouvé des activations du cortex moteur liées à la présentation de verbes d'action et ont également identifié une activation accrue aux alentours de la zone du mouvement MT/V5 lors de jugements d'actions par rapport à des jugements d'objets. Buccino et al. (2005) ont montré que les potentiels évoqués moteurs (MEPs) des mains et des pieds/jambes étaient modulés spécifiquement durant une tâche d'écoute de phrases relatives aux actions de ces différentes parties du corps (des données comportementales confirmées par une étude complémentaire en TMS).

Willems, Hagoort et Casasanto (2010) ont mis en évidence une activation du cortex prémoteur gauche pour les droitiers et du cortex prémoteur droit pour les gauchers, lorsque les sujets devaient juger des verbes d'actions réalisées avec la main par rapport à des verbes non manuels ; la simulation évoquée durant la compréhension du langage serait donc spécifique en fonction de notre propre corps et de nos propres expériences (Willems et al., 2010 ; Hauk & Pullvermuler, 2011).

- *Actions impliquant des objets*

De plus, les actions impliquant un objet mettraient en jeu également le même réseau, lorsque celles-ci sont décrites verbalement (Kemmerer et al. 2008). Yang et al., (2011) ont notamment montré que des verbes d'actions manuelles impliquant ou non un objet activent les mêmes régions principalement latéralisées à gauche : le gyrus frontal moyen/précentral (BA 6/9), GFi (BA44/45), le cortex préfrontal dorsolatéral (BA46/10) et les lobules pariétaux inférieur et supérieur (BA40/7) mais aussi le cervelet. Cependant, les verbes faisant référence à l'utilisation d'objets génèrent une plus grande activation du lobule pariétal supérieur, du gyrus frontal moyen et du gyrus temporal moyen postérieur, ainsi qu'une plus grande connectivité entre ces régions. Tous ces travaux confirment également l'implication d'un système sensorimoteur fronto-pariétal dans la compréhension des actions au niveau verbal.

- **Données neuropsychologiques**

Les travaux concernant des données neurospychologiques sont plus contradictoires, même si certains d'entre eux semblent confirmer l'existence de liens entre système moteur et compréhension langagière. Il a notamment été montré que la dégénérescence des neurones du système moteur (dans le cas d'une sclérose latérale amyotrophique ou d'une maladie héréditaire) pouvait affecter la compréhension des verbes d'actions (Bak et Hodges, 2003 ; Bak et al., 2006). Une étude sur des patients parkinsoniens (Boulenger et al., 2008) révèle une absence d'effet d'amorçage pour les verbes d'actions, reflétant une activation anormale des aires motrices et prémotrices, lorsque les patients n'ont pas de traitement dopaminergique alors que cet effet est présent chez des sujets contrôles et chez des patients traités. Plusieurs autres travaux recensent des déficits dans la compréhension langagière d'actions chez des patients présentant des lésions au niveau pariétal et prémoteur (Aziz-Zadeh & Damasio, 2008 ; Kemmerer et Tranel, 2003 ; Saygin et al., 2004). Rodríguez-Ferreiro et al (2009) ont montré que la capacité de patients parkinsoniens à dénommer des actions était spécifiquement altérée (par rapport à des objets). Ces différents travaux suggèrent que l'intégrité du système moteur est importante pour la compréhension langagière (Jirak et al., 2010).

- **Eléments verbaux responsables de la simulation sensorimotrice**

Même s'il est maintenant admis que la compréhension du langage implique une simulation mentale perceptivo-motrice, des travaux, dont les résultats sont plus controversés, ont cherché à déterminer précisément quels étaient les composants langagiers responsables de cette simulation dans une phrase et si des aires cérébrales spécifiques étaient impliquées dans la compréhension des différents éléments qui composent une phrase. Associée cette fois au langage, on retrouve la même problématique que celle qui, pour les stimuli visuels, opposait l'hypothèse visuelle et l'hypothèse d'appariement direct : certains indices (objets, corps, mouvements…) sont-ils essentiels et/ou suffisants à la simulation sensorimotrice et impliquent-ils des aires cérébrales spécifiques ? Dans le cadre du langage, les mots et/ou les verbes sont-ils les seuls composants essentiels ou les structures grammaticales peuvent-elles également contribuer à cette simulation ? Existe-t-il une différence entre les concepts d'objets (généralement désignés par des noms dans le langage) et les événements (généralement désignés par des verbes et des phrases dans le langage). En effet, interpréter une phrase va bien au-delà d'une simple simulation sensori-motrice et requiert l'intégration de multiples aspects, notamment des aspects syntaxiques, thématiques et de connaissances à propos du monde réel.

- *Abstrait vs Concret*

Intuitivement, il paraît évident que les représentations lexicales se développent pendant l'acquisition du langage grâce aux associations qui sont faites entre le signifiant et le signifié, c'est à dire entre les mots et les aspects perceptivo-moteurs des objets et situations auxquels ils font référence (Bergen et Wheeler, 2010). Mais les associations lexicales ne seraient pas forcément suffisantes, ainsi Bergen et al. (2007) ont constaté que des mots concrets décrivant un mouvement vers le bas ou vers le haut généraient des activations du système de la vision uniquement lorsqu'ils étaient intégrés dans une phrase décrivant un mouvement réel et non pour des phrases décrivant un mouvement métaphorique (e.g « les prix ont grimpé »). Aziz-Zadeh et al. (2006) constatent une différence similaire des activations somatotopiques du cortex prémoteur pour des phrases décrivant des actions concrètes impliquant le pied, la main ou la bouche vs des phrases utilisant les mêmes effecteurs mais dans un contexte métaphorique (ex : serrer les dents, saisir le sens). Tettamanti et al. (2005) font aussi référence à des activations spécifiques du cortex prémoteur pour des phrases d'actions par rapport à des phrases abstraites. Cependant, ces résultats sont controversés puisque d'autres auteurs (Glenberg et al., 2008 ; Rüschemeyer, Brass, & Friederici, 2007 ; Raposo, Moss, Stamatakis, & Tyler, 2009) ont montré des activations motrices similaires pour des stimuli langagiers abstraits ou concrets, ce qui paraît confirmer les hypothèses des théories les plus radicales de la

cognition incarnée, à savoir que le traitement des concepts abstraits implique le système sensorimoteur au même titre que les concepts concrets. Certains auteurs (Aziz-Zadeh et Damasio, 2008) sont plus nuancés et reconnaissent que ces mêmes réseaux cérébraux pourraient être impliqués, dans un premier temps, car on ferait initialement appel à des représentations « imagées » pour appréhender le sens d'une métaphore mais qu'une fois celui-ci bien établi, cela ne nécessiterait plus l'intervention du même réseau cérébral, ce que semblent confirmer des études qui montrent des zones cérébrales distinctes intervenant pour le traitement de nouvelles métaphores ou des métaphores plus classiques (Bottini et al., 1994; Mashal et al., 2007; Schmidt et al., 2007).

- *Verbes vs Noms*

Par ailleurs, les verbes et les noms pourraient être traités différemment dans le cerveau, en fonction de leur degré d'abstraction : les différents niveaux de dérivation de la signification littérale d'un mot pourraient conduire à des patterns d'activations cérébrales variés (Jirak et al., 2010 ; Newman et al., 2009). Bien que plusieurs travaux montrent des patterns d'activations similaires impliquant les zones motrices à la fois pour des mots et des verbes d'action (Bedny et al. 2008 ; Hauk et al, 2004; Pulvermüller et al, 2005 ; Kable et al., 2002 ; Boulenger et al., 2008 ; Canessa et al. 2008 ; Gennari et al., 2007 ; Gronau et al. 2008 ; Liljeström et al. 2008 ; Meister & Iacoboni 2007 ; Chao et Martin, 2000; Saccuman et al, 2006 ; Tyler et al., 2001; Vigliocco & Kita, 2006; Warburton et al., 1996), d'autres travaux signalent des patterns d'activations différents pour les noms et les verbes. Le gyrus temporal moyen postérieur et/ou le GFi gauches montreraient une activation plus importante pour les verbes (Davis et al., 2004; Fiez et al., 1996; Perani et al., 1999; Shapiro et al., 1991, 2005, 2006; Tranel et al., 2005; Tyler et al., 2004; Yokoyama et al., 2006) alors que les régions temporales inférieures gauches seraient plus fortement impliquées dans le traitement des noms (Shapiro et al., 2005; Shapiro et al., 2006). Pour Newman et al. (2009), le rôle du GFi (BA 44) reste avant tout associé au traitement syntaxique des verbes, puisque ces auteurs évoquent une activation spécifique de cette aire associée à des effets de facilitation et d'amorçage syntaxique (activation que les auteurs n'ont pas retrouvée pour les mots). Mais pour d'autres auteurs (Vigliocco et al., 2011), les différences de patterns d'activation entre noms et verbes ne seraient pas nécessairement liées à des différences syntaxiques mais plutôt à des différences sémantiques. Dans un important travail de revue, Vigliocco et al. (2011) recensent tous les travaux (comportements, neuropsychologie, neurophysiologie, TMS…) sur les différences de traitement entre noms et verbes. Ce travail met en évidence un certain nombre de problèmes méthodologiques liés au fait que dans ces études, sont souvent confondus des aspects sémantiques (objets vs actions) et grammaticaux (noms vs verbes) et qu'il y a un amalgame entre les études qui portent sur des traitements simples (mots isolés) et des traitements plus complexes (par exemple dans le cadre de l'intégration sémantique au sein d'une phrase) et il en

résulterait des données parfois contradictoires. Pour Vigliocco et al., (2011), si on tient compte de ces biais, le tableau est assez clair : il existe en effet une séparation au niveau neural entre le traitement des objets (noms) et des mots désignant des actions (principalement les verbes), confirmée par les études lésionnelles, en TMS et en neuroimagerie. D'après la synthèse de Vigliocco, le cortex temporal inférieur gauche serait impliqué dans la dénomination et la représentation des connaissances concernant les objets alors que le cortex préfrontal serait, quant à lui, plutôt recruté par tout ce qui concerne les actions. Il existerait cependant une zone engagée plus spécifiquement dans des processus d'intégration et de traitement morphosyntaxique, indistinctement pour les verbes et les noms ; elle se situerait au niveau du GFi gauche et du gyrus frontal moyen. Mais pour ces auteurs, les aspects grammaticaux ne seraient pas l'élément responsable de l'organisation cérébrale distincte des connaissances concernant les noms et les verbes. Les traitements des mots appartenant à différentes classes grammaticales (noms et verbes) seraient effectués par un système procédural combinatoire partagé et où la distinction noms-verbes refléterait en fait des différences d'accessibilité sémantique (Vigliocco et al , 2011 ; Martin & Chao, 2001; Martin, 2007; Cappa, 2008). Ces effets entre noms et verbes seraient d'autant plus marqués dans des tâches demandant un traitement plus approfondi (typiquement des tâches à forte contrainte sémantique ou lors de la compréhension d'un mot intégré dans un contexte).

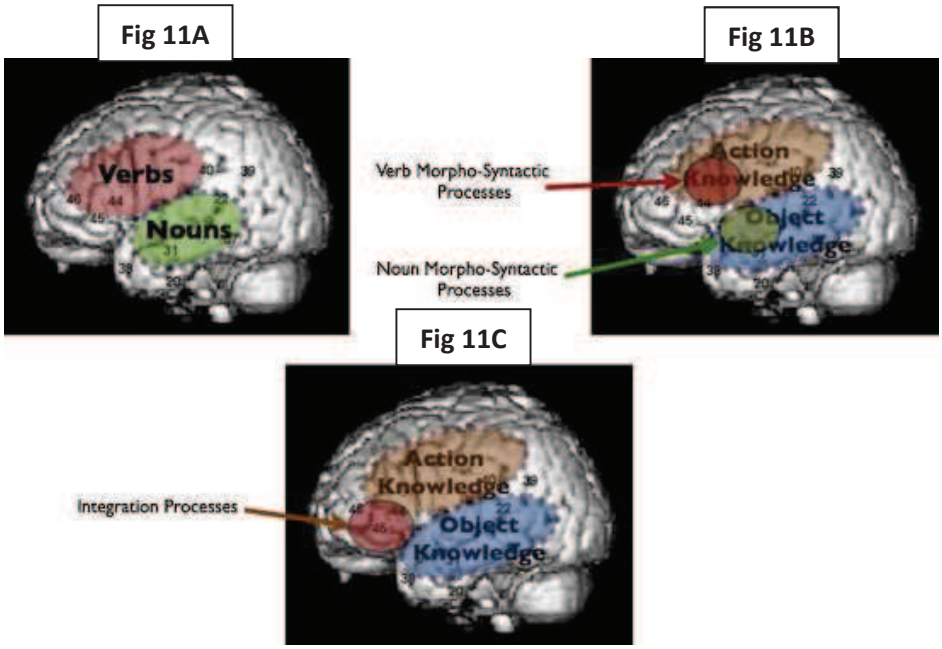


Fig 11) Vue schématique des principaux modèles neuronaux du traitement des noms et des verbes
Selon le premier modèle (11A), les noms et les verbes sont représentés dans des réseaux neuronaux partiellement distincts (noms =aires temporales gauches et verbe = aires frontales inférieures gauches). Selon le second (11B), ce ne sont pas les noms et les verbes qui sont traités dans des réseaux distincts, mais plutôt les processus morpho-syntaxiques qui sous-tendent ces deux catégories. Selon le dernier modèle (11C), il n'y aurait pas une séparabilité neurale noms-verbes, mais plutôt une distinction entre les mots se rapportant à des actions et ceux se rapportant à des objets ; les processus d'intégration seraient les mêmes pour ces deux catégories. **Adapté de Vigliocco et al., 2011**

Il existerait un système neural commun impliquant majoritairement le GFi (mais non limité à l'aire de Broca) et si ce système paraît parfois être engagé plus fortement dans le traitement morphosyntaxique des verbes que des noms, c'est parce que les verbes imposent des exigences de traitement plus importantes que celui des noms dans la plupart des cas (l'effet peut d'ailleurs être inversé expérimentalement en utilisant des noms qui demandent plus de ressources de traitement que des verbes, Siri et al., 2008).

Liljeström et al (2008) présentent des résultats assez similaires en IRMf à la fois pour la dénomination d'objets et d'actions (noms et verbes) même s'ils décrivent un réseau commun plus étendu avec des implications frontales mais aussi pariétales et temporo-occipitales bilatéralement. Leurs résultats suggèrent qu'indépendamment de la dénomination verbes-noms, les actions évoquent une activité accrue, par rapport aux objets, dans le cortex temporal postéro-moyen bilatéralement, la jonction temporo-pariétale et le cortex frontal gauches. La dénomination d'objets lorsqu'ils sont représentés dans le contexte plus global d'une action augmente l'activité du cortex pariétal postérieur gauche, ce qui laisse supposer que présenter l'objet dans le contexte de l'action met plus fortement l'accent sur les propriétés motrices associées à l'objet. Encore une fois, ces différences entre classes grammaticales semblent plutôt représenter la variabilité de contenu sémantique qui module la richesse de la simulation sensorimotrice impliquée dans la compréhension de ces stimuli. Xu et al., (2005) ont d'ailleurs démontré que la complexité de l'intégration sémantique augmente de façon proportionnelle en fonction de l'utilisation de mots, phrases ou récits : plus l'intégration sémantique est complexe, plus il y a une expansion de l'activité dans les zones linguistiques périsylviennes vers les aires frontales et temporales (pour les mots et les phrases). Traiter le sens d'un récit narratif recrute un plus grand nombre d'aires notamment le précuneus, le cortex préfrontal médian et des aires temporo-pariéto-occipitales.

- *Verbes et simulation sensorimotrice*

En tout cas, l'information contenue dans les verbes semble cruciale pour la simulation sensorimotrice. En contradiction avec les résultats qui suggèrent que les verbes impliqueraient plus spécifiquement les aires frontales (Vigliocco et al., 2011), d'autres travaux suggèrent que le gyrus postérieur temporal moyen de l'hémisphère gauche jouerait un rôle particulièrement important dans le traitement des verbes et dans la simulation d'actions complexes, cette région se situant à la limite de la jonction temporo-pariétale et juste en avant de l'aire MT/V5 (Tettamanti et al, 2005 ; Kable et al., 2002). Des activations de l'aire MT/V5 ont été observées lorsque l'on doit accéder à l'information sur une action contenue dans un verbe isolé (Kable et al. 2009). Kemmerer et al. (2008) ont montré que les verbes désignant une action activaient des structures prémotrices, mais que les verbes

impliquant surtout une notion de mouvement généraient des activations plus spécifiques des aires postérieures motrices, à proximité de MT/V5. Ces différents résultats suggèrent qu'en plus d'une simulation des propriétés motrices d'une action, les propriétés perceptivo-visuelles (telles que le mouvement) peuvent également être simulées à travers le traitement du langage (Kemmerer et al., 2008). Wallentin et al. (2005) ont trouvé des activations du gyrus postérieur temporal moyen gauche aussi bien pour des phrases dont le contenu est vraiment dynamique (verbes d'actions associés à personnage animé) que pour des phrases dont le mouvement est fictif (objet inanimé associé à un verbe d'action, par exemple : « le chemin arrive dans le jardin »). Pour ces auteurs, ces résultats indiqueraient que même dans un contexte où le mouvement n'est pas induit par une phrase décrivant une action, le sujet se crée une carte cognitive à partir de la phrase lue et que cette carte serait analogue à celle créée par le système perceptuel lorsque le sujet interagit réellement avec l'environnement. Il y aurait une connexion inhérente entre les représentations mentales de l'espace et du mouvement ; l'activité du gyrus temporal postéro-moyen (même dans les phrases où un objet inanimé est associé à un verbe d'action) pourrait être due au fait que le sujet explore activement l'espace mental élaboré à partir de la phrase lors du processus de compréhension. Kemmerer et Gonzalez-Castillo (2010) s'inspirent de la théorie linguistique « Two-Level Theory of verb meaning » (Levin, 1993) pour décrire les liens entre systèmes moteurs (en particulier le MNS) et compréhension des verbes. Selon cette approche, les verbes ont deux niveaux sémantiques distincts : le premier reflète le caractère unique de chaque verbe (la «racine ») alors que l'autre niveau correspond à une représentation plus stricte, partagée par tous les verbes d'une classe donnée (« modèle de structure d'événement » ou « noyau thématique »). Le sens d'un verbe repose sur l'association d'une racine particulière et d'un modèle : par exemple, les verbes de changement d'état (casser, sécher, faire fondre…) impliquent tous une transformation physique sur un objet (même modèle) mais ils varient en fonction de leur « racine » qui est spécifique au sens de chaque verbe et qui va permettre de préciser plus spécifiquement l'état. Pour Kemmerer et Gonzalez-Castillo (2010), la « racine » des verbes pourrait donc être associée aux propriétés somatotopiques des neurones miroirs des cortex moteur primaire et prémoteur car les représentations qui sont associées à la racine du verbe sont concrètes et souvent modalité-spécifiques, sans dépendre d'aspects grammaticaux , alors que le niveau du « modèle » reflète une structure logique abstraite commune à tous les verbes d'une classe donnée. En effet, ce dernier est composé principalement de prédicats schématiques et des variables pour les arguments, de ce fait il est dépendant des propriétés grammaticales de tous les verbes de cette classe. Le niveau « modèle » serait donc plutôt associé à des neurones miroirs aux propriétés plus schématiques, en particulier à l'aire BA 44 du GFi. Un intéressant travail en IRMf de James & Maouene (2009) confirme l'importance des régions sensorimotrices dans le traitement des verbes mais aussi dans leur apprentissage, puisque l'écoute passive de verbes d'actions recrute les aires

motrices du cortex frontal (SMA, cortex prémoteur et prémoteur dorsal), chez des enfants d'âge pré-scolaire (4-6 ans), alors qu'on ne retrouve pas ce type d'activations pour des adjectifs. Les liens fonctionnels entre les aires motrices et les aires linguistiques semblent donc précoces et importants pour le développement de la compréhension du langage mais aussi pour la compréhension des actions effectuées par autrui (Maouene, Hidaka & Smith, 2008 ; James & Maouene, 2009) ; certains travaux semblent également indiquer que les enfants apprennent et comprennent mieux les verbes lorsqu'ils réalisent eux-mêmes les actions plutôt que lorsqu'ils les observent (Childers & Tomasello, 2006).

- *Influence des facteurs grammaticaux sur les représentations sensorimotrices*

Si la différence de patterns d'activations Noms/Verbes est liée à des aspects sémantiques (Vigliocco et al., 2011) en revanche, au sein d'une catégorie grammaticale, certains facteurs purement grammaticaux pourraient influencer nos représentations. Pour Bergen et Wheeler (2010), le contexte plus large de la phrase qui entoure le verbe semble être un élément clé qui détermine le contenu perceptivo-moteur simulé mentalement : la simulation mentale ne serait pas initiée seulement grâce aux associations entre représentations lexicales et activités des systèmes perceptifs et moteurs ; la grammaire aurait un rôle de premier plan dans cette simulation car l'aspect grammatical modifierait la partie d'un événement qui est simulée en apportant plus ou moins de détails à la simulation et ce, de la même façon que l'on peut faire un focus mental sur différentes parties d'une scène simulée (Denis & Kosslyn, 1999; Mellet et al., 2002 ; Borghi, Glenberg, & Kaschak, 2004). Ainsi, des phrases utilisant les mêmes mots mais ne différant que sur des marqueurs grammaticaux peuvent modifier la simulation. Un aspect grammatical particulièrement important semble être celui associé au verbe : l'utilisation d'un temps progressif (verbes au présent, au gérondif présent ou passé) focaliserait la simulation sur l'action en cours alors qu'une forme verbale au passé (de type imparfait ou plus que parfait) engendrerait plutôt une simulation du résultat de l'action, un état fini (ce que confirment plusieurs résultats comportementaux et des données EEG ; Madden & Zwaan, 2003; Ferretti, Kutas, & McRae, 2007; Bergen et al., 2007 ; Madden & Therriault, 2009 ; Bergen et Wheeler, 2009 ; Madden & Ferretti, 2009 ; Magliano & Schleich 2000 ; Madden-Lombardi, Jouen et al., in prep). Les mots fourniraient le contenu de la simulation mentale mais la grammaire participerait aussi à cette simulation car elle permet d'assembler les différents contenus associés aux mots pour pouvoir créer une représentation cohérente (attribuer les rôles aux différents acteurs d'une scène par exemple). Mais la grammaire permettrait aussi de moduler les propriétés secondaires de la simulation mentale en focalisant la simulation sur certains aspects d'une scène (ex : action en cours ou action terminée, en fonction du temps du verbe employé dans la phrase).

- *Autres facteurs langagiers susceptibles d'influencer la simulation*

La simulation sensorimotrice à partir des phrases d'actions serait également modulée en fonction d'un certain nombre de facteurs intrinsèques à l'action, qui sont plus ou moins bien représentés à travers des éléments langagiers tels que la capacité de se mettre à la place de l'agent qui effectue l'action selon qu'elle est décrite à la 1ère ou à la 3ème personne (Gianelli et al., 2011), ou encore le but d'une action et la compatibilité d'une action avec ce but (Scorolli et Borghi, 2009) ; interviennent aussi des intentions de communication impliquées dans une action (Enrici et al., 2011), le contexte de l'action et l'intégration d'une situation à un contexte narratif plus important (Xu et al 2005), le codage des représentations spatiales dans un événement langagier (Chatterjee et al, 1999)… Tous ces facteurs (que nous ne détaillerons pas ici) moduleraient l'activité d'un système fronto-pariétal impliqué dans la sémantique de l'action, reflétant toute la complexité de l'intégration des multiples aspects langagiers (sémantique, thématique, grammaticaux), ce qui contribue à expliquer la variabilité parfois observée dans les résultats.

C) Travaux comparatifs stimuli verbaux/non verbaux

Comme nous avons pu le voir à travers différents travaux utilisant soit du matériel verbal, soit du matériel non verbal, un certain nombre de régions similaires semblent impliquées dans la compréhension des phrases et des images, même si leurs rôles précis restent incertains et dépendants de différentes approches théoriques. On peut cependant se demander s'il existe un réseau sémantique unique, commun aux phrases et aux images ou si ces deux types de stimuli ont recours à des systèmes et des processus différents.

- **Modèles théoriques**

  - *Des systèmes distincts pour traiter mots et images : le modèle du double codage de Paivio*

Certains modèles (Paivio, 1971, 1991; Shallice, 1988; Warrington & Shallice, 1984 ; Warrington & McCarthy, 1994) s'orientent vers l'idée qu'il existe des représentations conceptuelles distinctes pour les modalités d'entrées visuelles (objets) et verbales (mots). En particulier la théorie du double codage de Paivio défend l'idée que la cognition humaine s'est développée pour traiter simultanément des objets et des événements verbaux et non-verbaux, donnant ainsi naissance à deux systèmes symboliques indépendants (le système imagé et le système verbal). Le système verbal traiterait directement les inputs linguistiques, tout en ayant une composante symbolique par rapport aux inputs non-verbaux, qui eux seraient traités uniquement par le système non-verbal. Les deux systèmes sont supposés être distincts au niveau fonctionnel et structurel, quoique interconnectés par le biais de connexions référentielles. Il faut préciser que cette conception de représentations faisant appel à des réseaux neuronaux distincts repose principalement sur des études de cas neuropsychologiques, plusieurs cas de dissociations neuropsychologiques observées dans des tâches de jugement sémantique impliquant des images et mots ayant conforté les auteurs de l'existence de sous-systèmes distincts pour le langage et les images (Warrington, 1975 ; Ferreira, Giusiano, Ceccaldi, & Poncet, 1997; Warrington & McCarthy, 1994; Warrington & Shallice, 1984 ; Saffran, Coslett, Martin & Boronat, 2003). Le modèle du double codage de Pavio a notamment été développé dans le but d'expliquer la supériorité de rappel verbal observée pour les dessins d'objets par rapport à des mots (Denis, 1975 ; Florès, 1987). Les dessins d'objets seraient traités à la fois dans leurs aspects figuratifs et dans leurs aspects verbaux par les deux systèmes (verbal et imagé) et ainsi, lors de la récupération en mémoire, les sujets disposeraient de deux traces mnésiques au lieu d'une, ce qui expliquerait cet effet facilitateur.

- *Système sémantique commun aux deux modalités : les modèles de Caramazza et collègues*

Cependant, d'autres modèles ont rejeté cette approche, affirmant que toutes les voies de traitement convergeraient vers un seul ensemble de représentations conceptuelles, système commun aux deux modalités (Caramazza, Hillis, Rapp, et Romani 1990 ; Lambon Ralph, Patterson, & Hodges, 1997; Lambon Ralph, Graham, Patterson, & Hodges, 1999; Rapp, Hillis, & Caramazza, 1993). Ces auteurs réinterprètent les dissociations neuropsychologiques non comme résultant d'un trouble touchant l'une ou l'autre des composantes d'un système sémantique double, mais plutôt en termes de perturbation de l'accessibilité à un système sémantique commun (theory of privileged accessibility, Caramazza, 1990). Ils insistent aussi sur le fait que les tâches verbales et non verbales ne sont pas toujours comparables en termes de complexité, d'autres facteurs pertinents pouvant conduire à des dissociations plus apparentes que réelles : la dissociation observée résulterait plutôt d'atteintes à des niveaux pré-sémantiques nécessaires pour accéder au système conceptuel plutôt qu'à une atteinte de ce système lui-même (Bright et al., 2004). L'hypothèse de Caramazza (Organised Unitary Conceptual Hypothesis, OUCH, 1990) suppose que la perception des caractéristiques saillantes d'un objet présenté visuellement permettrait un accès privilégié à sa représentation sémantique, même si celle-ci est la même que pour le mot désignant cet objet. Ainsi, dans cette théorie, alors que la représentation sémantique d'un objet présenté visuellement et celle correspondant à sa description verbale sont identiques, la procédure par laquelle cette représentation est accessible peut varier : un mot présenté visuellement ou auditivement activera un lexique associé, qui, à son tour, activera les propriétés sémantiques définissant son sens, alors qu'un objet présenté visuellement accédera directement à ces mêmes propriétés sémantiques. Néanmoins, le modèle sémantique unitaire n'est pas en mesure d'expliquer pleinement tous les effets de dissociations neuropsychologiques et certains auteurs présentent une vision intermédiaire permettant de réconcilier les deux conceptions d'un système sémantique unitaire ou multiple (Bright et al., 2004 ; Plaut, 2002) : il existerait un système sémantique commun, activé indépendamment de la modalité de présentation, mais certaines sous-parties de ce système conserveraient une spécialisation relative à la modalité d'entrée, ainsi les déficits pourraient être plus ou moins spécifiques.

- **Travaux comportementaux**

Au niveau comportemental, un certain nombre de travaux semblent suggérer qu'un système sémantique commun pourrait intervenir, au moins en partie (Madden & Zwaan, 2006 ; Kaschak, Madden et al, 2005). On trouve notamment un priming intermodalité entre mots et images d'objets (Bajo, 1988 ; Kircher et al., 2009; Sass et al., 2009) mais aussi entre mots et scènes visuelles plus

complexes (Reinitz et al., 1989). Des données en neuropsychologie semblent aussi confirmer cette vision : on a ainsi observé des troubles similaires des représentations conceptuelles pour les images et pour les mots chez des patients atteints de démence sémantique et de la maladie d'Alzheimer, le degré d'information conceptuelle obtenue à partir des inputs verbaux étant un bon prédicteur de la qualité des connaissances sémantiques générées à partir d'une image (que ce soit dans des tâches de compréhension ou de dénomination) (Lambon et al. (1999) ; Hodges et Patterson, 1996 ; Hodges et al., 1996; Lambon Ralph et al., 1997).

- **Travaux en neuroimagerie**

Bien sûr, les techniques de neuroimagerie restent un moyen idéal pour explorer la neuroanatomie fonctionnelle d'un système conceptuel commun car contrairement à une approche basée sur les lésions, elles permettent d'explorer directement l'activité neuronale dans le cerveau sain en fonction de la modalité de présentation des stimuli. Des études en neuroimagerie ont confirmé l'existence des effets de priming, observés au niveau comportemental : non seulement il existe bien un priming inter-modalité (images/mots) où le gyrus frontal inférieur droit jouerait un rôle essentiel (Kircher et al., 2009) ; mais lorsque l'on compare les activations liées aux effets de priming pour des mots et des images séparément, on obtient la mise en évidence d'un réseau bilatéral impliquant les régions fronto-temporo-pariétales et similaire pour les deux modalités. Cependant, seules quelques études en neuroimagerie ont cherché à déterminer les bases neurophysiologiques communes à la compréhension de stimuli verbaux et non-verbaux en comparant directement les deux types de stimuli dans des tâches sémantiques.

  - *Différences de traitement sémantique entre images et mots*

Plusieurs études utilisant la TEP (Vandenberghe et al., 1996 ; Moore & Price ; 1999, Devlin et al., 2002 ; Bright, Moss & Tyler ; 2004) se sont intéressées à la question de la différence de traitement sémantique entre images et mots en contrastant l'activité durant des tâches sémantiques réalisées avec des mots ou images. Les deux types de stimuli génèrent des activations sémantiques spécifiques : situées au niveau de la partie postérieure du sillon temporal inféro-postérieur gauche pour les images alors que les mots impliquent le sillon temporal supérieur, le gyrus temporal antérieur moyen et le sillon frontal inférieur gauches (Vandenberghe et al, 1996); chez Bright et al. 2004, on retrouve le même type d'activations, en temporal pour les mots et en occipitotemporal pour les images, sauf qu'elles sont bilatérales. Dans les différentes études, les auteurs ont cependant trouvé un large réseau commun majoritairement latéralisé à gauche, impliquant le gyrus frontal inférieur, le cortex temporal inféro-moyen avec une implication particulière du gyrus fusiforme et

parahippocampique. L'étude de Vandenberghe et al. (1996) a mis en évidence un réseau plus étendu, impliquant également la jonction temporo-pariétale et le gyrus supérieur occipital gauches.

Fig 12A

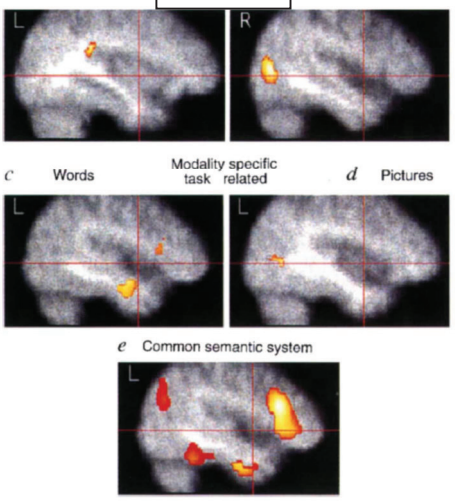


Fig 12B

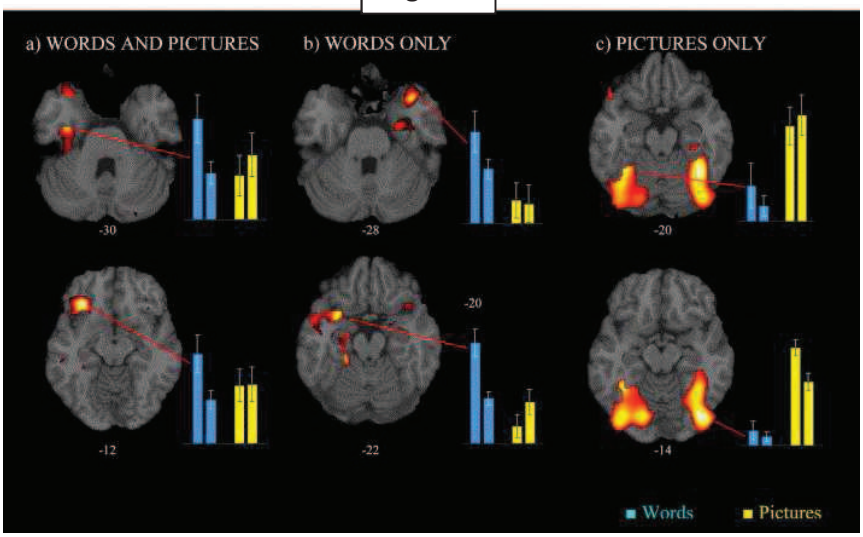


Fig 12) Différences et similitudes de traitement sémantique pour mots et images, illustrées à travers deux études
Dans la Fig 12A, les 4 figures du haut représentent les effets spécifiques pour les mots (gauche) et pour les images (droite). La figure du bas représente le réseau sémantique commun. Dans la Fig 12A, les activations communes pour mots et images sont à gauche, pour les mots, au milieu et pour les images, à droite.
Adapté de Vandenberghe et al., 1996 et Bright et al., 2004

Ces deux études concluent que ces données vont dans le sens d'un système sémantique unitaire, même si pour Bright et al. (2004), ce réseau serait hiérarchisé et sous la dépendance des régions antérieures du cortex temporal inférieur. Les lobes temporaux seraient impliqués dans le traitement à la fois des mots et des images, mais ce qui déterminerait plutôt leur implication serait le niveau de spécificité auquel un objet est traité. Ainsi les activations postéro-temporales, mises en évidence lors de la présentation des images, refléteraient plutôt des étapes pré-sémantiques du traitement visuel selon ces auteurs.

Dans une tâche de rappel de paires d'images et de mots, Buckner et al (1996) ont aussi mis en évidence un réseau commun, même si celui-ci recoupe à la fois des aspects mnésiques, épisodiques et sémantiques. Dans ce réseau, on observe des activations robustes et étendues entre les deux modalités : elles sont situées principalement au niveau du cortex frontal bilatéral (operculum) et du cortex cingulaire antérieur, du cortex pariétal postérieur également bilatéralement, du cortex préfrontal antérieur droit, des cortex moteur et prémoteur ainsi que du cervelet. Cette voie commune serait utilisée pour accéder et maintenir les représentations conceptuelles. Chee et al., (1999) ont cherché à comparer les activations générées en IRMf par des tâches sémantiques impliquant des mots écrits en anglais, des mots écrits en caractères chinois et des images. Ils observent que les aires activées conjointement par les caractères chinois et les images se situent

dans l'hémisphère gauche au niveau des régions préfrontales (BA9, 44 et 45), de la partie postérieure du cortex temporal, du gyrus fusiforme et des régions pariétales. Ils observent un pattern très similaire quand ils comparent les mots anglais et les images ainsi que les caractères chinois et les mots anglais. Il existe donc un réseau sémantique commun pour le traitement des caractères chinois, des mots anglais et des images, même s'il y a quelques différences spécifiques aux différents stimuli, principalement dans l'intensité du signal. La nature idéographique des caractères chinois est supposée rendre le concept relativement transparent pour le lecteur (Smith, 1985; Wang, 1973) et par conséquent, peut-être plus proche du traitement des stimuli images. Pourtant les résultats de Chee et al (1999) suggèrent que le traitement sémantique des caractères chinois est plus proche (dans l'intensité du signal) de celui des mots en anglais que des images.

Van Doren et al. (2010) ont comparé les activations IRMf pour des images et des mots présentés à la limite du seuil perceptif via une tâche requérant un niveau de traitement sémantique minimum. Les résultats de cette étude ont mis en évidence un réseau sémantique commun pour les images et les mots où les aires impliquées étaient activées de façon plus importante quand les stimuli était perçus au-delà du seuil subliminal. Ce réseau, latéralisé à gauche, recrute des régions occipito-temporo-pariétales et frontales inférieures. Au sein de ce réseau, le gyrus frontal inférieur et le sillon occipito-temporal correspondraient à des régions sémantiques amodales, en revanche d'autres régions dépendraient de la modalité (le sillon temporal supérieur et l'operculum frontal gauches pour les mots ; le gyrus occipital moyen bilatéralement et le sillon collatéral gauche). Reinholz & Pollmann (2005) ont étudié les activations en IRMf lorsque les sujets regardaient passivement des images ou des mots désignant des visages ou des bâtiments (église, château, usine…). Pour la présentation des visages et des constructions, les auteurs évoquent des activations catégories-sélectives au niveau du cortex occipito-temporal ventral (gyri fusiforme et parahippocampique), activations cohérentes avec la mention dans la littérature de l'existence d'aires spécialisées pour le traitement des visages (FFA) et des constructions/lieux (PPA). La présentation des mots correspondants ne fait pas intervenir les mêmes régions catégorie-spécifiques, elle génère des activations principalement situées au niveau du gyrus occipital inférieur gauche, à la limite du sillon occipitotemporal latéral et donc proches des régions du traitement de la forme visuelle des mots (VWFA). Comme le précisent eux-mêmes les auteurs, cette étude porte avant tout sur les différences entre catégories et n'est pas nécessairement conçue pour mettre en évidence les zones activées conjointement, les stimuli étant présentés brièvement et la tâche ne requérant pas nécessairement un traitement conceptuel approfondi. Initialement les mots écrits seraient traités comme une catégorie distincte d'objets (au même titre que les visages et les constructions). Ainsi les différentes activations décrites dans cet article, pour les images de visages et de constructions et pour les mots écrits, refléteraient une étape de traitement

comparable pour les différents stimuli impliquant différentes aires spécifiques du lobe occipital ; cette étape de traitement ne serait pas nécessairement sémantique.

Adamo et Ferber (2009) se sont, quant à eux, intéressés à la comparaison mots/images d'objets manipulables (outils) dans le but de mieux comprendre les interactions entre la voie dorsale de l'action et la voie ventrale de l'identification des objets, les objets manipulables étant représentés dans les deux voies à la fois pour identifier leur fonction et l'action qu'ils peuvent permettre. Pour Adamo et Ferber (2009), la voie ventrale et la voie dorsale interagissent via des mécanismes de sélection attentionnelle et de consolidation, mis en évidence par le fait qu'un objet présenté visuellement dont la forme est cohérente avec la fonction d'un autre objet vu précédemment (ex : clou/marteau) bénéficie d'une plus grande attention (marquée par un blink attentionnel réduit et une plus grande P300). Mais pour les auteurs, c'est là que se trouve la limite entre mots et images puisqu'ils ne reproduisent pas ces résultats lorsque les stimuli sont présentés sous la forme de mots. Pour eux, l'aspect purement sémantique de la relation d'action entre deux objets, représentée à travers des mots, n'est pas suffisante pour augmenter les processus attentionnels comme c'est le cas avec des objets présentés visuellement, car les propriétés visuelles de ces derniers permettent de faire des affordances sur l'action. Ils affirment que les objets manipulables doivent être présentés sous la forme d'images pour engager les zones de la voie dorsale qui traitent de la fonction d'un objet : le processus de reconnaissance des objets effectué par la voie ventrale provoquerait un traitement de l'action dans la voie dorsale, l'activation de cette voie entraînant en retour une plus grande sélection attentionnelle et une meilleure consolidation des objets qui correspondent à l'action affordée. Cette interaction entre voie ventrale et voie dorsale disparaitrait lorsque l'association entre les éléments est d'ordre sémantique et non physique, et donc non propice à une affordance de l'action. Par conséquent, les outils nécessiteraient un traitement spécial dans le cerveau en raison de leur caractère fonctionnel, ce qui est nécessaire pour que cette fonction soit immédiatement perceptible et puisse influencer la sélection attentionnelle et la consolidation.

Dans un intéressant travail, Shinkareva et al. (2011) ont cherché à étudier les bases cérébrales communes des représentations associées à des mots et à des images, en combinant les résultats IRMf obtenus chez des sujets en train de regarder ces différents stimuli présentés rapidement et une méthode de classification de patterns grâce à un algorithme (classifieur) entraîné à reconnaître les patterns d'activités cérébrales associés soit à des images soit à des mots. Les stimuli utilisés (mots et images) étaient de deux types : des outils et des habitations. A partir de cet apprentissage, le modèle était capable d'identifier la catégorie d'image (outil ou habitation) que le sujet avait vue à partir des activations cérébrales associées aux mots et inversement, il pouvait identifier la catégorie du mot lu à partir des patterns cérébraux associés aux images (que les données viennent du même sujet ou

d'un autre sujet). Pour les auteurs, cela correspond à une démonstration empirique que mots et images partagent des représentations communes.

- *Régions du système sémantique commun aux mots et aux images*

Dans l'ensemble, plusieurs études (Shinkareva et al., 2011; Bright et al., 2004; Chee et al., 2000; Gallagher et al., 2000; Gates & Yoon, 2005; Vandenberghe et al., 1996) sont d'accord pour admettre qu'ils existent des représentations sémantiques communes pour les mots et les images, même s'il y a un certain degré de spécialisation lié au format de présentation. Les régions principalement impliquées dans ce traitement amodal seraient le cortex occipitotemporal ventral (Buckner et al., 2000), la partie postérieure du gyrus temporal moyen gauche (Chertkow et al., 1997; Vandenbulcke et al., 2007; Kircher et al., 2009), le pôle temporal antérieur (Hodges et al., 1992; Vandenberghe et al., 1996; Rogers & McClelland, 2004), le sillon frontal inférieur (Wagner et al., 1997), le gyrus frontal inférieur dans sa partie antérieure (Goldberg et al., 2007) et le gyrus frontal moyen (Demb et al., 1995; Vandenberghe et al., 1996). Au niveau du cortex temporal postérieur, les gyri fusiforme et temporaux moyen et inférieur seraient particulièrement importants pour le traitement conceptuel : ils seraient activés aussi bien par des images et des mots mais aussi durant la compréhension de phrases présentées auditivement (Davis & Johnsrude 2003, Giraud et al. 2004, Rodd et al. 2005) ; plus le degré de compréhension augmente et plus l'activité dans les régions temporales postérieures gauches est importante (Martin, 2007).

- *Stimuli complexes*

Cependant, la plupart de ces travaux portent sur des stimuli relativement simples (mots et objets isolés), très peu de travaux se sont réellement intéressés à la sémantique d'évènements plus complexes (scènes, événements, phrases...) et plus « écologiques » c'est-à-dire plus proches des situations auxquelles nous pouvons être confrontés dans le monde réel. De plus, la grande variabilité des tâches et des stimuli employés risquent de masquer certains effets, les activations rapportées dans la littérature n'étant pas toujours consensuelles.

Objets et actions

Kable et al. (2002) se sont intéressés à la comparaison directe entre objets et actions, via une étude IRMf de l'activité du cerveau pendant que les sujets effectuaient une tâche d'appariement conceptuel utilisant soit des images, soit des mots. Pour les auteurs, il est nécessaire de distinguer les entités (telles que les humains, les objets, les animaux) auxquelles on se réfère généralement à

travers des mots dans le langage, tandis que les événements décrivent ce qui arrive aux entités, ils peuvent se traduire par des actions (c'est à dire des modifications spatiales et temporelles appliquées aux entités) et sont généralement désignés dans le langage par le biais de verbes et par des relations syntaxico-thématiques. Les événements d'actions, impliquant un mouvement à travers l'espace, devraient générer plus d'activations dans les zones cérébrales impliquées dans le traitement du mouvement ou le traitement spatial. Les résultats de Kable et al. (2002) confirment cette hypothèse puisqu'en comparaison avec les objets, l'accès à la connaissance des actions à travers les images et les mots met en jeu un réseau bilatéral incluant les cortex pariétal et temporal (principalement le gyrus temporal moyen postérieur), même si les activations sont légèrement différentes pour les deux types de stimuli. En ce qui concerne les images, on note une activité accrue de façon bilatérale dans les régions MT / MST et à proximité du cortex temporal latéral, alors que pour les mots, les zones activées sont situées antérieurement et dorsalement par rapport à la zone MT, dans la partie postérieure des gyri temporal moyen et supérieur gauches. Pour Kable et al. (2002), le cortex occipito-temporal latéral contiendrait une mosaïque de régions cérébrales impliquées dans des processus divers allant de la perception des mouvements réels des objets dans le monde jusqu'à la conceptualisation des mouvements décrits par les verbes d'actions, ce qui ferait de cette voie occipito-temporal latéral un candidat parfait pour traiter les caractéristiques perceptives et conceptuelles des événements d'actions alors que la partie ventrale du cortex occipito-temporal traiterait les caractéristiques perceptives et conceptuelles des entités.

Phrases et Scènes visuelles

Les travaux qui qui se sont efforcés de comparer directement l'observation d'actions et la lecture de phrases d'actions (Aziz-Zadeh et al., 2006; Aziz-Zadeh et al., 2008 ; Aziz-Zadeh & Damasio, 2008) constatent des activations proches, mais ne se recoupant pas totalement, au niveau des aires prémotrices et préfrontales, les activations pour la lecture des phrases étant plus antérieures et médiales que celles qui concernent l'observation d'actions. Cependant, dans l'étude d'Aziz-Zadeh et al. (2006) on note une forte congruence de l'organisation somatotopique du cortex prémoteur, en fonction de l'effecteur employé, à la fois pour les images et le langage. De plus, la plupart de ces études ont comparé des contrastes (tâches vs baseline) mais sans chercher à cerner spécifiquement les régions communes, par exemple avec une analyse de conjonction.

Cependant, ce n'est pas le cas du travail de Tremblay et al. (2011) qui ont utilisé des analyses de conjonction pour explorer les activations communes à l'observation de vidéos d'actions/écoute de phrases d'actions (e.g « Je règle le réveil »), mais aussi observation d'images d'objets/écoute de phrases liées à des objets (e.g « Le réveil est fait de plastique »). Un résultat assez surprenant de

cette étude est que l'activation du cortex prémoteur ventral associée au traitement des phrases d'action ne recoupe pas celle associée à l'observation des actions ; alors que pour les objets, le recoupement est plus important, le prémoteur étant d'ailleurs plus fortement impliqué par les stimuli liés aux objets. Tremblay et al. (2011) ne remettent pas en cause l'implication du cortex prémoteur ventral dans la compréhension du langage (comme le suggère le reste de leurs résultats qui montrent des activations des cortex moteur et prémoteur pour diverses tâches langagières) mais, néanmoins, ils s'opposent à la notion d'une simulation forte commune à l'observation des actions et à la compréhension des phrases d'actions.

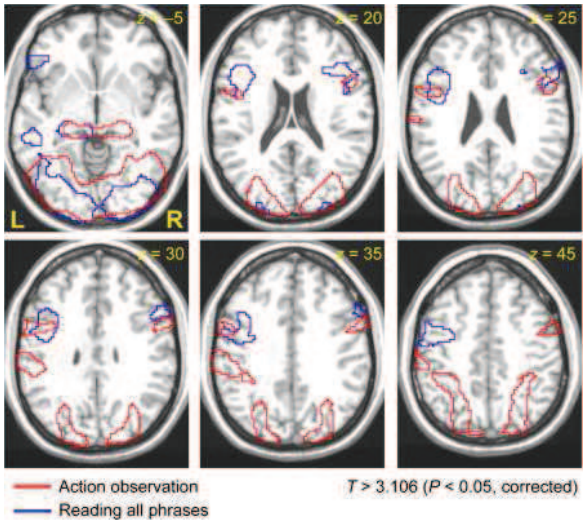


Fig 13) Superposition des activations associées à l'observation d'actions et de celles associées à la lecture de phrases d'action
Les activations entre les deux conditions se recoupent partiellement, principalement au niveau des aires prémotrices et préfrontales.
Adapté de Aziz-Zadeh et al., 2006

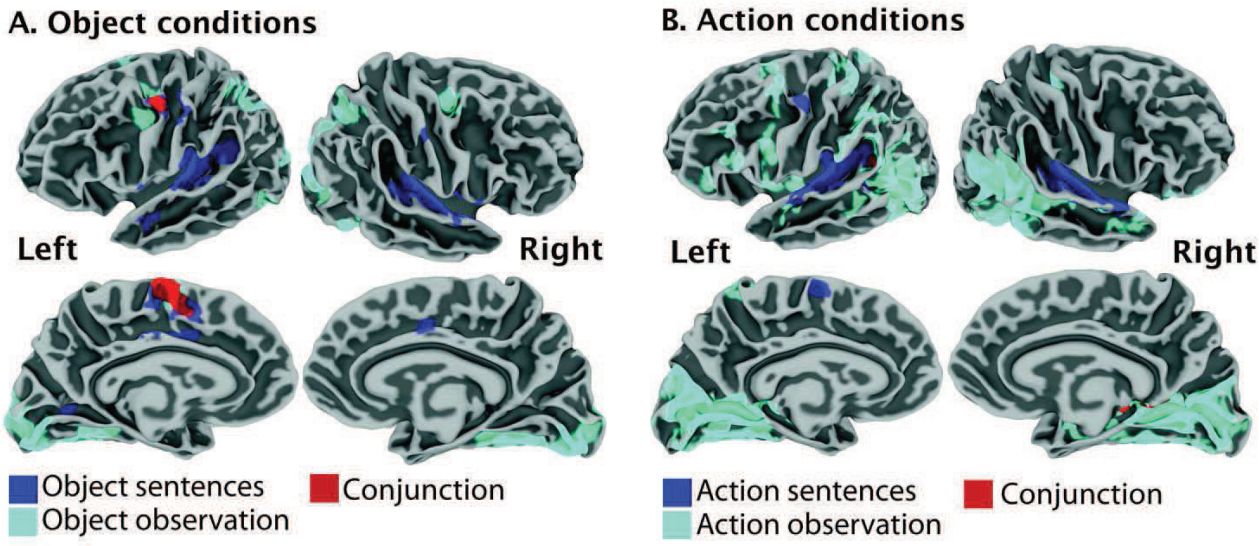


Fig 14) Activations communes et distinctes à l'écoute de phrases et à l'observation d'images/vidéos
Dans la figure de gauche, les résultats sont ceux correspondant à : des phrases (bleu), des images (turquoise) représentant des objets ou la conjonction des deux conditions (rouge). Dans la figure de droite, les résultats sont ceux correspondant à : des phrases (bleu), des vidéos (turquoise) représentant des actions ou la conjonction des deux conditions (rouge).
Adapté de Tremblay et al., 2011

Mais ces résultats contrastent avec ceux issus de la méta-analyse de Grèzes & Decety (2001) qui ont montré un recoupement important des régions cérébrales impliquées dans différents processus cognitifs de haut niveau liés à l'action (l'exécution, la simulation mentale, l'observation et la compréhension, la verbalisation d'actions), ces processus engageant la SMA, le cortex prémoteur dorsal, le gyrus supramarginal, le lobe pariétal supérieur. Ces résultats diffèrent aussi de ceux de Hamzei et al. (2003) qui ont utilisé une analyse de conjonction et qui trouvent des activations communes aux systèmes de la reconnaissance/exécution de l'action et de production du langage, avec des activations au niveau du gyrus frontal inférieur (aire de Broca), du gyrus précentral et du cortex pariétal. Si cette étude met effectivement en évidence des activations communes intéressantes, d'autant plus que les auteurs comparent directement des tâches de langage, de compréhension mais aussi d'exécution d'actions, les résultats de cette étude sont à nuancer compte tenu de certains choix méthodologiques. Dans les conditions où les sujets doivent observer une action, ce sont toujours les mêmes actions qui sont employées avec 6 acteurs différents (condition ACTIVE=boire dans une tasse ; condition PASSIVE=se détendre dans un fauteuil) ; de même dans la tâche de mouvement, c'est toujours le même geste qu'ils produisent (condition MOVE, action de saisir). Dans la tâche impliquant le langage, les sujets doivent générer silencieusement un verbe face à des noms présentés visuellement (condition VERB, 12 noms utilisés). Il est difficile de généraliser les résultats de cette étude alors qu'ils portent sur un nombre très restreints de stimuli et que les tâches impliquant langage et actions ne sont pas vraiment comparables. De plus, les analyses reposent sur le regroupement de résultats provenant de deux études IRMf légèrement différentes (avec seulement 6 sujets dans chaque étude).

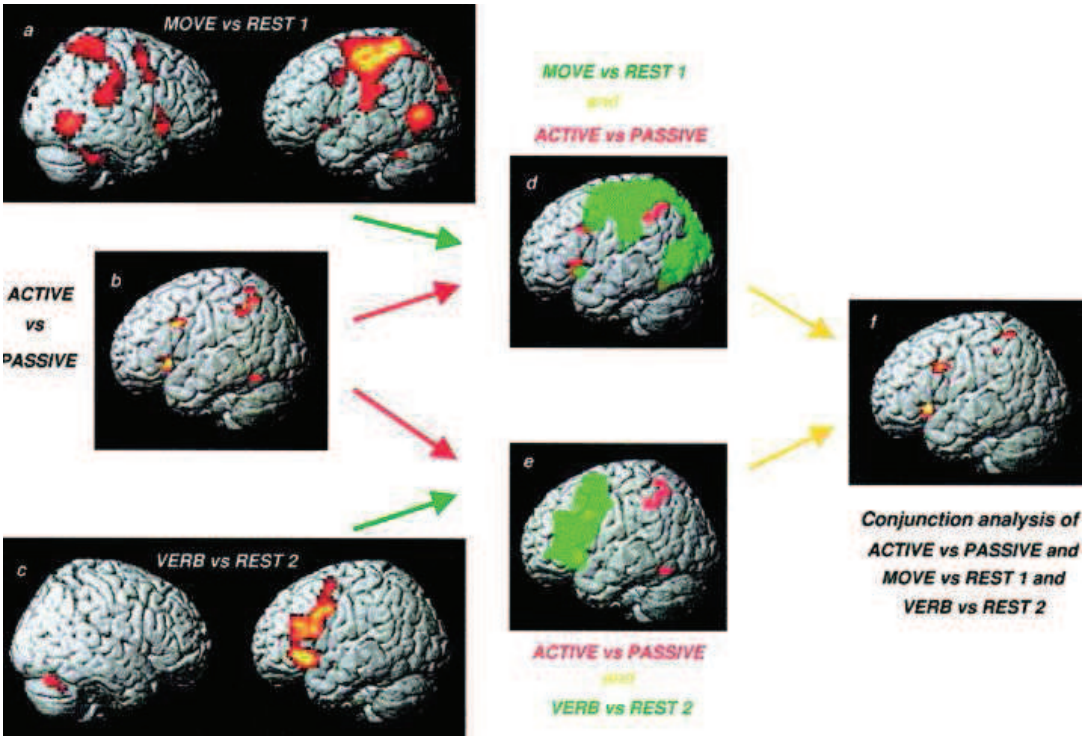


Fig 15) Activations communes aux systèmes de reconnaissance/exécution de l'action et de production du langage. Adapté de Hamzei et al., 2003

Un travail intéressant (Péran et al., 2010) compare plusieurs conditions expérimentales afin de mieux comprendre les bases communes aux représentations sémantiques des verbes d'actions et aux représentations mentales des actions elles-mêmes. Pour cela, ces auteurs ont essayé de voir si un même item (l'image d'un objet) génèrait les mêmes activations dans des tâches distinctes impliquant le langage (génération de mots d'action à partir de l'objet) ou une simulation de l'action (une simulation mentale ou tâche d'imitation). Péran et al (2010) montrent que ces différentes conditions activent un réseau fronto-pariétal commun avec de fortes implications motrices et prémotrices principalement latéralisées à gauche (GFi, LPI, SMA, Gyrus précentral mais aussi le gyrus fusiforme, le lobe temporal inféro-moyen et le lobule pariétal supérieur).

Comme nous l'avons déjà indiqué précédemment, la variabilité des résultats observés entre études ou entre les conditions peut s'expliquer par de nombreux facteurs : stimuli simples, variabilité des stimuli et des tâches utilisés, types d'analyses… Mais les différences observées pourraient aussi être liées au fait qu'il n'y a pas tout à fait le même niveau d'information entre des stimuli verbaux et non-verbaux. Certains auteurs ont suggéré notamment que l'implication du MNS pourrait être différente dans la cadre du langage, par rapport à des stimuli visuels, en raison de la nature abstraite de certains stimuli langagiers : le même système pourrait être impliqué, mais pas tout à fait avec la même intensité (Fischer & Zwaan, 2008). Buccino et al. (2005) se sont intéressés à cette question en comparant l'activité cérébrale associée à la présentation d'images et de phrases. Ils suggèrent que, dans le cas des images, il y a moins d'ambiguïté à propos des détails de l'action qu'il faut simuler, ce qui produirait un effet de facilitation lié à la vision. Dans le cas des descriptions verbales, il y a souvent un manque de précision qui pourrait générer un effet perturbateur voire une inhibition. L'ambiguïté du langage conduirait à une grande diversité de représentations possibles d'actions, elles rentreraient alors en compétition. Les différences de patterns d'activations cérébrales pourraient également refléter des étapes de traitement pré-sémantiques avec un accès spécifique, différent selon la modalité (Bright et al. 2004) ; en revanche les étapes ultérieures du traitement sémantique pourraient faire appel à un système de représentations communes (Lambon Ralph et al., 1999 ; Bright et al., 2004 ; Enrici et al., 2011), ce que pourraient mettre en évidence des tâches supposant un traitement sémantique plus profond. Ainsi, le travail d'Enrici et al. (2011) va au-delà de l'étude de la simulation sensorimotrice impliquée dans la compréhension de l'action, puisqu'ils étudient le décodage de l'intentionnalité de l'action soit à travers des images où l'intentionnalité de l'action est soulignée de façon évidente par une phrase (« S'il te plaît passe-moi la bouteille), soit à travers des images où elle est implicite à travers une gestuelle utilisée par les personnages de la scène (pointage en direction de l'objet). Les résultats de cette étude montrent qu'un large réseau commun est engagé pour traiter les intentions de communication, indépendamment de la modalité, et que ce

réseau recrute le precunéus, le STS postérieur et le TPJ bilatéralement, ainsi que le cortex préfrontal médian. Pour Enrici et al (2011), il y aurait une équivalence de la reconstruction mentale des intentions, même si les différents indices de communication ont été acquis via des indices propres à chaque modalité.

Ainsi, on pourrait retrouver ce réseau fronto-pariétal dans de nombreux processus cognitifs où il est nécessaire d'effectuer une analyse sémantique profonde pour saisir pleinement la situation (quelle est la situation, qui sont les personnages impliqués, quels sont leurs relations, leurs intentions, leurs croyances…), un système à forte composante sociale auquel nous nous référerions quotidiennement face aux situations de la vie courante.

I.3.*2*) « Où est le système sémantique ? »

Ce titre emprunté à un article de Jeffrey Binder (Binder et al. 2009) reflète bien toute la complexité inhérente à l'identification et à la localisation d'un système sémantique, d'autant plus si on envisage que ce système est commun aux stimuli verbaux/non verbaux. Les résultats de cette méta-analyse de 120 études en imagerie fonctionnelle (TEP et IRMf) semblent montrer un certain nombre de régions consensuelles, latéralisées à gauche. Il faut noter cependant que cette méta-analyse ne prend en considération que des études où les stimuli étaient verbaux, le propos des auteurs étant de savoir comment les connaissances conceptuelles étaient organisées et accessibles et non de savoir comment les objets sont reconnus. Ainsi on retrouve des implications du lobe pariétal postéro-inférieur, du gyrus temporal moyen, des gyri fusiforme et parahippocampique, des cortex préfrontaux dorsolatéral et ventromédial, du gyrus frontal inférieur et le gyrus cingulaire postérieur ; ce sont des régions que nous avons déjà mentionnées longuement comme impliquées dans la simulation sensori-motrice à la base de la compréhension des actions, mais qui pourraient aussi être recrutées dans de nombreuses autres fonctions cognitives.

Comment s'y retrouver face à la variabilité des fonctions cognitives associées à ces différentes régions ? La compréhension sémantique repose-t-elle uniquement sur cette simulation sensorimotrice ? Cela signifie-t-il que les régions à la base de cette simulation forment le réseau sémantique proprement dit ? Quels peuvent-être les rôles plus spécifiques des différentes régions impliquées et comment interagissent-elles conjointement ?

- **Les simulations sensori-motrices sont-elles nécessaires à la compréhension sémantique ?**

De nombreux auteurs se sont interrogés sur le rôle des activations sensorimotrices : sont-elles nécessaires à la compréhension ou correspondent-elles à un épiphénomène (Jirak et al, 2010 ; Kiefer et al., 2008 ; Kiefer & Pullvermuler, 2012 ; Desai et al., 2010 ; Meyer & Damasio, 2009, Van Dam et al., 2012 ; Pulvermueller & Fadiga, 2010 ; Willems & Casasanto, 2011 ; Mahon & Caramazza, 2008) ? Cette question est cruciale car elle permet de répondre partiellement aux interrogations sur la localisation du système sémantique et sur la nature de la compréhension. Les études concernant l'imagerie mentale montrent bien que les simulations sensori-motrices ne correspondent pas à des processus bottom-up automatiques de traitement de tout stimulus visuel, puisque l'on obtient des représentations corticales similaires à celles obtenues dans des tâches perceptives même en l'absence de stimuli présentés visuellement. En particulier, Clichy et al., (2012) ont comparé directement les activations cérébrales liées à la perception d'images et celles impliquées dans une tâche d'imagerie mentale ; ils ont montré que ces deux processus partageaient des représentations corticales similaires dans les aires de haut niveau du cortex visuel ventral.

Cependant, dans le cadre de la compréhension, on ne sait pas si ces simulations sensorimotrices soutiennent la compréhension ou si, au contraire, ce sont des activations qui lui sont concomittantes ou postérieures, reflétant des processus d'imagerie visuelle ou motrice. Des images mentales pourraient provenir de l'activité des neurones des cortex associatifs d'ordre supérieur et les activations des cortex sensori-moteurs ne feraient que refléter des processus non fonctionnels, ne permettant pas d'accéder à la façon dont le sens est représenté (Mahon & Caramazza, 2008). Cette question est particulièrement pertinente dans le cas du langage où la simulation pourrait simplement correspondre à un enrichissement du traitement lexico-sémantique, postérieur à la compréhension du mot, ce qui signifierait que le système moteur ne supporterait pas les représentations mentales lexico-sémantiques (Jirak et al., 2010). Cette hypothèse est soutenue par les travaux qui ne sont pas parvenus à mettre en évidence des activations du système moteur dans un contexte verbal (Raposo et al., 2009 ; Rueschemeyer et al., 2007). Il faut préciser que ces travaux impliquaient le plus souvent un verbe intégré dans un contexte métaphorique ou des conditions verbales un peu ambiguës (par exemple : Raposo et al., 2009 expressions métaphoriques « to kick the bucket »; ou Rueschemeyer et al., 2007, qui ont utilisé des verbes complexes avec une racine motrice : ex : « comprehend » qui contient « prehend »). Les travaux de Sato et al (2008) semblent indiquer que le recrutement des aires sensorimotrices n'interviendrait que lorsqu'une représentation sémantique explicite est requise par la tâche. Mais le fait qu'on ne retrouve pas d'activation automatique des systèmes sensori-moteurs ne nécessite pas forcément que ces systèmes ne soient pas impliqués, pour Van Dam et al., (2012) cela pourrait refléter le fait que les représentations « incarnées » sont flexibles , la flexibilité

étant caractérisée par cette présence ou absence d'activation dans nos systèmes perceptifs et moteurs. Pour ces auteurs, il est possible que le degré de contribution d'une région sensorimotrice dépende du contexte de récupération des caractéristiques conceptuelles. C'est ce que semble confirmer leur étude où des mots, dont les propriétés visuelles et motrices sont essentielles à la constitution d'un concept correspondant, impliquent différemment les régions visuelles et motrices selon que les objets ont des propriétés visuelles et/ou motrices. Dans cette expérience IRMf, les auteurs ont utilisé des mots fortement associés à des actions spécifiques tels que « agrafeuse » ou à des couleurs spécifiques « robe de mariée » ainsi que des mots associant à la fois des propriétés de couleur et de motricité, « balle de tennis » ou « gant de boxe ». Mais les auteurs ont aussi noté des activations différentes au sein des aires motrices (pIPS, LPI et MTG gauches) selon que l'on demandait aux sujets de se focaliser sur l'apparence (quelle est la couleur d'une balle de tennis ?) ou sur l'action (comment utiliser une balle de tennis ?) : les régions motrices étaient plus fortement activées dans le second cas. Or cette différence n'a pas été observée en réponse à des mots n'ayant aucune propriété d'action ou seulement des propriétés liées à la couleur, les résultats ne peuvent donc pas être liés à des aspects linguistiques mais plutôt aux contraintes contextuelles de la situation pour laquelle on a besoin d'accéder à une représentation.

Comme nous l'avons déjà signalé, plusieurs données neuropsychologiques semblent opposer des arguments forts à l'encontre d'une vision de type épiphénoménale puisque des lésions du système sensori-moteur peuvent perturber voire empêcher la compréhension d'actions, que ce soit pour des images ou dans le cadre du langage (Caramazza et al., 1990 ; Lambon et al., 1997; Lambon et al., 1999; Rapp et al., 1993 ; Hodges et al., 1996; Bak et Hodges, 2003, 2004 ; Bak et al. ; 2006, Boulenger et al., 2008 ; Kemmerer & Tranel 2003; Tranel et al. 2003; Silveri & Ciccarelli 2007; Grossman et al. 2008 ; Fazio et al. 2009 ; Rodríguez-Ferreiro et al., 2009). Ces données montrent un lien fonctionnel fort, mais restent cependant controversées.

Une façon plus classique d'étudier le rôle des représentations sensorimotrices est supportée par les études portant sur des informations temporelles précises concernant le traitement des stimuli (en MEG, EEG ou éventuellement en IRMf), car elles apportent des éléments de réponse importants à propos du rôle fonctionnel des représentations sensorimotrices. Ces études qui portent sur le décours temporel montrent généralement des activations très précoces des aires sensorimotrices (<200 ms voire même <100 ms) après la présentation du stimulus (Pulvermüller, 2005, 2008 ; Pulvermüller et al. 2001 ; Hauk & Pulvermüller 2004; Pulvermüller, Shtyrov, et al., 2005; Borreggine & Kaschak 2006; Boulenger et al. 2006; Zwaan & Taylor 2006; Boulenger et al., 2012 ; Scorolli & Borghi 2007). Cette activation différentielle précoce est difficile à concilier avec l'hypothèse que les activations sensorimotrices reflètent des phénomènes d'imagerie motrice postérieurs à la

compréhension ; elle est plus cohérente avec l'idée qu'il existe un lien causal et que la simulation sensorimotrice est responsable, au moins en partie, de la compréhension (Glenberg & Gallese, 2012 ; Gallese, 2008).

Les méthodes permettant d'interférer avec des processus cognitifs fournissent une approche encore plus directe car elles mettent en évidence un lien fonctionnel immédiat entre les systèmes sensorimoteurs et les connaissances conceptuelles (protocoles utilisant la TMS, mais aussi les paradigmes comportementaux tels que les tâches d'interférence, par exemple entre motricité et processus de plus haut niveau, langage, sémantique). Plus convaincants sont, en effet, les résultats qui montrent que stimuler le cortex moteur par TMS influence les performances dans des tâches conceptuelles spécifiquement pour les stimuli représentant des actions (Devlin & Watkins, 2007; Pulvermüller et al., 2005; Buccino et al., 2005 ; Ganis et al., 2000 ; Glenberg et al. 2008). De même, les expériences utilisant l'amorçage ou des paradigmes d'interférence mettent en évidence que les activations liées à des représentations sensorielles ou à des actions modulent les performances dans des tâches conceptuelles ultérieures (Vermeulen et al., 2008 ; Kiefer et al., 2011; Helbig et al., 2006; Myung et al., 2006 ; Witt et al., 2010; Fischer & Zwaan, 2008). Par exemple, observer un film représentant une action en tant qu'amorce facilite l'accès conceptuel à un objet cible présenté secondairement (Helbig et al., 2010). Pour Kieffer & Pullvemüller, ces résultats démontrent que les représentations sensorielles et motrices jouent un rôle fonctionnel dans le traitement conceptuel et excluent la possibilité que l'activation de représentations modalité-spécifiques dans des tâches conceptuelles résultent d'un épiphénomène, reflétant des processus associatifs non pertinents par rapport à la tâche. Pour certains auteurs, les représentations lexico-sémantiques seraient stockées dans les aires sensorimotrices (Pullvemüller, 2005; Pullvemüller & Fadiga, 2010).

- **<u>un réseau anatomico-fonctionnel étendu : importance des zones de convergence et de l'architecture des fibres de matière blanche</u>**

Comme nous l'avons déjà largement abordé, les systèmes sensorimoteurs ainsi que d'autres régions corticales et sous-corticales, semblent essentiels pour la compréhension sémantique, mais ces dernières seraient également impliquées dans d'autres fonctions cognitives. Les rôles de ces différences régions et la façon dont elles interagissent restent peu clairs ; d'autant plus que les découvertes de ces 20 dernières années dans le champ de la cognition incarnée remettent en cause les modèles du cerveau qui prédominaient jusque-là, via une interaction des systèmes sensorimoteurs et langagiers.

Une approche qui devient de plus en plus prégnante aujourd'hui est celle qui postule que le fonctionnement du cerveau en tant que réseau serait essentiel et qu'il existerait un certain nombre de zones de convergence (Damasio, 1989a ; 1989b; Damasio & Damasio, 1994) qui pourraient être engagées dans différents processus. Selon Damasio (1989), «*il est évident que la perception de la réalité extérieure (…) est une affaire neurophysiologique multi-sites*. » Ces zones de convergence fourniraient le lien neural qui permet la conversion automatique des co-activations non verbales en éléments narratifs et inversement. Ainsi, une même région du cerveau - une zone de convergence - pourrait intervenir dans différentes opérations cognitives en fonction de ses connexions au sein d'un réseau spécifiquement activé par une tâche donnée. C'est le fonctionnement en réseau qui permettrait l'émergence des processus cognitifs, mais une aire pourrait aussi avoir différents rôles fonctionnels car elle pourrait appartenir à plusieurs réseaux cérébraux en même temps. Il y aurait donc, en particulier, un réseau de la compréhension sémantique dont l'activation conjointe des différentes régions impliquées permettrait de créer et d'utiliser des représentations, les différentes régions de ce réseau pouvant également participer à d'autres fonctions cognitives. Ceci expliquerait qu'un certain nombre de régions engagées dans le traitement sémantique des actions sont aussi impliquées dans d'autres processus cognitifs (TOM, default mode network, mémoire épisodique…) L'organisation temporelle au sein de ce réseau (c'est à dire l'activité conjointe de populations de neurones distinctes) mais aussi la connectivité réelle, l'architecture de ce réseau, soutenue par les fibres de matière blanche seraient tout aussi importantes pour le fonctionnement de ce réseau. En effet si un tel réseau existe, il est probable que des voies de substance blanche spécifiques contribuent à l'infrastructure de ce réseau distribué et les connaissances neuroanatomiques du cerveau humain sont alors essentielles et indissociables des connaissances issues de la neuroimagerie fonctionnelle et des expériences comportementales.

Les dissections post-mortem, mais aussi les comparaisons que l'on peut faire avec l'anatomie du singe, sont pertinentes pour comprendre l'architecture de ces réseaux ; par ailleurs, l'imagerie par tenseur de diffusion (DTI), basée sur les propriétés de diffusion des molécules d'eau (Makris et al., 1997) fournit une méthode d'étude privilégiée pour analyser les voies de substance blanche chez l'être humain, in vivo. Un autre moyen explicite de tester l'hypothèse du fonctionnement cérébral en réseaux est apporté par des techniques d'analyses assez récentes qui permettent de vérifier la pertinence de l'activité d'une ou plusieurs régions cérébrales au sein d'un réseau, ces techniques postulent que la communication entre différentes régions cérébrales est reflétée dans la covariance des changements de l'activité cérébrale. Dans les modèles d'équations structurelles (SEMs), les covariances sont testées au sein d'un modèle anatomique restreint avec des hypothèses a priori sur la causalité des interactions neuronales (Penny, Stephan, Mechelli, & Friston, 2004; Büchel & Friston,

1997, 2000; MacIntosh & Gonzalez-Lima, 1991) ; les analyses DCM (dynamical causal modeling) explorent également de tels changements mais en fonction d'un input externe (Friston, Harrison, & Penny, 2003). D'autres approches, encore, se basent sur l'hypothèse que la communication efficace entre les aires du cerveau requiert des changements de l'activité qui soient cohérents dans le temps (Fries, 2005).

Ainsi, plusieurs régions dont nous avons parlé précédemment, pourraient effectivement jouer ce rôle de zones de convergence, sorte de « hubs » fonctionnels et anatomiques du réseau de la compréhension sémantique : il s'agirait particulièrement de la jonction pariéto-temporale, du cortex temporal mais aussi les régions préfrontales.

- *Régions pariéto-temporales*

En effet, depuis les premiers travaux de Vandenberghe sur les mots et les images (Vandenberghe et al 1996) jusqu'aux méta-analyses récentes sur le système sémantique (Binder & Desai 2011, Binder et al., 2009, Graves et al., 2010), la jonction temporo-pariétale, et en particulier le gyrus angulaire, apparaît comme une région cruciale pour la compréhension du sens. Il pourrait jouer ce rôle de « hub intégratif » car c'est un cortex associatif avec un rôle d'intégration multisensorielle qui intervient dans de nombreuses fonctions cognitives tels que la compréhension des buts et des intentions d'une action (Van Overwalle, 2009; Enrici et al., 2011 ; Frith & Frith, 2003), l'agentivité, la conscience de soi et de ses propres actions (Decety & Grèzes, 2006), l'imitation (Chaminade & Decety, 2002), la localisation dans l'espace (Ionta et al., 2011 ; Thirioux et al., 2010), dans la TOM (Saxe & Kanwisher, 2003) etc… Pour Jeffrey Binder (Binder & Desai, 2011), le gyrus angulaire (ainsi que la région plus large du TPJ) pourrait être une zone de convergence de haut niveau où les informations multimodalitaires sont assemblées en une représentation cohérente. Bien sûr, le TPJ aurait aussi une situation privilégiée au sein de la voie dorsale et en tant que hub, de par sa localisation anatomique à l'intersection des pôles pariétal et temporal et sa proximité avec de nombreuses aires adjacentes impliquées dans la compréhension des actions. Notamment, le STS qui est situé antérieurement et de façon adjacente au TPJ, dans le cortex temporal, aurait également un rôle clé dans la compréhension sémantique. Le STS apparaît comme « le caméléon des aires cérébrales » car cette région jouerait, comme le TPJ, un rôle dans de nombreux processus cognitifs : théorie de l'esprit et cognition sociale, intégration audiovisuelle, perception du mouvement, traitement de la parole… (Hein & Knight, 2008). Le STS est un important pivot parmi les sillons du lobe temporal et il est entouré de cortex adjacents tout aussi importants, les gyri supérieur et moyen (STG & MTG) ainsi que le gyrus angulaire à l'intersection postérieure avec le lobe pariétal inférieur. Le STS est également caractérisé par un pattern de connections anatomiques réciproques avec

d'autres régions du cerveau (Seltzer & Pandya, 1989, 1994 ; Keysers & Perrett, 2004; Petrides & Pandya, 2001 ; Barbas, 2000; Carmichael & Price, 1995).

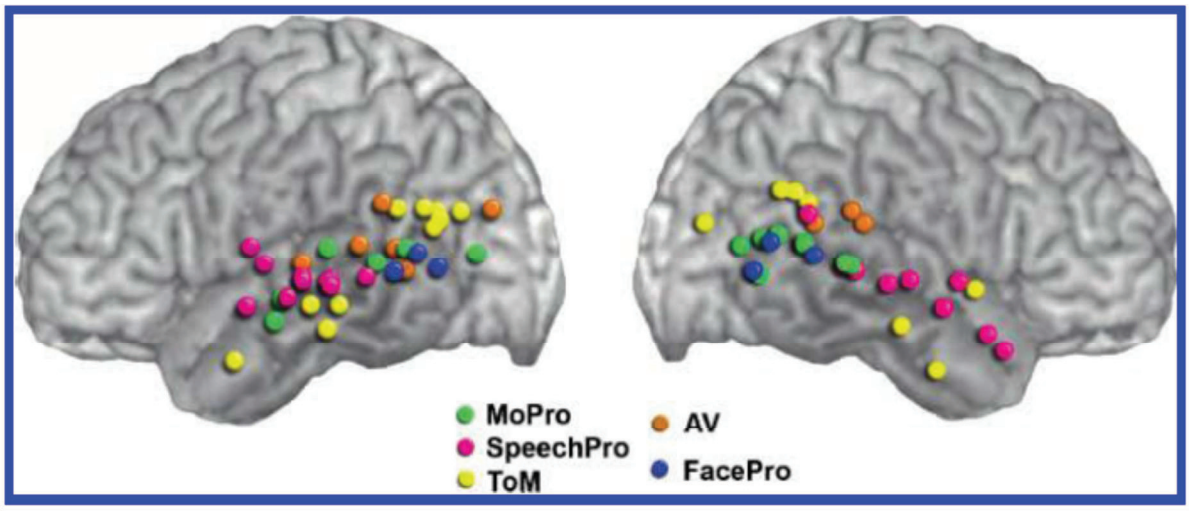


Fig 16) Foci d'activités issus de différentes études et représentant la variabilité des processus associés au STS
MoPro = traitement du mouvement; SpeechPro = traitement du langage ; ToM = theorie de l'esprit ; AV = intégration audiovisuelle; FacePro = traitement des visages.
Adapté de Hein & Knight., 2008

Classiquement, on considère que les différentes sous-régions du STS sont associées spécifiquement à des processus cognitifs distincts, mais pour Hein & Knight (2008), la variété des connexions réciproques entre certaines régions du STS et d'autres aires d'ordre supérieur pourrait être la base anatomique pour une hypothèse alternative. Au lieu d'une subdivision fonctionnelle stricte, la multifonctionnalité de la région STS pourrait dépendre des co-activations avec d'autres régions du cerveau, comme les cortex frontal, pariétal et temporal médial. L'activité de la région STS serait déterminée par la nature des interactions en réseau avec d'autres régions du cerveau activées conjointement. Par exemple, les co-activations avec les régions frontales ventrale et médiale, projetant vers le système limbique, pourrait déterminer le rôle de la STS dans TOM, alors que les co-activations avec les zones prémotrices lui conférerait son rôle dans le traitement du mouvement. C'est ce que semble confirmer la méta-analyse de Hein & Knight (2008) basée sur des études en IRMf. Même s'ils ont mis en évidence une division du STS en deux régions (une région plus antérieure, principalement associée au traitement de la parole ainsi que, secondairement, à la TOM et une région plus postérieure engagée dans le traitement des visages, l'intégration audio-visuelle, le traitement du mouvement et la TOM), leurs résultats montrent également que même les régions antérieures du STS les plus « spécialisées » ne sont pas uniquement reliées au traitement de la parole. Pour Hein & Knight, la différenciation antéro/postérieure du STS refléterait plutôt des différences de degrés de multifonctionnalité de ces régions plutôt qu'une subdivision fonctionnelle.

Au niveau temporal, certains considèrent le pôle temporal antérieur comme un « hub » intégratif en raison de son large rôle dans le traitement conceptuel (Simmons & Martin, 2009; Visser et al., 2010 ; Kiefer & Pulvermüller, 2012). Comme nous l'avons déjà mentionné, cette région serait considérée comme une zone de traitement conceptuel amodal et symbolique selon les défenseurs d'un système conceptuel amodal (McClelland & Rogers, 2003; Rogers et al., 2004) servant de « lieu de stockage » pour les représentations conceptuelles. Dans cette perspective, les simulations sensorielles et motrices ne seraient que des portes d'entrée et de sortie au système conceptuel amodal (Rogers et al, 2004), les systèmes sensoriels permettant d'accéder à la mémoire conceptuelle et les systèmes moteurs permettant d'exprimer les connaissances conceptuelles à travers les comportements. Mais Kiefer & Pulvermüller (2012) proposent une vision alternative, vision tout à fait compatible avec l'aspect multifonctionnel des aires temporo-pariétales et un fonctionnement en réseaux mais aussi avec les perspectives de la cognition incarnée. Les connaissances conceptuelles seraient « incarnées » et représentées dans les zones sensorielles et motrices tandis que le cortex temporal antérieur ne ferait que faciliter et influencer le traitement conceptuel dans ces zones mais ne stockerait pas les connaissances conceptuelles en soi.

Encore une fois, le modèle de Damasio des zones de convergence serait tout à fait approprié pour expliquer ce rôle du cortex temporal antérieur car cela donnerait la possibilité à cette région d'intégrer les différentes caractéristiques conceptuelles modalités-spécifiques dans un espace sémantique commun (Pulvermüller et al., 2010; Patterson et al., 2007; Kiefer, Schuch et al., 2007 ; Kiefer, Sim et al., 2007; Simmons & Barsalou, 2003). Cette intégration pourrait être réalisée grâce à des représentations supramodales de haut niveau indiquant quelles sont les caractéristiques sensorimotrices essentielles pour la constitution d'un concept et où elles sont représentées, ces représentations de haut niveau ne stockant pas le contenu sémantique en soi mais servant à guider la récupération des informations stockées en réactivant l'activité des aires motrices et sensorielles. Cette approche de « représentations supramodales » a déjà été utilisée pour discuter le rôle du cortex préfrontal dans la mémoire de travail (Kessler et Kiefer, 2005) ou de l'hippocampe dans la mémoire épisodique (McClelland et al, 1995).

- *Régions frontales*

Les régions frontales, quant à elles, seraient aussi impliquées dans de nombreux processus, principalement à forte composante sociale tels que le suivi de l'action, la perception des personnes et la mentalisation (Amodio & Frith, 2006). Dans le réseau temporo-pariéto-frontal, les régions frontales (et en particulier le cortex préfrontal médial) pourraient avoir un rôle de « chef d'orchestre » avec une influence spécifique sur les autres régions. Dans la compréhension, les

simulations des aires sensorimotrices seraient sous le contrôle des aires frontales et de leurs fonctions d'organisation structurelle et de planification (Lallée, Madden et al., 2010, Madden et al., 2010). Les régions frontales contribueraient à l'intégration sémantique car elles seraient responsables de l'intégration des informations en provenance des voies ventrales et dorsales, du « what » et du « where » (Zhu et al., 2012 ; De Borst et al. 2012). L'activité du cortex préfrontal ventro-latéral a souvent été associée à un rôle de contrôle top-down de l'activité en mémoire sémantique, qui permettrait de guider la recherche et la sélection d'informations conceptuelles sous la dépendance de régions temporales postérieures et pariétales (Martin, 2007 ; Gold et al. 2005). En particulier, il aurait un rôle de sélection quand il y a un choix à effectuer, une compétition entre plusieurs alternatives avec un rôle actif dans la sélection de la représentation la plus adaptée pour une meilleure compréhension (Thompson-Schill et al. 1999, Gough et al. 2005 ; Martin, 2007). Toutefois, le fonctionnement en réseau entre ces différentes régions resterait vraiment essentiel. Van Doren et al. (2010) ont cherché à modéliser la connectivité effective entre différentes régions de ce réseau fronto-pariétal par DCM. Ils ont montré que le modèle qui prédisait le mieux la récupération en mémoire d'informations sémantiques (pour des mots et des images) impliquait le gyrus frontal inférieur et le sillon occipito-temporal, car ces deux régions auraient des connexions privilégiées : le gyrus frontal inférieur constituerait ainsi une zone de convergence qui communiquerait avec les systèmes spécialisés pour les mots et images, via ses connexions avec le sillon occipito-temporal.

- *Connectivité anatomique*

En ce qui concerne la connectivité réelle de ces régions, des travaux à la fois chez l'humain et le singe soutiennent cette organisation en réseau fronto-temporo-pariétal : le TPJ et le STS seraient interconnectés mais aussi tous deux connectés aux aires frontales et aux régions prémotrices (Seltzer & Pandya, 1989, 1994 ; Keysers & Perrett, 2004; Petrides & Pandya, 2002 ; Barbas, 2000; Carmichael & Price, 1995 ; Decety & Grèzes, 2006 ; Rushworth et al., 2006; Koch et al., 2010 ; Uddin et al.,2010).

Si plusieurs études en neuroimagerie ont contribué aux connaissances sur les fonctions langagières, principalement en combinant IRMf et DTI (Anwander et al, 2007 ; Frey et al 2008 ; Makris et al 2005 ; Makris & Pandya, 2009 ; Friederici & Gierhan, 2012 ; Catani et al., 2005 ; Glasser & Rilling, 2008 ; Flöel et al., 2009 ; Harvey et al., 2013), à notre connaissance, aucune étude n'a cherché à déterminer l'architecture des fibres d'un réseau sémantique commun impliqué dans la compréhension ou la création de représentations à partir d'une image ou d'une phrase. En tout cas, les études déjà existantes réaffirment l'idée que l'architecture des fibres de matière blanche est importante fonctionnellement.

Ainsi, pour le langage, Catani et ses collègues (Catani et al 2005) ont utilisé le DTI pour examiner l'organisation du faisceau arqué de l'hémisphère gauche au sein du réseau périsylvien du langage. Ils ont identifié la voie classique reliant les zones de Broca et de Wernicke (le faisceau arqué), mais aussi une voie indirecte supplémentaire et jusqu'alors inconnue, passant par le cortex pariétal inférieur. Ils suggèrent que cette voie indirecte est impliquée dans les fonctions langagières sémantiques, alors que la voie directe est impliquée dans les fonctions phonologiques.

Friedereci et Gierhan (2012) se sont principalement intéressées aux voies de fibres qui soutiennent la différence entre fonctions sémantique et syntaxique du langage, mettant en évidence deux chemins dorsaux qui relient respectivement le cortex temporal et le cortex prémoteur d'une part (impliqué dans les processus de répétition) et le cortex temporal et la partie postérieure de Broca d'autre part (impliqué dans le traitement des processus syntaxiques complexes), tandis que ventralement on trouve le faisceau unciné et le faisceau inférieur fronto-occipital (qui interviennent dans les traitements des aspects sémantiques et des processus syntaxiques basiques).

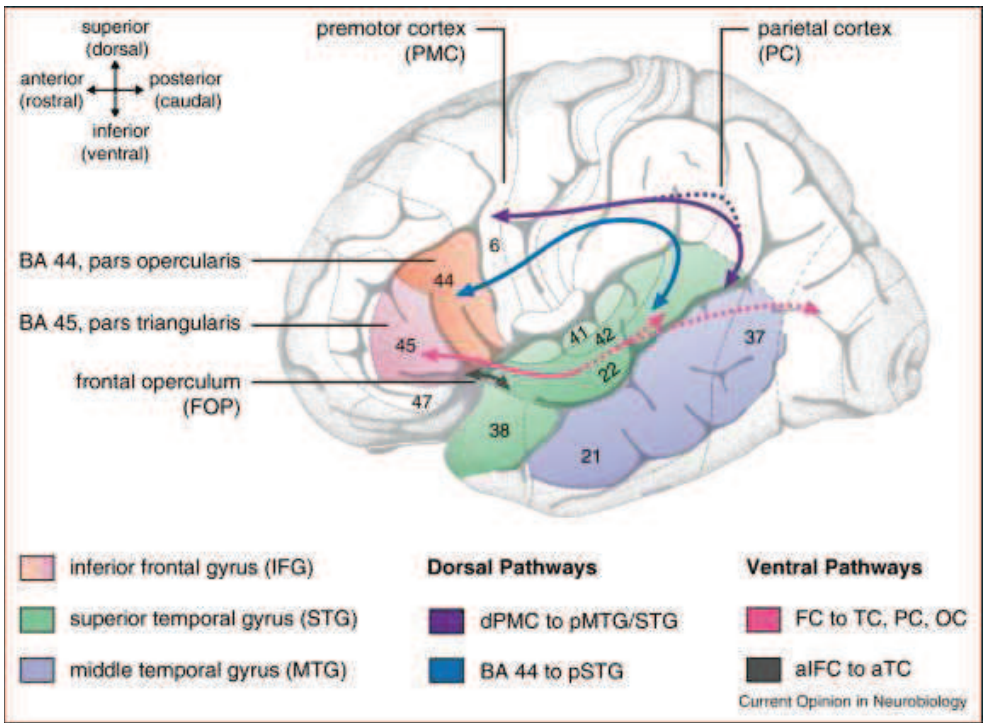


Fig 17) Représentation schématique des principales régions et principaux faisceaux du réseau langagier
La voie dorsale, reliant le cortex prémoteur dorsal (dPMC) au cortex temporal postérieur (pMTG / STG), implique le SLF III et/ou le SLF II et le SLF-tp ; la voie dorsale reliant BA 44 avec le STG postérieur implique l'AF. La voie ventrale reliant le cortex frontal (FC), c'est-à dire BA 45 et d'autres, avec le temporal (TC), le pariétal (PC), et le cortex occipital (OC), implique l'IFOF **;** la voie ventrale reliant le FC antéro-inférieur (aIFC), et le FOP, avec le TC antérieura (aTC), implique l'UF.
Adapté de Friederici et Gierhan, 2012

Harvey et al. (2013) ont étudié les rôles des faisceaux unciné, longitudinal inférieur et fronto-occipital inférieur dans la compréhension sémantique chez des patients aphasiques et des sujets sains contrôles. Ils ont mis en évidence que l'intégrité du faisceau unciné (qui serait touchée chez les aphasiques) et la force de connectivité fonctionnelle avec les régions que connecte l'unciné, sont de bons prédicteurs des performances des patients. Mais il est à noter que ce protocole porte avant tout sur des aspects verbaux de récupération d'un mot, et non sur une comparaison stimuli verbaux/non verbaux (même si les tâches sémantiques employées utilisent des images, Pyramids and Palm Trees test et auditory word-picture verification task). Les résultats de cette étude suggèrent que la déficience du contrôle sémantique dans la compréhension des mots, observée chez les patients aphasiques, résulte d'une communication appauvrie entre les régions que le faisceau unciné connecte, à savoir le gyrus frontal inférieur gauche (pars orbitalis) et le lobe temporal antérieur.

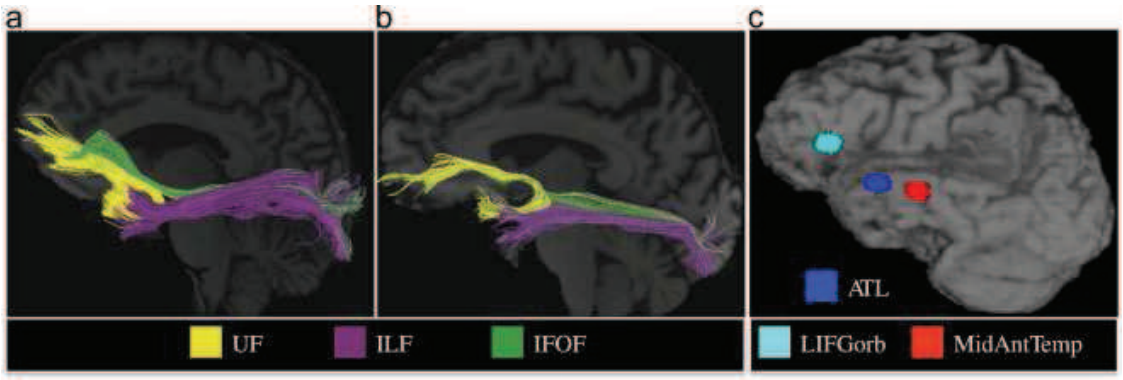


Fig 18) Représentation des rapports entre les faisceaux unciné (UF), longitudinal inférieur (ILF) et faisceau fronto-occipital inférieur(IFOF) ainsi que des régions fonctionnelles permettant d'initier la tractographie.
L'UF relie le gyrus frontal inférieur gauche pars orbitalis (LIFGorb) et le lobe temporal antérieur (ATL) ; l'ILF se termine dans l'ATL avec projections près du lobe temporal moyen antérieur (MidAntTemp) ; et l'IFOF se projette près du LIFGorb et du MidAntTemp.
Adapté de Harvey et al., 2013

Mais les résultats DTI présentent un certain nombre de contraintes méthodologiques qui limitent la précision et, par conséquent, l'interprétation que l'on peut donner à ces résultats. Les travaux de H. Duffau et ses collaborateurs qui combinent, chez des patients atteints de gliomes, données comportementales, stimulations électriques pré-opératoires, IRMf/DTI et dissections post-mortem, représentent une approche unique qui apporte des informations précieuses sur l'organisation anatomique et le fonctionnement des réseaux cérébraux, en particulier du système sémantique (Mandonnet et al., 2007 ; Sarrubo et al., 2011 ; Martino et al., 2009 ; Martino et al, 2010 ; Maldonado et al., 2012 …).

Ainsi, Mandonnet et al (2007) ont voulu vérifier l'hypothèse que le réseau sous-cortical soutenant les aspects sémantiques du langage (« voie ventrale sémantique ») serait constitué de 2 voies de

matière blanche, parallèles et distinctes (Duffau et al., 2002, 2005). Les résultats de cette étude (combinant données anatomiques lésionnelles des patients, IRMf et stimulation électrique corticale) semblent confirmer l'existence de ces deux voies parallèles au sein de l'hémisphère gauche : avec l'existence d'un faisceau direct, le faisceau occipito-frontal inférieur, qui connecte les aires temporales postérieures à la région orbitofrontale et l'existence d'un second faisceau indirect, le faisceau longitudinal inférieur, qui connecte les régions occipito-temporales postérieures au pôle temporal et serait ensuite relayé par le faisceau unciné, permettrant une connexion temporo-frontale. Cependant, alors que le premier faisceau direct serait crucial pour le traitement sémantique du langage (puisque la stimulation de ce faisceau génère des paraphasies sémantiques, Duffau et al., 2005), le second ne serait pas nécessairement indispensable (puisque les compétences langagières sont préservées durant la stimulation mais aussi après résection partielle). Le faisceau occipito-frontal inférieur est parfois considéré comme la voie sémantique « ventrale » proprement dite, car il connecte les aires verbo-sémantiques : depuis les aires postérieures du cerveau au aires frontales (GFi, cortex orbitofrontal latéral et cortex préfrontal dorsolatéral). Sarrubo et al (2011) ont combiné des données DTI et celles issues de dissections post-mortem pour caractériser plus finement l'anatomie et le rôle du faisceau fronto-occipital inférieur (IFOF) qui connecte les aires occipito-temporales au lobe frontal (cortex occipital, aires temporo-basales et LPS, Martino et al., 2009). Sarrubo et al (2011) ont mis en évidence l'existence de deux sous-faisceaux avec des terminaisons distinctes au niveau frontal : le premier est antéro-supérieur et se termine directement dans le GFi alors que le second, situé plus profondément, est divisé en 3 portions qui rejoignent le cortex orbitofrontal et le gyrus frontal moyen (portions postérieure, moyenne et antérieure).

L'IFOF correspondrait à un regroupement multi-fonctionnel de fibres, chacune de ces composantes anatomiques pouvant être impliquée dans des processus cognitifs distincts (Duffau et al. 2005; Henry et al. 2004; Martino et al. 2009). En particulier le faisceau antéro-supérieur et la composante postérieure du faisceau plus profond seraient impliqués dans des processus sémantiques. Par ailleurs, la composante moyenne du second faisceau jouerait un rôle dans l'intégration sensorimotrice et la composante antérieure dans des aspects plus émotionnels et comportementaux. La délimitation et le rôle de certains faisceaux restent mal connus, c'est notamment le cas du faisceau longitudinal moyen (MdLF) qui connecterait les régions pariétales et temporales et pourrait jouer un rôle dans le traitement sémantique (Menjot de Champfleur et al., 2012).

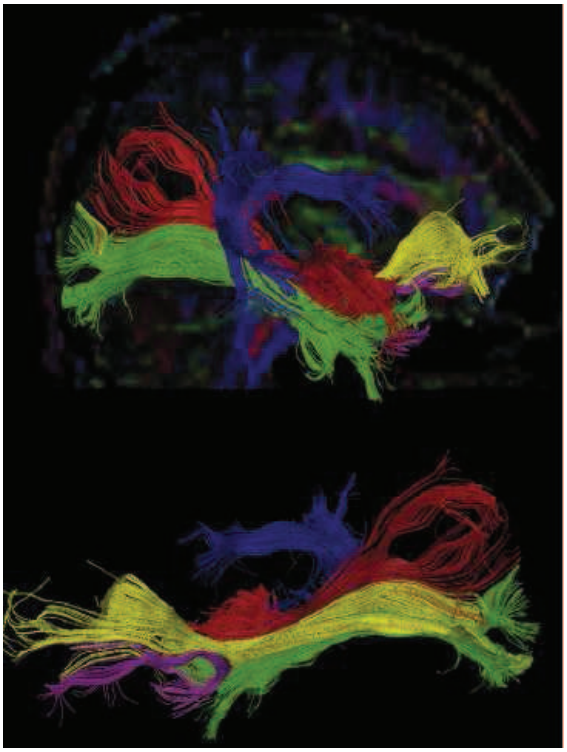


Fig 19) Vue globale des principaux faisceaux impliqués dans le traitement sémantique en vue latérale et médiale
En rouge : le MdLF, en jaune : l'IFOF, en bleu : le FA, en vert : l'ILF, en violet : le FU
Adapté de Menjot de Champfleur et al., 2012

Globalement, les notions de zones de convergence et de fonctionnement en réseaux permettent d'expliquer le rôle du réseau fronto-temporo-pariétal dans la compréhension mais aussi le fait que ce réseau soit associé à des processus cognitifs variés impliquant la perception/compréhension des actions, des processus de cognition sociale ou encore des aspects mnésiques épisodiques et/ou sémantiques. Elles fournissent également un support neuroscientifique à l'interaction corps/esprit qui a fait débat pendant des siècles. Dans son ouvrage « l'erreur de Descartes », A.Damasio illustre les interrelations étroites qu'entretiennent le corps et le cerveau pour percevoir et interagir dans le monde : le corps fournit le contenu fondamental des représentations mentales, car celui-ci constitue le cadre de référence de notre représentation du monde et de notre relation à ce dernier. Les représentations fondamentales du corps en train d'agir forment un cadre spatial et temporel stable, sur lequel les autres représentations pourraient s'appuyer car « *le corps fournit au cerveau davantage que ses moyens d'existence et que la modulation de ses activités. Il fournit un contenu faisant intégralement partie du fonctionnement mental normal* » (Damasio, 1994, 285). Ainsi, selon l'approche de Damasio et en opposition avec la pensée cartésienne, le fait d'exister précéderait celui de penser… *« Sum ergo Cogito* ».

CHAPITRE II : Partie Expérimentale

Préambule à la partie expérimentale

- **Création de protocoles expérimentaux : de la difficulté de normer des stimuli complexes**

Si nous voulons étudier le système sémantique, nous devons être très prudents sur la façon dont nous interrogeons ce système. Nous avons donc porté une attention particulière à la création des stimuli visuels et linguistiques utilisés dans nos expériences. Un certain nombre de problèmes méthodologiques se posent lorsque l'on veut réaliser un matériel destiné à tester et comparer des aspects verbaux et non-verbaux. Aussi, nous avons essayé de contrôler au mieux ces deux types de stimuli pour les différents protocoles que nous avons construits, en normant les stimuli Images ainsi que les stimuli Phrases, pour les rendre équivalents d'une modalité à une autre (Images vs Phrases), mais aussi équivalents entre eux au cours de différentes sessions ou conditions expérimentales et sur plusieurs aspects (nombre, genre et âge des personnages…) (cf. Description des protocoles IRMf, comportement et EEG pour plus de détails…).

Il existe un certain nombre de bases de données images permettant d'obtenir des stimuli et différents critères qui leur sont associés (familiarité, name agreement…), mais ces bases sont souvent très limitées (avec beaucoup moins d'entrées que celles portant sur des données langagières) et certains items sont fortement connotés culturellement et n'ont donc pas autant de signification pour des participants français (une batte de baseball, un casque de football américain…). Ces bases de données ne contiennent que l'évaluation de stimuli isolés de leur contexte (ex : objets, animaux) et non de stimuli plus complexes comme des scènes ou des actions. Dans le cadre de la base de données Lexique, Alario et Ferrand (1999) ont développé un set regroupant 400 images en noir et blanc, standardisées sur différentes variables (name agreement, image agreement, familiarité, complexité visuelle). Souhaitant utiliser des stimuli en couleur et plus complexes représentant des scènes de la vie quotidienne, nous avons sélectionné nos images sur la banque de données Getty images © dont les images respectent généralement un certain nombre de critères expérimentaux (valence émotionnelle, neutralité du contexte entourant l'action…) et sont facilement sélectionnables (moteur de recherche par thème, nombre de personnages, âge et sexe…). Les images choisies devaient être non ambigües et a priori facilement reconnaissables. Nous nous sommes inspirés du concept de « Name agreement » d'Alario et Ferrand (1999), qui prend en considération le degré d'accord des sujets avec le nom choisi pour un stimulus et l'avons appliqué aux images que nous avons nous-mêmes sélectionnées, mais plutôt sous la forme d'un critère de bonne

reconnaissance de l'action représentée par une image (Action Agreement). Pour chacun de nos protocoles, nous avons vérifié auprès de sujets pilotes qu'il y avait une bonne compréhension de la scène à travers ce critère d'action agreement (les sujets devaient décrire d'une phrase l'action représentée et nous avons comptabilisé ces descriptions en termes de scores et de pourcentages).

Toute la difficulté de la sélection et de la normalisation des stimuli a été rendue plus complexe par le fait qu'à la sélection d'images s'ajoutait la création de stimuli verbaux. Dans un second temps, nous avons donc créé des phrases décrivant des actions de la vie quotidienne, inspirées de nos stimuli images. Normer des stimuli langagiers pour une tâche pose toujours le problème d'une base de référence commune : comment sélectionner les stimuli et sur quels critères, quelle base de données recense les critères appropriés ? Là aussi, il existe différents outils (Brulex, basée sur les entrées du Petit Robert, Manulex ou Novlex sur la base de manuels scolaires…) sans référence commune entre eux ; il y a d'autres outils plus complets mais dans d'autres langues (la MRC Psycholinguistic Database, Celex). Par ailleurs les différentes bases ne comparent pas les mêmes aspects : fréquence, degré de familiarité, etc. Nous étions avant tout intéressés par l'évaluation de la fréquence des mots car nous souhaitions créer des jeux de phrases équivalents où des facteurs linguistiques, tels que la fréquence, ne viennent pas biaiser nos résultats. Nous avons choisi d'utiliser Lexique (Lexique 3 ; New, Pallier, Ferrand et Matos, 2001), une base de données qui fournit, pour 135 000 mots du français, leurs représentations orthographiques et phonémiques, la syllabation, la catégorie grammaticale, le genre et le nombre, les fréquences, les lemmes associés, etc. La base de données Lexique est sans doute en français, la base lexicale la plus complète actuellement et la plus adaptée au choix des items pour une étude en français. Cette base de données présente l'avantage de comporter un grand nombre d'entrées mais est aussi intéressante car les modalités de constitution de cette base sont originales et très documentées. Dans Lexique 3, la fréquence des mots est évaluée à la fois à partir de films et de livres ; ces deux évaluations présentent des avantages et des inconvénients. New, Brysbaert, Veronis & Pallier (2007) ont réalisé une étude sur la pertinence de l'usage de sous-titres de films pour estimer la fréquence des mots. Les fichiers de sous-titres étant très facilement accessibles sur Internet, ils représentent un moyen simple et très rapide de collecter des informations sur la fréquence des mots (beaucoup moins fastidieux qu'avec des corpus de textes). Les auteurs ont pu obtenir des informations de fréquence sur 52 millions de mots en français. Bien que cette méthode d'évaluation soit très prometteuse, elle pose quelques problèmes : en effet ce corpus est basé sur des sous-titres issus pour la majorité d'entre eux de films et de séries américains et anglais (46 millions de mots) et seulement pour une petite partie sur des sous-titres de films français et européens (1,9 et 2,5 millions de mots), ce qui représente un biais culturel certain (eg : les auteurs ont relevé une forte représentation du vocabulaire policier qui correspond au grand

nombre de séries policières produites aux USA). De plus, ce corpus présente un nombre important de références du langage parlé et les auteurs notent d'ailleurs un certain nombre d'hésitations et d'erreurs de prononciation, propres au langage parlé. Le critère fréquence livres, quant à lui, est basé sur un ensemble composé de 218 textes littéraires (romans) publiés entre 1950 et 2000, ce qui représente un corpus de 14,7 millions d'items. Bien sûr, un corpus littéraire soulève aussi des problèmes (problème d'actualisation de ces bases de données, langage soutenu vs langage oral, etc.), mais il reste cependant une bonne référence pour les stimuli langagiers, un corpus filmographique étant plus limité sur certains aspects (un film élude le côté descriptif : les objets sont présents dans le décor mais ne sont pas toujours cités dans le dialogue ou décrits alors que le livre fait une plus large part à la désignation et à la description d'objets ou d'êtres vivants).

Pour le choix de nos items langagiers de nos différents protocoles, nous avons donc utilisé les deux types de fréquences (fréquence livres et fréquence films). Nous avons choisi de n'utiliser que les fréquences des mots de classe ouverte (noms, verbes, adjectifs, adverbes…) et de moyenner l'ensemble de ces fréquences pour obtenir la fréquence pour une phrase. En revanche, nous avons choisi de ne pas tenir compte de la fréquence des déterminants (le, la, les, un, une…) ou des pronoms personnels (il, elle, ils, elles) car leurs fréquences sont tellement fortes qu'elles auraient noyé les subtilités liées aux mots de classe ouverte ; de plus leurs fréquences sont relativement équivalentes (le=la, un=une, il=elle). Nous décrirons plus en détails la sélection des stimuli pour chaque protocole dans les chapitres correspondants.

II.1) Partie Expérimentale 1 : utilisation de l'IRMf et du DTI pour mettre en évidence le réseau sémantique commun

II.1.*1*) Introduction

Nous avons largement décrit, dans le chapitre théorique précédent, comment la neuroimagerie, et en particulier l'IRMf, pouvait fournir des réponses sur la localisation des régions cérébrales impliquées dans le traitement sémantique. Un large réseau d'aires fronto-temporo-pariétales sensorimotrices et associatives semble mis en jeu dans ce processus, puisqu'on les retrouve de façon cohérente dans diverses études (Vandenberghe et al., 1996 ; Moore & Price ; 1999, Devlin et al., 2002 ; Bright, Moss & Tyler ; 2004 ; Chee et al. 1999 ; Shinkareva et al. 2011 ; Gallagher et al., 2000; Gates & Yoon, 2005 ; Kable et al. 2002 ; Aziz-Zadeh et al., 2006 ; Aziz-Zadeh & Damasio, 2008 ; Tremblay et al. 2011).

De même, la connectivité du réseau sémantique commence à être bien étudiée grâce au DTI. Cependant, elle se limite principalement au cadre du langage où l'on considère que ce réseau serait organisé selon deux voies (Hickok & Poeppel, 2000, 2004, 2007) : une voie « dorsale » phonologique constituée du faisceau arqué et de la troisième portion du SLF (SLFIII) (Makris et al., 2005) et une voie « ventrale », sémantique qui engagerait l'IFOF, l'ILF et l'UF (Duffau, 2008). Par ailleurs, le rôle de certains faisceaux reste mal connu, c'est notamment le cas du MdLF, un faisceau initialement décrit chez le singe, puis chez l'homme, qui connecterait la partie caudale du lobule pariétal inférieur au gyrus temporal supérieur (Schmahmann et al., 2007; Seltzer & Pandya, 1984 ; Makris et al., 2009 ; Menjot de Champfleur et al., 2012).

Bien que de nombreuses recherches aient abordé la notion de représentation conceptuelle dans le cadre du traitement des images, des mots ou des phrases (Binder & Desai, 2011), aucune étude à notre connaissance n'a cherché à combiner l'imagerie fonctionnelle et structurelle pour examiner les aires impliquées aussi bien dans la compréhension d'images que dans celle de phrases et pour essayer de déterminer les réseaux cérébraux de matière blanche qui pourraient constituer la structure soutenant ce système commun en reliant les principales zones d'intégration sémantique. Une question qui parait essentielle est celle de la voie sémantique : les faisceaux qui composent la voie sémantique du langage sont-ils les mêmes que ceux qui soutiennent la compréhension des scènes visuelles ?

L'objectif principal de ce travail est donc d'utiliser une approche combinant IRMf et DTI afin d'évaluer l'existence d'un système commun impliqué à la fois dans le traitement sémantique d'images et de phrases, représentant ou décrivant des personnes humaines en train d'effectuer une

action (la thématique de ce travail s'inscrivant pleinement dans le domaine de la cognition incarnée). Des analyses spécifiques (conjonction) seront utilisées pour caractériser les aires fonctionnelles recrutées dans la compréhension de ces deux types de stimuli (phrases et images). Puis nous nous intéresserons à la connectivité de ce réseau, en étudiant à la fois la connectivité fonctionnelle (à travers des analyses ICA) mais aussi la connectivité structurale en étudiant les faisceaux de matière blanche du réseau sémantique commun.

- **La méthodologie d'Imagerie par Résonance Magnétique fonctionnelle (IRMf)**

Rappelons que l'IRMf est une mesure indirecte de l'activité neuronale, permettant d'observer les activations cérébrales induites lors d'une tâche comportementale à travers des modifications hémodynamiques.

- *Principe de l'IRMf*

L'IRMf dérive de la méthode d'imagerie par résonance magnétique nucléaire (IRMN), elle-même basée sur le principe de la résonance magnétique nucléaire (RMN), un phénomène originellement utilisé en spectroscopie, pour l'étude des propriétés quantiques des noyaux atomiques (Hornak, 1996). L'IRM repose sur l'utilisation d'un champ magnétique puissant, généré par un aimant supraconducteur qui crée une magnétisation des tissus par alignement des moments magnétiques de particules élémentaires, les spins. Ces particules, ainsi placées dans un champ magnétique constant, sont soumises à l'application secondaire de champs magnétiques oscillants et plus faibles (dit champs radiofréquences) qui ont une orientation perpendiculaire à l'orientation du champ principal. Ces pulses de radiofréquences modifient légèrement l'alignement des spins et certaines particules émettent un signal détectable appelé RMN. La technique d'IRM repose principalement sur la réponse RMN du noyau d'hydrogène, le proton, présent en grandes quantités et différemment selon les tissus biologiques. La spécificité de l'IRM consiste à localiser précisément dans l'espace l'origine de ce signal RMN issu des protons, en appliquant des champs magnétiques non constants, des « gradients », qui vont induire des fréquences légèrement différentes en fonction de la position des atomes dans ces gradients.

- *Applications aux structures biologiques*

Toute activité cognitive nécessite un apport constant de glucose et d'oxygène au niveau des neurones impliqués dans cette activité, apport réalisé grâce à une hausse transitoire du débit sanguin cérébral (DSC ou regional Cerebral Blood Flow, rCBF). En effet, quand un groupe de neurones devient plus actif au sein de régions cérébrales, une vasodilatation locale des capillaires sanguins cérébraux crée un afflux de sang et donc d'oxygène vers ces régions. Même si le lien entre activité cérébrale et

augmentation du début sanguin avait été fait dès la fin du 19ème siècle (avec les travaux d'Angelo Mosso et de Roy et Sherrington), il a fallu attendre les travaux de Kety et Schmidt en 1948, pour vraiment comprendre le fonctionnement du débit sanguin cérébral en lien avec l'activité neuronale. Et ce n'est qu'avec la découverte de l'effet BOLD (pour Blood Oxygenation Level Dependent) dans les années 90 (Ogawa et al. 1990, 1992, 1998) que l'on a pu aboutir à l'IRMf telle que nous la connaissons aujourd'hui. Les neurones ne disposant pas de réserves d'énergie sous forme de glucose et d'oxygène, leur activation provoque un besoin en énergie auquel répondent les cellules sanguines par un processus appelé réponse hémodynamique. L'hémoglobine est une protéine possédant un atome de fer qui transporte l'oxygène ; elle a des propriétés magnétiques différentes selon qu'elle transporte de l'oxygène sous forme d'oxyhémoglobine ou qu'elle en a été débarrassée, constituant ainsi la désoxy-hémoglobine. Sous l'effet d'un champ magnétique (produit par l'IRM), la désoxy-hémoglobine va acquérir une aimantation dirigée dans le même sens que le champ magnétique, détectable par la machine IRM car la concentration de désoxy-hémoglobine est paramagnétique (alors que l'oxy-hémoglobine est diamagnétique et donc non active en IRM). Ce phénomène va produire une légère perturbation du champ magnétique et c'est cette différence de susceptibilité magnétique que l'IRMf va détecter. C'est donc grâce à cet effet BOLD, basé sur la variation des concentrations sanguines d'oxy et désoxy-hémoglobine dans les territoires veineux, que l'on peut mesurer indirectement l'activation neuronale.

En neuroimagerie fonctionnelle, on cherche généralement à connaître l'activité liée à une tâche comportementale en réalisant des cartes d'activations des variations du DSC et du niveau d'oxygénation local mesurées par la réponse RMN de l'hémoglobine et en les projetant sur des images anatomiques obtenues par IRM anatomique du cerveau. Les cartes d'activations observées en IRMf correspondent en fait à l'activité de groupements de voxels proches dont le signal va varier au cours du temps en fonction d'une tâche cognitive.

- *Avantages et inconvénients de la technique*

L'IRMf reste donc une technique relativement récente qui a beaucoup apporté aux neurosciences et qui présente des avantages certains ; il faut cependant toujours rester prudents quant à l'interprétation des résultats. Le principal avantage de l'IRMf est qu'elle possède une très bonne résolution spatiale (<2mm). Toutefois, elle présente des inconvénients, le principal désavantage étant de constituer une mesure indirecte de l'activité neuronale avec une mauvaise résolution temporelle, ce qui complique l'étude d'événements cognitifs brefs et ne permet pas d'accéder à des informations sur un décours temporel précis ou sur l'organisation dynamique des processus. En effet, le temps de latence de la réponse hémodynamique est relativement lent,

d'environ 2s (Kwong et al., 1992) par rapport à la rapidité des processus cognitifs (à partir de 100ms) et de la réponse neuronale (quelques ms). Cependant l'approche IRMf événementielle (event-related fMRI) permet un peu de palier la mauvaise résolution temporelle, en rendant possible le repérage de l'activité associée à la présentation de stimuli (même brefs) quand il sont répétés et séparés par des intervalles de temps irréguliers (Dale & Buckner et al., 1997, Dale & Buckner et al , 1998 ; D'Esposito et al., 1999).

- **La méthodologie d'Imagerie du Tenseur de Diffusion (DTI)**

L'IRM de diffusion, également appelée IRM DTI (Imagerie du Tenseur de Diffusion) est une technique basée sur l'imagerie par résonance magnétique (IRM) qui permet de mesurer de façon non-invasive la distribution des directions de diffusion des molécules d'eau dans le cerveau, l'eau étant un composant essentiel des tissus biologiques et les protons des molécules d'eau contribuant majoritairement au signal de résonance magnétique (Le Bihan, 2001; 2003 ; Mori, 2007).

Dans un milieu libre et sans obstacle, la diffusion des molécules d'eau s'opère de façon isotrope, selon les lois du mouvement brownien et on peut caractériser cette diffusion par un coefficient de diffusion. Dans les tissus biologiques, comme la température n'est pas nulle, l'agitation thermique provoque un mouvement de diffusion des molécules d'eau. Mais, dans les tissus, la diffusion est contrainte par des obstacles tels que les membranes biologiques, ce qui diminue la mobilité des molécules d'eau et par conséquent le coefficient de diffusion (car les molécules situées à proximité des obstacles ont un parcours moyen diminué). La diminution de la mobilité des molécules d'eau en présence d'un obstacle se fait dans la direction perpendiculaire à celui-ci et la diffusion devient anisotrope ; on parle alors de tenseur de diffusion. Dans la matière blanche cérébrale, qui est structurée en fibres constituées de regroupements d'axones, la diffusivité est de trois à six fois plus grande dans l'axe des fibres que dans le plan orthogonal. La mesure de l'axe de plus grande diffusivité donne donc une information sur l'orientation locale des fibres de matière blanche.

On peut rendre sensible le signal de résonance magnétique à la diffusion en appliquant des séquences IRM spécifiques qui font intervenir successivement deux gradients de champ magnétique courts et intenses, générant sur les protons un déphasage de précession relatif à leur position. Le premier gradient donne à la précession des protons une avance de phase proportionnelle à leur position sur l'axe du gradient tandis que le second (qui est opposé au premier) produit un retard de phase ; ainsi les protons qui se sont déplacés entre l'application des deux gradients subissent un déphasage non nul proportionnel à leur déplacement le long de l'axe des gradients. Il faut un imageur IRM assez puissant pour produire des déphasages mesurables car les protons ne se déplacent que sur de petites distances (un dizaine de microns). La diffusion des protons relative à

l'application des gradients provoque un déplacement suivant une certaine distribution de probabilité, d'autant plus large que la diffusion est importante, et la phase de précession suit cette même loi de probabilité. Le déphasage des protons conduit à une perte de cohérence du signal de résonance magnétique qui est d'autant plus atténué que la diffusion a été importante dans l'axe des gradients. On obtient ainsi des images pondérées en diffusion (DWI, diffusion weighted images) à partir desquelles on peut reconstruire un tenseur de diffusion en chaque point de ces images, la direction des gradients de diffusion étant différente dans toutes ces images. Calculer un tenseur revient à calculer les coefficients représentant les différents gradients via un système d'équations linéaires. On obtient ainsi des images du tenseur de diffusion (DTI, diffusion tensor images).

La tractographie est l'une des applications de l'imagerie de diffusion. Via des algorithmes de suivi de fibres, il est possible de reconstruire le trajet d'un faisceau de fibres nerveuses sur toute sa longueur, l'architecture des axones organisée en faisceaux parallèles facilitant la diffusion des molécules d'eau le long de leur axe. Via une approche probabiliste, on considère que la direction des fibres correspond à la direction de plus grande diffusivité, indiquée par le vecteur qui est associé à la plus grande valeur propre du tenseur de diffusion (Le Bihan, 2001; 2003 ; Mori, 2007).

II.1.*2*) Buts de l'expérience et hypothèses

Comme nous l'avons rappelé dans l'introduction, un nombre important de travaux suggèrent l'implication d'aires communes au traitement sémantique des stimuli verbaux et non-verbaux ; il y a cependant un manque certain de travaux cherchant à comparer les aires cérébrales impliquées dans la compréhension d'actions complexes qu'elles soient décrites dans une phrase ou présentées visuellement via une image. Le but de cette première expérience était donc d'étudier les corrélats neuroanatomiques sous-tendant un réseau cérébral commun impliqué dans la compréhension de stimuli plus complexes : des phrases et des images représentant des événements. Grâce à un protocole combinant IRMf et DTI, nous avons recherché l'existence d'un tel système, en examinant les activations et la connectivité cérébrales chez des sujets en train de lire des phrases ou de voir des images.

Nous avons construit notre protocole dans le but de comparer directement l'activité cérébrale associée aux traitements des images et des phrases par conjonction. En ce qui concerne l'analyse structurale par DTI, il avait été prévu que nous utiliserions les régions activées en IRMf, à la fois par les images et les phrases, comme sites d'intérêt pour étudier la connectivité anatomique. Pour mieux comprendre les implications fonctionnelles au sein de ce réseau commun, nous avions aussi prévu de compléter nos analyses IRMf par une étude de la connectivité fonctionnelle de ce réseau, basée sur une approche d'analyse en composantes indépendantes (ICA) grâce à une toolbox NetbrainWork

(Perlbarg et al., 2007, 2009 ; Perlbarg et Marrelec, 2008) qui permet d'observer des corrélations entre les décours temporels de différentes régions cérébrales au sein d'un réseau.

Notre première hypothèse est que la compréhension d'une scène d'action devrait activer les mêmes aires cérébrales (observables en IRMf), notamment des aires associatives fronto-pariétales mais aussi des aires motrices, quelle que soit la nature des stimuli, phrases ou images.

Notre seconde hypothèse concerne la connectivité anatomique et fonctionnelle au sein de ce réseau. Si un réseau sémantique commun existe indépendamment de la modalité, il doit être soutenu par une structure impliquant des faisceaux spécifiques de fibres de matière blanche (étudiable par DTI), faisceaux qui n'ont, jusque-là, pas été étudiés dans le cadre d'un réseau sémantique commun. De plus, l'activité de certaines aires cérébrales doit être au moins en partie corrélée d'une région à une autre pour expliquer le fonctionnement en réseau et l'émergence d'une représentation à partir des interactions entre différentes régions cérébrales (ICA - NetBrainWork).

Enfin, de façon un peu secondaire par rapport à notre objectif principal, notre troisième hypothèse concerne la notion de dynamisme et notre protocole s'inspire d'une expérience en EEG de Proverbio et al. (Proverbio, Riva & Zani, 2009) où les auteurs montraient que le niveau de dynamisme d'une action pouvait engager des aires cérébrales différentes. Dans ce protocole, 260 images d'actions avec un degré de dynamisme plus ou moins important étaient présentées aux sujets : des images très dynamiques (condition Dynamique : jouer au foot, courir) vs des images plus statiques (condition Statique : faire la vaisselle, jouer aux échecs). Ces auteurs observent une activité électrique significativement différente entre ces deux conditions ; mais les auteurs font aussi référence à une série de régions visuelles et prémotrices dont l'activité est plus fortement associée à la condition Dynamique, incluant l'aire MT/V5 droite et le gyrus fusiforme dans le cortex temporal inférieur (BA 19), l'aire du mouvement du corps (EBA extrastriate body aera) localisée dans le cortex occipitotemporal latéral (BA37), le STS gauche (BA38), les aires motrices (BA4) et prémotrices (BA6) gauches, le cortex cingulaire antérieur et postérieur (BA24/BA31) et le gyrus frontal inférieur et moyen (BA10/47). Ces résultats suggèrent que les actions contenant un mouvement implicite dynamique sont traitées différemment et activent plus les régions motrices, notamment le système des neurones miroirs. Sur la base de cette étude, nous avons aussi intégré une composante de dynamisme à nos stimuli verbaux et non-verbaux (en utilisant des actions dynamiques et des actions plus statiques). Si les observations de Proverbio et coll. (2009) sont confirmées en IRMf, nous devrions observer un réseau fronto-temporo-pariétal plus important pour les actions les plus dynamiques ou des activations différentielles (plus forte intensité, régions plus étendues, aires cérébrales supplémentaires).

II.1.3) Matériels et Méthodes

- **Sujets**

Dix-neuf sujets sains, droitiers, de langue maternelle française et ayant une vision normale ou corrigée ont participé à cette étude (11 hommes, âge moyen : 25,5 ans ; SD : 4,11). Aucun participant n'avait d'antécédents de troubles médicaux psychiatriques ou neurologiques. Tous les participants ont donné leur consentement éclairé avant l'examen IRMf et l'étude était approuvée par le comité d'éthique local.

- **Stimuli**

a) Création des Stimuli

Stimuli Images

Pour ce protocole, 200 images ont été sélectionnées (sur une banque d'images Getty images ©, comme mentionné dans le préambule) ; elles représentent toujours une personne seule en train de réaliser une activité impliquant le système moteur de façon plus ou moins dynamique. Ainsi ont été sélectionnées :

- 100 actions « dynamiques » : courir dans le jardin, marcher sur la plage, monter à l'échelle…
- 100 actions « statiques » : écouter de la musique sur son lit, regarder un tableau…

Nous avons sélectionné des images les plus neutres possible, sans valence émotionnelle forte, tout en en limitant les informations visuelles qui pourraient perturber l'action représentée (trop d'objets, autres personnages…).

Stimuli Phrases

A partir des images, 200 phrases descriptives ont été produites, respectant toujours la forme sujet, verbe, complément(s) (de lieu, d'objet…). Ont ainsi été créées :

- 100 phrases « dynamiques » : *« La petite fille court dans le jardin. », « L'homme grimpe à l'échelle. »*…
- 100 phrases « statiques » : *« La jeune femme hésite entre deux tee-shirts. », « Le petit garçon dort dans son lit. » …*

Les phrases ont été équilibrées selon un certain nombre de critères afin d'éviter tout influence de marqueurs linguistiques et grammaticaux : nombre de mots et de lettres, construction grammaticale

(type de verbes : transitif, intransitif, réflexif ; nombre de prépositions, d'adverbes et d'adjectifs), fréquence moyenne des phrases (basées sur les fréquence Livres et fréquence Films de la base Lexique, voir préambule) ; mais aussi en fonction d'autres critères expérimentaux tels que le sexe et l'âge du protagoniste décrit dans la phrase (voir Annexe 1).

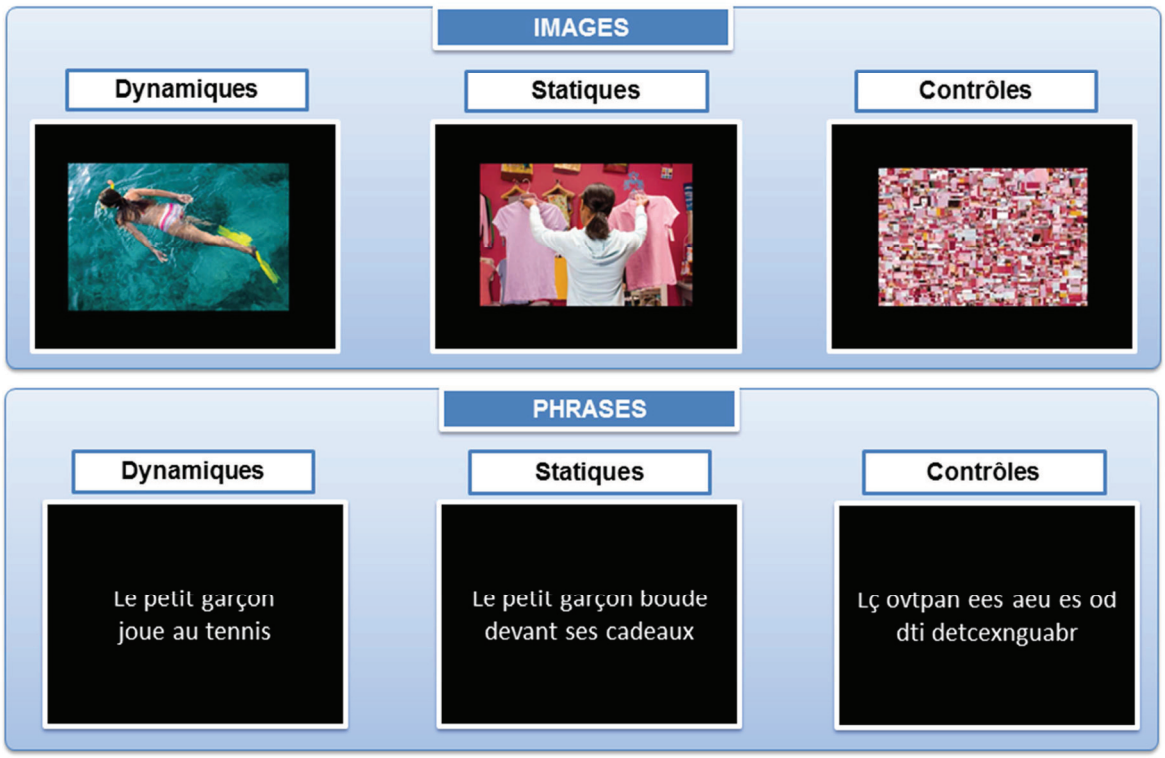


Fig 20) Exemple stimuli expérimentaux pour le protocole en IRMf

Les stimuli images ont été sélectionnés et répartis selon leur degré de dynamisme (actions plus statiques vs actions plus dynamiques). Les stimuli contrôles images sont des « scrambles » réalisés à partir des images d'origine afin qu'ils gardent certains propriétés (couleur, complexité) mais que les images d'origine ne soient plus reconnaissables (pas de traitement sémantique possible). Les stimuli phrases sont inspirés des images et les stimuli contrôles images sont aussi des « scrambles », constitués à partir des lettres issues de la phrase initiale.

Stimuli Contrôles

Pour soustraire les traitements cérébraux non sémantiques de bas niveau des analyses de contrastes ultérieures (essentiellement le traitement visuel et la lecture), nous avons choisi d'utiliser des stimuli contrôles pour les images et pour les phrases. Les deux types de stimuli contrôles ont été générés à partir des stimuli originaux.

- Les stimuli Contrôles Images correspondent à des « scrambles », images mixées à partir des images d'origine, ils gardent ainsi certaines propriétés des images de départ (couleurs, fréquences…).

- Les Contrôles Phrases correspondent également à des « scrambles » de lettres, dérivés des phrases d'origine : ils contiennent les mêmes lettres que les phrases d'origine et ont été séparés par des espaces pour les faire ressembler à des pseudo-mots et correspondre aux stimuli phrases en terme de nombre de « mots », nombre de lettres total, répartition voyelles/consonnes…

b) Création de 2 jeux de stimuli

L'ordre de présentation des sessions et des stimuli était évidemment randomisé pour chaque sujet, mais nous voulions éviter tout effet de « déjà vu » c'est-à-dire qu'une phrase puisse faire ressurgir l'image correspondante et inversement. Cette impression de « déjà vu » fausserait le protocole, en risquant d'impliquer trop fortement la mémoire dans une tâche de compréhension.

Nous avons créé deux versions du protocole (jeu A et jeu B). Chaque jeu comprend :

- 100 phrases : 50 phrases « dynamiques » et 50 phrases « statiques »
- 100 images : 50 « dynamiques » et 50 phrases « statiques »
- 50 contrôles phrases
- 50 contrôles images

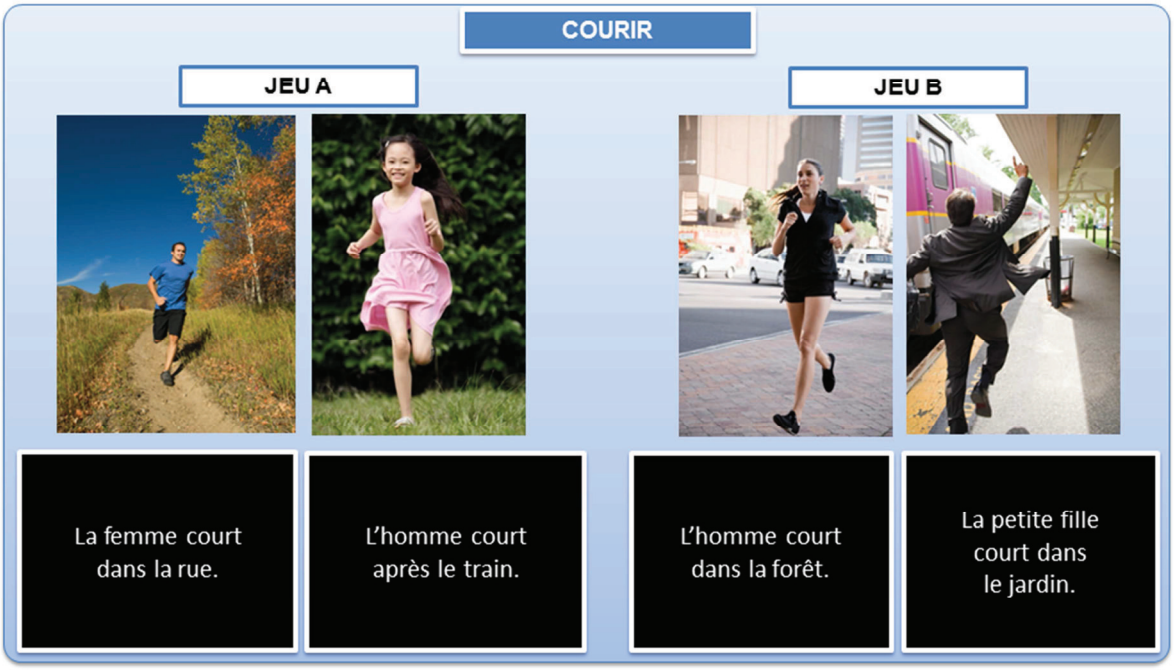


Fig 21) Création des 2 jeux expérimentaux

Dans un jeu expérimental (jeu A ou jeu B), une même action est représentée 4 fois (2 images et 2 phrases). Par exemple, pour le verbe COURIR, le jeu A comprend les deux images ci-dessus à gauche, et les phrases « *La femme court dans la rue*. » et « *L'homme court après le train*. ». Il faut noter que les phrases utilisées dans le jeu A correspondent aux images du jeu B et inversement. Ainsi, un même sujet ne verra pas les phrases correspondantes aux images, ce qui évite un effet de rappel entre phrases et images.

Comme l'illustre la figure 21), les 100 phrases du jeu A sont issues des 100 images du jeu B et inversement, les phrases du jeu B correspondent aux 100 images du jeu A. Les stimuli contrôles étant abstraits, nous avons utilisé les mêmes stimuli pour les deux jeux. Un sujet ne voit qu'une version du protocole, le jeu A ou le jeu B.

Nous avons volontairement limité le nombre d'actions utilisées, une même action est représentée 8 fois à travers les 2 jeux de données (cf exemple « courir » dans la figure 21). Toujours dans l'optique de limiter l'impression de « déjà-vu », nous avons choisi des personnages et des contextes les plus différents possible pour une même action (courir dans le jardin/courir après un train).

En raison de l'utilisation de 2 jeux de données, nous avons effectué des pré-tests de validation de nos stimuli (images et phrases) chez 20 sujets pilotes. Les résultats de ces pré-tests nous ont permis de répartir les stimuli entre les deux jeux de données (ces pré-tests sont décrits dans le paragraphe ci-après). Aucun sujet ayant participé à ces pré-tests n'a rapporté l'impression d'avoir vu des stimuli trop similaires entre phrases et images.

c) Validation et répartition des stimuli

Validation des Stimuli Images

Concernant les images, il a été demandé aux sujets pilotes de produire une phrase décrivant la situation afin de vérifier la bonne compréhension de nos stimuli selon un critère d' « Action agreement », s'inspirant du « name agreement d'Alario (cf préambule). Nous avons obtenu 96% de descriptions correctes.

Validation des Stimuli Phrases

Pour les phrases, il a été demandé aux sujets de donner un score de « représentabilité » c'est-à-dire que les sujets devaient noter, sur une échelle de 1 à 5, leur capacité à se représenter plus ou moins facilement une phrase (1=très difficile, 2 =difficile, 3 = ni difficile ni facile, 4 = facile, 5= très facile). Nous avons appelé ce critère « coefficient de représentabilité » (ou COR, nous reparlerons de ce critère ultérieurement). Ces tests nous ont confortés dans le choix de nos stimuli : nous avons obtenu 99% de bonnes représentations, avec une moyenne de 4,63 sur l'ensemble des phrases pour le COR (le plus bas score pour une phrase donnée moyenné sur l'ensemble des sujets étant de 3,88).

Validation du caractère dynamique

Enfin, les sujets devaient classifier les images et les phrases en terme de dynamisme sur une échelle de 1 à 5 (1 = très statique, 2 = statique, 3 = ni statique, ni dynamique, 4 = dynamique, 5 = très dynamique). Là aussi plus de 95 % des stimuli ont été répartis correctement par nos sujets pilotes.

Suppression des stimuli ambigus

Qu'il s'agisse d'images ou de phrases, ont été supprimés tous les stimuli ambigus, c'est-à-dire engendrant une mauvaise compréhension, une mauvaise classification dynamique/statique ou encore les stimuli trop difficiles à se représenter pour les phrases. Ils ont été remplacés par d'autres stimuli dans la version définitive du protocole.

Répartition en 2 jeux de données équivalents

Sur la base des pré-tests, nous avons réparti les stimuli pour que le jeu A et le jeu B soient équivalents. Nous avons équilibré les jeux de façon à ce que les différentes modalités (Images et Phrases) mais aussi les conditions (Statiques vs Dynamiques) contiennent des données similaires concernant le sexe et l'âge des personnages (nombre équivalent d'adultes vs enfants, de femmes vs d'hommes). Nous avons contrôlé le mode d'orientation (vertical ou horizontal) des images et pour les phrases, nous les avons présentées de façon à ce qu'elles aient une organisation générale comparable à celle des images (texte réparti sur 4 ou 5 lignes = mode vertical ; texte sur 2 lignes = mode horizontal). De plus, les phrases ont été réparties en tenant compte des marqueurs linguistiques et grammaticaux décrits précédemment (nombre de verbes réflexifs/transitifs, propositions relatives…).

Les stimuli contrôles n'ont pas fait l'objet de pré-test particulier mais ils ont également été répartis de façon équilibrée entre les deux jeux en termes d'orientation verticale ou horizontale (que ce soit pour les Contrôles Images ou Contrôles Phrases).

- **Protocole expérimental**

a) Sessions expérimentales

Comme le montre la figure 22), chaque jeu de stimuli comprend 4 sessions expérimentales : 2 avec des images (I1 et I2) et 2 avec des phrases (P1 et P2), les stimuli phrases et images n'étant jamais mélangés. Une session Images comporte : 25 images dynamiques (Id) 25 images statiques (Is) et 25 contrôles (Ic). Une session Phrases comporte : 25 phrases dynamiques (Id) 25 phrases statiques (Is) et

25 contrôles (Ic). Rappelons que lors de la passation en IRMf, la moitié des sujets a vu le premier jeu (Jeu A), et l'autre moitié le deuxième (Jeu B).

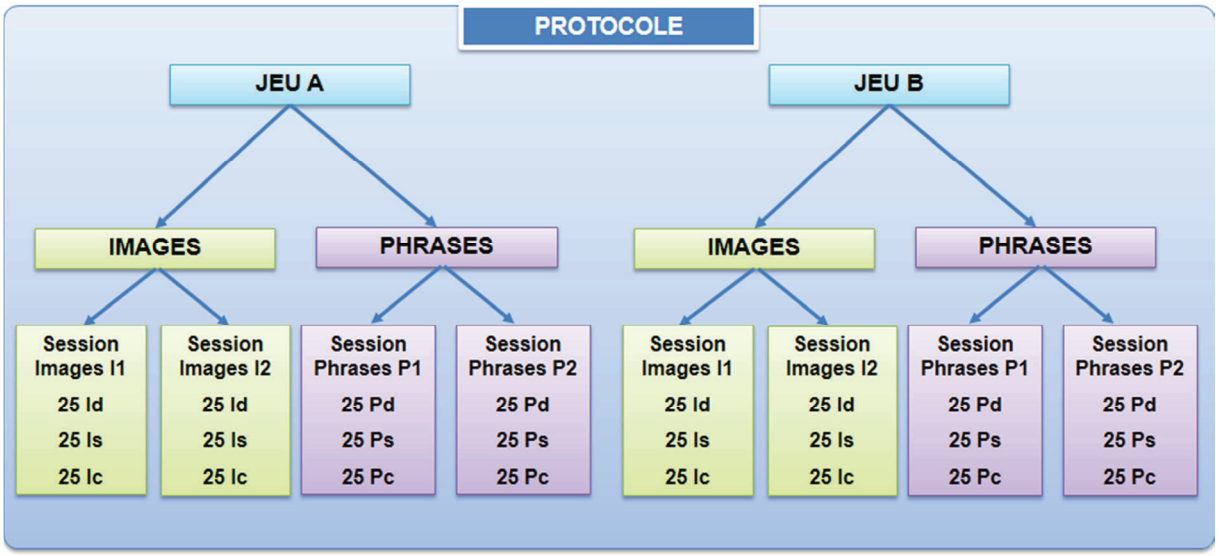


Fig 22) Protocole Expérimental IRMf

Les deux versions du protocole (Jeu A et Jeu B) comportent tous deux une condition Images et une condition Phrases réparties en deux sessions. Chaque condition, Images ou Phrases, contient 25 stimuli dynamiques, 25 stimuli statiques et 25 stimuli contrôles.

b) Procédure expérimentale

Pendant l'acquisition des images cérébrales, les 4 sessions expérimentales (I1, I2, P1, P2) étaient présentées dans un ordre aléatoire, chaque session durant environ 10 minutes. Au sein d'une session, l'ordre des stimuli était également randomisé. Nous avons utilisé le logiciel Optseq2 qui permet de générer des séquences de présentation optimisées afin d'obtenir de meilleurs contrastes durant les étapes d'analyses ultérieures. Une session expérimentale (phrase ou image) consiste en une alternance de stimuli des 3 conditions, répartis en 75 essais (25 dynamiques, 25 statiques et 25 contrôles).

Durant un essai, un point de fixation apparait tout d'abord au centre d'un écran noir pendant 500 ms ; il est suivi d'un stimulus phrase ou image (contrôles y compris) en fonction de la session expérimentale. Ce stimulus est présenté sur fond noir pendant 2 secondes. Un écran noir, avec point de fixation, lui succède ensuite (2 sec) et permet de marquer le passage à l'essai suivant, l'intervalle inter-stimuli (ITI) est variable en fonction d'un jitter pouvant aller de 0 à 6 secondes. Dans 10 % des essais, le stimulus (image, phrase ou contrôle) est suivi d'une question. Les sujets avaient pour consigne de regarder attentivement les stimuli pour pouvoir répondre aux questions (les sujets ne sont pas avertis à l'avance qu'un essai va contenir une question).

Fig 23) Illustration du protocole expérimental pour la condition IMAGES (en haut) et pour la condition PHRASES (en bas)
Déroulement d'un essai

Lors des essais avec question, il est demandé aux sujets s'ils ont déjà vu ce stimulus auparavant ; la question peut porter aussi bien sur les stimuli qui ont du sens que sur les stimuli contrôles. Cependant le but étant de maintenir constante l'attention des sujets, nous avons volontairement utilisé une question simple et dont la réponse est évidente si les sujets ont été attentifs. En effet, s'ils ont déjà rencontré le stimulus et que la réponse attendue est « oui », il s'agit du stimulus présenté à l'essai précédent (comme le montre l'exemple avec des stimuli Images, dans la figure 23). Si la réponse attendue est « non », le stimulus présenté dans la question est tout à fait nouveau et les sujets n'ont jamais pu le rencontrer au cours du protocole. Le nombre de bonnes réponses comportementales dépasse les 95%, ce qui indique que les sujets étaient attentifs.

c) *Dispositif expérimental*

Après avoir expliqué les consignes au sujet en dehors de l'IRM, celui-ci était installé dans la machine en vue de l'acquisition des images d'activation cérébrale par IRMf. Allongé le plus confortablement possible sur la table de la caméra IRM, le sujet tenait un boîtier de réponse dans chaque main (main gauche pour Non, main droite pour Oui). Un miroir fixé au dessus de sa tête permettait au sujet de regarder les stimuli visuels projetés à partir d'un projecteur situé à l'extrémitédu tunnel de l'IRM.

d) *Acquisition des données IRMf-DTI*

Les enregistrements IRMf et DTI ont été effectués avec un scanner 3T (Siemens Magnetom TrioTim) à NeuroSpin (CEA Saclay, Gif-sur-Yvette, France). Pour l'enregistrement IRMf, une séquence d'acquisition EPI (Echo Planar Imaging) en écho de gradient a été utilisée avec 38 gradients transversaux. Les coupes cérébrales, d'une épaisseur de 3mm, ont été acquises selon un mode entrelacé. Concernant les paramétres d'acquisition, le temps de répétition (TR) était de 2000 ms et le temps d'écho (TE) de 30ms ; l'angle de bascule de 76 degrés ; la largeur de bande de 2694 Hz par pixel ; le champ de vision de 192 mm2 ; la taille du voxel fonctionnel de 3.0 mm3 et il y a eu 300 répétitions.

Après les séquences d'acquisition IRMf, une séquence anatomique tridimensionnelle à haute résolution spatiale et pondérée en T1 a été acquise afin d'obtenir des images anatomiques du cerveau du sujet, pouvant être superposées aux images fonctionnelles lors de l'étape de coregistration spatiale (séquence T1 → TR/TE: 2300/2.98 ms; angle de bascule: 9 degrés; FOV: 256 mm2; épaisseur de coupe:1.10 mm, coupes sagittales).

A la fin d'une même session d'acquisition, 3 scans d'images de diffusion à 60 directions (19, 20 et 21) étaient enregistrées afin de pouvoir reconstruire les fibres de matière blanche (TR/TE: 13000/93 ms;

angle de bascule: 90 degrés; matrice d'acquisition: 128x128; FOV: 220 mm; épaisseur de coupe : 1.70 mm, coupes axiales, max. b-value: 1500 s/mm2).

e) *Post-test*

Après l'IRMf, un questionnaire a été proposé aux sujets ayant participé au protocole. Voulant vérifier la capacité à se représenter les scènes visuelles évoquées par des phrases chez les sujets qui venaient de passer l'expérience, nous leur avons présenté un questionnaire contenant 50 phrases issues de notre protocole. Parmi ces 50 phrases, 25 provenaient du jeu A et 25 provenaient du jeu B, ainsi les sujets avaient pu voir certaines de ces phrases mais il y en avait aussi qui leur étaient totalement inconnues. Il était demandé aux sujets de donner un score de « représentabilité » (équivalent à celui décrit pour la validation du matériel) en répondant à la question : arrivez-vous facilement à vous représenter/vous imaginer cette phrase ? Sur une échelle de 1 à 5, de 1= très difficile, 2=difficile, 3=intermédiaire, 4=facile, 5= très facile. Dans la question posée, il n'était pas précisé aux sujets quelle était la nature attendue de la représentation.

II.1.*4*) Analyses des données

Pour analyser les données acquises en IRM/DTI, différentes analyses ont été réalisées : les données IRMf ont été traitées grâce à des analyses d'activations (constrastes) et des analyses de connectivité fonctionnelle ; les données DTI ont été appréhendées au moyen d'analyses de tractographie pour déterminer la connectivité anatomique. Elles ont été complétées par des analyses portant sur les données comportementales issues du questionnaire présenté aux sujets durant le post-test. Avant de décrire les résultats associés à ces différentes méthodes, nous détaillerons les étapes de pré-traitements et d'analyses des données IRMf et DTI.

A) Analyses des données IRMf

L'analyse de données IRMf a été réalisée avec le logiciel Statistical Parametric Mapping (SPM, version SPM5), (Welcome Department of Imaging Neuroscience, London UK ; http://www.fil.ion.ucl.ac.uk/spm) (Friston et al., 2003 ; Friston, 2007) sous environnement Matlab (The Mathworks, Inc., Natick, MA, USA).

- **Etape de prétraitement spatial**

Cette étape de prétraitement permet de corriger au mieux les erreurs introduites par la technique IRM de recueil des données. Elle consiste en plusieurs opérations qui permettent de corriger les

images pour diminuer le bruit et les artéfacts dus au scanner (correction du décalage d'acquisition entre coupes, mouvements, orientation et distorsion spatiale). Elle est également essentielle pour permettre d'effectuer, dans un second temps, des analyses de groupe. En effet, les images doivent être recalées par rapport à un atlas de référence (Talairach), pour les rendre superposables, dans un même repère spatial.

Recalage Temporel ou « Slice Timing » :

La première opération est la correction du décalage temporel existant entre les coupes (Slice Timing) et d'une durée égale au temps de répétition (TR). Cette correction est nécessaire puisque lors d'une acquisition fonctionnelle les différentes coupes d'un même volume ne sont pas acquises simultanément mais successivement. La correction proposée par SPM consiste à « ramener », par interpolation temporelle, l'instant d'acquisition de toutes les coupes à un instant commun, qui est l'instant d'acquisition d'une des coupes du volume, qualifiée de coupe de référence. Dans notre protocole, la coupe de référence est la 19ème, les coupes ayant été acquises dans un mode entrelacé.

Réalignement :

La deuxième opération effectuée a été la correction du mouvement (Realign ou recalage) ; l'objectif de l'opération est de corriger les artefacts inévitables dus aux mouvements de la tête du sujet en réalignant la série de volumes fonctionnels sur le premier volume acquis (via des transformations affines rigides). Ce prétraitement permet d'éviter les « faux positifs » dans les cartes d'activation, souvent activés en périphérie du cerveau et qui correspondent à des voxels qui sont considérés comme activés alors qu'ils ne le sont pas en réalité.

Normalisation :

Nous avons ensuite effectué une opération de « coregistration » qui correspond à un recalage entre l'image anatomique et la moyenne des images fonctionnelles, puis les volumes fonctionnels ont été normalisés à l'aide d'une opération non-linéaire afin qu'ils puissent se confondre avec un volume de référence, le « template » (modèle anatomique MNI, Montreal Neurological Institute). Il existe en effet une grande variabilité anatomo-fonctionnelle interindividuelle. Ramener les cerveaux des sujets dans un cerveau moyen commun rend comparables les activations cérébrales observées dans la mesure où l'on peut alors déterminer les coordonnées des zones activées dans ce repère commun.

Lissage :

La dernière opération de prétraitement correspond à un lissage spatial (Smoothing), utilisé pour contourner les problèmes de corrélations spatiales (le signal acquis dans un voxel n'est pas indépendant du signal acquis dans les voxels voisins), notamment au moment de l'étape de seuillage

des cartes d'activation. Cette opération permet d'augmenter le rapport signal sur bruit. Un filtre istrotropic gaussien kernel, avec une banque de fréquence de 8 Hz a été utilisé pour cette étape.
C'est après le lissage spatial des volumes corrigés temporellement, réalignés et «normalisés» que l'on obtient une série d'images fonctionnelles, prête à être soumise à l'étape statistique de traitement.

- **Etape de traitement statistique**

Principe du General Linear Model (GLM)

Rappelons que l'IRMf permet d'observer les activations cérébrales induites par une tâche cognitive à travers les modifications hémodynamiques et l'estimation de l'effet BOLD. Mais il faut d'abord transcrire les données brutes afin de pouvoir répondre dans un second temps aux questions cognitives posées. Les données ont été analysées sur une base « voxel par voxel » utilisant les principes du modèle général linéaire (Friston et al., 1994) qui considère que le signal brut est une combinaison de plusieurs modèles expérimentaux.
Pour chaque voxel du cerveau, Y correspond au signal acquis au cours du temps (une fois que les données brutes ont subi les prétraitements), c'est à dire à la réponse hémodynamique mesurée dans ce voxel. On fait classiquement l'hypothèse que la réponse hémodynamique est linéaire. Le signal acquis dans un voxel alors que le sujet exécute un protocole peut-être représenté par un modèle Y, ou Y = condition A + condition B + condition C… On connaît la valeur du signal Y en q points, Y étant un vecteur de taille q (q désignant le nombre de scans ou volumes cérébraux acquis au total) ; spécifier le modèle consiste à définir la forme théorique de la réponse hémodynamique pour ces q points. La forme théorique est modélisée par des fonctions temporelles, appelées régresseurs qui sont eux aussi les vecteurs de taille q, regroupés en une matrice X composée de q lignes et d'un nombre de colonnes égal au nombre de régresseurs. Le modèle statistique, incluant les régresseurs, décompose le signal de la façon suivante : $Y = M + \beta1R1 + \beta2R2 + \beta3R3 + \beta4R4 + E$ où R correspond au régresseur pondéré par un coefficient de régression β qui représente en quelque sorte la contribution du régresseur associé à la variation du signal IRMf dans le voxel considéré, E correspondant à l'erreur résiduelle (partie des données non expliquées par le modèle, le bruit inévitable dû à des activations cérébrales parasites) et M désignant la valeur moyenne du signal, appelée régresseur constant ($\beta0R0$). Les coefficients β sont inconnus au départ, l'étape d'estimation des paramètres du modèle a pour objectif de déterminer leurs valeurs de telle façon que l'erreur résiduelle E soit minimale (variance minimale).

Application à notre protocole

Pour notre protocole, une matrice de régresseurs a été créée pour chaque participant, incluant les onsets des différents stimuli et des régresseurs séparés représentant les différentes conditions expérimentales : Images Statiques (Is), Images Dynamiques (Id), Images Contrôles (Ic), Phrases Statiques (Ps), Phrases Dynamiques (Pd), Phrases Contrôles (Pc) ; s'y ajoutent 6 régresseurs représentant les mouvements de la tête (rotations et translations). En utilisant le modèle linéaire général (GLM) (Friston et al., 1994), les réponses BOLD individuelles ont été modélisées par la fonction de réponse hémodynamique (HRF) et sa dérivée convoluée par une fonction delta. Les changements liés à la tâche dans le signal BOLD ont été estimés comme une combinaison linéaire des différents régresseurs représentant les différentes conditions de l'expérience. Deux modèles différents ont été implémentés : l'un pour tester nos hypothèses liées au facteur de Dynamisme et l'autre pour la comparaison des stimuli Images et Phrases. Dans le cadre de ces deux modèles, les analyses de 1er niveau (ou statistiques individuelles) ont permis de spécifier les différentes conditions expérimentales pour chaque participant et d'obtenir des contrastes permettant d'opposer ces conditions entre elles. Les mêmes contrastes ont été réalisés chez tous les sujets afin de pouvoir étudier ces effets sur l'ensemble du groupe grâce aux analyses de second niveau (contrastes T et analyse de conjonction). Ces analyses permettent d'obtenir des cartes paramétriques de l'activité du signal BOLD dans les voxels. La comparaison statistique, grâce aux constrates, de cartes paramétriques seuillées permet de créer des cartes d'activation et de visualiser l'activité associée à une condition expérimentale. Pour les différentes analyses (analyses de second niveau et conjonction), des cartes d'activations (t-map) ont été obtenues selon les paramètres spécifiés pour chaque analyse (cluster Ke et seuil de significativité p). Les coordonnées MNI, issues des clusters d'activations représentés sur ces cartes, ont été transformées en coordonnées Talairach (Talairach & Tournoux, 1988) grâce à une transformation utilisant la formule de Brett (http://www.brainmap.org/icbm2tal/). Ce sont principalement aux résultats de l'analyse de second niveau, sur le groupe, que nous nous intéresserons par la suite.

- **<u>Contrastes</u>**

<u>Contrastes Dynamiques/Statiques</u>

Analyse au 1er niveau

Dans un premier modèle, destiné à étudier spécifiquement les effets liés au degré de dynamisme des actions, les stimuli images et phrases ont été séparés selon leur dynamisme : les régresseurs utilisés étaient alors Pd (pour Phrases « dynamiques »), Ps (pour Phrases « statiques »), Id (pour Images Dynamiques) et Is (pour Images statiques) ; les contrastes destinés à souligner l'effet de dynamisme étant (Pd > Ps) et (Id > Is).

Analyse au 2ème niveau

Après avoir réalisé les contrastes (Pd > Ps et Id > Is) au niveau individuel, des analyses de groupe de second ordre ont été réalisées sur ces derniers (analyse de groupe à effets aléatoires, modèle « one-sample t-test ») pour étudier les activations communes aux 19 sujets. Les changements du BOLD liées à la tâche ont été extraites sur les résultats de groupe en utilisant un seuil FDR (false discovery rate) corrigé à p<.05 et un Ke=10, le Ke désignant le nombre de voxels minimal d'un cluster d'activation.

Contrastes Images/Phrases

Analyse au 1er niveau

Dans le second modèle, destiné à mettre en évidence l'activation cérébrale liée aux conditions images et phrases, les régresseurs utilisés étaient : Id (pour Images Dynamiques), Is (pour Images statiques), Ic (pour les « Contrôles images »), Pd (pour Phrases « dynamiques »), Ps (pour Phrases « statiques ») et Pc (pour les « Contrôles Phrases »). Cependant, le but étant d'extraire les changements du signal BOLD liés au traitement sémantique des phrases et des images, nous avons réalisé des contrastes entre les stimuli et leurs contrôles, sans prendre en considération la différence Statique/Dynamique, dans chaque modalité. Les contrastes étaient : [(Id + Is) – 2 Ic] pour les images et [(Pd + Ps) – 2 Pc] pour les phrases.

Analyse au 2ème niveau

Après avoir réalisé les contrastes Phrases vs Contrôles [(Pd + Ps) – 2 Pc] et Images vs Contrôles [(Id + Is) – 2Ic] au niveau individuel, des analyses de groupe de second ordre ont été effectuées sur ces derniers (analyse de groupe à effets aléatoires, modèle « one-sample t-test ») pour étudier les activations communes au 19 sujets. Les changements du BOLD liés à la tâche ont été extraits sur les résultats de groupe en utilisant un seuil FDR (false discovery rate) corrigé à p<.05 et un Ke=10, le Ke désignant le nombre de voxels minimal d'un cluster d'activation.

Analyse par conjonction

Afin de déterminer les structures cérébrales activées en commun durant le traitement des images et des phrases, nous avons choisi d'effectuer une analyse de conjonction (Nichols, 1999, 2005). Ce type d'analyse permet d'extraire des effets de groupe robustes car elle présente les avantages de l'analyse à effets aléatoires (elle peut refléter les aspects de l'anatomie fonctionnelle typique de la population sur laquelle le groupe a été échantillonné, Friston et al., 1999) ; mais surtout elle permet de réaliser la conjonction proprement dite, c'est à dire de définir quelles sont les zones communes activées chez tous les sujets ou entre différentes conditions d'intérêt. Les contrastes d'intérêt ont été définis individuellement puis une matrice commune tenant compte de tous ces contrastes a été

réalisée. Pour notre analyse de conjonction, nous avons utilisé les contrastes principaux de chaque sujet (Phrases vs Contrôles [(Pd + Ps) – 2 Pc] et Images vs Contrôles [(Id + Is) – 2 Ic]) définis durant l'étape d'analyse au 1er niveau. Nous avons choisi la technique de conjonction « Global Null Hypothesis » où l'on cherche à déterminer quels sont les voxels qui présentent des effets similaires à travers différents contrastes.

L'analyse de conjonction est représentée par la formule suivante : [(Pd + Ps) – 2 Pc] ∩ [(Id + Is) – 2 Ic], rejeter l'hypothèse nulle correspondant à conclure qu'un pixel est activé par toutes les conditions. Après identification des voxels individuels « activés » sur les cartes d'activations, nous n'avons retenu parmi eux que ceux qui formaient des groupes (clusters) composés d'un minimum de 10 voxels adjacents et dont l'intensité moyenne dépassait le seuil de significativité ($p<0.005$ non corrigé).

B) Analyses des données ICA

Pour compléter notre première série d'analyses sur les données IRMf, nous avons effectué une analyse de connectivité fonctionnelle sur nos données pré-traitées. L'outil NetBrainWork (Perlbarg et al., 2001 ; Perlbarg, 2008 ; Perlbarg, 2009) permet d'analyser des données IRMf en utilisant l'analyse en composantes indépendantes (ICA). Il apporte des informations essentielles sur la connectivité fonctionnelle via l'évaluation des corrélations entre les fluctuations du signal BOLD de différentes régions d'intérêt (ROIs) (Perlbarg et al., 2009 ; Perlbart, 2008 ; Coynel et al., 2010). L'utilisation de ces outils comporte deux étapes principales : l'identification automatique de larges réseaux cérébraux au niveau du groupe et le calcul des interactions fonctionnelles à l'intérieur ou entre ces réseaux de connectivité dynamique.

- **Identification des réseaux à l'échelle du groupe**

La première étape de l'analyse NetBrainWork correspond à une approche exploratoire (l'approche NEDICA pour Network Identification, Perlbarg, 2008).

Analyse en composantes spatiales

Les données sont tout d'abord traitées par l'intermédiaire d'une analyse en composante indépendante spatiale adaptée aux données IRMf (sICA, Perlbarg et al., 2008), elle conduit à l'identification automatique, indépendamment des onsets/stimuli, de réseaux anatomo-fonctionnels représentatifs d'une condition expérimentale au sein d'une population donnée de sujets. Dans notre cas, deux ICA spatiales ont été réalisées séparément : une pour les images (comprenant les 2 sessions Images) et une pour les phrases (comprenant les 2 sessions Phrases). Une fois l'ICA réalisée, les 40 composantes spatiales expliquant le mieux la variance sont extraites grâce à un algorithme d'ICA

infomax, puis normalisées par rapport aux z-scores et reprojetées dans un espace standard MNI grâce à des transformations spatiales non linéaires. Dans notre étude, un seuil de significativité de p=0.01 non corrigé a été appliqué ainsi qu'un seuil Ke (cluster extension) = 10 voxels.

Regroupement hiérarchique et détermination des classes d'intérêt (COIs)

Les 40 composantes spatiales sont ensuite analysées selon une étape de regroupement hiérarchique effectuée sur l'ensemble des sessions expérimentales pour une condition donnée (dans notre cas, session Image 1 + session Image 2 ou session Phrase 1 + session Phrase 2), ce qui permet d'obtenir des composantes indépendantes (ICs). Dans notre étude, pour une condition donnée, image ou phrase, 1520 composantes indépendantes ont ainsi été obtenues (19 sujets x 2 sessions par condition x 40 composantes spatiales/session = 1520 ICs). Les composantes indépendantes sont regroupées selon une approche basée sur la similarité spatiale aboutissant ainsi à un arbre de similarité (l'algorithme utilisé diminue la similarité intra-classe). Ensuite, les composantes similaires provenant de différentes sessions sont regroupées en un certain nombre de classes (classes d'intérêt COIs) qui sont les plus représentatives de la population à partir de l'arbre de similarité. Idéalement, chaque classe ne doit contenir qu'une seule IC provenant de chaque session. NetBrainWork propose une manière automatique de définir ces classes en utilisant un algorithme ad-hoc permettant d'optimiser deux critères pour chaque classe : le degré de représentativité (DR) et le degré d'unicité (DU). Pour une classe particulière, le DR correspond au nombre N_R de sessions distinctes qui contribuent à cette classe, divisé par le nombre total de sessions N_T (cependant, chaque session peut contribuer à une classe via plusieurs ICs) tandis que le DU est défini par le nombre de sessions qui contribuent à une classe à travers une (et une seulement) composante, divisé par N_R (Perlbarg et al. 2008). Une classe optimale serait donc caractérisée par un DR=1 et un DU=1, ce qui signifie qu'une classe optimale contient une et une seule composante issue de chaque session.

Dans NetBrainWork, une procédure de bootstrap très stricte est utilisée plusieurs fois pour vérifier la reproductibilité et la similarité de l'identification des classes en testant le niveau de confiance du regroupement des ICs et la robustesse de la détection des réseaux (Perlbarg et al., 2008).

Cartes d'activations des t-scores et sélection des régions d'intérêt (ROIs)

Pour chaque classe d'intérêt, une carte d'activations des t-scores (fixed-effect group map of t-scores) est ensuite calculée pour mettre en évidence les clusters de voxels ayant des z-scores hauts. Avec NetBrainWork, les pics de ces cartes d'activations sont automatiquement détectés et peuvent être utilisés pour définir des régions d'intérêt (ROI). Ces ROIs peuvent servir à définir un ou plusieurs

réseaux d'intérêts (NOIs) et elles serviront à analyser la connectivité fonctionnelle de ces réseaux dans les étapes d'analyses ultérieures.

- **Mesure de la connectivité fonctionnelle**

La deuxième grande étape d'analyse implémentée dans NetBrainWork, permet de calculer une mesure de l'interaction (connectivité fonctionnelle) entre les ROIs d'un ou plusieurs réseaux identifiés via les étapes précédentes. La connectivité fonctionnelle est généralement définie comme la corrélation entre le décours temporel des activités de différentes régions (Friston, 1994). L'outil NetBrainWork permet de réaliser deux types de corrélations : les corrélations globales représentent une mesure de l'interaction entre les décours temporels de deux régions, mais sans occulter l'influence possible d'autres régions, tandis que les corrélations partielles représentent la force de l'interaction directe spécifique entre deux régions, sans tenir compte de l'influence d'autres régions (Marrelec, 2008). La corrélation partielle exclut le hasard.

En ce qui concerne notre étude, à partir des étapes décrites précédemment, nous avons utilisé deux réseaux d'intérêt pour lesquels nous avons mesuré la connectivité fonctionnelle. Ces deux réseaux sont issus d'analyses sICA distinctes pour la condition Image et la condition Phrase. Chacun de ces réseaux est issu d'une sélection de ROIs provenant uniquement d'une et une seule classe d'intérêt, elle-même résultant de l'analyse sICA. Nous décrirons plus en détails les critères de sélection de ces réseaux dans la partie résultats.

C) Analyses des données DTI

Les différentes étapes de prétraitements et d'analyses ont été réalisées en utilisant conjointement les logiciels AFNI (Cox, 1996) et DTIQuery v1.1 (Sherbondy et al., 2005) selon une méthodologie de tractographie déterministe des fibres.

- **Prétraitements des données DTI**

Le prétraitement des données a été réalisé automatiquement avec le logiciel AFNI. Chaque image de diffusion pondérée a tout d'abord été alignée sur une image anatomique T1, image de laquelle le contour cranial a été retiré. A ensuite été effectuée une correction de l'orientation des gradients via des paramètres de mouvement angulaire (Leemans et Jones, 2009), suivi d'un calcul des tenseurs de diffusion. Parallèlement, l'image IRM T1 a été normalisée spatialement sur un template anatomique (Montreal Neurological Institute N27 brain) et la matrice de transformation affine à 12 paramètres, résultant de cette normalisation, a été sauvegardée. La résolution issue de l'espace natif a été

conservée pour les volumes de diffusion pondérés (1.72x1.72x1.70 mm) afin de pouvoir réaliser la tractographie des fibres sur l'ensemble du cerveau (pas de transformation en Talairach). Cela nous a permis de reconstruire l'ensemble de fibres pour tout le cerveau, ces fibres ont été sauvegardées dans une base de données, individuellement pour chaque sujet, afin de pouvoir les utiliser ultérieurement.

- **Tractographie déterministe des fibres dans l'espace propre de chaque sujet**

Une fois les prétraitement effectués, un tenseur dans l'espace propre de chaque sujet a ensuite été calculé grâce à DTIQuery v1.1 (Sherbondy et al., 2005), en utilisant un algorithme STT (streamline tracking algorithm) ainsi que différents paramètres optimisés (Basser et al., 2000; path step size = 1.0 mm, seed point spacing = 2.0 mm, fractional anisotropy termination threshold =0.15, angular termination threshold = 45 degrees, maximum pathway length = 300 mm, Euler's method for STT numerical integration). Puis La tractographie des faisceaux a été réalisée en 3 étapes, individuellement pour chaque sujet : 1) transformation des coordonnés des maximum locaux IRMf pour obtenir les coordonnées correspondantes dans l'espace natif propre de chaque sujet (AFNI) ; 2) utilisation de ces coordonnées pour placer des volumes d'intérêt (VOI) dans DTI query (radius = 5 ou 10 mm en fonction de l'anatomie et de la taille des régions d'intérêt fonctionnelles) ; 3) les faisceaux se situant à l'intersection avec un VOI ont été sauvegardés dans l'espace natif de chaque sujet en tant qu'image binaire (DTI query).

Nous avons utilisé les maximums locaux des clusters obtenus par l'analyse de conjonction ([(Pd + Ps) – 2 Pc] ∩ [(Id + Is) – 2 Ic]) comme volumes d'intérêt (voir la partie Résultats pour plus de détails). Bien que la transformation des coordonnées des maximums locaux soit une étape automatique implémentée dans AFNI, nous avons tenu compte de la plausibilité de la localisation anatomique de chaque ROI en prenant en compte l'anatomie propre de chaque sujet pour placer nos ROIs lors de l'étape suivante, qui utilise DTI query et qui permet une visualisation du cerveau et des fibres de chaque sujet.

- **Identification des faisceaux de fibres pour le groupe de sujets et analyses individuelles**

Ces étapes ont été réalisées avec le logiciel AFNI. Pour chaque sujet, les images binaires représentant l'espace natif et contenant l'ensemble des faisceaux de tractographie issus des différents VOI ont été transformées en données MNI via l'utilisation d'une matrice de transformation à 12 paramètres (issue de la procédure de normalisation spatiale de l'image MRI T1). Cette procédure a permis de générer une image « volume mask » pour chaque sujet (pour les faiceaux issus d'un VOI), utilisable dans un espace standard de coordonnées commun à tous les sujets. Les volumes des différents sujets

ont ensuite été sommés les uns aux autres pour créer une carte unique de volumes où la valeur entière de chaque voxel représente le nombre de sujets qui possèdent des faisceaux de tractographie passant par un voxel donné. Un seuil de significativité a été appliqué à ce volume afin de visualiser les voxels pour lesquels la tractographie met en évidence des connections communes chez au moins 10 sur les 19 sujets qui ont participé à cette étude. Si un faisceau n'est pas présent chez un individu, cela ne veut pas dire qu'il n'existe pas chez ce sujet ; nous mettons plutôt en cause les limitations inhérentes aux méthodes d'acquisition et d'analyses DTI (Mori, 2007). Etant donné la variabilité des fibres de matière blanche d'un individu à l'autre et les limitations techniques du DTI, la visualisation des faisceaux sur le groupe n'est pas toujours évidente : même si elle confirme que deux régions sont bien reliées, il faut parfois descendre à un seuil très bas pour pouvoir visualiser les faisceaux sur l'ensemble des sujets. C'est pourquoi, nous avons complété ces analyses de groupe par des analyses individuelles afin de mieux définir l'origine et le trajet de nos faisceaux d'intérêt et de vérifier qu'il s'agissait bien du même faisceau d'un sujet à l'autre. Pour ces analyses individuelles, nous nous sommes concentrés sur les faisceaux présents seulement chez 15 sur 19 de nos sujets.

II.1.*5*) Résultats

A) Résultats Partie I – résultats IRMf

Rappelons que nos différentes hypothèses concernaient les effets qui pouvaient être associés au degré de dynamisme des stimuli (images ou phrases) d'une part et d'autre part la mise en évidence d'un réseau sémantique commun impliqué dans la compréhension des phrases et des images.

- **Analyse de dynamicité**

Contrastes « Images dynamiques vs Images statiques » et « Phrases dynamiques vs Phrases statiques »

Nous avons choisi de présenter tout d'abord les effets liés au degré de dynamisme, effets qui devaient être mis en évidence par la comparaison de deux contrastes issus des analyses de second niveau : Images dynamiques vs Images statiques (Id>Is) et Phrases dynamiques vs Phrases statiques (Pd>Ps).

Contrairement à ce qui était attendu, nous ne retrouvons pas le réseau fronto-pariéto-temporal décrit par Proverbio et al. (2009). Comme le montre la figure 24, le contraste Id>Is (seuil FDR corrigé=0.05, Ke=15 voxels) génère une seule activation bilatérale située au niveau du cortex occipito-temporal latéral. Le maximum local du cluster situé dans l'hémisphère gauche se situe au

niveau du gyrus temporal moyen (BA39, en coordonnées de Talairach TAL -50 -72 12). Cette activité s'étend ventralement jusqu'en BA37 (TAL -48 -61 3) et médialement, jusqu'au niveau du gyrus occipital moyen (BA 19, TAL -39 -72 9). L'activité dans l'hémisphère droit est beaucoup plus limitée, elle est circonscrite au gyrus temporal moyen (BA37, TAL 50 -61 6).

Le contraste Pd>Ps (seuil FDR corrigé=0.05) ne permet pas de reproduire les résultats observés pour les images et ne fait émerger aucune différence entre les conditions Statiques et Dynamiques pour les phrases.

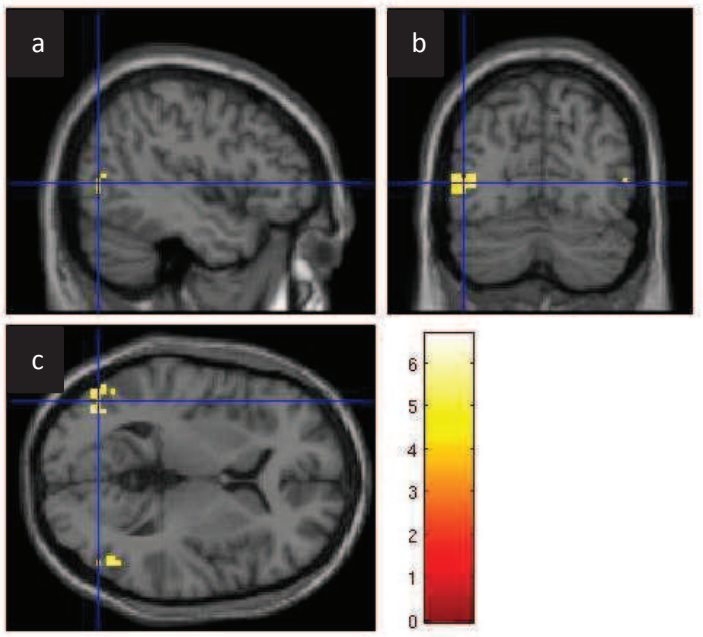


Fig 24) Activations cérébrales pour le contrastes Images dynamiques vs Images statiques (Id>Is)

Les résultats représentent les activations pour l'ensemble des sujets, projetées sur une image anatomique en coupes sagittale (a), frontale (b) et horizontale(c).

- **Analyses des aspects sémantiques**

Contrastes « Phrases vs Contrôles Phrases » ; « Images vs Contrôles Images » et Conjonction « Phrases/Images »

Rappelons qu'une analyse voxel par voxel a été réalisée sur l'ensemble des activations cérébrales, séparément pour le traitement des phrases (Phrases vs Contrôles Phrases) et des images (Images vs Contrôles Images), et ensuite conjointement sur les deux types de stimuli grâce à une analyse de conjonction (Nichols et al 2005) dans le but d'identifier un réseau commun impliqué dans la compréhension des images et des phrases.

- *Phrases vs Contrôles Phrases (Pd + Ps) – 2 Pc*

Nous avons identifié un réseau distribué entre les cortex frontal et temporal, montrant une activité plus importante pour les phrases que pour les contrôles composés de pseudo-mots, à un seuil FDR à p corrigé < 0.05 (Figure 25 et table 1, les coordonnées sont en Talairach). Le contraste des activations entre Phrases et Contrôles Phrases (Pd + Ps) – 2 Pc génère un important cluster de voxels activés bilatéralement dans les gyri temporaux moyen et supérieur, qui s'étend médialement au niveau du gyrus parahippocampique (BA29/36) et du cortex rétrosplénial (BA29, BA30), mais aussi dorsalement au niveau du gyrus angulaire (BA 39). Dans chaque hémisphère, ce cluster est très étendu et contient deux principaux pics d'activation, localisés en BA21 et BA39 ; dans l'hémisphère gauche, cette activation temporale se prolonge antérieurement, au niveau du gyrus temporal supérieur (BA38).

Aire anatomique	BA	TAL x, y, z	t-stat	Ke
(Pd + Ps) – 2 Pc		P < 0.05 (FDR)		
Gyrus Temporal Moyen HG	21	-53 -9 -10	7.60	3442
Gyrus Temporal Moyen HG	39	-42 -66 23	6.64	
Gyrus Parahippocampique HG	36	-33 -38 -8	6.63	
Gyrus Temporal Moyen HD	39	53 -69 26	3.89	132
Gyrus Temporal Supérieur HD	22	45 -46 11	3.34	
Gyrus Temporal Moyen HD	21	33 -40 8	2.89	
Gyrus Temporal Moyen HD	21	53 -7 -17	4.39	82
Gyrus Temporal Moyen HD	21	50 -15 -7	3.03	
Gyrus Temporal Supérieur HG	38	-50 14 -18	4.79	45
Gyrus Frontal Moyen HG	6	-39 3 52	4.05	43
Gyrus Frontal Inférieur HG	46	-56 29 7	3.41	21
Gyrus Frontal Inférieur HG	45	-56 27 18	3.23	
Gyrus Frontal Médian HG	11	-9 52 -13	3.16	11
Gyrus Frontal Médian HD	11	3 40 -17	3.13	12
Noyau caudé HD		6 15 8	3.93	131
Noyau caudé HD		9 9 -3	3.56	

Table 1) Régions anatomiques présentant une activation significative lorsque les phrases sont contrastées avec les stimuli contrôles phrases (Pd + Ps) – 2 Pc

Au niveau frontal, un petit site d'activation englobe la partie du lobe frontal inférieur gauche (BA 46-45) appartenant au réseau périsylvien classiquement impliqué dans la compréhension des phrases. On trouve également un cluster d'activation bilatéral, plus restreint, au niveau du cortex orbitofrontal (BA11). Ce contraste révèle aussi une activation d'une région du cortex prémoteur, au niveau du gyrus frontal moyen (BA6), formant un petit cluster (Ke=43) dont le maximum local se situe à -39, 3, 52 (coordonnées de Talairach). Au niveau subcortical, on relève une implication du noyau caudé, à travers un cluster d'activation assez important (Ke=131) qui s'étend ventralement et dorsalement.

- *Images vs Contrôles Images (Id + Is) – 2 Ic*

Le contraste (Id + Is) – 2 Ic, opposant les images et leurs contrôles, révèle un large réseau occipito-préfrontal (FDR, p corrigé < 0.05) qui recoupe partiellement les régions impliquées dans la compréhension des phrases (Figure 25 et table 2, les coordonnées sont en Talairach).

Aire anatomique	BA	TAL x, y, z	t-stat	Ke
(Id + Is) – 2 Ic P < 0.05 (FDR)				
Gyrus Temporal Inférieur HG	19	-42 -76 -1	20.86	11670
Gyrus Occipital Moyen HD	39/19	50 -72 9	18.38	
Gyrus Temporal Moyen HD	39	50 -69 17	16.55	
Noyau caudé HG		-2 13 -5	2.37	
Noyau caudé HD		3 13 -5	2.91	
Gyrus Frontal Médian HD	11	0 40 -17	6.08	1050
Gyrus Frontal Médian HG	10	-6 62 19	4.91	
Gyrus Frontal Médian HD	10	9 61 5	4.29	
Gyrus Frontal Inférieur HG	45	39 19 21	2.73	12
Gyrus Précentral HG	4	-42 -6 53	3.46	77
Gyrus Précentral HG	6/4	-50 -10 39	2.55	
Gyrus cingulaire HG	24	-3 -4 39	3.21	70
Gyrus cingulaire HG	31	-12 -27 40	2.79	

Table 2) Régions anatomiques présentant une activation significative lorsque les images sont contrastées avec les stimuli contrôles images (Id + Is) – 2 Ic

On trouve une activité significative importante et bilatérale (Ke=11670) : les maximums locaux sont principalement localisés dans le cortex visuel et associatif de l'hémisphère droit, incluant le gyrus fusiforme et les gyri occipitaux moyens (BA19, BA37, BA39) ; mais l'activation de cet important cluster s'étend bilatéralement et antérieurement au niveau du précuneus, du cortex temporo-pariétal (gyrus angulaire BA39 et BA40), des régions para-hippocampiques (BA35, BA36) et du cortex rétrosplénial (BA29,BA30). Une activation plus limitée (Ke= 70) se situe au niveau du cortex cingulaire antérieur et postérieur (BA 24, BA 31). L'activité liée aux images est aussi significativement plus importante au niveau du cortex préfrontal, formant trois régions distinctes d'activation situées dans le gyrus préfrontal moyen (BA6), le gyrus préfrontal inférieur (BA45-46) et bilatéralement dans le cortex orbitofrontal (BA11). On retrouve également un petit locus d'activation sous-cortical situé au niveau du noyau caudé.

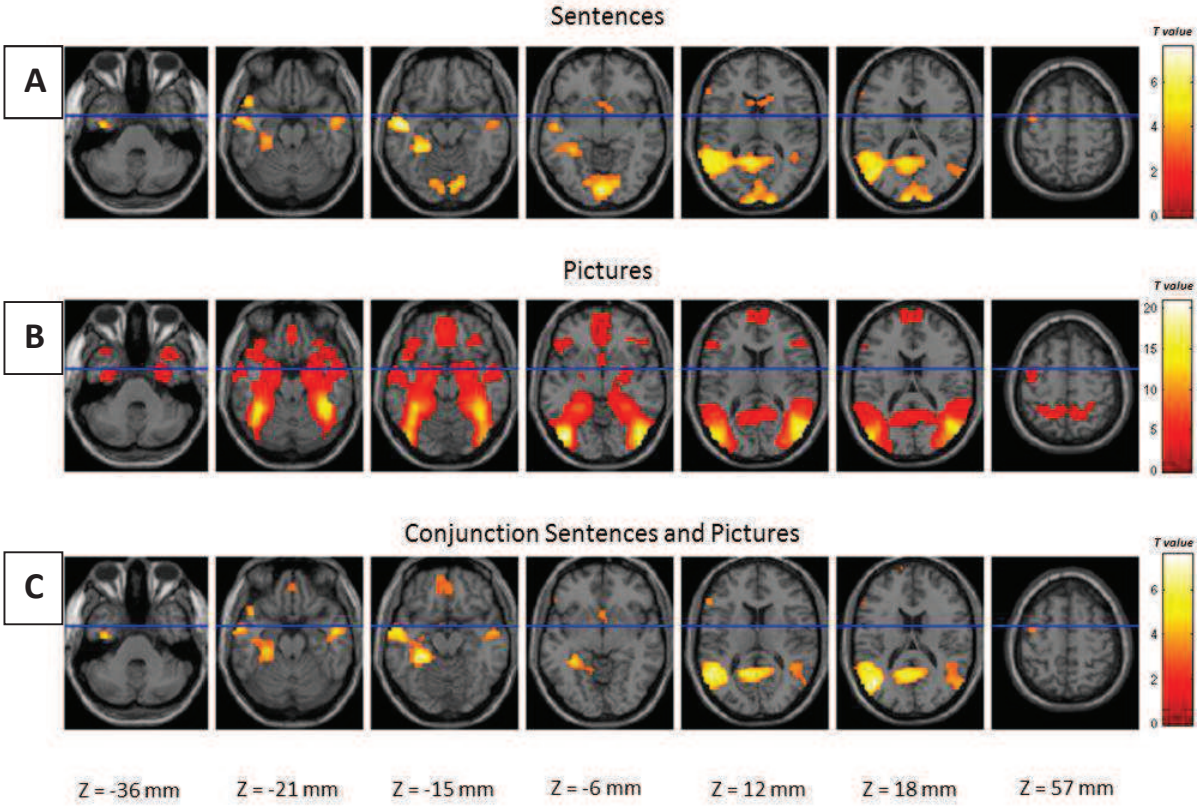


Fig 25) Comparaison des contrastes visant à mettre en évidence le réseau sémantique

A : Phrases dynamiques vs Phrases statiques (Pd>Ps) ; B : Images dynamiques vs Images statiques (Id>Is) ; C : Conjonction Phrases et Images (P>Pc) ∩ (I>Ic). Les résultats représentent les activations pour l'ensemble des sujets, projetées sur une image anatomique à différents niveaux de coupe horizontale.

L'analyse de conjonction regroupant les contrastes concernant les phrases et les images [(Pd + Ps) – 2 Pc] ∩ [(Id + Is) – 2 Ic] révèle un large réseau distribué au niveau fronto-temporo-pariétal à un seuil non corrigé p <.005 (figures 25 et 26 ; Table 3). Une première zone d'activations communes aux images et aux phrases se situe au niveau du cortex frontal gauche, formant trois clusters

d'activation : dans le cortex orbito-frontal (BA11) avec une activation bilatérale du cortex orbital médian (Ke= 27) qui s'étend rostralement et dorsalement au niveau du gyrus dorsal préfrontal gauche (Ke = 19), et une activation plus latérale (Ke = 21) se situant quant à elle au niveau du gyrus frontal inférieur gauche et impliquant les aires de Brodmann BA 46 et BA 45 jusqu'à la limite avec BA 47. On retrouve cette activation au niveau du Gyrus frontal inférieur pour l'HD, mais elle est restreinte à l'aire BA47 de ce côté (TAL 32 37 -9, Ke=15). On trouve également une activation commune pour les images et les phrases situées au niveau du cortex prémoteur, avec un maximum local d'activation situé dans le gyrus préfrontal moyen (BA6) en -42, 0, 53 (coordonnées de Talairach TAL).

Dans les régions postérieures, au niveau des lobes temporaux et pariétaux, la conjonction met en évidence de larges zones d'activation (Ke>100). On trouve notamment un cluster important (Ke= 1019) qui peut être subdivisé en deux sous-clusters : le premier se situe au niveau de la région hippocampique gauche (TAL -33,-39,-12, BA35-36) et inclut, ventralement, le gyrus fusiforme (BA37) ; le deuxième cluster qui est plus étendu et plus médial, implique les deux hémisphères, les pics principaux étant situés dans le cortex retrosplenial (BA29, BA30) au niveau de TAL -9,-55,17 à gauche et à TAL 12,-51,15 à droite. Cette activation retrospléniale bilatérale est très étendue, puisqu'elle recoupe l'hippocampe (BA29), le gyrus cingulaire (BA 23,30) dorsalement, voire même le précuneus (BA19) pour la partie la plus haute de cette activation.

Au niveau du lobe temporal, un pattern d'activation bilatéral et symétrique est concentré au niveau du gyrus temporal moyen dans l'aire BA39 et l'activité s'étend antérieurement et bilatéralement jusque dans l'aire BA 21, avec un maximum d'activation à TAL -54,-6,-15. On distingue aussi un cluster plus petit (Ke=19) au niveau du gyrus temporal supérieur (BA38, TAL -45, 18, -24), localisé plus antérieurement et seulement pour l'hémisphère gauche. Ainsi, les principales activations communes aux phrases et aux images se situent au niveau de la jonction temporo-pariétale (TPJ) avec une activation étendue, plus marquée dans l'hémisphère gauche (Ke=651) que dans l'hémisphère droit (Ke=76). Bien que le pic d'activation se situe au niveau du gyrus temporal moyen (BA39) pour les deux hémisphères, le cluster d'activation a une taille vraiment importante et s'étend dorsalement jusqu'à la partie postéro-ventral du gyrus angulaire adjacent et du lobule pariétal inférieur.

La conjonction [(Pd + Ps) – 2 Pc] ∩ [(Id + Is) – 2 Ic] met aussi en évidence un petit locus d'activation dans la partie droite médioventrale du noyau caudé (TAL 3, 12, -6), l'activation de ce cluster s'étend jusqu'au noyau caudé de l'HG.

Aire anatomique	BA	TAL x, y, z	t-stat	Ke
[(Pd + Ps) – 2 Pc] ∩ [(Id + Is) – 2 Ic] P < 0.005 (unc.)				
Gyrus Temporal Moyen HG	39	-42 -66 23	6.64	651
Gyrus Temporal Supérieur HG	22	-53 -52 14	5.72	
Gyrus Parahippocampique HG	35/36	-33 -38 -8	6.63	1019
Gyrus cingulaire postérieur (RSC) HG	30/23	-9 -55 17	5.52	
Gyrus cingulaire postérieur (RSC) HG	29/30	-15 -49 8	5.06	
Gyrus Temporal Moyen HD	21	53 -7 -17	4.39	76
Gyrus Temporal Moyen HD	39	53 -69 26	3.89	166
Gyrus Temporal Moyen/Supérieur HD	22	45 -46 11	3.34	
Gyrus Temporal Supérieur HG	38	-45 16 -21	3.84	19
Gyrus Frontal Inférieur HG	45/46	-56 29 7	3.41	21
Gyrus Frontal Médian HG	11	-9 52 -13	3.16	19
Gyrus Frontal Médian HD	11	3 40 -17	3.13	27
Gyrus Frontal Moyen HG	6	-42 0 53	3.00	10
Noyau caudé HD		3 11 -6	3.04	13

Table 3) Régions anatomiques présentant une activation significative pour la conjonction (P>Cp ∩ I>Ci)

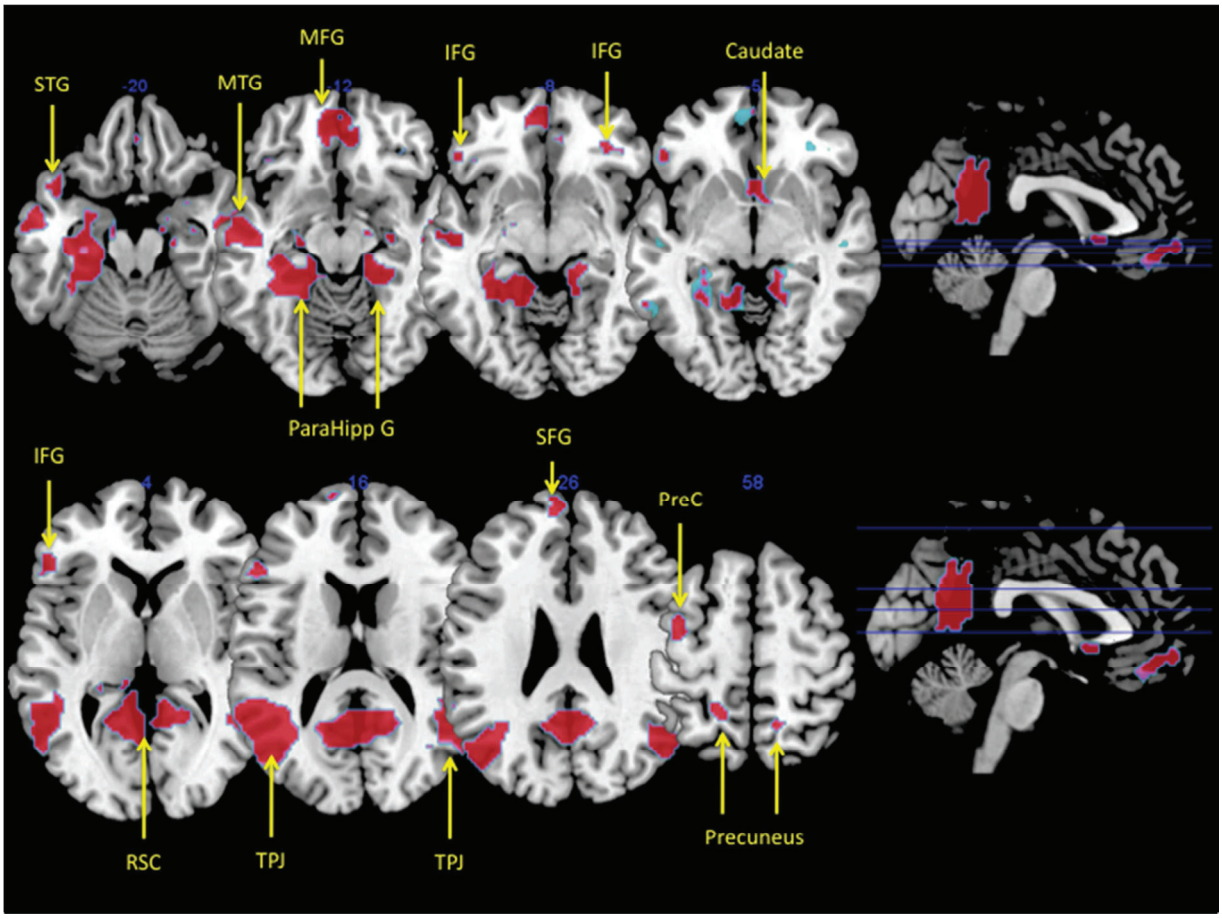


Fig 26) Pattern d'activations cérébrales pour la conjonction entre les conditions Phrases et Images ($p < 0.005$ uncorrected). Cette analyse met en évidence une vaste activation fronto-temporo-pariétales, impliquant le complexe rétrosplenial (RSC), le gyrus frontal inférieur (IFG), la jonction pariéto-temporale bilatéralement (TPJ), le gyrus frontal supérieur (SFG), le gyrus précentral (PreC), le gyrus frontal supérieur (STG), le gyrus frontal moyen (MTG) et le gyrus frontal médian (MFG).

- **Synthèse des résultats IRMf**

Contrairement à Proverbio et al. (2009), nos analyses ne nous ont pas permis de retrouver un réseau fronto-temporo pariétal qui serait différemment engagé dans le traitement des stimuli dynamiques et statiques. Par contre, pour les images dynamiques, nous notons quelques activations très spécifiques du cortex occipito-temporal dont les coordonnées (maximum local pour l'HG -50 -72 12 et pour l'HD -50 -61 6) sont très proches de celles rapportées dans la littérature pour désigner l'aire MT/V5, notamment ceux de Kourtzi et al. (2001, TAL 42.9 ± 5.5, -72 ± 4.7, 4.4 + 4.3 pour l'HG et -45.8 ± 6.3, -67.4 ± 7.1, 0.2 ± 6.9 pour l'HD). Nos résultats sont donc cohérents avec les travaux qui montrent que des images fixes d'actions entraînent une activation des aires qui traitent le mouvement. En revanche, nous n'avons pas reproduit ce résultat avec les phrases, ce qui est plutôt surprenant puisque différents travaux ont mis en évidence des activations proches de l'aire MT avec des stimuli verbaux (Kable et al. 2009 ; Tettamanti et al, 2005 ; Kable et al., 2002).

Par contre, nos analyses sur les aspects sémantiques (en particulier l'analyse de conjonction) ont permis de mettre en évidence un certain nombre de régions activées de façon semblable pour la compréhension des images et des phrases. Il semblerait que le réseau décrit initialement par Proverbio et al. (2009) (incluant l'aire MT/V5 droite et le gyrus fusiforme dans le cortex temporal inférieur (BA 19), l'aire du mouvement du corps (EBA extrastriate body aera) localisée dans le cortex occipito-temporal latéral (BA37), le STS gauche (BA38), les aires motrices (BA6) et prémotrices (BA4) gauches, le cortex cingulaire antérieur et postérieur (BA24/BA31) et le gyrus frontal inférieur et moyen (BA10/47)) corresponde plutôt au réseau sémantique que nous avons observé puisque ces régions sont impliquées à la fois dans la compréhension des images et des phrases. Par ailleurs, il semblerait que stimuli statiques et dynamiques (images et phrases) entraînent des activations similaires de ce réseau étendu puisque la comparaison des conditions statiques et dynamiques entraîne peu ou pas de différence. La simulation ne se limiterait pas aux aspects moteurs des actions : tout événement impliquant une personne humaine activerait ce réseau pour permettre la représentation de cet événement et ainsi d'accèder à la compréhension sémantique de la situation. Ainsi, la simulation n'est pas limitée aux aires sensorimotrices et impliquerait un large réseau fronto-temporo-pariétal.

B) Résultats Partie II – ICA

- **Sélection des ROIs et des NOIs**

Pour rappel, l'analyse de connectivité fonctionnelle permet d'obtenir une mesure de l'interaction entre des régions d'intérêt (ROIs) d'un ou plusieurs réseaux (NOIs) formés à partir des classes (COIs) résultant de l'analyse ICA. Comme nous l'avons précisé dans la partie descriptive sur les analyses ICA avec NetBrainWork, nous avons effectué les différentes étapes de cette analyse séparément pour les sessions Images et Phrases. Nous avons ainsi déterminé deux COIs et caractérisé deux réseaux fonctionnels distincts (un pour chaque condition, toutes les ROIs sélectionnées pour composer ces réseaux étant issues respectivement de ces deux COIs). Rappelons que le but de cette analyse était d'évaluer la connectivité fonctionnelle au sein du réseau commun aux images et aux phrases, mis en évidence par l'analyse de conjonction en IRMf. C'est pourquoi, nous avons cherché spécifiquement, parmi les classes issues de l'ICA, s'il existait des réseaux correspondants et nous avons volontairement sélectionné nos NOIs pour leur cohérence avec les résultats obtenus à partir de l'analyse de conjonction en IRMf. Ainsi, parmi les réseaux obtenus, nous avons été capables d'identifier, pour chaque condition, des réseaux très similaires au réseau IRMf mais aussi très similaires entre eux. Nous désignerons maintenant ces réseaux sous le nom de « Réseau Images » et

« Réseau Phrases ». Comme en témoignent la superposition des cartes d'activation (t-mpas) sur la Figure 27) ainsi que le tableau de correspondance des pics d'activation des différentes ROIs (Table 4, ROIs issues de la conjonction et ROIs qui composent nos deux NOIs), le chevauchement des activations est presque parfait et on retrouve les mêmes aires cérébrales impliquées à partir des différentes analyses. En plus de leur correspondance anatomique avec le réseau IRMf, ces deux réseaux obtenus par analyse ICA présentent un bon niveau de confiance sur la base des critères DR et DU des classes dont ils sont issus. Le « réseau Images » a été défini à partir d'une classe unique dont les valeurs sont DR=0.95 et DU=0.91 ; le « réseau Phrases » a été défini à partir d'une classe unique dont les valeurs sont DR=0.95 et DU=0.97.

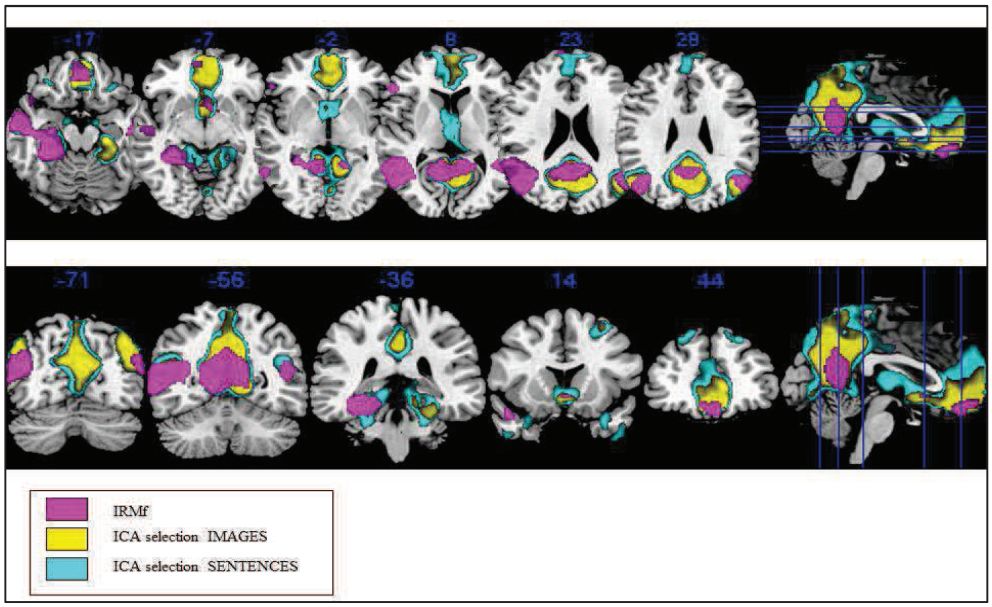


Fig 27) Recoupement des t-maps issues de l'analyse de conjonction en IRMf (rose), de l'analyse ICA pour le réseau IMAGES (jaune) et de l'analyse ICA pour le réseau PHRASES (turquoise)

Plus précisément, les réseaux issus des analyses ICA recrutent tout deux largement et bilatéralement les régions temporo-pariétales avec une activation plus étendue dans l'hémisphère gauche (comme nous l'avons identifié dans l'analyse IRMf). A gauche, le cluster d'activation situé au niveau temporo-pariétal implique une large portion de gyrus temporal moyen qui s'étend antérieurement depuis BA21 et médialement au niveau du gyrus parahippocampique jusqu'à BA39 postérieurement où il rejoint le gyrus angulaire. Dans l'hémisphère gauche, on trouve une activation spécifique du gyrus temporal supérieur (BA38).

ROIs	Conjonction	Réseau Images	Réseau Phrases
Gyrus Temporal Moyen BA 39 HG	-42 -66 23	-50 -60 20	-56 -59 18
Gyrus Temporal Moyen BA 39 HD	53 -69 26	52 -64 28	52 -64 31
Gyrus Temporal Moyen BA 21 HG	-54 -6 -15	-57 -13 -14	-58 -11 -14
Gyrus Temporal Moyen BA 21 HD	53 -7 -17	60 -6 -18	60 -3 -18
Gyrus Temporal Supérieur BA 38 HG	-45 16 -21	-37 14 -26	-31 9 -23
RSC HG	-9 -55 17	-8 -57 14	-4 -60 23
RSC HD	12 -51 15	9 -58 14	6 -60 22
Gyrus Frontal Médian BA 11	-9 52 -13	-4 49 -13	-1 43 -16
Gyrus Frontal Inférieur BA45/46 HG	-56 29 7	-52 28 11	-55 28 10
Noyau Caudé	3 11 -6	1 9 -7	0 10 -7

Table 4) Comparaison de la localisation des régions issues de la conjonction IRMf et des analyses ICA NetBrainWork pour les Images et pour les Phrases (coordonnées de Talairach)

A droite, les activations temporo-pariétales sont plus limitées mais on retrouve également, sous la forme de 2 clusters, un recoupement au niveau de la région du TPJ (BA39) et de la portion plus antérieure du gyrus temporal moyen (BA21) à la fois pour les 2 réseaux issus de l'ICA et pour l'analyse de conjonction. Une large activation de la région rétrospléniale (BA29/30) s'étendant jusqu'au précuneus (BA19) est également commune à nos deux réseaux. Les aires frontales sont aussi impliquées à la fois dans le réseau Images et dans le réseau Phrases au niveau du gyrus frontal inférieur gauche (BA45/BA46) et du gyrus frontal médial bilatéralement (BA11) avec une activation assez étendue de part et d'autre de la scissure inter hémisphérique. On retrouve également des activations au niveau sous-cortical avec une activation du noyau caudé et des régions adjacentes. Nous avons donc sélectionné les maximums locaux issus de ces différentes régions pour chaque réseau (réseau Images et réseau Phrases) afin de définir les régions d'intérêt nécessaires aux calculs des corrélations. Après réduction de la taille des ROIs (sphères de 10 mm de radius centrées sur les

maximum locaux), on retrouve des données excessivement proches de celles obtenues par l'analyse de conjonction comme le montre la comparaison des maximums locaux issus des deux analyses, rapportée dans la table 4.

- **<u>Corrélations à partir des ROIs issues des « Réseau Images » et « Réseau Phrases »</u>**

Rappelons qu'en utilisant la toolbox NetBrainWork, notre objectif était d'étudier la connectivité fonctionnelle au sein des deux réseaux obtenus par ICA, en calculant le degré de corrélation des décours temporels entre les différentes régions de ces réseaux. Ainsi, nous pouvons comparer la connectivité fonctionnelle des réseaux « Phrases » et « Images » et essayer d'en déduire des processus communs impliqués dans la compréhension des images et des phrases ou du moins de mieux comprendre le fonctionnement du réseau sémantique commun.

L'outil NetBrainWork permet tout d'abord d'obtenir un graphe des corrélations, représentation graphique de la force des interactions entre différentes régions au sein d'un réseau (les corrélations représentées dans ce graphe correspondent aux corrélations globales). D'après cette interprétation graphique où nous n'avons gardé que les corrélations supérieures à 0.20, on peut voir qu'un certain nombre de régions semblent interagir conjointement de façon assez similaire pour le réseau Images et pour le réseau Phrases. La figure 28 illustre la similarité de la connectivité fonctionnelle de ces deux réseaux : les mêmes régions cérébrales sont reliées fonctionnellement par des forces d'interactions équivalentes entre le réseau Phrases et le réseau Images.

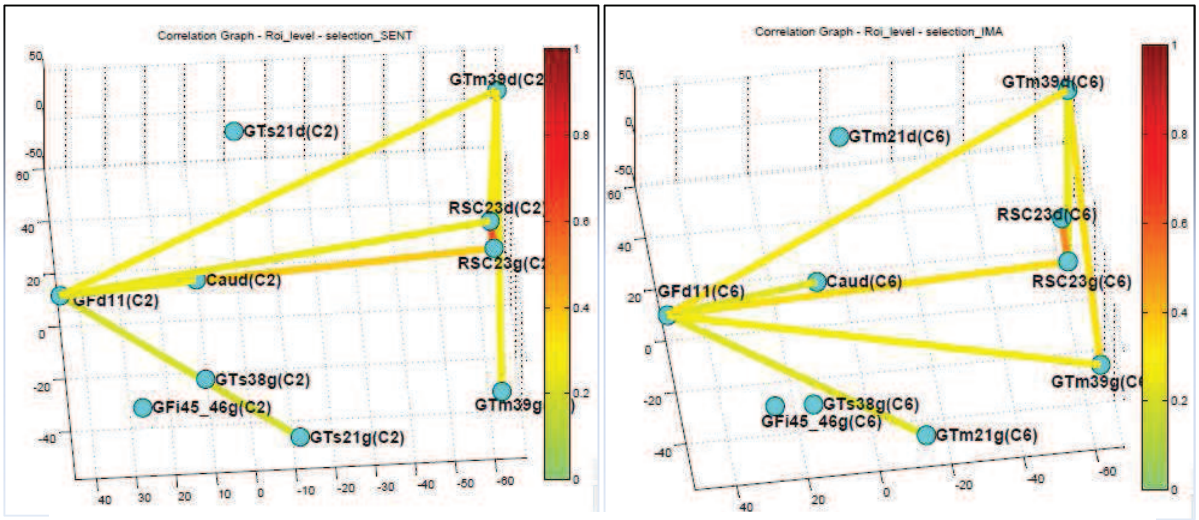


Fig 28) Graphes de corrélations pour les réseaux Phrases (à gauche) et Images (à droite)
Ces graphes permettent de visualiser les corrélations des décours temporels au sein de ces deux réseaux et d'observer une tendance similaire entre les deux réseaux, à ce seuil.

Nous avons ensuite calculé les corrélations (partielles et globales) de ces régions afin de voir quelles corrélations étaient significatives. Les corrélations partielles sont plus précises et nous ne présenterons que les résultats de ces corrélations ici ; cependant nous avons aussi calculé les corrélations globales qui, pour rappel, intègrent aussi l'influence des autres régions sur la connectivité fonctionnelle sans que cette influence soit dissociable. De ce fait, elles illustrent aussi l'importance du fonctionnement cérébral en tant que réseau (voir Annexe 2).

Toutes les corrélations partielles présentées dans la figure 29 sont significatives à p=.01, elles confirment que les principales régions impliquées pour les images et les phrases sont les mêmes : les mêmes régions sont corrélées avec un niveau de corrélation et un ordre d'importance équivalents. A partir de ces corrélations partielles, un ensemble de régions communes semblent importantes pour le traitement sémantique des images et des phrases : en particulier les interactions entre le RSC gauche et droit (BA29/30), le RSC et le cortex préfrontal médian (BA11), la jonction temporo-pariétale gauche et son opposé controlatéral (BA39), le cortex frontal médian (BA11) et le noyau caudé.

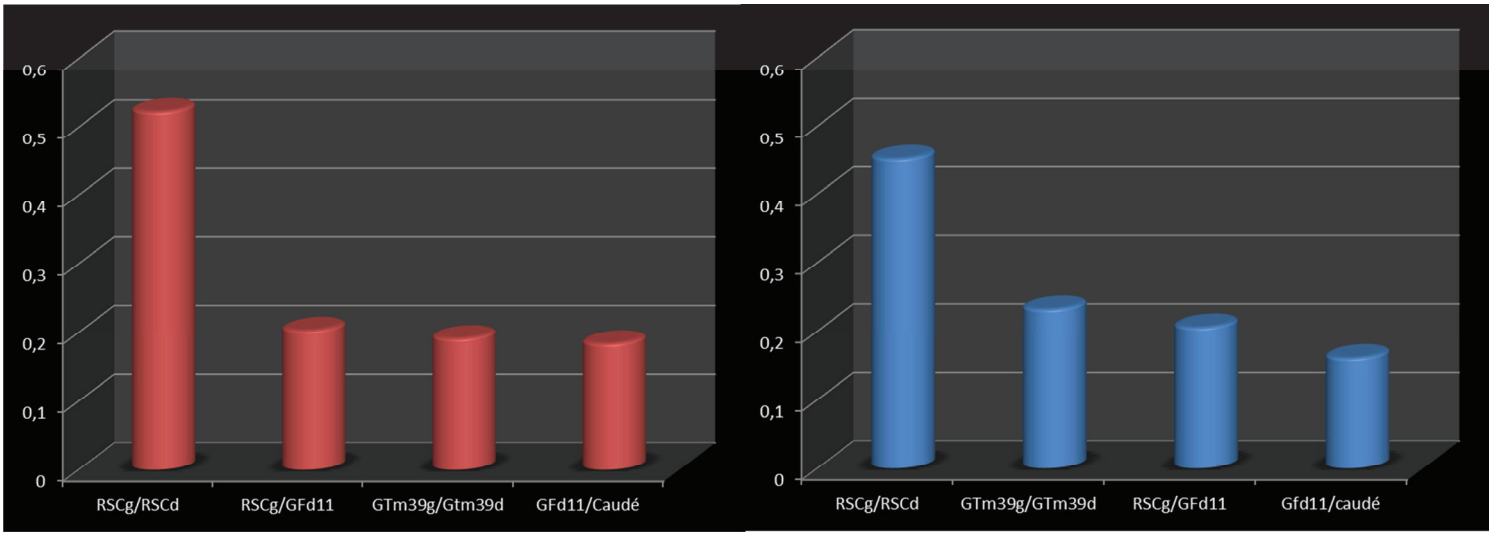


Fig 29) Représentation des corrélations partielles significatives (p=0.01) pour les réseaux Phrases (à gauche) et Images (à droite)
(IC Intégration Réseau Phrases [0.2923 ; 0.2940] et IC Intégration Réseau Images [0.2633 ; 0.2649])

- **<u>Synthèse des résultats ICA</u>**

Les analyses ICA permettent de confirmer la description du réseau sémantique pariéto-temporo-préfrontal, commun à la compréhension des images et des phrases et surtout de mettre en évidence des interactions fonctionnelles entre les différentes régions cérébrales qui le composent. Les corrélations globales fournissent l'aspect global de ce réseau et réaffirment l'importance d'un fonctionnement en réseau. Par ailleurs, les corrélations partielles nous ont permis de préciser plus

finement la connectivité fonctionnelle au sein de ce réseau. Il apparaît que, pour images et phrases, les interactions de plusieurs régions soient particulièrement importantes : il s'agit du RSC, du Gyrus temporal moyen (TPJ BA39), du Gyrus frontal médian et du noyau caudé.

C) Résultats Partie III – DTI

Avec l'analyse tractographique, nous nous sommes efforcés d'identifier les faisceaux sous-tendant la connectivité structurale du réseau sémantique et de les comparer avec les faisceaux décrits dans la littérature pour mieux comprendre leurs rôles.

- **Délimitation des régions d'intérêt permettant la tractographie**

Les maximums locaux des principales ROIs issues de l'analyse de conjonction IRMf ont été utilisés comme points d'ancrage (« seeds regions ») pour initier la tractographie en DTI (voir table 5). Lorsque l'activation d'un cluster était particulièrement importante ou qu'elle recoupait plusieurs régions anatomiques, nous avons préféré définir le centre de masse du cluster d'activité (grâce à l'outil Matlab Marsbar) ou des sous-maximums locaux.

Nous avons utilisé une approche « deux ROIs » pour initier la tractographie : pour chaque sujet, les faisceaux de connexion ont été reconstruits en indiquant deux régions d'ancrage pour délimiter un faisceau (en utilisant les logiciels AFNI et DTIquery comme décrit dans la partie Matériels et Méthodes). Cette approche a été appliquée à toutes les ROIs issues de la conjonction, pour tester la connectivité de toutes nos régions deux par deux. Seuls les résultats des régions pour lesquelles nous avons trouvé une connectivité consistante avec au moins une autre ROI seront présentés (connexion présente chez 15 sujets au minimum, nombre moyen de fibre du faisceau >10).

ROIs	Coordonnées ROI	Définition de la ROI	Max local principal cluster d'origine	Taille de la ROI
TPJ Gyrus Temporal Moyen BA 39 HG	-47 -55 17	Centre de masse	-42 -66 23	20 mm3
TPJ Gyrus Temporal Moyen BA 39 HD	48 -53 18	Centre de masse	53 -69 26	20 mm3

Gyrus Temporal Moyen BA 21 HG	-54 -6 -15	Coordonnées d'origine	-54 -6 -15	15 mm3
Gyrus Temporal Moyen BA 21 HD	53 -7 -17	Coordonnées d'origine	53 -7 -17	15 mm3
Gyrus Temporal Supérieur BA 38 HG	-45 16 -21	Coordonnées d'origine	-45 16 -21	15 mm3
Gyrus Frontal Médian BA 11 HG	-9 52 -13	Coordonnées d'origine	-9 52 -13	15 mm3
Gyrus Frontal Médian BA 11 HD	3 40 -17	Coordonnées d'origine	3 40 -17	15 mm3
Gyrus Frontal Inférieur BA45/46 HG	-56 29 7	Coordonnées d'origine	-56 29 7	15 mm3
Gyrus Frontal Inférieur BA47 HG	-39 28 -14	Maximum local secondaire	-56 29 7	15 mm3
Gyrus Frontal Inférieur BA47 HD	32 37 -9	Coordonnées d'origine	32 37 -9	15 mm3
Gyrus Précentral BA6 HG	-42 0 53	Coordonnées d'origine	-42 0 53	15 mm3
RSC BA29/30 HG	-9 -55 17	Coordonnées d'origine	-9 -55 17	15 mm3
RSC BA29/30 HD	12 -51 15	Coordonnées d'origine	12 -51 15	15 mm3
Precuneus HG	-18 -52 60	Maximum local secondaire	-9 -55 17	15 mm3
Precuneus HD	12 -48 62	Maximum local secondaire	12 -51 15	15 mm3
Noyau Caudé HG	-3 14 -5	Maximum local secondaire	3 11 -6	10 mm3
Noyau Caudé HD	3 11 -6	Coordonnées d'origine	3 11 -6	10 mm3

Table 5) Création des ROIs permettant d'initier la tractographie en DTI
Les coordonnées utilisées pour créer les ROIs (coordonnées de Talairach), leurs origines et la taille des ROIs sont précisées dans ce tableau.

Les principales connexions que nous avons mises en évidence sont des connexions intrahémisphériques cortico-corticales reliant les régions pariétales et temporales aux régions frontales ; la plupart de ces connexions recoupent celles généralement désignées comme étant la voie ventrale sémantique dans le cadre du langage (Duffau, 2008). Mais nous avons également mis en évidence des connexions interhémisphériques et cortico-sous corticales.

- **Résultats principaux : Connectivité intrahémisphèrique des faisceaux de la voie « ventrale » sémantique**

  - *Connexions fronto-pariétales*

Le TPJ est connecté bilatéralement aux régions frontales par un faisceau qui se divise en deux branches pour rejoindre le cortex orbito-frontal latéralement, au niveau du gyrus frontal inférieur (BA45/46/47) et médialement, au niveau du gyrus frontal médian (BA11). De par son trajet et son orientation antéro-postérieure, ce faisceau correspond à la description faite dans la littérature du faisceau fronto-occipital inférieur (IFOF), un faisceau qui ne semble exister que chez l'homme (Catani, 2007). En effet, l'IFOF connecte le cortex frontal inféro-latéral et dorso-latéral au cortex temporal postérieur et à la partie ventrale du lobe occipital (Catani et al., 2002 ; Catani, & Thiebaut de Schotten, 2008). Les analyses de groupes permettent de visualiser ce faisceau pour l'ensemble des sujets (Figure 30), mais nous préférons nous baser sur les données individuelles pour le décrire plus précisément (Figure 31).

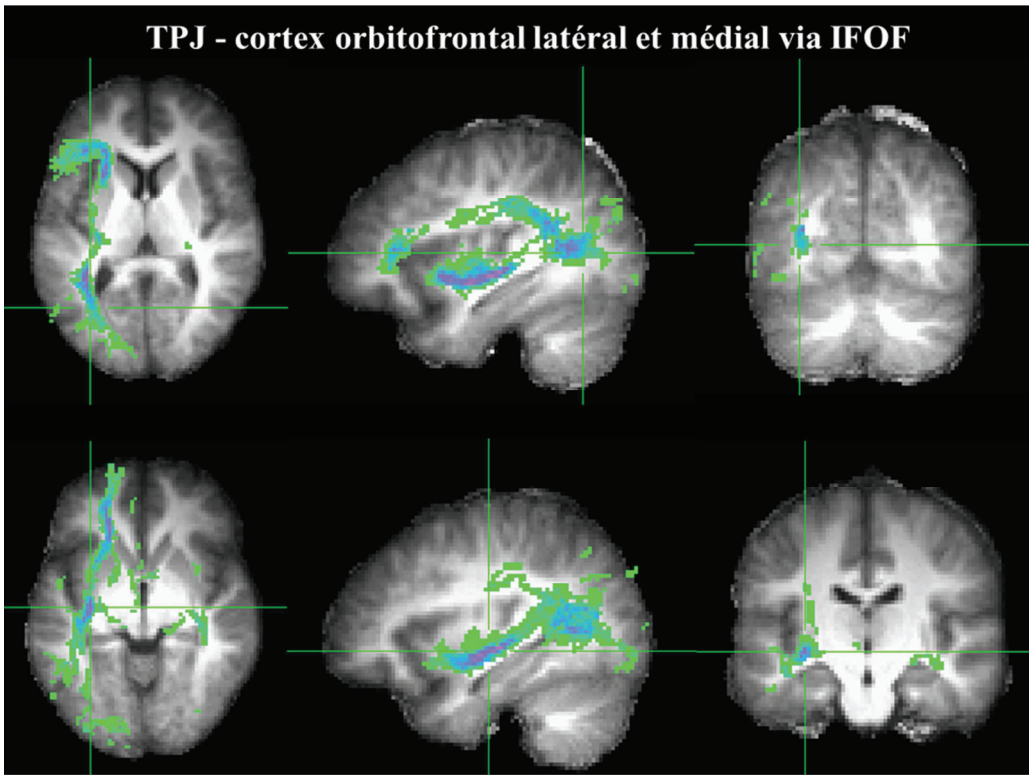


Fig 30) Figure issue de l'analyse sur le groupe de sujets, permettant de visualiser le trajet du faisceau IFOF reliant le TPJ au cortex orbitofrontal latéral dans l'hémisphère gauche.
(en vue horizontale, sagitale et frontale)

Dans les deux hémisphères, la terminaison et le trajet de ce faisceau sont consistants chez plus de 15 sujets. Dans notre étude, une partie des fibres de ce faisceau connecte les régions postérieures à la fois à la partie dorsolatérale (BA 46) et la partie ventrolatérale (BA45) du cortex préfrontal, comme nous le montre la localisation de notre ROI à l'intersection de BA45 et BA46. Cependant, il est intéressant de noter que ce faisceau passe encore plus directement par notre ROI située, plus ventralement, au niveau de la pars orbitaris (BA47). Le faisceau reliant le TPJ à la ROI GFi BA45/46 n'est composé que de 20,35 fibres en moyenne et n'est présent que chez 14 sujets ; en revanche le faisceau reliant le TPJ à la ROI située en BA47 est présent chez 16 sujets et a un nombre moyen de fibres égal 35,87 pour l'hémisphère gauche. C'est encore plus marqué pour l'hémisphère droit où le faisceau entre les ROIs TPJ et GFi BA47 comporte en moyenne 80 fibres pour les 16 sujets chez qui il est présent. Rappelons que même si ces deux ROIs (GFi BA47 HG et GFi BA47 HD) recoupent une même région anatomique pour chaque hémisphère, elles sont basées sur les potentiels locaux des activations IRMf et donc pas nécessairement situées tout à fait au même endroit.

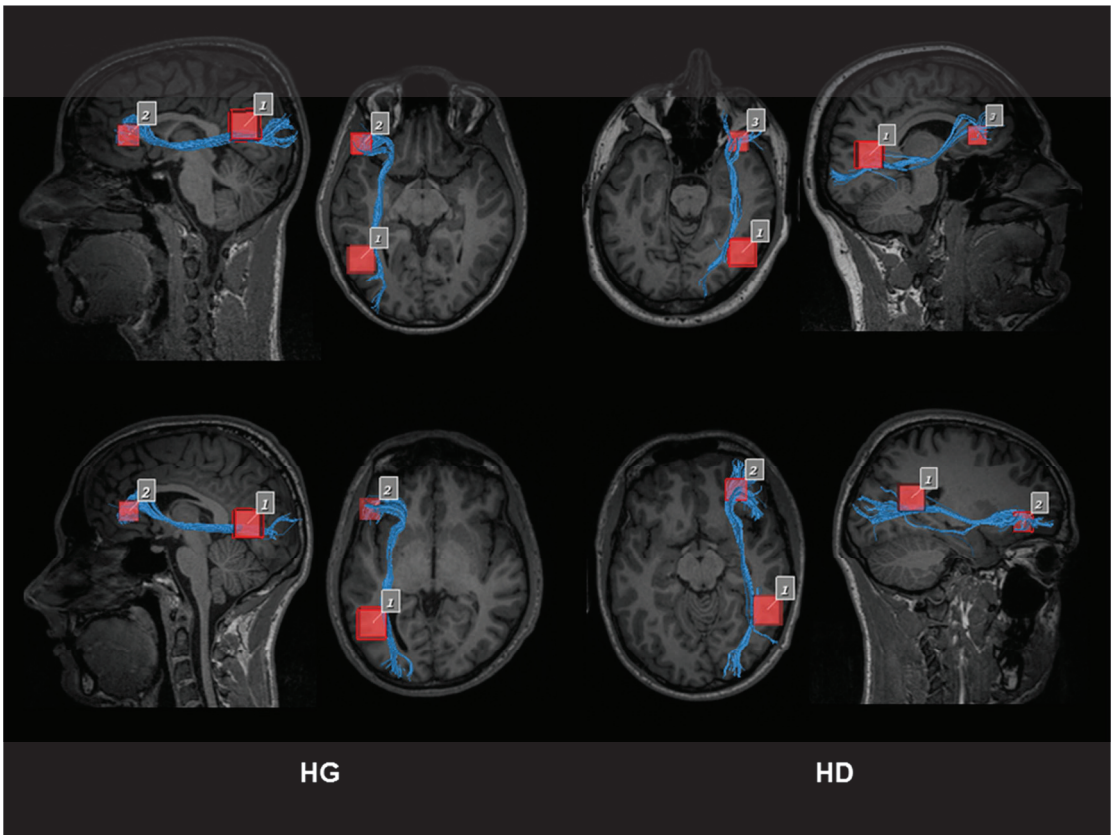


Fig 31) Tractographie du faisceau fronto-occipital inférieur (IFOF)
Chez 4 sujets sains, pour l'hémisphère gauche (figures de gauche) et l'hémisphère droit (figures de droite)

Comme le montre la figure 32, Il semble que l'origine frontale de la partie latérale de ce faisceau se fasse pleinement dans la pars opercularis où cette origine est caractérisée par une boucle de fibres

qui arrivent latéralement au niveau de cette région. Quelques fibres plus médiales appartenant à l'IFOF prennent également naissance au niveau du gyrus frontal médian (BA11) : elles sont plus médiales et moins nombreuses (nombre moyen de fibres=17). Cette branche préfrontale secondaire de l'IFOF est présente chez 15 sujets pour l'hémisphère gauche, on ne la retrouve pas dans l'hémisphère droit (plus précisément elle n'est présente que chez 8 sujets, probablement en raison de la position plus postérieure du pic de la ROI située dans GFm BA11 de l'hémisphère droit).

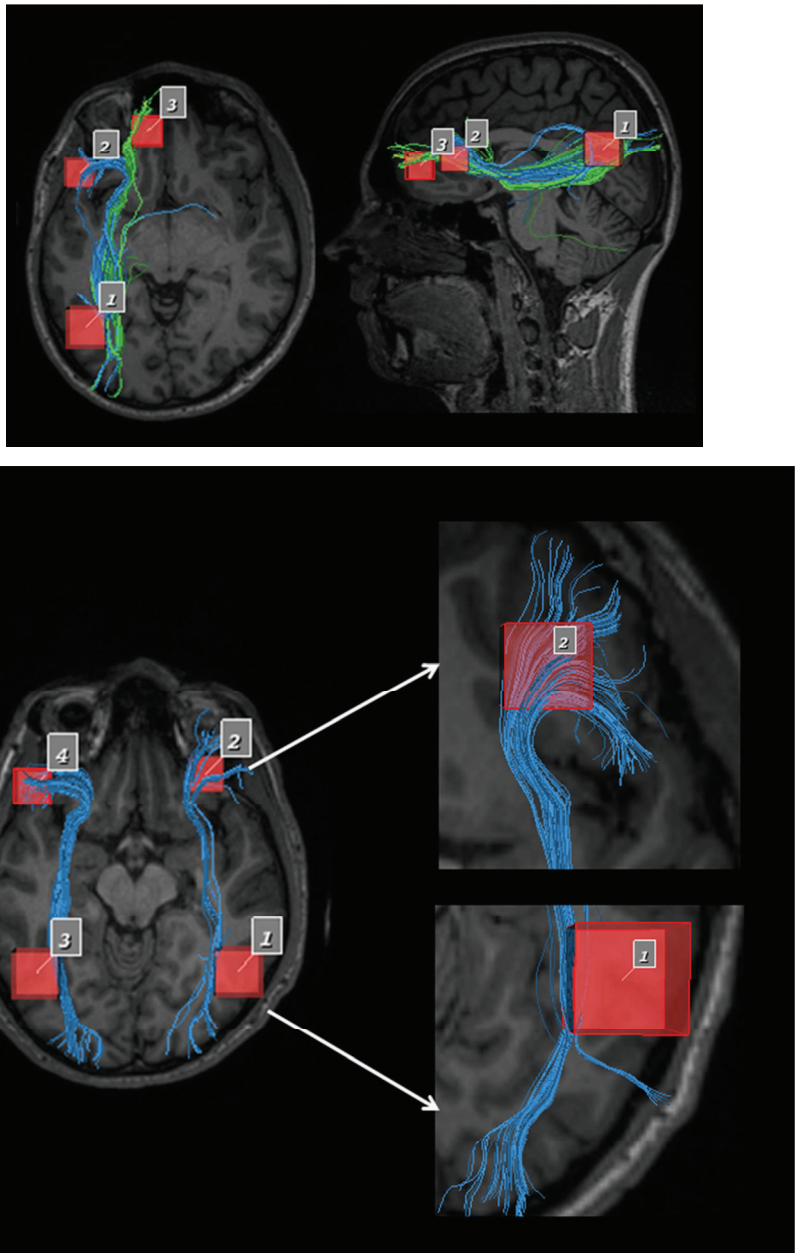


Fig 32) Terminaisons du faisceau fronto-occipital inférieur (IFOF)
Figure du haut : Origines frontales latérales (bleu) et médiales (vert) de l'IFOF chez 1 sujet.
Figure du bas : Trajet bilatéral, origines frontales et terminaisons pariéto-occipitales de l'IFOF chez 1 sujet

A partir des régions préfrontales, les fibres se dirigent ensuite vers l'arrière, en direction des régions temporales. A la jonction des lobes frontaux et temporaux, le faisceau devient plus étroit à son passage au niveau de la partie antérieure de la capsule externe. Dans la littérature, on considère que les terminaisons postérieures de l'IFOF se situent au niveau des régions temporales moyenne et postérieure, ainsi qu'au niveau des régions occipitales inférieures. Elles se traduisent par des radiations se terminant au niveau des gyri temporaux moyen, inférieur et fusiforme pour le lobe temporal, ainsi qu'au niveau du gyrus lingual pour le lobe occipital (Catani et al., 2002). Dans notre cas, étant donné la position de nos deux régions d'intérêt, on s'aperçoit que l'IFOF traverse la partie la plus médiale de notre ROI TPJ avant, effectivement, de projeter en arrière au niveau du cortex occipital inférieur.

- *Connexions fronto-temporales*

Les régions temporales antérieures sont elles aussi reliées aux régions frontales médiale et latérale (BA11 et BA47, Figures 33 et 34). Cette connexion se fait par l'intermédiaire du faisceau unciné, ce qui est là aussi tout à fait cohérent avec les données de la littérature où le faisceau unciné est décrit comme un faisceau d'association ventral qui connecte la partie antérieure du lobe temporal au cortex orbitofrontal et latéral. (Stuss & Knight, 2002 ; Catani et al., 2002 ; Catani, & Thiebaut de Schotten, 2008).

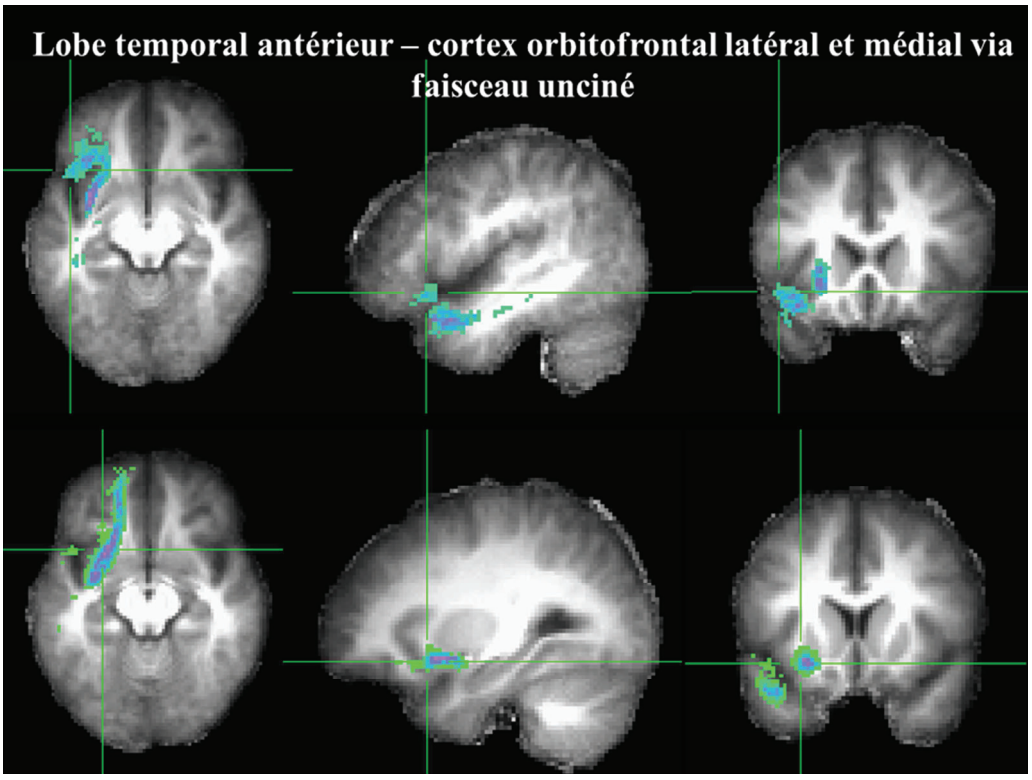


Fig 33) Figure issue de l'analyse sur le groupe de sujets, permettant de visualiser le trajet du faisceau unciné reliant le lobe temporal antérieur au cortex orbitofrontal latéral et médian dans l'hémisphère gauche.
(en vue horizontale, sagitale et frontale)

Chez 18 de nos 19 sujets, on retrouve la forme caractéristique du faisceau unciné (forme de C concave en avant, nombre moyen de fibres=26,7). Dans notre étude, la ROI temporale qui nous a permis d'initier la tractographie du faisceau unciné est le Gyrus temporal Supérieur (BA38 HG), les autres ROIs temporales (Gyrus temporal moyen BA21 HG et BA21 HD) ne permettant pas de mettre en évidence ce faisceau. C'est aussi pour cette raison que nous n'avons pu délimiter ce faisceau que pour l'hémisphère gauche, n'ayant pas observé d'activation du gyrus temporal supérieur (BA38) à droite en IRMf. L'origine temporale du faisceau unciné se situe donc dans la partie la plus rostrale du gyrus temporal supérieur.

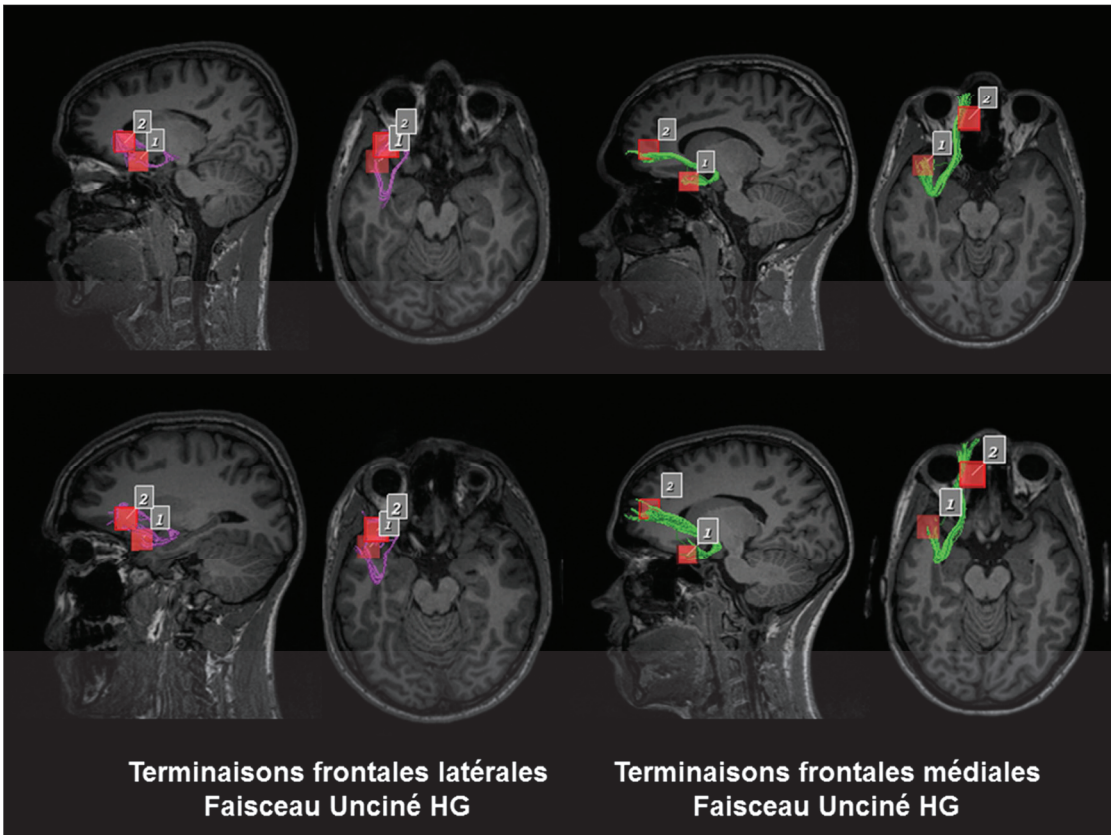


Fig 34) Tractographie et terminaisons frontales latérales ou médiales du faisceau unciné (UF)
Chez 4 sujet sains, pour l'hémisphère gauche (figures de gauche)

Comme pour l'IFOF, les terminaisons frontales du faisceau unciné (UF) se font également via deux branches bifurcatrices, l'une se dirigeant latéralement vers le gyrus frontal inférieur (pars opercularis, BA47) et l'autre médialement vers le gyrus frontal médian (BA11) où ses fibres se confondent avec celles de l'IFOF, après le passage de la capsule externe. Ces terminaisons frontales sont illustrées dans la figure 35). Dans notre étude, on retrouve la branche frontale latérale du faisceau unciné chez 18 de nos sujets, avec un nombre de fibre moyen égal à 26,6 fibres, tandis que la branche frontale

médiale est présente chez seulement 16 sujets mais avec un nombre de fibres moyen de 47,93 fibres.

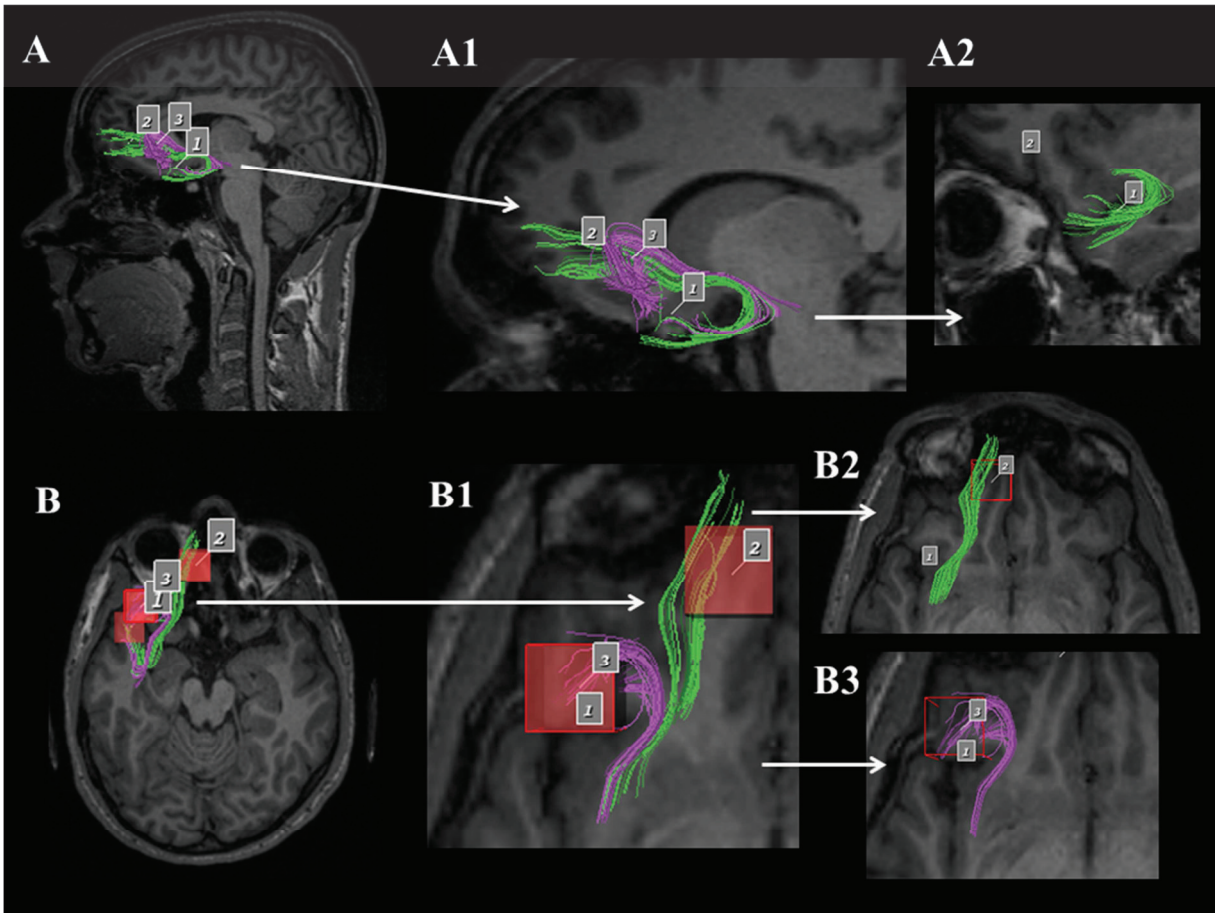


Fig 35) Rapports des composantes fronto-latérales et fronto-médiales du faisceau unciné (UF)
Rapports des composantes fronto-latérales et fronto-médiales de l'UF, chez 1 sujet, en coupes sagittales (A et A1) et en coupes horizontales (B et B1). Détails des origines/terminaisons de l'UF au niveau de la matière grise : origine temporale en coupesagittale (A2), terminaison fronto-médiale (B2) et terminaison fronto-latérale (B3) en coupe horizontale (les ROIs sont masquées pour mieux visualiser la localisation

- *Connexions pariéto-temporales*

Les connexions pariéto-temporales ont été tracées entre les ROIs TPJ BA39 et GTS BA21 pour initier la tractographie bilatéralement entre les régions pariétales et temporales. Ces deux régions sont directement reliées par un gros faisceau de fibres qui suit en profondeur le sillon temporal supérieur le long du gyrus temporal moyen (BA21) comme le montrent les figures 36 et 37.

La tractographie de ce faisceau semble assez consistante dans les deux hémisphères, puisque l'on retrouve ce faisceau, ayant une origine et une terminaison semblables chez nos 19 sujets : dans l'hémisphère gauche, avec un nombre moyen de fibres = 154 (soit 0,20% de l'ensemble des fibres du

cerveau) et dans l'hémisphère droit, avec un nombre moyen de fibres=188 (soit 0,24% de l'ensemble des fibres du cerveau).

L'origine au niveau temporal de ce faisceau se fait très nettement au niveau du Gyrus temporal Supérieur (BA21, Figure 38) car lorsque l'origine de la tractographie est située plus antérieurement et plus ventralement en utilisant la ROI B38 HG comme point d'ancrage, on obtient un faisceau de fibres beaucoup moins important (36 fibres en moyenne, soit 0,5% de l'ensemble des fibres du cerveau). Cette connexion secondaire suit cependant le même trajet, dans la profondeur du STS, et est également présente chez nos 19 sujets pour l'hémisphère gauche. Même si quelques fibres se prolongent jusqu'à la partie la plus antérieure du lobe temporal (BA38), la majorité des fibres semblent donc se terminer au niveau du Gyrus temporal supérieur (BA21).

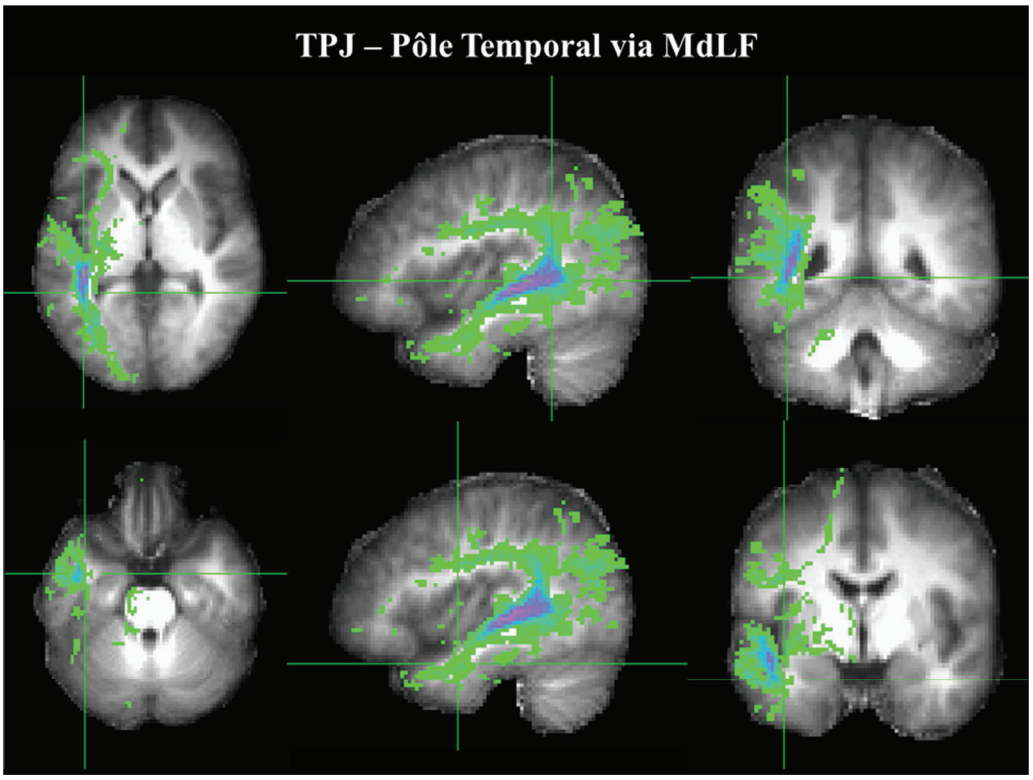


Fig 36) Figure issue de l'analyse sur le groupe de sujets, permettant de visualiser le trajet du faisceau longitudinal moyen (MdLF) reliant le TPJ au GTs.
(en vue horizontale, sagittale et frontale)

En ce qui concerne les terminaisons postérieures de ce faisceau, bien que la majorité des fibres se terminent dans le TPJ, une partie des fibres semblent se prolonger au niveau occipital supérieur (Figure 38). Les coordonnées utilisées pour la ROI TPJ dans notre étude (-47 -55 17 pour l'hémisphère gauche et 48 -53 18 pour l'hémisphère droit) sont très proches des coordonnées présentées comme correspondant à la localisation du gyrus angulaire dans un grand nombre d'études (e.g., Calvert et al. 1997 (49, −50, 17) ; Small et al. 1996 (−47, −58, 19); Fiez et al. (1995) (40, −54, 19); Demonet et al.

1992 (−48. −52, 28); Howard et al. 1992 (−50, −48, 8); Price et al. 1994 (−52, −38, 24); Bookheimer et al. (1995) (−44, −66, 32)) et en particulier celles utilisées par Horwitz et al. (1998) (±44, −54, +24) comme étant les coordonnées à l'intersection des différentes études susnommées. Ce faisceau apparaît donc unir la partie antérieure du lobe temporal (BA21) jusqu'à la jonction pariéto-temporale (BA39/22), à proximité du gyrus angulaire.

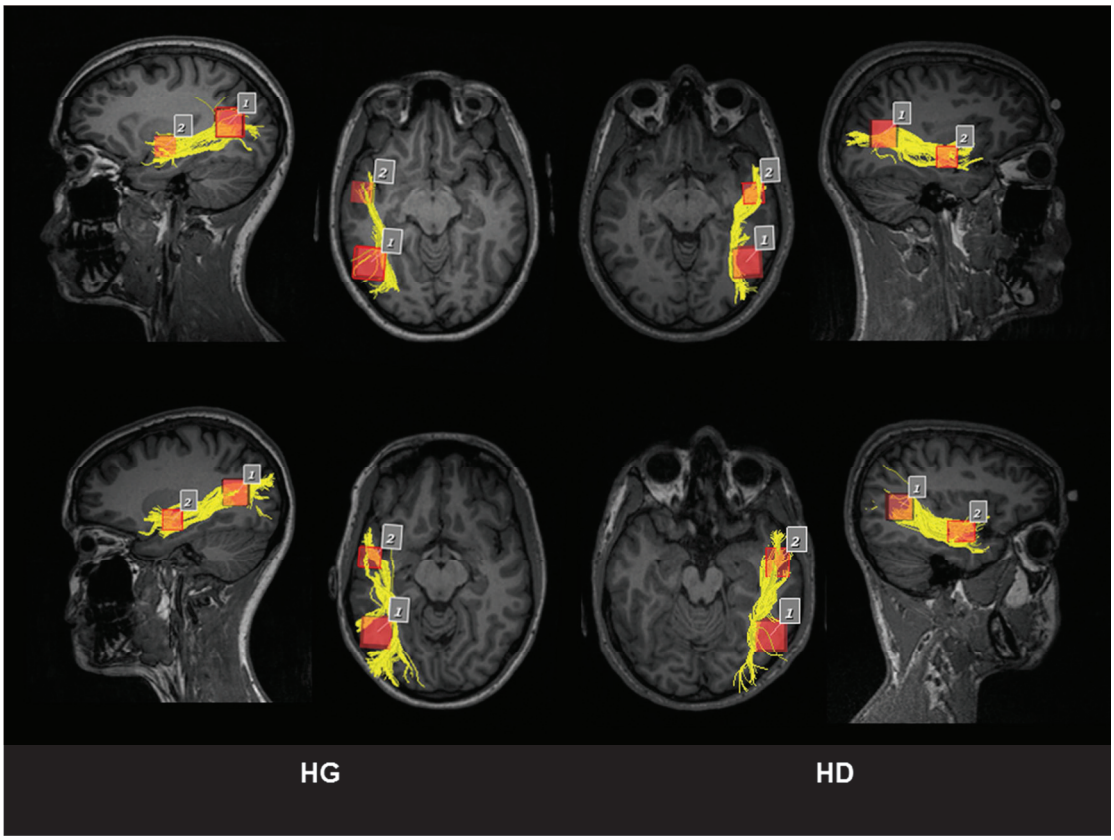


Fig 37) Tractographie du faisceau longitudinal moyen (MdLF)
Chez 4 sujets sains, pour l'hémisphère gauche (figures de gauche) et l'hémisphère droit (figures de droite)

De par la localisation des ROIs utilisées pour la tractographie et par l'orientation de ses fibres, ce faisceau semble correspondre au faisceau longitudinal moyen (MdLF), tel qu'il est décrit dans la littérature. Le MdLF est un faisceau rostrocaudal reliant le cortex temporal antérieur et le lobe pariétal ; originellement mis en évidence chez le singe (Seltzer & Pandya, 1984 ; Schmahmann et al. 2007), son existence a récemment été confirmée chez l'homme grâce à des travaux en DTI et des dissections post-mortem (Makris et al., 2008 ; Turken & Dronkers, 2011 ; Maldonado et al., 2013). Les études en DTI suggèrent que le MdLF pourrait se terminer dans le TPJ, plus précisément au niveau du gyrus angulaire (Makris et al., 2008), même si les études de dissections ne semblent pas confirmer ces résultats et postulent que ce faisceau se prolongerait également postérieurement vers les portions supérieures du lobe occipital (Maldonado et al., 2013). Ce faisceau n'est pas toujours facile à mettre en évidence en DTI car son trajet se confond partiellement avec les autres faisceaux

adjacents : faisceaux longitudinaux inférieur/supérieur, faisceau arqué et faisceau fronto-occipital inférieur. Nous verrons un peu plus loin comment nous avons essayé de caractériser au mieux ce faisceau en comparant sa trajectoire et ses origines par rapport aux autres faisceaux obtenus dans cette étude.

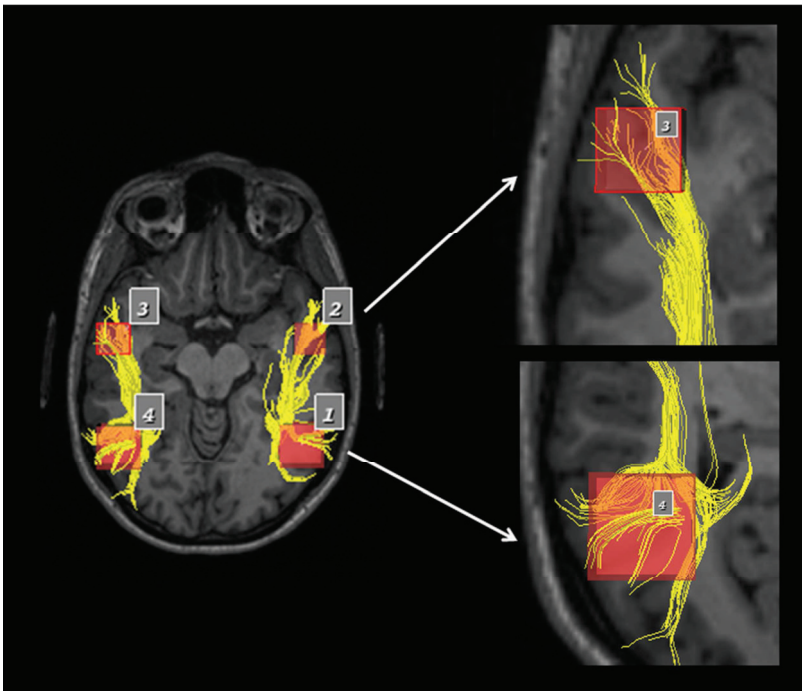


Fig 38) Terminaisons temporales et pariétales du faisceau longitudinal moyen (MdLF)

- **Résultats complémentaires**
  - *Connectivité du TPJ*

Le TPJ apparaissant comme une zone de convergence cruciale au sein du réseau sémantique, nous avons évalué la connectivité de cette région en plaçant une ROI unique dans cette région pour initier la tractographie (pour l'hémisphère gauche et l'hémisphère droit).

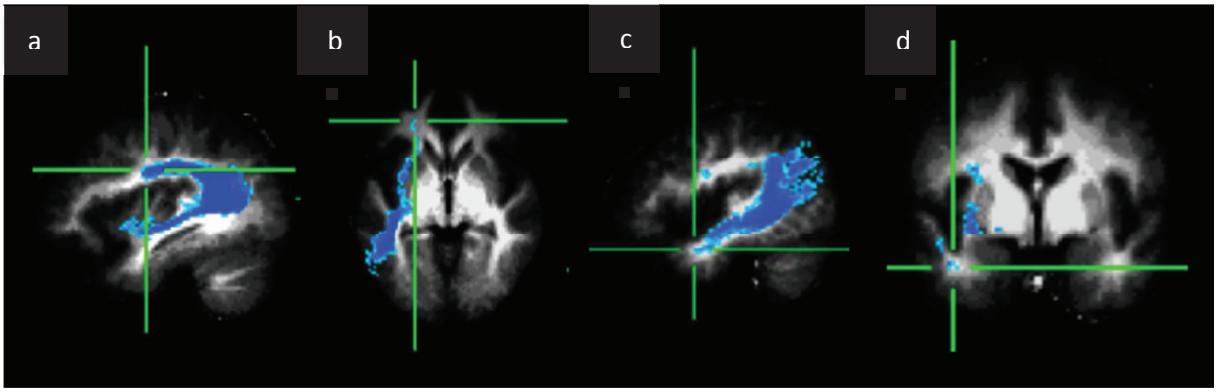


Fig 39) Tractographie issue du TPJ révélant un faisceau à 3 branches
Composante dorsale reliant le prémoteur (vue sagittale, a) ; composante médiale vers le cortex préfrontal (b, vue horizontale) ; composante ventro-latérale vers les régions temporales antérieures (c, vue sagittale) ; visualisation de la hauteur de ces 3 branches (d, vue coronale).

Les analyses de groupe mais aussi les analyses individuelles révèlent un réseau à 3 branches émanant du TPJ (Figure 39), que l'on retrouve pour la plupart des sujets et dans les deux hémisphères. Ce réseau à 3 branches est caractérisé par une composante ventro-latérale, une composante médiale et une composante dorsale. La composante dorsale atteint les régions prémotrices, en rejoignant a priori le faisceau longitudinal supérieur ou le faisceau arqué (Catani & Thiebaut de Schotten 2008). Des travaux antérieurs sur le langage (Ellmore et al. 2010) ont montré des terminaisons du faisceau arqué dans des régions postérieures, lesquelles étaient localisées à x = -46, y = -43, z = -7 en coordonnées de MNI, ce qui est différent de la localisation de notre ROI qui est plus postérieure et plus dorsale (-47 -57 15 en MNI). Ceci suggère que la composante fasciculaire dorsale issue du TPJ passerait par un faisceau additionnel, adjacent et postérieur au faisceau AF/SLF (des analyses plus fines, notamment post-mortem, pourraient confirmer cette hypothèse). La composante médiale de ce réseau à 3 branches, issu du TPJ, rejoint le pôle frontal dans sa partie ventrale, par l'intermédiaire du faisceau fronto-occipital inférieur (comme nous l'avons déjà décrit précédemment). Tandis que la composante ventro-latérale rejoint le pôle temporal : comme nous l'avons mentionné auparavant, il y a une ambiguïté sur le fait que les fibres de cette composante ventro-latérale appartiennent au faisceau longitudinal inférieur (Catani & Thiebaut de Schotten, 2008) ou au faisceau longitudinal moyen (Makris et al., 2008 ; Turken & Dronkers, 2011 ; Maldonado et al., 2013).

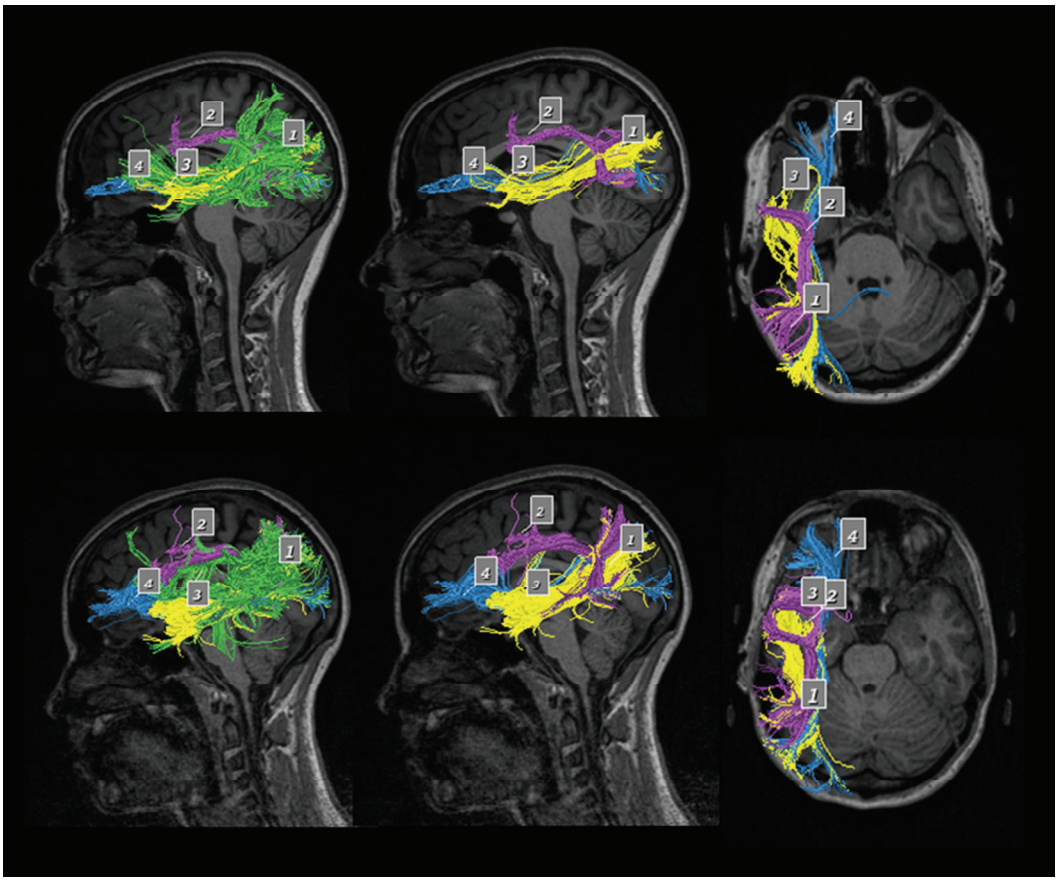


Fig 40) Caractérisation des différents faisceaux issus du TPJ de l'HG, pour 2 sujets individuels
Figure de gauche : connectivité totale issue du TPJ (vue sagittale) ; Figures du milieu et de droite : caractérisation plus spécifique des différents faisceaux issus du TPJ : en jaune (MdLF), en violet (faisceau arqué ou composante postérieure à l'AF/SLF), en bleu (IFOF)
(vues sagittale et horizontale)

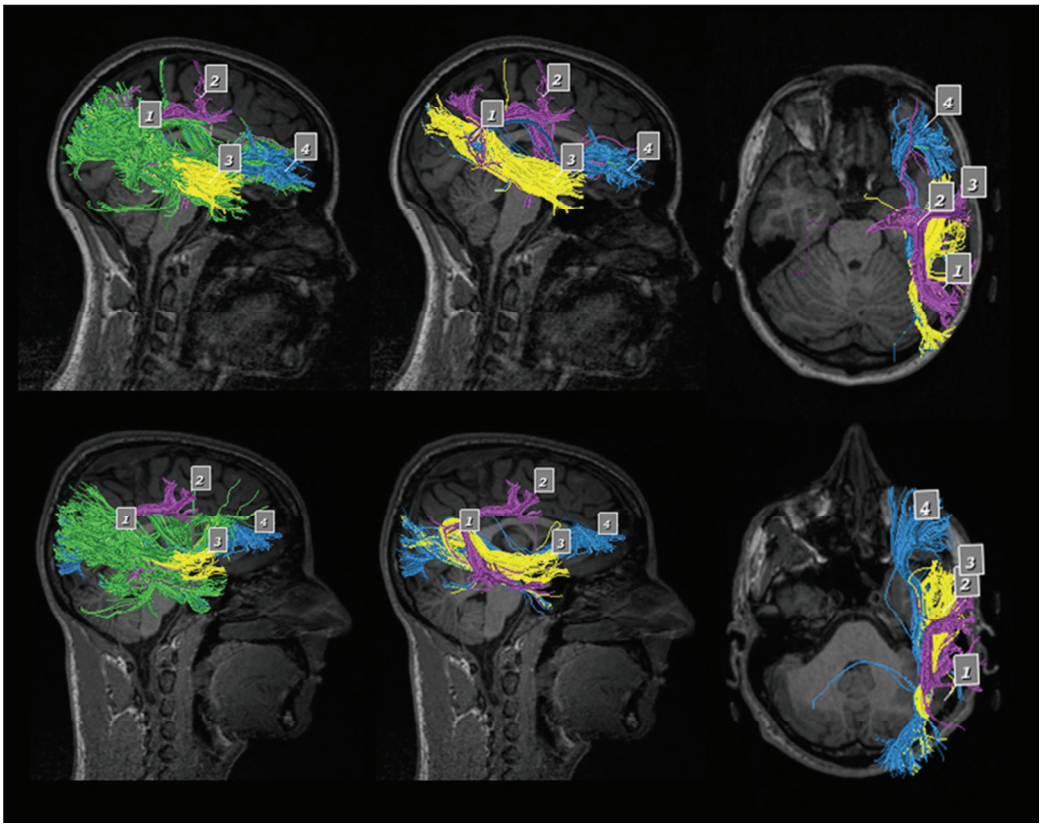


Fig 41) Caractérisation des différents faisceaux issus du TPJ de l'HD, pour 2 sujets individuels
Figure de gauche : connectivité totale issue du TPJ (vue sagittale) ; Figures du milieu et de droite : caractérisation plus spécifique des différents faisceaux issus du TPJ : en jaune (MdLF), en violet (faisceau arqué ou composante postérieure à l'AF/SLF), en bleu (IFOF)
(vues sagittale et horizontale)

L'analyse tractographique à partir du TPJ (gyrus angulaire) met en évidence le rôle clé de cette région pariéto-temporale, identifiée dans notre étude comme carrefour, ou « hub », entre les régions prémotrices, temporales et frontales, comme l'illustrent les figures 40) et 41).

- *Caractérisation du MdLF à travers la comparaison des trajets des différents faisceaux*

Afin de nous assurer de la pertinence de la caractérisation des différents faisceaux que nous avons pu observer, nous avons comparé leurs trajets respectifs en lien avec ce qui est rapporté dans la littérature. Nous souhaitions principalement apporter des éléments complémentaires à la description de ces faisceaux afin de lever l'ambiguïté sur l'identité du faisceau reliant les ROIS pariétales et temporales (MdLF/faisceau longitudinal inférieur), montrant que le faisceau que nous présumons être le MdLF a une trajectoire distincte des autres faisceaux considérés dans ce travail. Seuls les faisceaux de l'HG seront considérés puisque nous n'avons pas pu obtenir tous les faisceaux correspondant dans l'HD.

Nous avons tout d'abord comparé les trajets du faisceau présumé être le MdLF (entre TPJ et BA21) et du SLF (issu de la région TPJ et qui rejoint les aires prémotrices). La figure 42 illustre les rapports de ces deux faisceaux.

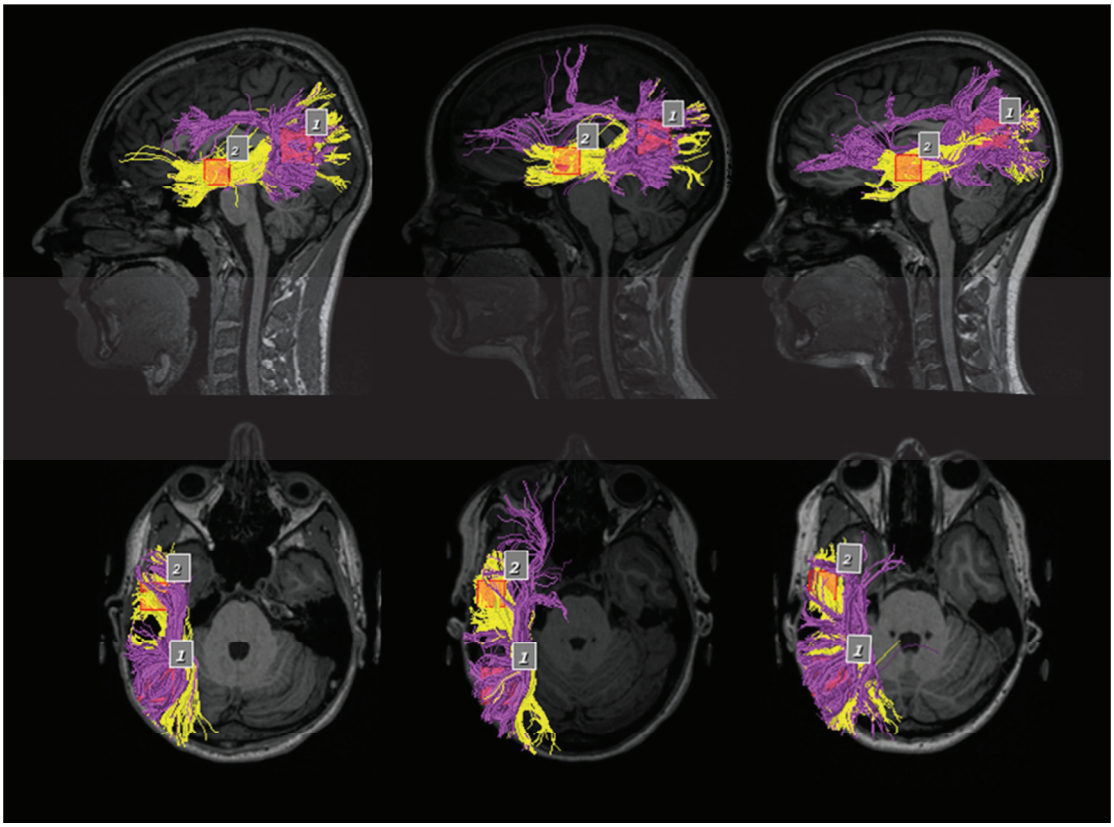


Fig 42) Rapports du MdLF (jaune) et du SLF (en violet)
chez 3 sujets en coupes sagittale et horizontale

Comme nous pouvons le constater, au niveau des régions pariétales, la trajectoire du MdLF est située en dedans du SLF/faisceau arqué et son trajet est beaucoup plus étendu, que ce soit ventralement vers les régions temporales ou dorsalement vers les régions occipitales. Le SLF, après avoir transité au niveau du gryus angulaire dans la région pariétale, vient se terminer également au niveau du STG mais sa terminaison se fait à la partie postérieure du STG alors que celle du MdLF se fait à la partie antérieure et moyenne du STG.

Par ailleurs, en comparant les trajets du MdLF et du faisceau unciné (Figure 43), on s'aperçoit que la terminaison au niveau temporal du faisceau unciné est plus antérieure et légèrement plus médiale que celle du MdLF ; très clairement, ces deux faisceaux ne connectent pas les même régions anatomiques (BA21 pour le MdLF et BA38 pour l'unciné dans notre étude).

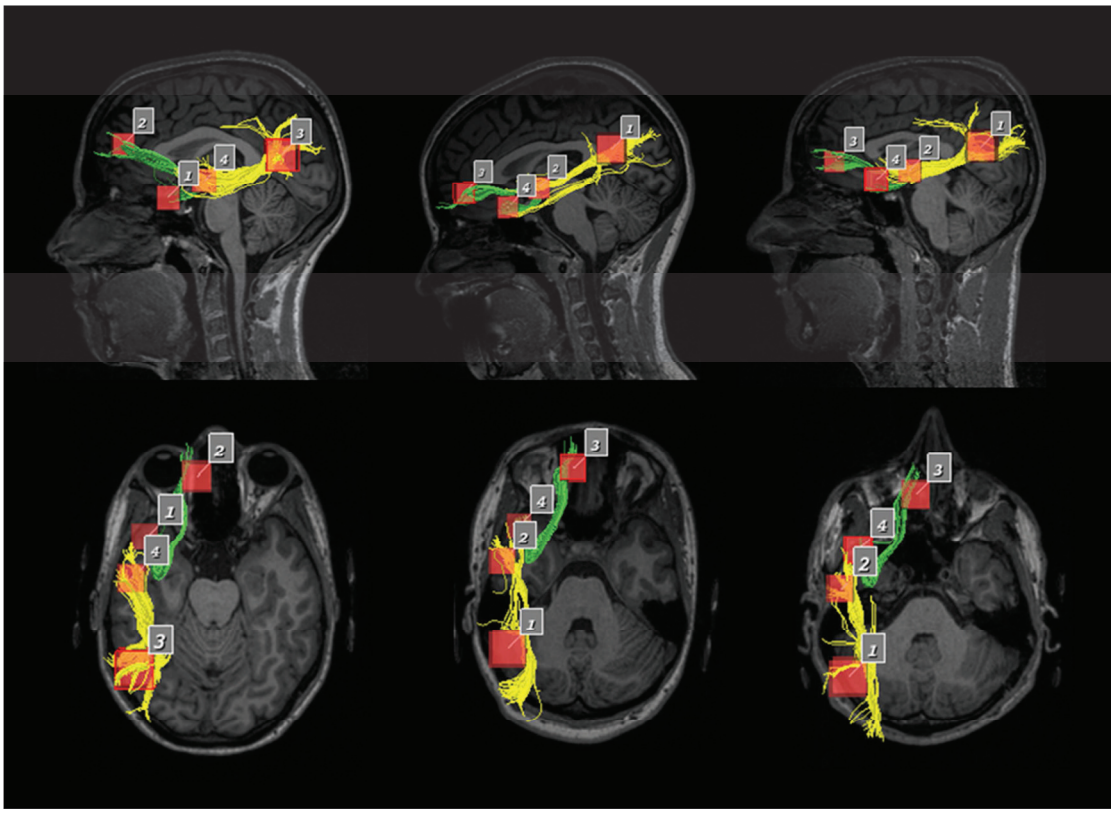


Fig 43) Rapports du MdLF (jaune) et de l'UF (en vert)
chez 3 sujets en coupes sagittale et horizontale

Les fibres du MdLF et de l'IFOF se croisent au niveau du stratum sagittal mais le MdLF est situé latéralement par rapport à l'IFOF dans sa partie antérieure (au niveau temporal), puis il surcroise l'IFOF et se porte vers le haut et vers le dedans alors que les deux faisceaux se dirigent postérieurement vers les régions occipitales (Figure 44).

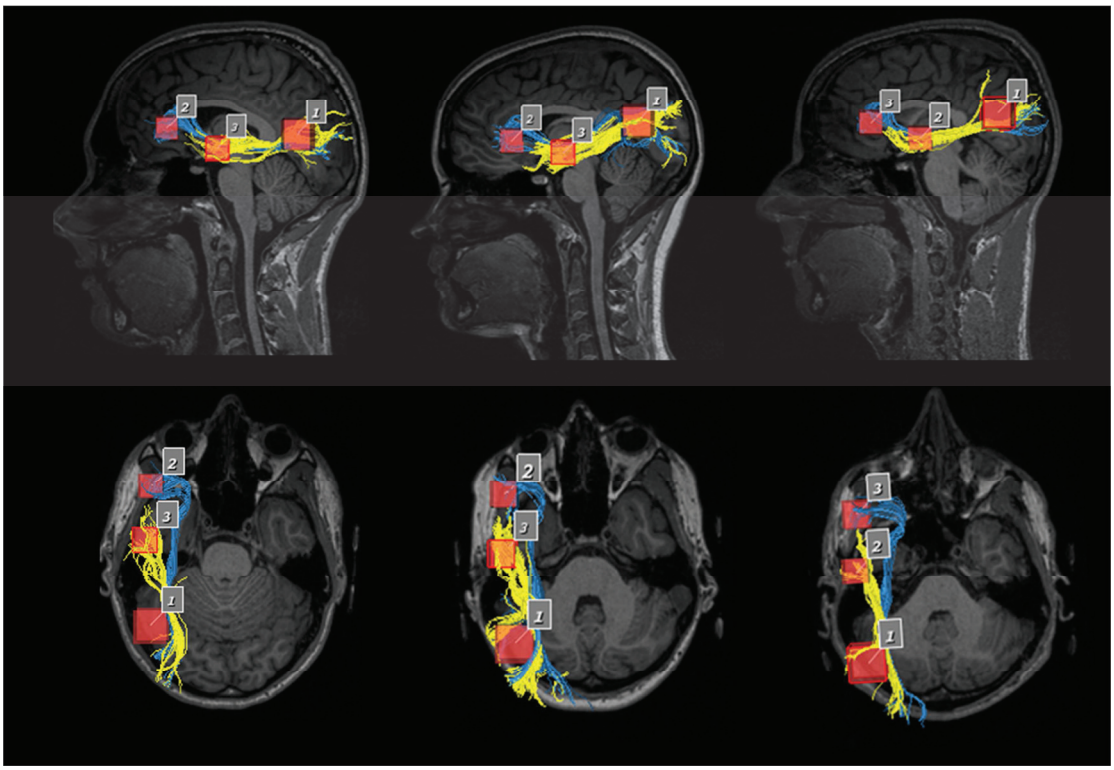


Fig 44) Rapports du MdLF (jaune) et de l'IFOF (bleu)
chez 3 sujets en coupes sagittale et horizontale

Enfin, nous avons voulu déterminer si nous pouvions séparer le MdLF du faisceau longitudinal inférieur (ILF). Pour ce faire, nous nous sommes inspirés de l'étude de Menjot de Champfleur et al. (2012) où, bien que la méthode de définition des ROIs soit différente, deux ROIs étaient utilisées, pour initier la tractographie de l'ILF : une ROI définie dans la partie antérieure du lobe temporal et une ROI située au niveau occipital (la même que celle utilisée pour la tractographie de l'IFOF dans cette étude). Nous avons choisi d'ajouter une autre ROI qui n'était pas présente dans nos ROIs d'origine afin d'initier la tractographie de l'ILF. Cette ROI a été placée pour tous les sujets à la terminaison du faisceau IFOF, que nous avions pu observer en plaçant nos ROIs au niveau frontal (BA47/BA11) et au niveau temporo-pariétal (BA 39 TPJ). Pour la seconde ROI, nous avons choisi d'utiliser la ROI située à la partie la plus antérieure du lobe temporal (GTs BA38) car c'est aussi celle qui nous a permis de définir le faisceau unciné dans notre propre étude comme dans le travail de Menjot de Champfleur et al. (2012), où la même région temporale est utilisée pour l'unciné et l'ILF. Par ailleurs, nous avons aussi vu que le faisceau supposé être le MdLF était bien mieux caractérisé par une terminaison, au niveau temporal, légèrement plus postérieure et latérale (GTs BA21) puisque seulement quelques fibres se projettent au niveau de BA38.

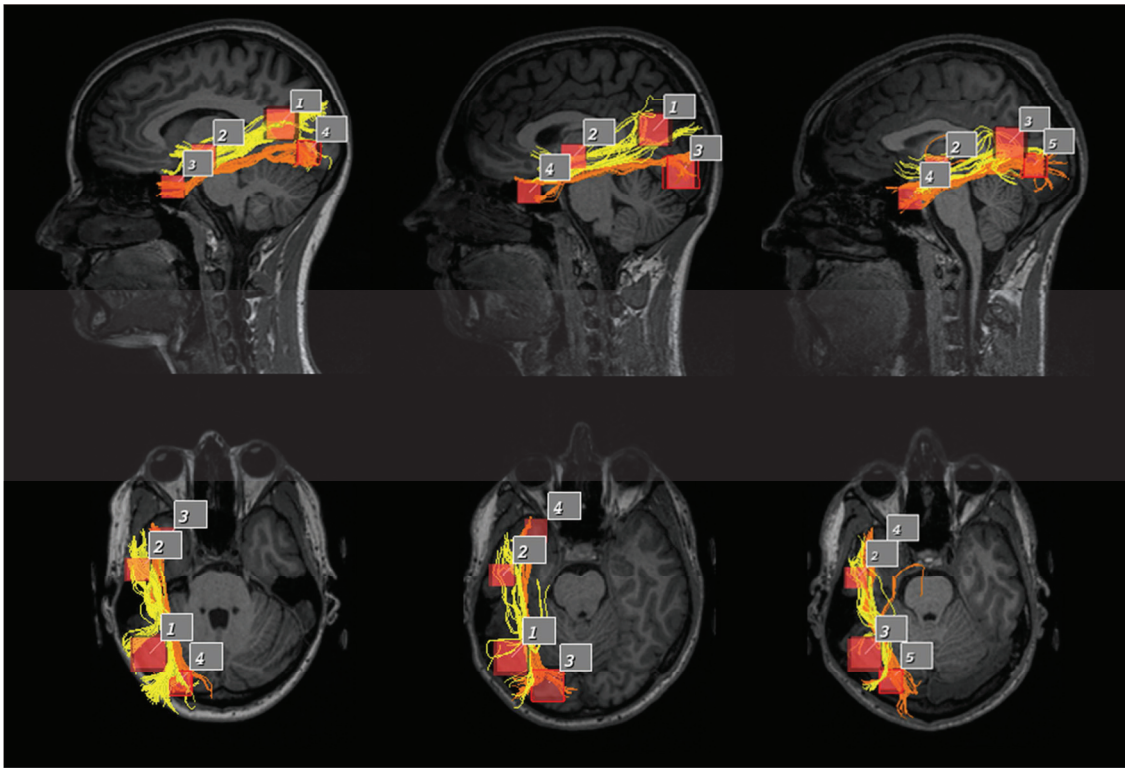


Fig 45) Rapports du MdLF (jaune) et de l'ILF (orange)
chez 3 sujets en coupes sagittale et horizontale

Il semble que cette méthode de caractérisation pour séparer MdLF et ILF soit plutôt pertinente, puisque nous avons effectivement trouvé deux faisceaux distincts chez 16 de nos sujets. Comme dans l'étude de Menjot de Champfleur et al. (2012), la course de ces faisceaux est parallèle et l'ILF est toujours situé en position ventrale par rapport au MdLF, comme le montre la figure 45.

- *Connectivité inter-hémisphérique*

N'oublions pas que notre réseau implique des activations bilatérales en IRMf, avec en particulier une composante postérieure située au niveau des TPJ et une composante antérieure avec des activations bilatérales du cortex orbitofrontal médian. La communication inter hémisphérique semble donc importante pour ce réseau. Cette communication pourrait être rendue possible, au sein de notre réseau, par deux faisceaux appartenant au corps calleux : postérieurement, au niveau du tapetum et antérieurement au niveau du forceps minor (figure 46).

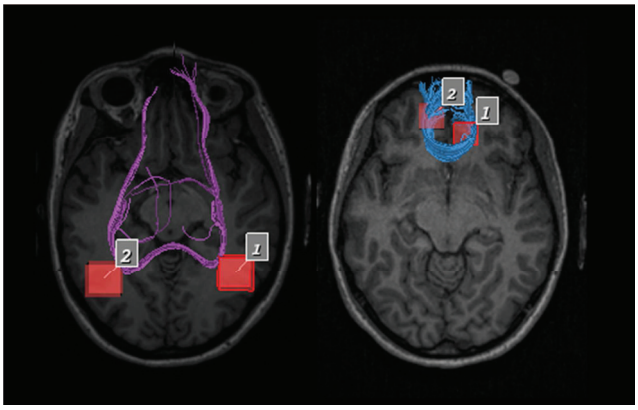


Fig 46) Connectivité interhemisphérique postérieure via le tapetum (à gauche) et antérieure via le forceps minor (à droite)
chez 1 sujet en coupe horizontale

En effet, les fibres qui transitent au niveau du tapetum relient médialement les deux ROIs situées dans le TPJ, elles permettraient ainsi un transit interhémisphérique postérieur de l'information entre ces deux « hubs » (Turken & Dronkers, 2011). Cependant, on ne retrouve ce faisceau distinctement que chez seulement 12 de nos sujets et il semble ne rejoindre que la partie très médiale des ROIs temporo-pariétales. Toutefois, les limitations dues au DTI pour séparer les croisements de fibres rend difficile la caractérisation de ce faisceau, car les fibres de celui-ci se confondent avec d'autres projections plus longues, à destination du cortex préfrontal.

En revanche, la connexion inter-hémisphérique antérieure reliant les deux ROIs des Gyri frontaux médians (GFm BA 11 HG et GFm BA 11 HD) est plus nette et forme un faisceau de fibres courtes chez nos 19 sujets, faisceau que nous identifions comme faisant partie du forceps mineur du corps calleux par sa forme facilement reconnaissable de « U » en avant. Ces fibres forment un faisceau

relativement dense chez tous nos sujets, avec une moyenne de 188 fibres pour ce faisceau (soit 0.24% de l'ensemble des fibres du cerveau). Le cortex orbitofrontal médian pourrait donc être une zone de convergence où l'information serait relayée entre les deux hémisphères par l'intermédiaire des fibres du forceps mineur.

- *Connectivité cortico-striatale*

Enfin, nous avons également mis en évidence une connectivité cortico-striatale au sein de ce réseau puisque le cortex orbitofrontal médian (BA11) est relié bilatéralement au noyau caudé (tête du caudé) par des faisceaux de fibres courtes (Figure 47). On retrouve ce faisceau cortico-striatal chez les 19 sujets et dans les deux hémisphères même s'il apparaît plus dense à gauche (61 fibres en moyenne pour l'HG et 22 fibres pour l'HD). Ces résultats sont cohérents avec les projections striato-frontales décrites dans la littérature et qui relient la tête du noyau caudé au cortex préfrontal (Robinson et al., 2012 ; Kamali et al., 2010), projections également décrites comme particulièrement denses (Afifi et Bergman, 1998).

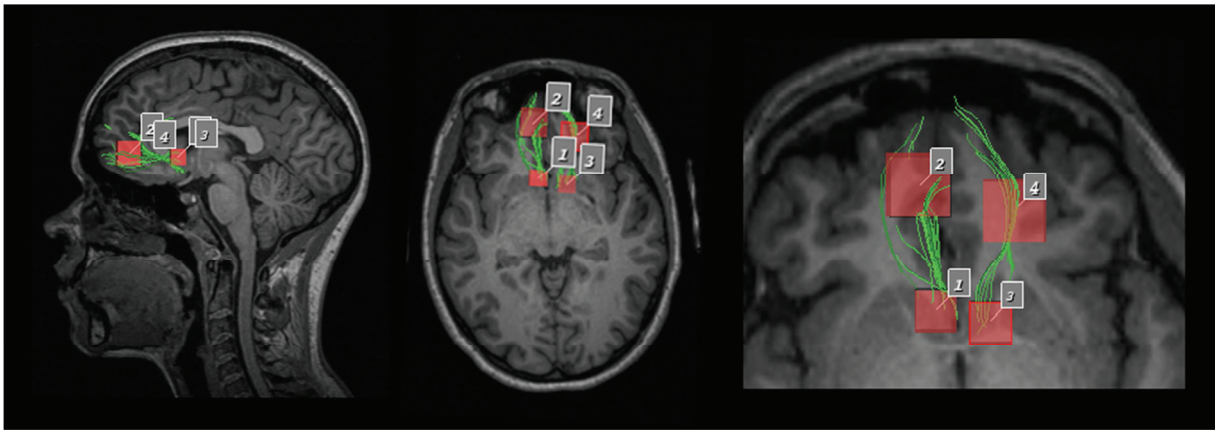


Fig 47) Projections bilatérales striato-frontales entre le noyau caudé et le cortex orbitofrontal médian (BA11) chez 1 sujet en coupes sagittale et horizontale

- **Synthèse résultats DTI**

Grâce à ces analyses DTI, nous avons montré que le réseau sémantique mis en évidence par IRMf était soutenu par une architecture de matière blanche cortico-corticale mais aussi cortico-striatale. Plus précisément, la connectivité de ce réseau est soutenue par l'IFOF qui assure les connexions fronto-pariétales, le faisceau unciné qui relie les régions temporales antérieures au cortex préfrontal et probablement le MdLF qui unit les régions pariétales et temporales. Nous avons montré que le TPJ semblait être une zone de convergence particulièrement importante au sein de ce réseau puisque, de par sa connectivité, il interconnecte les différentes zones du réseau fronto-temporo-pariétal (temporales, motrices, frontales). Ce réseau est aussi soutenu par une connectivité

interhémisphèrique, en avant et en arrière du corps calleux ainsi que par des connexions fronto-striatales.

D) Résultats Partie IV : résultats comportementaux

- **Mise en évidence du coefficient de représentabilité**

Cette dernière partie s'articule autour d'un résultat que nous avons obtenu en faisant passer un questionnaire post-test aux sujets ayant participé au protocole IRMf-DTI. Rappelons qu'il était demandé aux sujets de donner un score de « représentabilité » pour 50 phrases, sur une échelle de 1 à 5, de 1= très difficile, 2=difficile, 3=intermédiaire, 4=facile, 5= très facile.

Au fur et à mesure des passations, deux populations se sont dégagées parmi les sujets : certains sujets répondaient très rapidement et donnaient presque systématiquement la note 5 ; d'autres, au contraire, prenaient leur temps pour répondre et la difficulté qu'ils éprouvaient face à certaines phrases se reflétait dans les réponses par un pattern de notes variées. A partir de ces premiers résultats, nous avons défini la notion de Coefficient de Représentabilité (COR=Coefficient of Representability) et de bons/mauvais « représentateurs ». La figure 48 représente la répartition des scores entre les deux populations. Sans aucun jugement de valeur dans la notion de bon ou mauvais « représentateurs », ces différences entre les individus relèveraient plutôt de stratégies cognitives différentes, au même titre que certains ont une meilleure mémoire visuelle, sont meilleurs en calculs mental ou en géométrie dans l'espace.

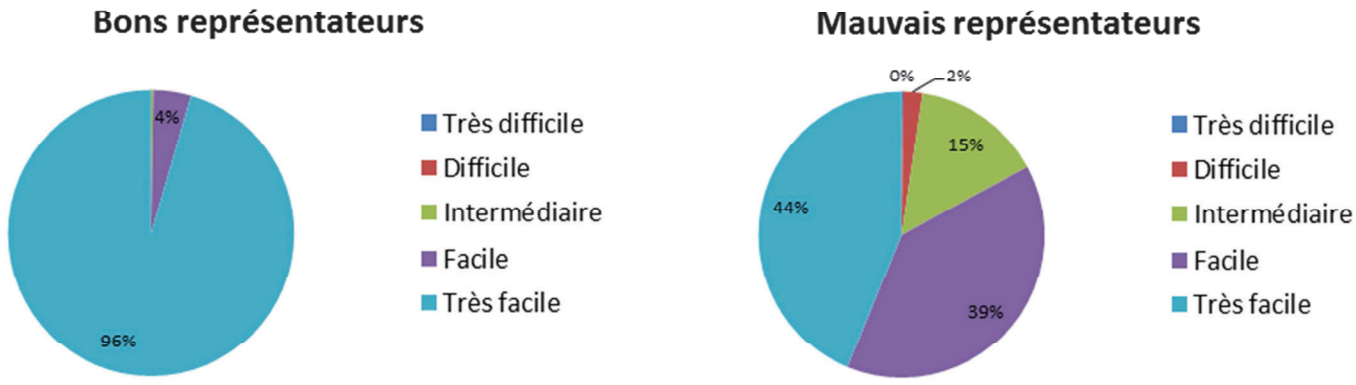


Fig 48) Répartition des réponses COR entre les bons et les mauvais représentateurs

Nous avons cherché à valider statistiquement ces observations. Des tests statistiques très significatifs (Student=5,95743E-05 ; Test F=2,98396E-05) nous ont permis de constater que les 2 groupes de sujets étaient significativement dissociables (8 « bons représentateurs » vs 11 « mauvais

représentateurs »). Il y a une réelle différence statistique entre les deux groupes, illustrée par la figure 49 : la moyenne du groupe des « bons représentateurs » est de 4,95 et l'écart type est peu dispersé (0.05) alors que la moyenne des « mauvais représentateurs » est plus basse (4,24) et l'écart-type plus dispersé (0,36). Nous avons vérifié la pertinence de notre classification en bons et mauvais représentateurs en effectuant, pour chaque sujet, un test statistique de comparaison de la moyenne par rapport au score 5 (unilatéral : H0 : moyenne = 5 et H1 : moyenne < 5). Si H0 n'est pas rejetée, le sujet est considéré comme un « bon représentateur ». Cette analyse a confirmé que notre classification des sujets en deux groupes était représentative à p=0.01.

	Bons «Représentateurs »	Mauvais «Représentateurs »
	5	4,04
	5	4,38
	5	4,32
	4,9	3,86
	4,9	4,56
	5	3,46
	4,92	4,66
	4,9	4,06
		4,34
		4,64
		4,34
MOY	**4,95**	**4,24**
ETY	**0,05**	**0,36**

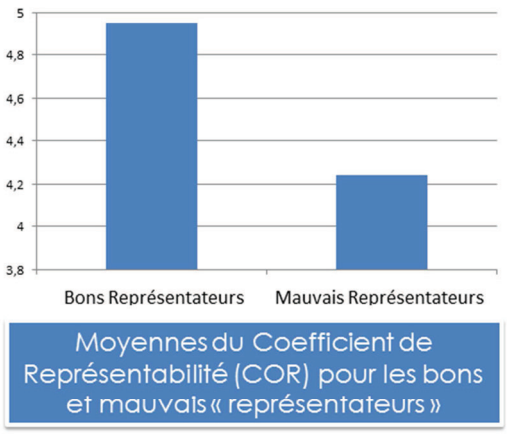


Fig 49) Scores COR et moyennes du COR pour les bons et les mauvais représentateurs

- **<u>COR et données de neuroimagerie</u>**

Nous avons ensuite voulu savoir si ces différences comportementales pouvaient se refléter dans nos données de neuroimagerie. Si la mise en relation de ces résultats avec les données IRMf n'a pas apporté grand-chose (voir Annexe 3), ce que nous avons trouvé par rapport au DTI est beaucoup plus intéressant.

Nous avons commencé par une simple étude de contrastes sur les données IRMf en tenant compte de la séparation en deux groupes selon les capacités de représentation des sujets (contraste « bons » vs « mauvais représentateurs ». Ces résultats n'ont pas montré de différence particulière au niveau des activations IRMf. Cette absence de différence au niveau fonctionnel est peut-être due au fait que les deux groupes de sujets ne contenaient pas un nombre assez important d'individus (8 bons et 11

mauvais représentateurs), des analyses sur un plus grand nombre de sujets seraient à envisager. Une analyse complémentaire (analyse statistique ANOVA sur les β, pourcentage de changement du signal pour les ROIs issues de l'analyse de conjonction) a donné quelques résultats pour le noyau caudé et le gyrus temporal supérieur (résultats consultables en annexes).

Par ailleurs, nous avons mis en lien nos données issues des analyses individuelles DTI et nos résultats comportementaux. Nous avons cherché s'il existait des corrélations significatives (logiciel Statistica) entre la quantité de fibres d'un faisceau et le score de représentabilité COR, pour tous les faisceaux dont nous avions initié la tractographie par une approche à deux ROI (même ceux dont les résultats n'ont pas été présentés ici car les faisceaux n'étaient présents que chez moins de 15 sujets). Plutôt que d'utiliser le nombre de fibres d'un faisceau, nous avons préféré utiliser le ratio de fibres (rapport entre le nombre de fibres de ce faisceau et le nombre total de fibres pour le cerveau d'un sujet).

Parmi tous les faisceaux que nous avons testés, un seul faisceau présente une corrélation positive significative, qui est de 0.56 ($F = 7.34$, $p = 0.014$), avec le COR : il s'agit du faisceau entre la ROI TPJ BA 39 de l'HG et la ROI GTm BA21 de l'HG, le faisceau que nous présumons être le MdLF. Plus le MdLF comprend de fibres, meilleures sont les capacités de représentation du sujet (traduites par le COR) (Fig 50).

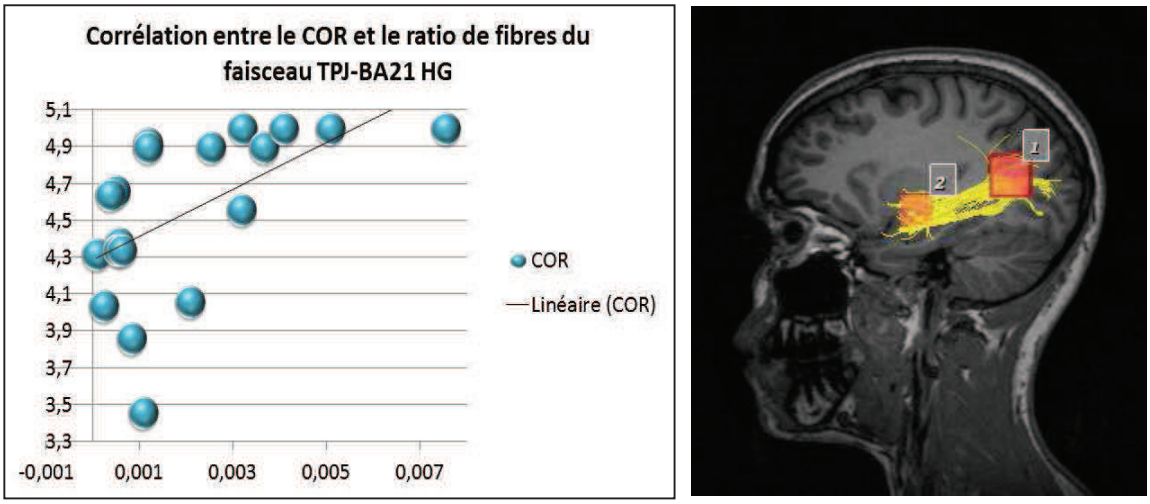


Fig 50) Corrélation entre le COR et le ratio de fibres pour le MdLF

Fait intéressant, la plupart des autres corrélations sont bien inférieures (voire proches de zéro) et aucune d'entre elles n'est significative, comme le montre la table 6.

TPJ-BA21 HG	MdLF HG	**0,566421053**
TPJ-BA21 HD	MdLF HD	0,305756133
TPJ-BA47 HG	IFOF term latérale HG	-0,13887198
TPJ-BA47 HD	IFOF term latérale HD	0,007467559
TPJ-BA11 HG	IFOF term médiale HG	0,257383504
BA38-BA47 HG	IFOF term médiale HD	0,2157083
BA38-BA11 HG	unciné	-0,246802939
BA11-caudate HG	cortico-strital HG	-0,0868633
BA11-caudate HD	cortico-strital HD	-0,082671471

Table 6) Corrélation entre le COR et les différents faisceaux mis en évidence par la Tractographie

Nous avons essayé de caractériser un peu mieux les différences existantes pour ce faisceau entre bons et mauvais représentateurs en effectuant des tests t-Student entre les différents paramètres liés à ces faisceaux (nombre de fibres, FA moyen, FA médian, FA max). Le FA (pour anisotropie fractionnelle) est un indice très utilisé pour interpréter les données de DTI : il caractérise le manque d'uniformité des mouvements de diffusion dans les différentes directions de l'espace (milieu isotrope, le FA tend vers 0 et milieu anisotrope, le FA tend vers 1). La substance blanche présente naturellement un certain degré d'anisotropie et cet indice est très employé en pratique clinique car en cas de processus lésionnel, la perte d'anisotropie du milieu se traduit par une chute de la FA. Chez le sujet sain, certains auteurs condidèrent que le FA peut permettre d'évaluer la qualité de la communication entre deux régions cérébrales (Thiebaut de Schotten et al., 2012). Dans notre étude, à part le nombre de fibres, le seul autre paramètre où il existe une différence significative est le FA max qui est plus important chez les bons représentateurs. On peut envisager qu'une partie des fibres du MdLF aurait un FA plus important chez les bons représentateurs uniquement, ce qui pourrait expliquer leurs meilleures capacités de représentation par une meilleure connectivité entre les régions temporales et pariétales.

	Nb fibres	FA moyen	FA médian	FA max
Bons	237,81	0,46	0,40	0,55
Mauvais	72,8	0,45	0,45	0,49
Test T	0,0099	0,3185	0,5228	0,0070

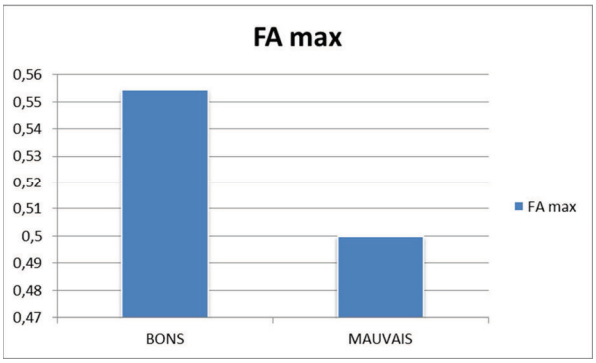


Fig 51) Différences caractérisant le MdLF entre bons et mauvais représentateurs (nombres de fibres et FA)
Les valeurs significatives au test de Student sont représentées en rouge

II.1.*6*) Conclusion et perspectives

Dans ce protocole, nous avons présenté à des sujets sains des phrases et des images, décrivant ou représentant des événements riches mettant en scène des humains, dans le but d'identifier les substrats neuronaux de la représentation du sens, indépendamment de la modalité de présentation. Ce travail s'inscrivant dans le domaine de la cognition incarnée, nous faisions l'hypothèse que cette approche nous permettrait de caractériser les réseaux sensori-moteurs qui soutiennent la compréhension de ce type de représentations (Desai et al 2010). Nous avons effectivement retrouvé une large implication des systèmes sensorimoteurs et ce, de façon commune pour le traitement des images et des phrases. Bien sûr, il y a des différences entre images et phrases, mais surtout en termes d'extension des patterns d'activités (les patterns sont plus étendus pour les images, principalement au niveau des aires visuelles en occipital). A l'instar de certains auteurs (Buccino et al., 2005 ; Bright et al. 2004), nous pensons que ces différences de patterns d'activations cérébrales pourraient refléter des étapes de traitement pré-sémantiques différentes selon la modalité et que dans le cas des images, il pourrait y avoir moins d'ambiguïté à propos des détails de l'action qu'il faut simuler, ce qui produirait un effet de facilitation lié à la vision. Cependant, l'analyse de conjonction que nous avons effectuée révèle un large réseau d'aires fronto-temporo-pariétales que nous identifions comme étant le système sémantique commun aux deux modalités. De façon tout à fait cohérente avec la littérature (même avec les travaux les plus précoces sur la comparaison images/mots, cf Vandenberghe et al. 1996), nous avons pu observer l'implication d'un large réseau commun majoritairement latéralisé à gauche, impliquant le gyrus frontal inférieur et le gyrus frontal médian, le cortex temporal inféro-moyen ainsi que le gyrus parahippocampique, la jonction temporo-pariétale bilatéralement, le gyrus précentral gauche... Nos résultats semblent corroborer plus fortement les résultats des travaux qui montrent un recoupement important des régions cérébrales impliquées dans la compréhension des deux types de stimuli (par exemple : Grèzes & Decety, 2001 ; Hamzei et al., 2003) et vont à l'encontre de ceux qui s'opposent à la notion d'une simulation commune à l'observation des actions et à la compréhension des phrases d'actions (ex : Tremblay et al. 2011). Ces résultats semblent montrer l'implication d'une composante motrice importante qui est associée au MNS puisqu'on retrouve plusieurs des régions décrites comme contenant des neurones miroirs chez l'homme : le cortex prémoteur (BA6), l'aire de Broca (GFi pars triangularis BA45) et notre activation située au niveau du TPJ gauche s'étend jusqu'au LPI (LPI, aire dans laquelle on reconnaît classiquement l'existence de neurones aux propriétés miroirs chez l'homme).

Cependant, nos résultats montrent l'activation d'un réseau cortico-striatal beaucoup plus étendu, ce qui laisse penser que la composante sensorimotrice ne serait qu'un représentant de la simulation à la base de la compréhension sémantique commune des images et des phrases. C'est pourquoi, nous ne

nous limiterons pas à cette simulation sensorimotrice pour interpréter nos résultats ; nous présenterons, au contraire, une vision plus large de la cognition incarnée où différentes régions pourraient jouer des rôles spécifiques dans le processus de simulation.

- **Compréhension et Modèles mentaux**

Les premiers modèles de la cognition incarnée (Barsalou, 1999) faisaient référence aux modèles mentaux de Johnson-Laird, modèles qui ont été décrits dans le contexte de la compréhension de textes et qui reposent sur un haut niveau de cognition (Johnson-Laird, 1983). Ces modèles mentaux peuvent être construits à partir d'inputs visuels et/ou linguistiques, et ils permettent de se représenter l'état du monde, une capacité de simulation qui est à la base du raisonnement (Johnson-Laird, 2010). Pour Barsalou, cette notion de simulation perceptuelle fournit un cadre réaliste pour la façon dont les personnes construisent le sens, comme dans les modèles mentaux de Johnson-Laird. Ces conceptualisations « incarnées » feraient intervenir une simulation basée sur 4 types d'informations particulières : (a) la perception des personnes et des objets, (b) des actions, (c) l'introspection et (d) création d'une scène visuospatiale (Barsalou, 2009). Pour Johnson-Laird et ses collègues (Johnson-Laird, 2010 ; Knauff et al., 2003), ce sont ces modèles mentaux et les simulations correspondantes qui fourniraient la base à une cinquième perspective permettant le raisonnement (e). Dans cette perspective élargie, une capacité de compréhension incarnée ne serait pas limitée à la simulation sensorimotrice, mais inclurait ces dimensions supplémentaires et recruterait un réseau plus large. Étonnamment, on observe que les régions du cerveau qui, dans notre étude, présentent une activité importante pour les phrases et les images correspondent à des systèmes neuronaux impliqués dans ces cinq dimensions fonctionnelles.

(a) Aires temporo-pariétales et perception des objets et des personnes

Dans notre étude, les phrases et les images correspondent à des événements impliquant des personnes humaines et souvent des objets. Ces événements ont généré des activations très étendues au niveau du lobe temporal postérieur (GTm BA39), des activations qui s'étendent jusqu'à la jonction temporo-pariétale TPJ. Ce sont des aires qui ont été fortement associées à la mémoire sémantique (Binder & Desai 2011) y compris celle liée aux personnes et aux objets. Le lobe temporal a souvent été associé au traitement conceptuel des objets puisqu'on retrouve généralement une perte multimodale des connaissances concernant les objets chez les patients atteints de démence sémantique, chez qui cette région est touchée (Hodges, et al., 1992 ; Mummery et al., 2000). Toute une série de régions situées entre la partie plus antérieure du lobe temporal et le TPJ ont des propriétés associées à la perception et à la représentation du corps ou de parties du corps : l'EBA (Downing et al., 2001, 2006; Urgesi et al., 2004 ; Astafiev et al., 2004), la FFA (Sergent et al., 1992 ;

Kanwisher et al., 1997). Ces différentes régions temporo-pariétales pourraient faire le lien entre les éléments perçus ou représentés et les concepts stockés en mémoire sémantique.

(b) *Représentation des actions au niveau pariétal et prémoteur*

Le réseau sémantique, observé dans notre étude, implique des activations étendues comprenant le cortex pariétal sensorimoteur (les aires 37, 39 et 40), le cortex prémoteur. Comme nous l'avons déjà largement mentionné, l'idée qu'une représentation mentale de l'action recrute des zones motrices a été évoquée dans de nombreux contextes. Decety (Decety et al, 1994) a démontré que les aires corticales motrices sont activées lors de l'imagerie mentale de l'action. Nous avons déjà évoqué la découverte des neurones miroirs au sein des cortex prémoteur et pariétal, qui codent à la fois nos propres actions et les actions observées effectuées par autrui (di Pellegrino et al., 1992 ; Buccino et al., 2004, Rizzolatti & Sinigaglia 2010) et les liens entre MNS et langage, à travers le fait que ce système participait également à la représentation du sens d'une action dans le langage (Tettamanti et al., 2005). Nos résultats sont cohérents avec les activations pariétales et prémotrices décrites dans la littérature et l'interprétation actuelle des théories de la cognition incarnée (Barsalou, 1999 Barsalou, 2009) qui suggèrent que la simulation incarnée sous-tend non seulement la compréhension du langage, mais aussi les aspects les plus profonds de traitement conceptuel et de cognition sociale.

(c) *Réseau du mode par défaut (Default Mode Network) et Introspection*

Le résultat le plus surprenant est peut-être le chevauchement important de notre réseau sémantique avec le « Default Mode Network » (Andrews-Hanna et al., 2010 ; Buckner et al., 2008 ; Raichle et al., 2001 ; Raichle et Snyder, 2007). Dans un important travail de méta-analyse des études d'activations cérébrales liées à la mémoire autobiographique, la prospection, la navigation, la théorie de l'esprit et le mode par défaut, Spreng et al. (Spreng et al., 2009) ont déterminé qu'un même système neuronal était impliqué dans l'ensemble de ces processus, processus précédemment considérés comme distincts, confirmant ce que d'autres travaux avaient pressenti (Buckner & Carroll, 2007). Ce système fait intervenir le lobe temporal médian, le précuneus, le cortex cingulaire postérieur et notamment le cortex rétrosplénial ainsi que la jonction temporo-pariétale et le cortex préfrontal médian. En général, l'activation du réseau du mode par défaut est associée à des tâches qui encouragent le sujet à se tourner vers une mentalisation interne, y compris des souvenirs faisant appel à la mémoire autobiographique, une réflexion sur son avenir, la théorie de l'esprit ou encore une prise de décision auto-référentielle et affective (Andrews-Hanna et al, 2010). Le fait que ce réseau partagé par des processus aussi différents (mémoire, navigation, TOM…) soit le même que celui du Default Mode Network laisse penser que le rôle de ce réseau serait la capacité à se représenter soi-même (Buckner & Carroll, 2007) ou un mécanisme permettant de construire la représentation d'une scène (Hassabis

& Maguire, 2007). Ces auteurs estiment que ce réseau pourrait permettre l'auto-projection : la capacité à se projeter mentalement dans le moment présent en une simulation d'un autre temps, de lieu ou de perspective, ce qui expliquerait en particulier les liens entre ce réseau et mémoire autobiographie (Spreng, 2009). Nos résultats indiquent effectivement que la compréhension des scènes illustrant des activités humaines (par le biais des phrases ou d'images) recrute également en partie ce système puisqu'un certain nombre d'aires cérébrales sont les mêmes que celles impliquées dans le Default Mode Network (cortex préfrontal médian, TPJ, RSC, précuneus et lobe temporal médian). Cela offre de nouvelles perspectives à la notion de compréhension incarnée. Pour Bar (Bar, 2009), comprendre nécessite d'intégrer le stimulus « en cours de compréhension » à un système proactif de mémoire/prédictions. En effet, pour cet auteur, la reconnaissance (que nous pouvons étendre à la compréhension) consiste à déterminer « A quoi cela ressemble ? » (« What is this like ?»), ce qui nécessite un lien proactif avec des situations existant en mémoire. Bar a suggéré que mentaliser durant les périodes de « repos » pourrait être le résultat d'une simulation mentale d'événements possibles, qui créerait des sortes de « souvenirs » pouvant être utilisés ultérieurement pour guider les comportements. Ce qui est intéressant dans notre étude, c'est qu'à la fois entrées linguistiques et imagées pourrait servir de déclencheurs pour activer ces processus de simulation, processus qui initieraient la compréhension.

(d) Représentation de la perspective spatiale au niveau du RSC

Le réseau sémantique, mis en évidence dans cette étude, comprend également une activation important du RSC qui est très intéressante du point de vue de la compréhension de scènes visuelles représentées en images ou décrites verbalement. Si, comme nous l'avons mentionné ci-dessus, le RSC appartient à un réseau plus large d'aires associées à de nombreuses fonctions cognitives supérieures (TOM, navigation…), le RSC a été notamment identifié comme jouant un rôle important dans des processus de la visualisation et la manipulation de scènes visuospatiales (Vann et al., 2009). La construction d'une représentation de la scène en cours de compréhension (à partir d'une phrase ou d'une image dans notre étude) nécessiterait la génération d'une scène visuospatiale. Ce serait le RSC (et probablement aussi l'hippocampe, région également activée bilatéralement dans notre étude) qui permettrait de générer mentalement et de manipuler une scène complexe, ou encore de gérer la cohérence de cette dernière.

(e) Raisonnement et structure temporelle sous la dominance des régions frontales

Plus qu'un traitement passif, la compréhension reposerait sur des processus constructifs permettant d'aboutir au raisonnement. Johnson-Laird et ses collègues (Knauff et al., 2003) ont étudié comment l'imagerie mentale pouvait faciliter le raisonnement par déduction sur différents aspects de relations

visuo-spatiales décrites dans des phrases (relations visuospatiales faciles à envisager spatialement et visuellement « au dessus / en dessous », faciles à envisager visuellement et difficiles à envisager spatialement « plus propre/plus sale »…). Ils ont observé qu'un réseau ressortait, malgré ces différentes conditions d'imagerie : il implique des activations dans les cortex pariétaux, temporaux et frontaux. Ces résultats sont proches des nôtres puisqu'on retrouve, parmi les activations mentionnées dans cette étude, une implication des gyri temporaux moyen et supérieur gauches (BA21-BA38) ainsi que des gyri frontaux moyen et inférieur gauches (BA46-BAA47). Fait intéressant, nous avons également observé des activations supplémentaires au niveau du cortex préfrontal médial (BA11). Dans le langage, cette aire a été associée aux traitements des aspects liés à la structure grammaticale (Kuperberg et al., 2008), au fait de combiner l'information de plusieurs phrases pour obtenir un sens cohérent au niveau du discours (Xu et al., 2005) et du traitement de la structure temporelle des événements (Krueger et al., 2009a, Krueger et al., 2009b). Ceci est cohérent avec le modèle de compréhension proposé dans d'autres travaux réalisés par notre équipe où un système langagier cortico-striatal permet d'extraire les rôles thématiques et d'activer une simulation au niveau des aires sensorimotrices temporo-pariétales (Dominey et al., 2006 ; Hinaut & Dominey, 2013), tous ces éléments étant sous la dépendance des aires frontales et de leur rôle de planification structurelle (Lallée et al., 2010 ; Madden et al., 2010).

- **Connectivité et zones de convergence**

Un aspect commun aux tâches qui nécessitent traitement sémantique, cognition sociale, mémoire épisodique, construction d'une scène, et connaissances de soi, c'est la récupération des connaissances conceptuelles. Binder et Desai (Binder & Desai 2011) suggèrent qu'une partie de la raison pour laquelle des processus si divers pourraient faire appel au même système d'aires fronto-temporo-pariétales, c'est que ces processus reposent tous sur la récupération d'un contenu conceptuel. Ils affirment que la mémoire sémantique est soutenue par la convergence de l'information à travers de larges régions des cortex associatifs temporal et pariétal inférieur, ce que semble confirmer nos résultats de connectivité en DTI. Dans notre étude, de par leur connectivité anatomique et/ou fonctionnelle, plusieurs régions semblent avoir un rôle clé au sein du réseau sémantique (si ce n'est un rôle de zone de convergence). Il s'agit de la jonction temporo-pariétale, de la région temporale antérieure, du RSC, du cortex préfrontal médial et au niveau sous-cortical, du noyau caudé.

- *Jonction temporo-pariétale*

Depuis les premiers travaux de Vandenberghe sur les mots et les images (Vandenberghe et al 1996) jusqu'aux méta-analyses récentes sur le système sémantique (Binder & Desai 2011, Binder et al., 2009, Graves et al., 2010), la jonction temporo-pariétale, et en particulier le gyrus angulaire, apparaît comme une région cruciale pour la compréhension du sens. Il pourrait jouer ce rôle de « hub intégratif » car c'est un cortex associatif avec un rôle d'intégration multisensorielle : Binder suggère que le gyrus angulaire (ainsi que la région, plus large, du TPJ) pourrait être une zone de convergence de haut niveau où les informations de plusieurs modalités convergent pour former une représentation cohérente. C'est ce qui lui donnerait la possibilité d'intervenir dans de nombreuses fonctions cognitives liées à l'action (la compréhension des buts et des intentions d'une action, l'agentivité, la conscience de soi et de ses propres actions, l'imitation, la localisation dans l'espace, etc…). Le TPJ aurait aussi une situation privilégiée en tant que hub, de par sa localisation anatomique à l'intersection des pôles pariétal et temporal et sa proximité avec de nombreuses aires adjacentes impliquées dans la compréhension des actions. Conformément à cette interprétation, nos résultats montrent que le TPJ semble être une importante zone de convergence comme en témoignent les nombreuses voies de substance blanche qui connectent cette région aux autres régions qui nous intéressent, notamment un réseau majeur à trois branches issues du TPJ (le TPJ est lié au pôle temporal par le MdLF, au pôle frontal via l'IFOF et au cortex prémoteur via le faisceau longitudinal supérieur).

Nos données issues de l'analyse ICA confirment également l'importance de cette région au sein du réseau sémantique. Par ailleurs, ces analyses apportent un autre élément de réponse intéressant puisqu'elles montrent l'importance de l'activation bilatérale de cette région : en effet, en plus d'être l'une des seules régions corticales activées bilatéralement avec le gyrus temporal moyen et le cortex préfrontal médian dans notre réseau principalement latéralisé à gauche, l'analyse ICA révèle qu'il existe une corrélation significative entre les décours temporels des régions TPJ situées dans les deux hémisphères (GTm39 HG et GTm39 HD). Il semble donc que la coordination entre ces deux régions, homologue controlatéral l'une de l'autre, soit importante pour le réseau sémantique. Notre étude ne nous permet pas actuellement de déterminer si TPJ gauche et droit ont des rôles distincts au sein du réseau sémantique (ce que laissent pourtant penser nos résultats montrant l'existence d'une corrélation significative entre le nombre de fibres des faisceaux reliant spécifiquement le TPJ et le GTm BA21 de l'hémisphère gauche) ; il sera intéressant de chercher à déterminer ces rôles dans nos prochains protocoles. Ainsi, un certain nombre de travaux suggèrent que le TPJ droit aurait un rôle spécifique dans le traitement des intentions et des états mentaux attribuables aux autres individus (Saxe & Kanwisher, 2003 ; Enrici et al., 2011 ; Carter et al., 2011 ; Sommer et al., 2007). Il ferait partie

d'un sous réseau plus large, incluant le cortex frontal médial et la partie postérieure du STS droits, ce réseau serait impliqué dans diverses tâches telles que le suivi de l'action, la détection et perception des protagonistes d'une action et des processus de mentalisation (Mason & Just, 2006, 2009 ; Amodio & Frith, 2006).

- *Région temporale antérieure*

Comme nous l'avons déjà indiqué dans l'introduction, certains auteurs considèrent le pôle temporal antérieur comme un « hub » intégratif en raison de son large rôle dans le traitement conceptuel (Simmons & Martin, 2009 ; Visser et al., 2010 ; Kiefer & Pulvermüller, 2012). Dans notre étude, la connectivité de cette région semble particulièrement importante puisque les régions temporales (GTs BA21 et GTm BA38) sont reliées aux régions frontales médiale et latérale (BA11 et BA47) par l'intermédiaire du faisceau unciné ; mais surtout le TPJ est relié au gyrus temporal supérieur (BA21) par le MdLF. Ce faisceau paraît avoir un rôle important dans le langage (Makris et al., 2009 ; Saur et al., 2008) et nos données semblent effectivement montrer que le MdLF a un rôle dans la compréhension sémantique puisque la quantité de fibres qui composent ce faisceau est corrélée avec la capacité des sujets à se représenter une phrase (COR). Nous discuterons plus en détail le rôle de ce faisceau ci-après.

La région temporale antérieure a longtemps été considérée comme un « hub » sémantique amodal. Cependant, nous préférons la vision proposée par Kiefer & Pulvermüller (2012), vision plus compatible avec la notion de fonctionnement en réseaux impliquant des zones de convergence multifonctionnelles, mais aussi avec les perspectives de la cognition incarnée. Selon ces auteurs, les connaissances conceptuelles seraient « incarnées » et représentées dans les zones sensorielles et motrices, tandis que le cortex temporal antérieur ne ferait que faciliter et influencer le traitement conceptuel dans ces zones, sans stocker les connaissances conceptuelles proprement dites. Cette interprétation est intéressante, surtout par rapport à nos données comportementales : peut-être que les mauvais représentateurs ont plus de difficulté à se représenter une phrase en raison d'une connectivité moins bonne entre les régions associatives pariétales et le cortex temporal antérieur (connectivité moins bonne qui se traduirait par un nombre de fibres moins important). Chez ces sujets, la facilitation s'effectuerait moins bien au niveau temporal et le traitement conceptuel, au niveau des zones sensorielles et motrices, en serait moins efficace.

- *Cortex rétrosplénial*

Nous avons déjà mentionné le rôle du RSC dans la création d'une scène visuo-spatiale, mais il apparaît aujourd'hui que cette région interviendrait également dans de nombreuses fonctions

cognitives de haut niveau (Vann et al., 2009), en particulier mnésiques, notamment avec la découverte qu'un dysfonctionnement de cette région était présent dès les stades précoces de la maladie d'Alzheimer. Un certain nombre de recherches ont montré que le RSC et l'hippocampe étaient activés dans des tâches de rappel des expériences passées et aussi lorsqu'il s'agissait d'imaginer des expériences personnelles futures plausibles ou des expériences fictives (Szpunar et al., 2007 ; Okuda et al., 2003 ; Botzung et al. 2008 ; Hassabis et al., 2007). Mais le RSC serait aussi impliqué dans des aspects de navigation spatiale, d'apprentissage et de mémoire épisodique (Vann et al., 2009) ; relier ces différents domaines pourrait apporter un éclairage nouveau sur le rôle du RSC. Une étude a utilisé l'IRMf pour examiner l'activation du cerveau lors de la récupération d'événements autobiographiques, d'épisodes tirés d'un film, ainsi que d'événements précédemment imaginés (Summerfield et al., 2009). Tout d'abord, les activations du RSC (BA29/30) sont clairement distinctes de celles des autres zones cingulaires postérieures (BA 23 et 31) ; ces dernières sont plus actives lors de la récupération d'événements réels que d'événements imaginés. En revanche, les RSC gauche et droit sont activés par les événements autobiographiques, réels ou imaginaires, ayant impliqué le participant personnellement. Il est possible que lors de la récupération d'événements autobiographiques (réels ou imaginaires), les participants doivent constamment mettre à jour leur « point de vue à la première personne » en fonction de l'évolution spatiale de l'environnement au fur et à mesure que l'événement se déroule. En revanche, la, perspective de l'observateur est « détachée » par rapport à la plupart des films et nécessite moins d'auto-référencement par rapport à un contexte spatial spécifique. Cette interprétation pourrait correspondre à la fonction proposée du RSC qui est d'effectuer une mise à jour entre cadres de références égocentriques (centré sur soi) et allocentriques (centré sur le monde). Une suggestion qui a été faite serait que le RSC utilise des informations actuelles pour traduire les représentations allocentriques (hippocampe) en représentations égocentriques (pariétales), ce processus serait nécessaire pour reconstituer une scène en mémoire ou en imaginer une nouvelle (Burgess et al., 2001 ; Byrne et al., 2007). Ce rôle du RSC permettrait de faire un lien entre ses fonctions de navigation, son implication dans la mémoire épisodique, la réflexion, la capacité à imaginer le futur… Cependant, les fonctions du RSC restent mal connues (souvent en raison d'une délimitation anatomique délicate).

Nos analyses ICA semblent, en tout cas, confirmer le rôle important de cette région au sein du réseau sémantique puisqu'il apparaît que la connectivité fonctionnelle entre le RSC gauche et le RSC droit est très importante, à la fois pour des phrases et pour des images. La capacité à faire émerger une représentation visuo-spatiale cohérente proviendrait peut-être de cette inter-coordination entre RSC gauche et droit. En ce qui concerne la connectivité réelle de cette aire, certains travaux en DTI (Greicius et al. 2009) ont mis en évidence des connections robustes entre le RSC et le cortex

temporal médial (hippocampe) et entre le RSC et le cortex préfrontal. Nos analyses DTI ne nous ont pas permis de retrouver ces connexions RSC-hippocampes et RSC-préfrontal, nous les avons parfois observées mais seulement chez un petit nombre de sujets. Ces éléments indiquent que nous devrons tenter de mieux caractériser la localisation de nos ROIs situées dans le RSC pour des analyses ultérieures de la connectivité de ces régions.

- *Projections entre le cortex préfrontal et le noyau caudé*

Le circuit neuronal reliant le cortex préfrontal et le striatum (noyau caudé et putamen) jouerait un rôle dans le contrôle cognitif (Casey et al., 2002). Le fait que, dans notre étude, on retrouve à la fois une connectivité anatomique et fonctionnelle entre ces deux régions serait cohérent avec les rôles parfois attribués à ces régions. L'activation du noyau caudé a été associée à la vérification d'informations et avec un rôle plus large de recherche et d'appariement avec des informations particulières stockées en mémoire (Marques et al., 2009). Le noyau caudé est aussi impliqué dans les processus de contrôle du langage (Friederici, 2006). Dans le réseau temporo-pariéto-frontal, les régions frontales (et en particulier le cortex préfrontal médial) joueraient un rôle de « chef d'orchestre » avec une influence spécifique sur les autres régions, notamment la simulation au niveau des aires sensorimotrices dans la compréhension, qui serait sous le contrôle des aires frontales et de leurs fonctions d'organisation structurelle et de planification (Lallée et al., 2010, Madden et al., 2010). L'activité du cortex préfrontal ventro-latéral a souvent été associée à un rôle de contrôle top-down de l'activité en mémoire sémantique, qui permettrait de guider la recherche et la sélection d'informations conceptuelles sous la dépendance de régions temporales postérieures et pariétales (Martin, 2007 ; Gold et al., 2005). La connectivité préfrontal-noyau caudé pourrait donc assurer ces mécanismes de contrôles, à la fois sur la recherche d'informations pertinentes en mémoire sémantique mais aussi sur la simulation sensorimotrice.

- *Voie ventrale sémantique*

Enfin, il faut noter que la connectivité que nous avons relevée pour le réseau commun, correspond tout à fait à la voie ventrale sémantique qui est généralement mise en évidence à travers les études en DTI (Mandonnet et al., 2007 ; Catani et al., 2002 ; Saur et al., 2008 ; Duffau, 2008 ; Menjot de Champfleur et al., 2012). En effet, dans notre étude, les principaux faisceaux de matière blanche impliqués sont l'IFOF, l'UF, l'ILF ou MdLF (la distinction de ces deux faisceaux et l'appartenance du MdLF à la voie ventrale étant controversées). En tout cas, nos résultats montrent que les faisceaux classiquement désignés comme composantes de la voie ventrale sémantique du langage sont aussi impliqués dans la compréhension sémantique des stimuli Images.

Rappelons que la voie sémantique ventrale du langage présente deux composantes : une voie directe qui connecte les aires temporales postérieures à la région orbitofrontale par l'intermédiaire de l'IFOF et une voie indirecte qui connecte la région occipito-temporale au pôle temporal par l'ILF, puis le pôle temporal aux aires basifrontales par le faisceau unciné. Cette seconde composante indirecte ne semble pourtant pas indispensable au langage puisque elle est compensée après résection ou lors des stimulations pre-opératoires (Mandonnet et al., 2007), ce qui suggère l'existence de réseaux de suppléance parallèles et bilatéraux. La voie dorsale, essentiellement phonologique, serait constituée du faisceau arqué et de la troisième portion du SLF (SLFIII) (Makris et al., 2005). Elle unirait, quant à elle, l'aire de Broca au gyrus supramarginalis.

Les rapports du MdLF par rapport aux voie ventrale et dorsale du langage restent encore mal connus et discutés (Menjot de Champfleur et al., 2012). Comme le montre le travail de Menjot de Champfleur et al. (2012), le trajet est bien distinct de ceux des autres faisceaux des voies ventrale et dorsale du langage ; en particulier l'ILF et le MdLF sont bien distingués. La terminaison du faisceau longitudinal inférieur se fait dans les gyrus moyen, inférieur et fusiforme au niveau temporal ; celle du faisceau unciné est plus médiale et se projette dans le pôle temporal, l'uncus, le gyrus hippocampique et l'amygdale ; la terminaison du MdLF se fait à la partie antérieure et moyenne du gyrus temporal supérieur, au-dessus, en dehors et en arrière des projections temporales de l'ensemble des faisceaux de la voie ventrale (Menjot de Champfleur et al., 2012 ; Catani et al., 2002).

Pour Markris et al. (Makris et al., 2009), il est possible que le MdLF puisse participer au réseau du langage de par sa localisation entre le GTS et le LPI : il jouerait le rôle de voie alternative entre le gyrus angulaire (« aire de Geschwind ») et le gyrus temporal supérieur (« aire de Wernicke »), suppléant ainsi au rôle dévolu au faisceau arqué au sein de la voie dorsale du langage. Pour Saur et al. (Saur et al., 2008), le MdLF serait impliqué dans les deux voies dorsale et ventrale. Dans l'hypothèse de Saur et al., le MdLF serait uni au faisceau arqué à la partie postérieure du STG, ce qui constituerait un relais pour atteindre la région prémotrice après une course dans la profondeur du lobe pariétal. En ce qui concerne la voie ventrale sémantique, elle serait constituée du MdLF, du faisceau de la capsule extrême et de l'ILF dans cette hypothèse. Ces résultats sont en désaccord avec ceux de Menjot de Champfleur et al., (2012) qui vont à l'encontre de cette hypothèse d'une participation du MdLF à la voie dorsale puisque le faisceau est systématiquement distingué du faisceau arqué lors de la tractographie et qui montrent également que le MdLF est bien discriminé des autres composants de la voie ventrale, et en particulier de l'IFOF.

Pour Menjot de Champfleur et al. (2012), la situation du MdLF (qui relie le gyrus temporal supérieur dans ses portions antérieures et moyennes au lobule pariétal inférieur) plaide pour une participation

du MdLF au réseau du langage, en particulier dans la voie ventrale sémantique. Cependant, le rôle du MdLF ne serait pas nécessairement essentiel à la compréhension du langage : les résultats d'une étude clinique (de Witt Hamer et al., 2010), lors de chirurgies de patients en condition éveillée, confirment la présence de sites langagiers au niveau du STG et du MTG ; cependant la stimulation électrique directe (SED) du faisceau longitudinal moyen (MdLF) dans sa partie antérieure n'induit aucune perturbation du langage. Il n'y a pas non plus de troubles du langage post-opératoire. Les auteurs concluent que si le MdLF de l'HG participe au traitement du langage, le rôle de ses fibres antérieures ne serait pas essentiel. Dans l'hémisphère mineur, il pourrait avoir un rôle à jouer dans les processus attentionnels.

Nos résultats apportent un élément de réponse intéressant sur le rôle, encore très débattu, du MdLF. Nos données semblent tout d'abord confirmer l'existence de ce faisceau, distinct des autres faisceaux mis en évidence dans cette étude (et ce, même si la méthode de définition de nos ROIs manque de précision pour délimiter ce faisceau). Par ailleurs, la mise en relation de nos données comportementales (COR) avec le ratio de fibres composant le MdLF semble confirmer son rôle dans le traitement sémantique puisque plus le nombre de fibres qui composent ce faisceau est important, meilleures sont les capacités des sujets à se représenter une phrase. Le rôle de ce faisceau pourrait donc soutenir la capacité à se créer une représentation mentale cohérente à partir d'une phrase, capacité qui pourrait faciliter la compréhension mais ne serait pas forcément nécessaire. Le fait qu'on ne retrouve pas cette corrélation pour les fibres de ce faisceau dans l'HD confirme aussi que le rôle du MdLF pourrait être différent pour l'hémisphère droit.

Nous préférons tout de même rester prudents en ce qui concerne la mise en évidence de ce faisceau dans notre étude car la technique employée pour déterminer les ROIs (basée sur nos données IRMf) n'est pas forcément la plus adaptée pour délimiter précisément des faisceaux anatomiques. Par ailleurs, les données concernant la connectivité fonctionnelle (ICA - corrélations) ne sont pas tout à fait en adéquation avec nos données DTI : on ne retrouve pas forcément une connectivité anatomique pour les régions qui semblent montrer une connectivité fonctionnelle importante, même si, bien sûr, corrélation n'est pas synonyme de causalité. En effet deux sites distants présentant une importante corrélation peuvent présenter plusieurs types de connectivité : ils peuvent être en connexion directe, tous les deux sous l'influence d'un troisième site, ou enfin tous deux exercer une influence conjointe sur un troisième site. De plus, même si deux sites sont connectés, la corrélation ne nous donne pas d'informations sur la direction de la connexion. Des analyses ultérieures seront nécessaires pour mieux comprendre comment relier connectivités structurale et fonctionnelle dans notre étude. Nous pourrons compléter nos analyses par d'autres données de connectivité fonctionnelle (notamment un modèle DCM sur nos données IRMf), mais

surtout par des analyses plus précises sur nos données DTI. Ainsi, nous pourrons déterminer certains faisceaux que nous n'avons pas observés de manière systématique (notamment à partir du RSC) et comprendre comment ils peuvent participer au réseau sémantique. Enfin, nous pourrons effectuer des analyses statistiques plus précises sur les données issues de nos faisceaux (FA, longueur de fibres…), elles nous permettront peut-être également de mieux comprendre le rôle des différents faisceaux, en particulier dans une perspective d'asymétrie droite-gauche.

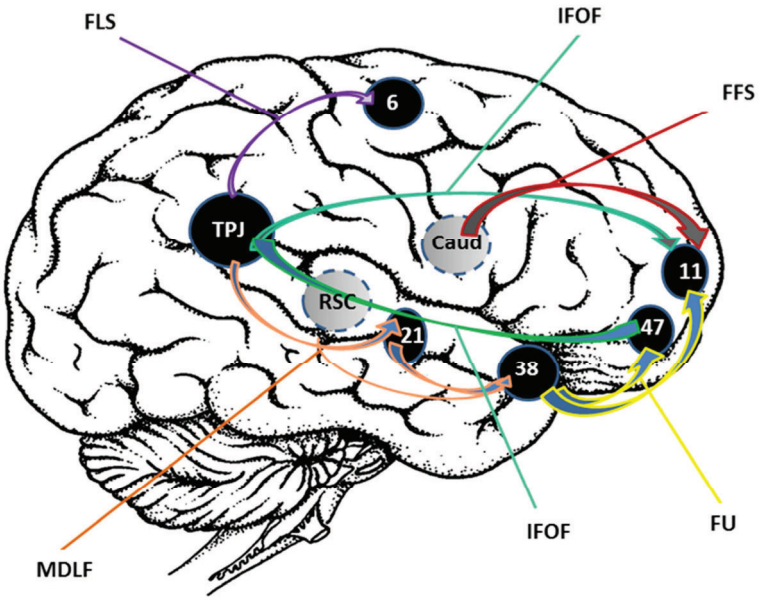


Fig 52) Régions cérébrales et faisceaux du système sémantique impliqué dans la compréhension des stimuli verbaux et non verbaux

Régions cérébrales : TPJ=Jonction temporo-pariétale, RSC= cortex rétrosplenial, BA21=Gyrus temporal moyen, BA 38 = Gyrus temporal moyen, BA6 = Gyrus précentral, BA47 = Gyrus frontal inférieur, BA11= Gyrus Frontal Médian et Caud = noyau caudé

Faisceaux de matière blanche : MdLF = faisceau longitudinal moyen, IFOF = faisceau fronto-occipital inférieur, FLS = faisceau longitudinal supérieur, FU = faisceau unciné, FFS = faisceau fronto-striatal

Synthèse des résultats Partie Expérimentale 1 – IRMf et DTI

Dans ce protocole, nous souhaitions identifier les substrats neuronaux de la représentation du sens, indépendamment de la modalité en présentant à des sujets sains des phrases et des images, décrivant ou représentant des événements riches. Les résultats issus de ce protocole en IRMf/DTI ont permis de mettre en évidence un réseau sémantique commun impliqué à la fois dans la compréhension des images et des phrases. Loin de se limiter à une activation sensorimotrice, ce réseau ferait intervenir des aires intégratives multimodales distribuées, notamment les régions pariéto-temporales et fronto-striatales. Les différentes analyses que nous avons réalisées ont permis d'établir la connectivité fonctionnelle et anatomique de ce réseau. Les aires multimodales qui forment ce réseau (le TPJ, le RSC, région temporale antérieure et cortex préfrontal) auraient un rôle crucial en tant qu'hubs intégratifs, rôle soutenu par une connectivité anatomique impliquant la voie ventrale sémantique à travers les faisceaux IFOF, UF et MdLF. La mise en relation de nos données comportementale (COR) avec le nombre de fibres qui composent le faisceau MdLF semble indiquer que ce dernier pourrait jouer un rôle important dans la capacité qu'on les sujets à se représenter une phrase.

II.2) Partie Expérimentale 2 : Expérience comportementale pour évaluer les différences individuelles dans la capacité à se représenter une phrase

II.2.*1*) Introduction

Un certain nombre d'études en neuroimagerie se sont intéressées aux différences interindividuelles, mais elles ont surtout porté sur des capacités et/ou des populations très spécifiques. Ainsi plusieurs études ont montré que chez les personnes ayant des aptitudes particulières dans un domaine (par exemple les musiciens pour la perception auditive, les danseurs professionnels pour la perception des mouvements ou encore les sommeliers en ce qui concerne l'odorat), les aires cérébrales engagées n'étaient pas les mêmes que chez des sujets naïfs (Onishi et al., 2001 ; Calvo-Merino et al., 2010 ; Bangert et al., 2006 ; Stewart et al., 2008 ; Castriota-Scanderberg et al., 2005). En plus de ces différences fonctionnelles, on trouve des différences au niveau morphologique. Par exemple, Maguire et al. (2000, 2003, 2006) se sont intéressés à la fabuleuse capacité qu'ont les chauffeurs de taxi londoniens à construire et à utiliser une carte mentale de la ville de Londres ; ils ont montré que ces derniers présentaient des différences concernant le volume de matière grise des régions hippocampiques par rapport à des sujets contrôles. Cela laisse supposer que ces régions du cerveau peuvent être réorganisées selon des mécanismes de plasticité cérébrale à l'âge adulte, ce qui donnerait à ces chauffeurs la capacité d'acquérir une représentation spatiale d'un environnement vaste et complexe.

Ces études sur des populations très spécifiques sont de bons modèles pour montrer comment l'expérience peut « sculpter » les processus cérébraux et le cerveau lui-même : les connexions neurales changent avec l'entraînement et l'expertise. Cependant, peu d'études se sont intéressées aux différences individuelles chez des populations non spécifiques, notamment en ce qui concerne les capacités de représentation. C'est pourtant une question importante qui est commune à toute évaluation d'aspects cognitifs impliquant des représentations : comment peut-t-on évaluer les expériences sensorielles d'un individu et surtout les comparer d'un individu à un autre (Bartoshuk, 2000 ; Bartoshuk, Duffy et al., 2004 in Wine and the Brain, Goode, 2007) ; est-ce que deux personnes qui perçoivent la même scène ont une même représentation de cette scène ? Dans l'étude de la perception visuelle d'actions, un certain nombre de travaux ont montré que l'activité des aires prémotrices était plus importante face à des actions familières que face à des actions non familières, suggérant un rôle de l'expertise sur les effets de résonance motrice, où à la fois la familiarité visuelle et motrice joueraient un rôle important (Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005; Calvo-Merino, Grèzes, Glaser, Passingham, & Haggard, 2006; Cross, Hamilton, & Grafton, 2006 ; Calvo-Merino et al., 2010). Wagner et al. (2011) ont observé que, chez des fumeurs en train de

regarder des films représentant l'action de fumer, certaines régions du réseau fronto-pariétal étaient plus activées que chez les non-fumeurs : les fumeurs se représenteraient spontanément l'action de fumer alors que les non-fumeurs ne le feraient pas. Il existe donc une variabilité des représentations d'un individu à l'autre, dépendante du mode de vie et de l'environnement de chacun. Cela semble tout à fait normal puisque si l'on revient à la définition originelle de la cognition incarnée de Barsalou (Barsalou, 1999), les concepts résultent de nos expériences sensorielles, motrices, sociales et émotionnelles et traiter un concept correspond à réactiver une trace multimodale de ces expériences vécues. Puisque ces concepts sont élaborés à partir d'une expérience qui nous est propre, ils font appel à la fois à des composantes sémantiques plus ou moins généralisées, mais aussi à des souvenirs personnels qui appartiennent à notre mémoire autobiographique. Bien sûr, nous avons tous été confrontés à des situations de la vie quotidienne probablement assez similaires d'un individu à un autre mais étant donné que nous ne pouvons pas partager directement et pleinement nos expériences sensorielles, comment être sûrs que ce que l'on perçoit, vit, ressent est de la même nature d'un individu à un autre ? Si la familiarité avec certaines situations, l'entraînement et l'expertise semblent être des facteurs susceptibles de déterminer ce que l'on est « capable » de se représenter (cf fumeurs qui se représentent l'action de fumer alors que les non-fumeurs ne le font pas), il est aussi tout à fait envisageable que les individus n'aient pas tous la même capacité à se représenter une situation. De la même manière que certains auront une meilleure mémoire, des capacités différentes en calcul mental ou en géométrie dans l'espace, certaines personnes pourraient avoir une habileté plus grande à se représenter et/ou comprendre une situation. Ces questions sont d'autant plus pertinentes dans le cadre du langage où, en l'absence d'un support visuel équivalent à celui d'une scène imagée, l'ambiguïté est plus grande lorsqu'il s'agit de se créer une représentation. Il paraît d'autant plus évident, dans le cadre du langage, que les individus ne sont pas forcément égaux pour se créer une représentation à partir d'une phrase.

De nombreux travaux se sont intéressés aux différences individuelles en ce qui concerne la lecture et la compréhension de phrases, mais la plupart de ces travaux se situent dans une perspective développementale puisqu'ils portent sur des enfants qui apprennent ou ont appris à lire récemment. En effet, lecture de mots et capacité de compréhension sont des compétences qui sont étroitement liées, en particulier chez les jeunes lecteurs où un manque de maîtrise de la lecture de mots peut limiter la capacité à comprendre un texte (Gough, Hoover, et Peterson, 1996). Néanmoins, ces compétences peuvent être dissociées (ce qui montre que ces compétences relèvent de mécanismes différents) : certains enfants ont des difficultés spécifiques pour la lecture des mots tandis que d'autres ont des difficultés spécifiques pour la compréhension de textes (de Jong et van der Leij, 2002 ; Oakhill, Caïn, et Bryant, 2003 ; Rack, Snowling, & Olson, 1992 ; Cain & Oakhill, 2004,

2006 ; Cain, Oakhill & Lemmon, 2004). Plusieurs hypothèses cherchent à expliquer ces différences dans les capacités de compréhension, à la fois chez les enfants mais aussi chez les adultes ; nous les aborderons dans la partie discussion de ce chapitre car notre approche est un peu différente : nous ne cherchons pas à évaluer les différences interindividuelles au niveau de la compréhension, mais au niveau des capacités de représentation (la facilité à obtenir une représentation à partir d'une phrase). Sans obliger les sujets à avoir une représentation « imagée » quand nous leur demandons de se représenter une phrase, il est certain que la consigne de la tâche utilisée fait sûrement appel à des processus qui sont plus de l'ordre de l'imagination et nous ne pouvons affirmer que ce processus de représentation correspond à la compréhension elle-même. De même que nous ne pouvons dire que cette représentation est équivalente au phénomène de simulation sensorimotrice, telle qu'elle est décrite par la cognition incarnée, et qui est supposée se produire dans le cerveau lors de la compréhension d'une phrase ou d'une image. Cette thématique de recherche n'a jamais vraiment été abordée dans d'autres travaux, tant il est difficile d'évaluer les représentations qui restent un phénomène vraiment subjectif. Il n'empêche que l'étude des différences individuelles, concernant cette capacité à se représenter une phrase, reste intéressante et susceptible de nous aider à mieux comprendre comment sont structurées nos représentations et si elles peuvent être de nature différente d'un individu à l'autre.

Pour nous, la mise en place de ce protocole comportemental répondait avant tout à différentes interrogations liées à nos autres protocoles expérimentaux. En effet, nous avons défini le Coefficient de Représentabilité (COR) lors des passations de notre expérience IRMf-DTI (voir Chapitre II.1) ; rappelons que le COR est un score obtenu en demandant aux sujets, après passation, de noter leur facilité à se représenter des phrases (sur une échelle allant de 1 à 5 : de 1 = Très difficile à se représenter à 5 = Très facile à se représenter). De cette analyse, nous avons extrait deux populations distinctes : les sujets qui éprouvaient une grande facilité à se représenter la quasi-totalité des phrases (note maximum la plupart du temps) et que nous avons qualifiés de « bons représentateurs » et les sujets qui avaient plus ou moins de difficultés à se représenter les phrases (ensemble de scores variés) et que nous avons désignés sous le terme « mauvais représentateurs ». Grâce à ce protocole comportemental, nous avons voulu savoir si ces différences entre « bons » et « mauvais représentateurs » se confirmaient chez un nombre plus important de sujets et si elles étaient associées à des marqueurs comportementaux particuliers (temps de réaction, temps de lecture…).

Par ailleurs, en lien avec les réflexions développées pour notre protocole EEG (que nous détaillerons dans le prochain sous-chapitre II.3), nous voulions également déterminer quelle pouvait être l'influence du contexte induit par une première phrase sur la représentation d'une seconde phrase

selon que celle-ci est cohérente avec la 1ère ou non. Par cette expérience comportementale, nous souhaitions évaluer si des effets de cohérence contextuelle pouvaient se refléter sur des marqueurs comportementaux (TR, temps de lectures, scores…) en comparant ces données pour deux phrases présentées successivement et pouvant avoir ou non un lien de cohérence.

Plus qu'une partie expérimentale à part entière, ce chapitre fait donc la transition entre nos protocoles IRMf et EEG et sert avant tout de base de réflexion pour des travaux ultérieurs, notamment en neuroimagerie.

II.2.*2*) Buts de l'expérience et hypothèses

L'objectif de cette expérience comportementale était de quantifier les capacités de représentabilité d'une phrase à travers l'étude de 3 marqueurs comportementaux : les scores de représentabilité (COR) ; les réponses motrices à travers les temps de réaction (TR, temps pour répondre à la question posée à la fin de chaque phrase, question demandant d'évaluer le niveau de représentabilité pour cette phrase) et les temps de lecture (TL, pour chaque mot des phrases présentées en modalité visuelle).

A travers ces différents marqueurs, nous souhaitions mettre en évidence les effets de plusieurs facteurs (effet de cohérence contextuelle : Suite/NonSuite et effet de groupe : bons/mauvais représentateurs). Cette étude avait aussi pour objectifs de préciser la notion de COR sur un plus grand nombre de sujets, de vérifier la pertinence du COR sur les autres marqueurs comportementaux (TR et TL) et de mettre en lien ces résultats avec nos données issues de la neuroimagerie.

Enfin nous avons choisi d'utiliser des phrases présentées à la fois en modalités visuelle et auditive, afin de savoir si la modalité de présentation pouvait avoir une influence sur les scores de représentabilité et/ou les TRs (et par conséquent sur les groupes bons/mauvais représentateurs).

Notre première hypothèse concerne la cohérence sémantique entre deux stimuli et le fait que celle-ci puisse avoir une influence sur les réponses motrices, la lecture et les scores de représentabilité, par l'intermédiaire d'une sorte de priming représentationnel. Nous avons émis l'hypothèse que lorsque deux phrases sont sémantiquement cohérentes, il y aurait un effet de facilitation et les TR seraient plus rapides pour la seconde phrase tandis que dans le cas où les deux phrases ne sont pas sémantiquement cohérentes, on aurait un effet inverse et les TR seraient ralentis entre la seconde et la première phrase. Cet effet de cohérence influencerait de la même façon les scores de représentabilité attribués par les sujets pour les deux phrases successives.

Notre seconde hypothèse concerne les effets de groupes liés aux capacités de représentation des sujets (bons vs mauvais représentateurs). Nous postulons tout d'abord qu'il existe différents profils de sujets et que l'on peut retrouver cette catégorisation à travers des différences dans les résultats comportementaux (TR et scores). Ces différents groupes pourraient avoir des stratégies représentationnelles différentes qui se refléteraient dans leurs résultats comportementaux : les « bons représentateurs » seraient moins sensibles aux effets facilitateur et perturbateur liés à la cohérence des stimuli car ils auraient un niveau de représentation constant et très bon ; en revanche les « mauvais représentateurs » seraient plus dépendants de la cohérence des stimuli au sein d'un contexte global.

Par ailleurs, nous souhaitions savoir si le temps de lecture pouvait expliquer ces différences entre bons et mauvais représentateurs : peut-être que les bons représentateurs lisent plus rapidement et ont un accès au sens plus rapide, ce qui expliquerait une différence dans les temps de réaction.

De même, pour la comparaison modalité auditive/visuelle, nous souhaitions savoir si des données comportementales pouvaient confirmer l'existence d'un système amodal (on s'attend alors à une indépendance face à la modalité de présentation) ou si, au contraire, l'une des deux modalités pouvait favoriser une accessibilité plus rapide au sens et expliquer, en partie, les différences entre bons et mauvais représentateurs. Peut-être que ces différences ne s'expliquent que parce que les sujets doivent lire une phrase et que donc en modalité auditive, la différence entre bons et mauvais représentateurs disparaîtra. De même pour les effets de cohérence contextuelle, il est possible que la cohérence/l'incohérence soient plus facilement appréhensibles dans l'une ou l'autre des modalités.

II.2.*3*) Matériels et Méthodes

- **Sujets**

Un groupe de 31 sujets sains, de langue maternelle française et ayant une vision normale ou corrigée ont participé à cette étude (16 hommes, âge moyen : 22,12 ans ; SD : 5,65). Aucun participant n'avait d'antécédent de trouble médical psychiatrique ou neurologique majeur.

- **Stimuli**

a) Sessions et Conditions

Sessions

Dans cette expérience, 100 couples de phrases sont présentés aux sujets. L'expérience est composée de 2 sessions :

- une en modalité auditive, la session « AUDITION » qui contient 50 couples de phrases à écouter. Pour les phrases en modalité auditive, les données comportementales recueillies sont le COR et le TR.
- une en modalité visuelle, la session « LECTURE » qui contient 50 couples de phrases à lire. Pour les phrases en modalité visuelle, les données comportementales recueillies sont le COR, le TR et le TL.

Conditions

Que ce soit en lecture ou en audition, chaque session comprend 2 conditions :

- la condition Suite (S), qui se compose de 25 couples de phrases sémantiquement reliées (phrase 1 S et phrase 2 S). Exemples de Phrases 1 S et 2 S pour la condition Suite : *« La femme fait la queue à la boulangerie. »/ « Elle achète une baguette. » ; « L'homme attend le train sur le quai. »/ « Il monte dans le wagon. »*
- la condition NonSuite (NS), qui se compose de 25 couples de phrases non reliées sémantiquement (phrase 1 NS et phrase 2 NS). Exemples de Phrases 1 NS et 2 NS pour la condition NonSuite : *« L'homme lit le menu dans le restaurant. »/ « Il répare la voiture. » ; « La jeune femme introduit la clé dans la serrure. »/ « Elle égoutte le riz dans l'évier. »*

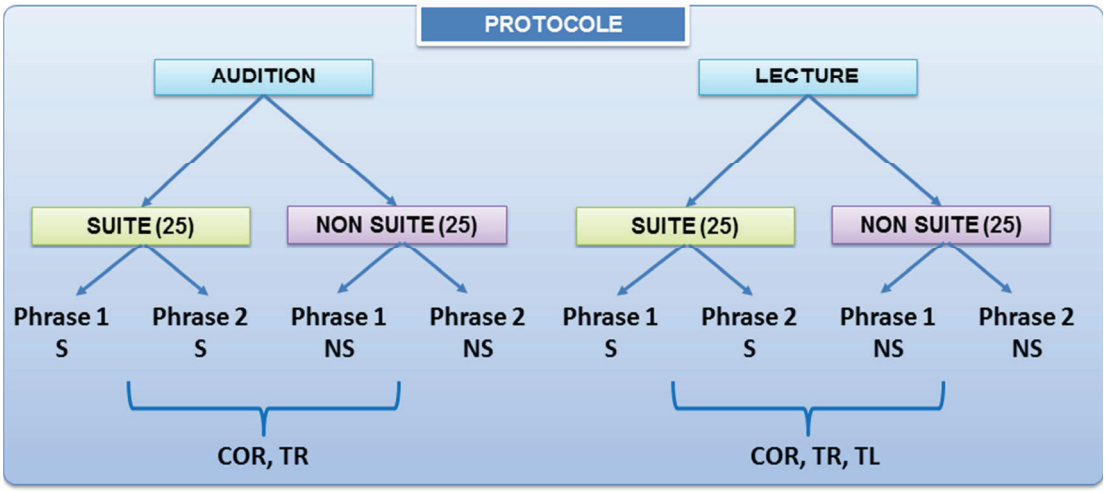


Fig 53) Protocole expérimental comportemental
Le protocole est composé de 2 sessions : audition et lecture. Chaque session comprend 2 conditions : la condition Suite (25 couples de phrases) et la condition Non Suite (25 couples de phrases)

b) Création de deux versions du protocole et validation des stimuli

Comme pour notre expérience IRMf, nous avons créé deux versions du protocole qui ont été présentées de façon aléatoire aux sujets : la moitié d'entre eux ont eu la « version 1 » et l'autre moitié, la « version 2 ». Les stimuli des sessions LECTURE et AUDITION sont inversés dans chaque version : les phrases AUDITION de la version 1 correspondent aux phrases LECTURE de la version 2 et inversement. Ainsi chaque phrase a un exemplaire en modalité LECTURE et un en modalité AUDITIVE mais un même sujet n'est pas confronté deux fois à la même phrase. Cependant l'ensemble des phrases existant dans les deux modalités, les conditions LECTURE et AUDITION sont similaires lorsque l'on considère les données sur l'ensemble des sujets.

Comme pour nos autres protocoles, les stimuli ont été répartis pour former des sessions équivalentes (AUDITION/LECTURE) grâce aux critères de la base Lexique (FreqFilms et FreqLivres), il a aussi été vérifié que chaque session contenait une répartition proche (homme/femme, nombre de lettres, nombre de mots… voir Annexe 4). Les mêmes vérifications ont été effectuées pour les conditions SUITE/NON SUITE et les stimuli ont été pré-testés chez une dizaine de sujets naïfs. Les phrases en modalité auditive ont été enregistrées (voix féminine) et nettoyées avec le logiciel Audacity pour enlever tous les artefacts pouvant être liés à l'enregistrement (suppresion du bruit résiduel, lissage de la hauteur de certaines phrases, réglage de l'intonation et de la vitesse de certains mots/phrases).

c) Procédure expérimentale

Le protocole a été conçu grâce au logiciel E-Prime 2.0 (E-prime Psychology Software Tools Inc. Pittsburgh, USA) qui gère la présentation et la randomisation des stimuli au sein d'une session. Durant deux sessions indépendantes, le sujet lit ou écoute des phrases puis il doit produire un jugement de représentabilité (de 1 à 5) pour chaque phrase qu'il voit ou écoute. Le sujet est averti que les deux phrases d'un même essai pourront évoquer deux événements dont la continuité est sémantiquement logique ou non. Selon qu'il a la version 1 ou 2 du protocole, le sujet commence par la session AUDITION ou par la session LECTURE. Pendant toute la durée de l'expérience, le sujet est assis face à un ordinateur portable de type PC, un casque sur les oreilles et il répond grâce au clavier de l'ordinateur. Chaque session débute par une petite période d'entraînement afin de s'assurer que le sujet a bien compris le fonctionnement de l'expérience et la tâche.

Un essai correspond systématiquement à la présentation successive de deux phrases. Que ce soit en modalité auditive ou visuelle, chaque essai commence par une croix de fixation de couleur verte (indiquant que c'est le début d'un nouvel essai), le sujet appuie lui-même sur un bouton (barre

espace) quand il est prêt à lancer l'essai. S'il est en modalité auditive, la première phrase enregistrée se lance automatiquement une fois que le sujet a appuyé ; une fois la phrase terminée, apparaît une diapositive sur laquelle on demande au sujet d'estimer son niveau de représentation de cette phrase : « quel est votre degré de représentation ? Appuyez sur les boutons de 1 à 5 ». La question est assez succincte, mais il a été bien expliqué au sujet avant l'expérience et pendant l'entraînement à quoi correspondaient les scores (1= très difficile ; 2=difficile ; 3=ni difficile, ni facile ; 4=facile ; 5= très facile). Le sujet doit alors produire un premier jugement en répondant par l'intermédiaire de 5 boutons de réponses (1,2,3,4,5 correspondant aux touches G,H,J,K,L d'un clavier azerty, des étiquettes avec chiffres ont été collées sur ces touches). Une fois que le sujet a produit son 1er jugement, la question disparaît pour faire place à un symbole > de couleur noir, indiquant au sujet qu'il entendra la deuxième phrase dès qu'il aura appuyé sur la barre d'espace : l'écoute de la seconde phrase se lance alors automatiquement et comme pour la 1ère phrase, le sujet doit produire un jugement de représentabilité à la fin de cette écoute. Il passe ensuite à l'essai suivant en appuyant sur la barre d'espace.

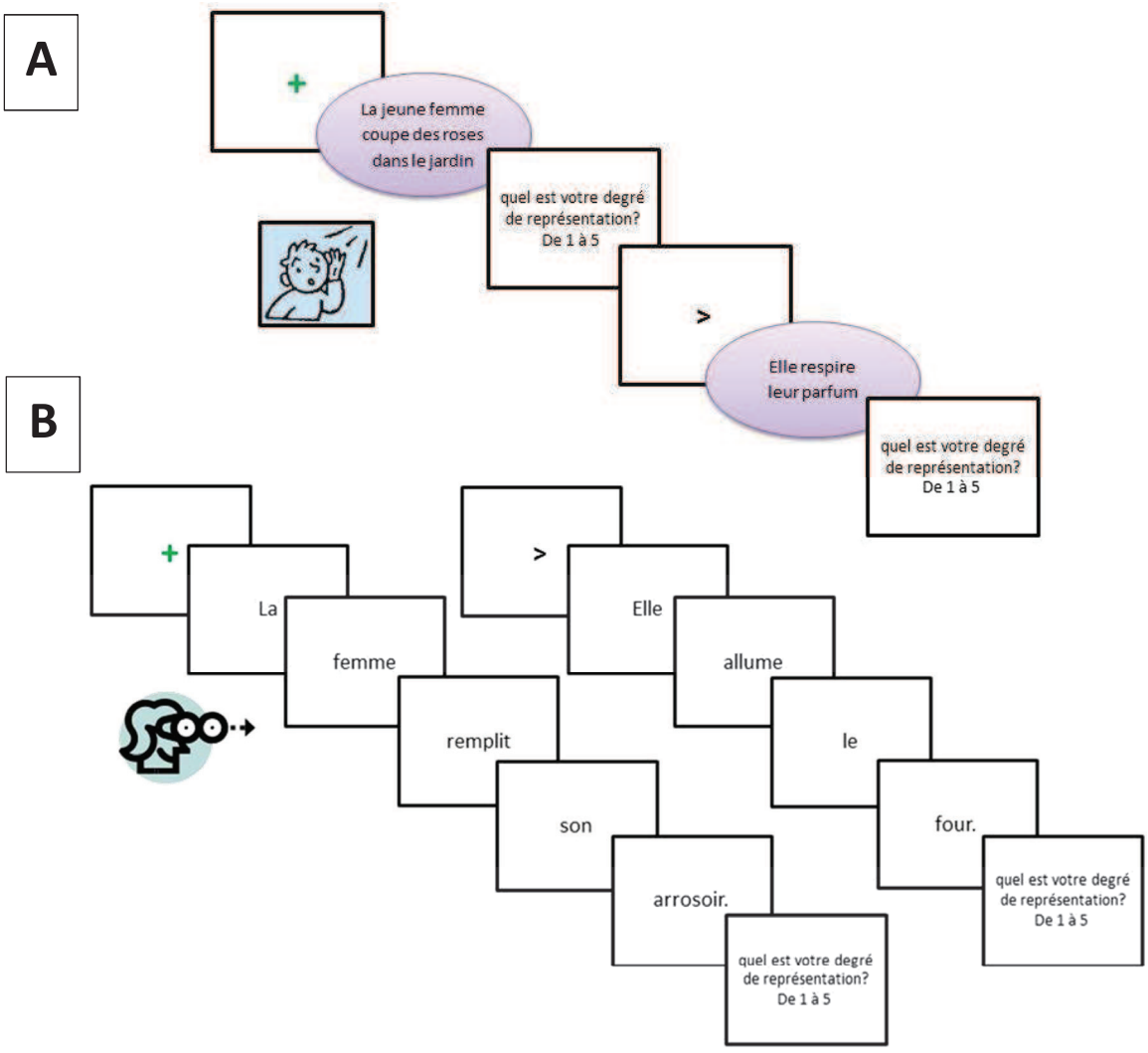


Fig 54) Déroulement d'un essai pour les sessions AUDITION (A) et LECTURE (B)
Dans l'exemple ci-dessus, les deux stimuli font partie de la condition Suite pour la session Audition, et de la condition Non Suite pour la session Lecture

Dans la condition LECTURE, la présentation successive des stimuli et la tâche sont les mêmes, la seule différence réside dans la lecture des phrases qui sont présentées mot à mot (un mot est inscrit au centre de l'écran, en noir sur fond blanc, temps de présentation : jusqu'à ce que le sujet appuie sur un bouton pour passer au mot suivant), c'est au sujet de faire défiler les mots de chaque phrase en appuyant sur la barre d'espace. Une fois qu'il a fait défiler toute la phrase (phrase 1 ou phrase 2), comme pour la condition AUDITION, le sujet doit produire un jugement de représentabilité. La durée totale de l'expérience dépend du rythme et des performances de chacun, mais elle dure environ 40 minutes.

A la fin de la session expérimentale, un questionnaire a été distribué aux sujets afin de mieux appréhender leurs réponses comportementales. Dans ce questionnaire, il est demandé aux sujets s'ils ont éprouvé des difficultés particulières à effectuer la tâche dans l'une ou l'autre des modalités et de décrire la nature de leurs représentations (voir questionnaire en annexe 5).

II.2.*4*) Analyses des données

a) Paramètres comportementaux

Dans les analyses statistiques que nous avons effectuées, les paramètres comportementaux étudiés sont :

- le coefficient de représentabilité (COR) qui correspond aux scores donnés par les sujets à la fin de la présentation de chaque phrase pour évaluer leur degré de représentabilité (Phrase 1 et Phrase 2).
- les réponses motrices (TR) qui représentent le temps mis pour répondre à la question « Quel est votre degré de représentabilité ? » posée à la suite de la présentation d'une phrase (Phrase 1 et Phrase 2).
- les temps de lecture (TL), voir paragraphe ci-après pour plus de précisions sur ce paramètre.

Le COR, TR et les temps de lecture (TL) sont utilisés comme paramètres comportementaux issus des sessions LECTURE ; alors que seuls le COR et le TR sont utilisés pour les sessions AUDITION puisque ce sont les seuls paramètres disponibles pour cette modalité. Le COR et le TR proviennent des données comportementales brutes issues de l'expérience, mais les TL ont nécessité quelques traitements préliminaires afin d'être exploitables.

Pré-traitements sur les données en lecture

Nous voulions que l'analyse des temps de lecture prenne en compte la longueur des mots et l'influence que ce facteur peut avoir sur les temps de réponse, le temps de réponse correspondant ici au moment où le sujet appuie sur le bouton pour faire défiler un mot et passer au mot suivant. Pour chaque sujet, nous avons réalisé une analyse de régression entre le temps de réponse pour un mot isolé et sa longueur. Pour un sujet donné, nous avons ainsi calculé la différence des scores obtenus par rapport aux valeurs attendues prédites par l'analyse de régression pour chaque mot. Les temps de réponse, résultant de cette étape de calcul, qui étaient supérieurs ou inférieurs à 3 écarts-types ont été exclus. Ce sont donc sur ces données « temps de réponse » pondérées par la longueur des mots que nous avons ensuite effectué nos analyses statistiques. Nous désignerons ces données par « temps de lecture » (TL) à partir de maintenant.

b) Facteurs d'intérêt et statistiques réalisées sur les données issues du protocole expérimental

- **Méthode des k-moyennes**

Dans un premier temps, nous avons concentré nos analyses sur la mise en évidence de différents « profils de représentateurs » parmi les sujets, ces profils ont été déterminés à partir d'une analyse de classification des COR par la méthode des k-moyennes.

L'objectif de l'algorithme des k-moyennes consiste à trouver la meilleure "partition" pour regrouper un certain nombre d'objets en k classes aussi différentes que possible les unes des autres, la procédure utilisée permettant de répartir les objets en différentes classes afin de minimiser la variance intra-classe et de maximiser la variance inter-classe. A l'issue de l'analyse de classification par la méthode des k-moyennes, les moyennes (de chaque classe sur chacune des dimensions) ont été examinées pour estimer dans quelle mesure nos k classes sont différentes. Idéalement, nous devrions obtenir des moyennes très différentes pour presque toutes les dimensions utilisées dans l'analyse, voire même toutes. La valeur du F dans l'analyse de variance réalisée sur chacune des dimensions constitue une autre indication pour savoir si la dimension respective discrimine bien les classes. Bien qu'un test de significativité soit inadapté, nous pouvons observer les résultats de l'analyse de la variance afin de comparer les moyennes sur chaque dimension (c'est-à-dire, les mesures des scores pour chaque condition) entre les différentes classes (les groupes de sujets). Dans notre étude, cette méthode de classification a été utilisée principalement à des fins exploratoires pour déterminer les différentes classes que nous avons utilisées pour effectuer les analyses statistiques anova.

- **Analyses ANOVA**

Les analyses statistiques des différents paramètres COR, TR et TL ont été réalisées par ANOVA à mesures répétées. Les facteurs d'intérêt, correspondant aux différents paramètres expérimentaux de notre protocole, sont :

- le facteur inter-sujet : « Groupe » déterminé à partir des différents profils COR des sujets
- les facteurs intra-sujet : « Modalité » qui permet la comparaison des sessions LECTURE et AUDITION, « Ordre » qui permet la comparaison de la phrase 1 et de la phrase 2 et « Suite » qui représente la comparaison entre les phrases de la condition Suite et celles de la condition Non Suite.

Trois analyses ANOVA ont été réalisées. Le niveau de significativité était déterminé avec un intervallle de confiance à 95%. Toutes les analyses statistiques ont été effectuées avec le logiciel Statistica (Statistica 10, Statsoft).

- *Analyses des scores de représentabilité en fonction du COR*

Ce modèle correspond à une analyse de variance sur les scores de représentabilité en tenant compte de la classification en 3 groupes selon leur COR

- *Variable dépendante* : scores de représentabilité donnés par les sujets pour chaque phrase.
- *Facteurs d'intérêt intra-sujets* : Le plan d'expérience comporte 3 facteurs intra-sujets, chacun présentant deux valeurs : le facteur Modalité (audition et vision) ; le facteur Suite (Suite et NonSuite) ; le facteur Ordre (phrase 1 et phrase 2).
- *Facteur d'intérêt inter-sujets :* Le plan comporte un facteur inter-sujet : Groupe Représentateurs à 3 modalités (G1=Mauvais, G2=Intermédiaires, G3=Bons).

- *Analyses des TR en fonction du COR*

Nous avons ensuite réalisé une analyse de variance sur les temps de réaction en tenant compte de la classification en 3 groupes selon le COR

- *Variable dépendante* : TR au moment de la tâche de notation pour chaque phrase
- *Facteurs d'intérêt intra-sujets* : Le plan d'expérience comporte 3 facteurs intra-sujets, chacun présentant deux valeurs : le facteur Modalité (audition et vision) ; le facteur Suite (Suite et NonSuite) ; le facteur Ordre (phrase 1 et phrase 2).
- *Facteur d'intérêt inter-sujets :* Le plan comporte un facteur inter-sujet : Groupe Représentateurs à 3 modalités (G1=Mauvais, G2=Intermédiaires, G3=Bons).

- *Analyses des TL en fonction du COR*

- *Variable dépendante* : temps de lecture pour chaque mot de l'ensemble composé des phrases 1 et 2
- *Facteurs d'intérêt intra-sujets* : Le plan d'expérience comporte 2 facteurs intra-sujets le facteur Suite (conditions Suite et Non Suite) ; le facteur MOT qui comporte 11 modalités (mot 1 à mot 11, le mot 1 correspondant au 1er mot de la première phrase, le mot 7 au mot de la seconde phrase et le 11ème mot au dernier mot de la seconde phrase, voir le paragraphe ci-après pour plus de détails)
- *Facteur d'intérêt inter-sujets :* Le plan comporte un facteur inter-sujet : Groupe Représentateurs à 3 modalités (G1=Mauvais, G2=Intermédiaires, G3=Bons).

Malgré l'analyse de régression que nous avons réalisée pour contrôler les effets dus à la longueur des stimuli, certaines phrases contiennent plus de mots que d'autres. Pour avoir un nombre de stimuli équivalents sur lesquels effectuer les statistiques, nous avons choisi de n'utiliser que le nombre de mots communs à toutes les phrases (même si les phrases sont en réalité plus longues), à savoir 6 mots pour la Phrase 1 et 5 mots pour la Phrase 2 (dont systématiquement le dernier mot de la phrase 2). Comme elles n'ont pas le même nombre de mots, nous avons considéré la Phrase 1 et la Phrase 2 comme un ensemble pour effectuer l'analyse statistique (analyse de la variance sur les temps de lecture en tenant compte des groupes de sujets déterminés précédemment), mais il est toujours possible de dissocier les effets qui sont liés aux mots de la phrase 1 ou 2 puisque l'on sait que les 6 premiers mots correspondent à la première phrase et les 5 derniers à la seconde phrase.

II.2.5) Résultats

A) Mise en évidence de profils COR différents

L'analyse de classification par méthode des k-moyennes décrite précédemment nous a permis d'obtenir 3 groupes de sujets :

- le groupe des « mauvais représentateurs » (classe 1) dont les profils de notes semblent plus variables et sont fortement influencés par le lien contextuel (Suite/Non Suite) qui existe entre deux phrases, quelle que soit la modalité.

- le groupe des « représentateurs intermédiaires » (classe 2) qui montrent un pattern similaire aux groupes des mauvais représentateurs concernant l'influence du lien contextuel, mais avec des scores généralement plus hauts.

- le groupe des « bons représentateurs » (classe 3) dont les scores, dans toutes les conditions, sont proches de la valeur maximale (5), ce qui laisse supposer que ces sujets sont très peu influencés par la cohérence contextuelle entre deux phrases et ce, quelle que soit la modalité.

Pour simplifier, à partir de maintenant nous désignerons ces groupes de la façon suivante : G1=classe1=Mauvais Représentateurs ; G2=classe2=Représentateurs Intermédiaires ; G3=classe 3 = Bons Représentateurs. Les performances de chaque groupe de sujets, pour chaque session (Audition=Aud ou Lecture=Vis), condition (Suite/Non Suite) et ordre des stimuli (Stimulus 1/Stimulus 2) sont représentées dans la Figure 55.

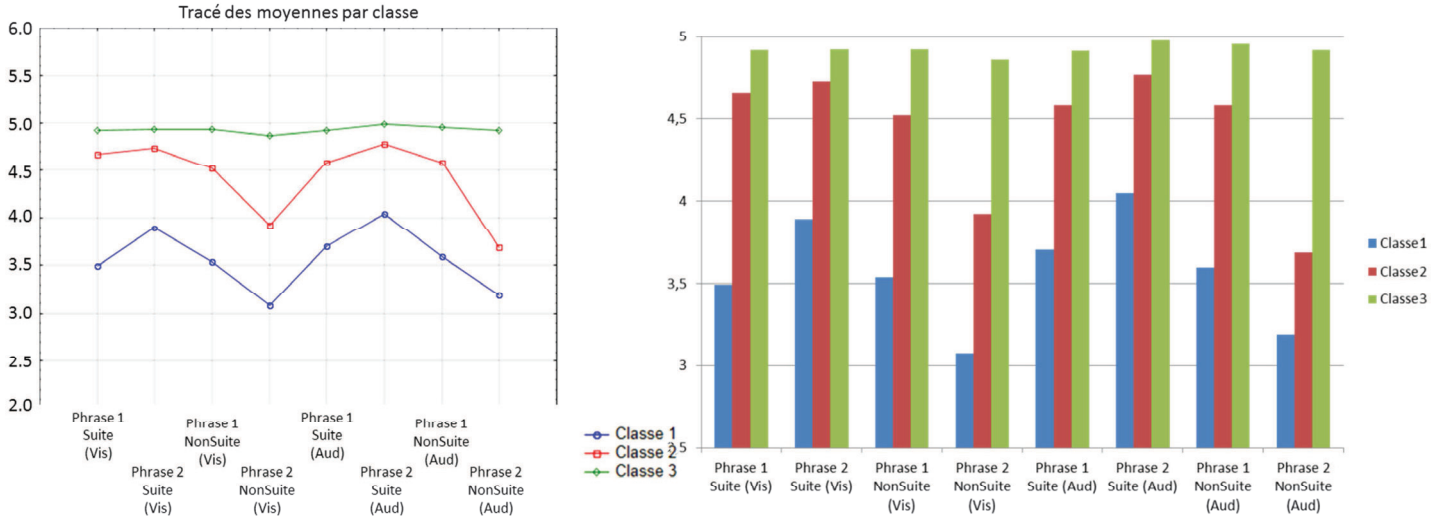


Fig 55) Représentations graphiques de la répartition des sujets en 3 classes selon la méthode des k-moyennes
Tracé des moyennes par classe (Figure de gauche) et histogramme des moyennes du COR des différents groupes, pour chaque condition (Figure de droite). Classe 1 = G1 = mauvais représentateurs (en bleu), Classe 2 = G2 = représentateurs intermédiaires (en rouge), Classe 3 = bons représentateurs (en vert)

Nous avons effectué une analyse de la variance complémentaire, afin de comparer les moyennes des scores pour chaque condition entre les différents groupes de sujets. Les résultats de cette analyse confirment que la catégorisation en 3 classes est pertinente par rapport à l'ensemble des conditions, puisque eu égard à l'importance des valeurs F et de leurs niveaux de significativité, toutes les conditions apparaissent déterminantes dans la constitution des classes (table 7).

Condition	Moyenne			F	p
	G1	G2	G3		
Phrase 1 Suite (Vis)	3,5	4,6	4,9	60.00	< 0,0001
Phrase 2 Suite (Vis)	3,9	4,7	4,9	31.37	< 0,0001
Phrase 1 NonSuite (Vis)	3,5	4,5	4,9	49.23	< 0,0001
Phrase 2 NonSuite (Vis)	3,1	3,9	4,8	26.98	< 0,0001
Phrase 1 Suite (Aud)	3,7	4,6	4,9	30.23	< 0,0001
Phrase 2 Suite (Aud)	4,0	4,7	4,9	16.00	< 0,0001
Phrase 1 NonSuite (Aud)	3,6	4,6	4,9	22.23	< 0,0001
Phrase 2 NonSuite (Aud)	3,2	3,7	4,9	22.41	< 0,0001

Table 7) Analyse de la variance réalisée sur les différentes classes

B) Résultats des analyses ANOVA

- **Effets simples**

  - *Effet principal du facteur Groupe (G1, G2, G3)*

- COR : L'analyse statistique confirme qu'il existe un effet significatif du groupe sur le COR ([$F_{(2, 28)}$ = 83.59, p< 0,0001], avec les moyennes suivantes G1 (3,5) < G2 (4,3) < G3 (4,9). Cet effet est indépendant de de la condition (Suite/Non Suite) et de l'ordre de présentation (Phrase 1 vs Phrase 2). (Voir figure 56-A).

- TR : L'analyse statistique montre qu'il existe un effet significatif du groupe sur les TRs ([$F_{(2, 28)}$ = 7.14, p<0.003], les TR sont les plus longs pour G1 (1426,99 ms) > G2 (1217,55 ms) > G3 (705,62 ms), indépendamment de la condition (Suite/Non Suite) et de l'ordre de présentation (Phrase 1 vs Phrase 2). (Voir figure 56-B).

- TL : En revanche, il n'y a pas d'effet global du facteur Groupe sur les temps de lecture ([F= 0.16, p<0.85]. Les trois groupes lisent à une vitesse équivalente.

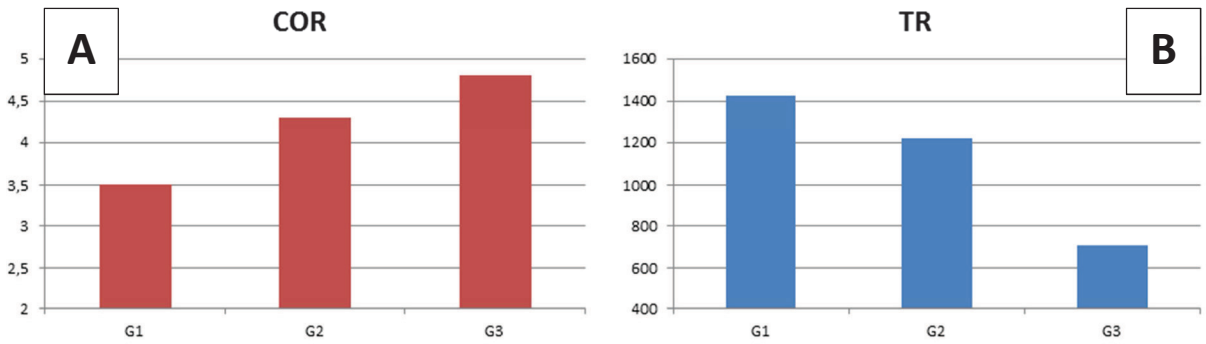


Fig 56) Moyennes du COR pour les 3 groupes (A) et Moyennes du TR pour les 3 groupes (B)
G1=mauvais représentateurs, G2=représentateurs intermédiaires, G3=bons représentateurs

- *Effet principal du facteur Modalité (Lecture vs Audition)*

Les différentes analyses ANOVA que nous avons réalisées n'ont révélé aucun effet de la modalité, il n'y a pas de différence entre les sessions AUDITION et LECTURE que ce soit pour le COR ou le TR.

- COR : Les scores donnés par les sujets sont indépendants du mode de présentation des phrases, auditif ou visuel ([$F_{(2, 28)}$ = 0.442, p=0.51]).

- TR : Les réponses motrices (TR) lors de la tâche de notation sont équivalentes, que les phrases soient lues ou entendues ([$F_{(2, 28)}$ = 0.606, p=0.44]).

- *Effet principal du facteur Ordre (Phrase1 vs Phrase 2)*

- COR : Il n'y a pas de différence significative entre Phrase 1 et Phrase 2 en ce qui concerne les scores COR. Que la phrase soit présentée en première ou en seconde position n'influence pas le score de représentabilité donné pour cette phrase ([$F_{(2, 28)}$ = 3.48, p=0.07).

- TR : Il n'y a pas non plus d'effet significatif sur les TR : les réponses motrices lors de la tâche de notation sont équivalentes pour la phrase 1 et pour la phrase 2 ([$F_{(2, 28)}$ = 0.03, p=0.86]).

- *Effet principal du facteur Suite (Suite vs NonSuite)*

- COR : Il y a un effet significatif du facteur Suite sur le COR ([$F_{(1, 28)}$ = 25.03, p<0,0001]) marqué par des CORs significativement plus bas pour la condition Non Suite (3,99) que pour la condition Suite (4,34). (Voir figure 57-A)

- TR : On trouve également un effet significatif du facteur Suite sur le TR ([$F_{(2, 28)}$ = 28.43, p<0.0001]) qui révèle que les TR sont plus rapides pour la condition Suite (1085 ms) et plus lents pour la condition Non Suite (1261 ms). (Voir figure 57-B)

- TL : Il n'y a aucun effet significatif lié au facteur Suite sur les TL, il n'y a pas de différence dans les temps de lecture entre les conditions Suite et Non Suite pour l'ensemble des mots présentés ([F= 0.0068, p<0.93]).

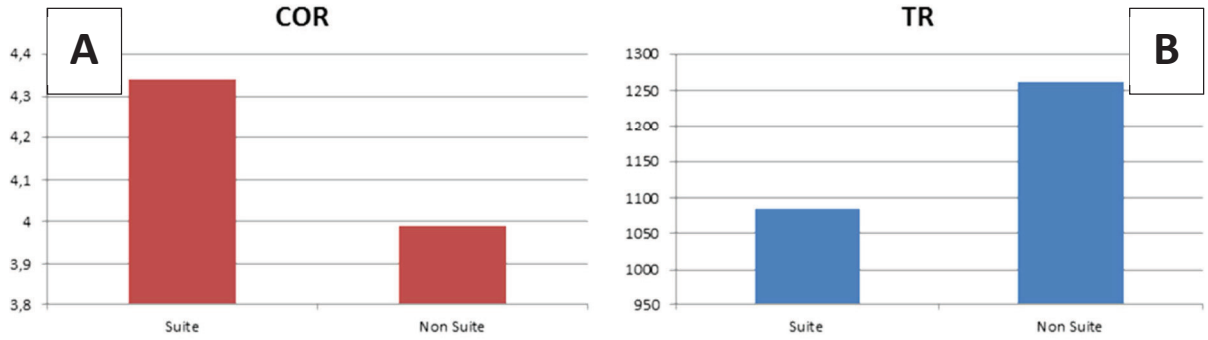


Fig 57) Moyennes du COR pour les conditions Suite et Non Suite (A) et Moyennes du TR pour les conditions Suite et Non Suite (B)

- **Effets d'interactions**

  - *Interaction Suite x Ordre*

- COR : L'interaction significative Suite x Ordre ([$F_{(2, 28)}$ = 18.97, p<0.0001]) révèle un effet sur les scores de représentabilité COR pour les conditions Suite et NonSuite en fonction de l'ordre de présentation de la phrase (phrase 1 vs phrase 2) et quelle que soit la modalité de présentation (voir figure 58-A). Dans la condition Suite, le score de la phrase 2 est plus grand (4,45) que celui de la phrase 1 (4,23) alors que dans la condition Non Suite, le score de la phrase 2 est plus petit (3,78) que celui de la phrase 1 (4,21). Les scores de la Phrase 1 sont équivalents pour les conditions Suite et non Suite alors que les scores de la Phrase 2 pour les mêmes conditions sont significativement différents. Un test post-hoc de type Bonferroni, effectué sur l'ensemble des facteurs, confirme que les effets de cette interaction sont significatifs indépendamment de la modalité (voir figure 58-B). Pour la modalité visuelle, la valeur du COR diffère significativement entre la phrase 1 et la phrase 2 pour la condition Suite (p=0,004) ainsi qu'entre la phrase 1 et la phrase 2 pour la condition Non Suite (p<0,0001). De même en modalité auditive, la valeur du COR diffère significativement entre la phrase 1 et la phrase 2 pour la condition Suite (p=0,0008) ainsi qu'entre la phrase 1 et la phrase 2 pour la condition Non Suite (p<0,0001).

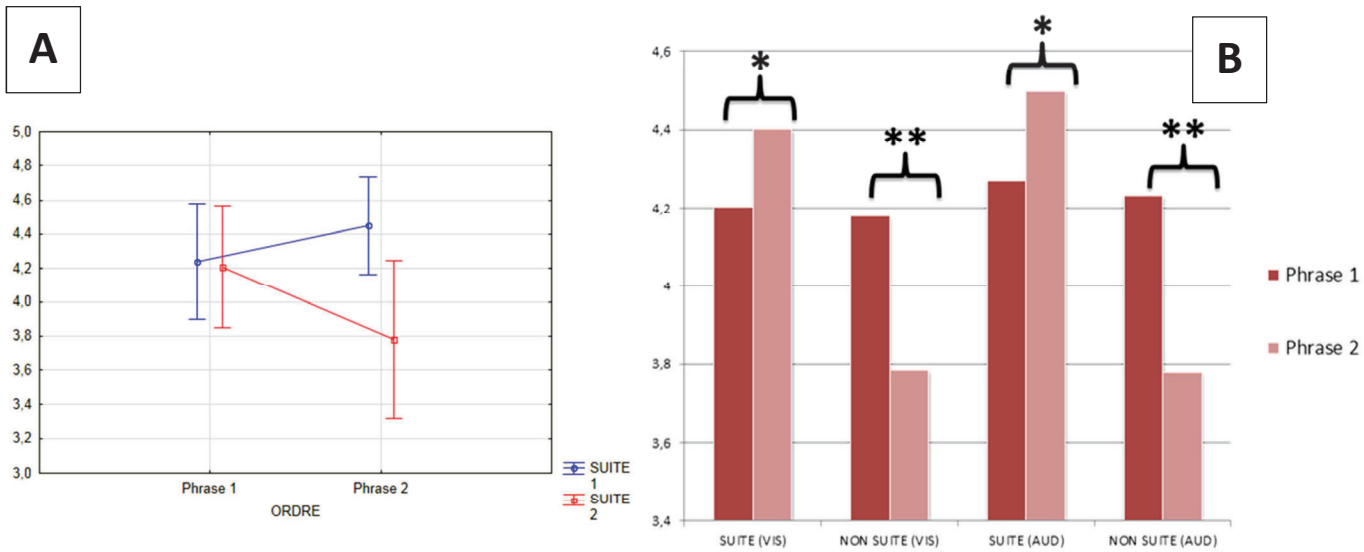


Fig 58) Illustrations de l'effet d'interaction Suite x Ordre sur le facteur COR
La figure A représente l'interaction elle-même : dans la condition Suite, le score de la phrase 2 est plus grand (4,45) que celui de la phrase 1 (4,23) alors que dans la condition Non Suite, le score de la phrase 2 est plus petit (3,78) que celui de la phrase 1 (4,21). Suite 1 : condition Suite ; Suite 2 : condition Non Suite
La figure B représente les moyennes du COR sur l'ensemble des sujets pour la phrase 1 et pour la phrase 2 (pour les conditions Suite/Non Suite et pour les deux modalités, auditive et visuelle). Les différences significatives sont représentées par une * (test de Bonferroni)

- TR : Là aussi, l'interaction Suite x Ordre se révèle significative ([$F_{(2, 28)}$ = 34,47, p<0.0001] : les TR sont plus lents pour la phrase 2 (1334 ms) par rapport à la phrase 1 (1188 ms) dans la condition NonSuite et plus rapides pour la phrase 2 (990 ms) par rapport à la phrase 1 dans la condition Suite (1180 ms) ; les TR pour les phrases 1 Suite vs Non Suite (1180 ms – 1188 ms) ne diffèrent pas statistiquement alors que pour les phrases 2, la différence Suite/non Suite (990 vs 1334) est significative (voir figure 59-A). Un test post-hoc de type Bonferroni, confirme que l'on retrouve les mêmes effets, indépendamment de la modalité, comme l'illustre la figure 59-B. Pour la modalité visuelle, la valeur moyenne du TR diffère significativement entre la phrase 1 et la phrase 2 pour la condition Suite (p=0,04) ainsi qu'entre la phrase 1 et la phrase 2 pour la condition Non Suite (p=0,01). De même en modalité auditive, la valeur moyenne du TR diffère signicativement entre la phrase 1 et la phrase 2 pour la condition Suite (p=0,0001) ainsi qu'entre la phrase 1 et la phrase 2 pour la condition Non Suite (p=0,05).

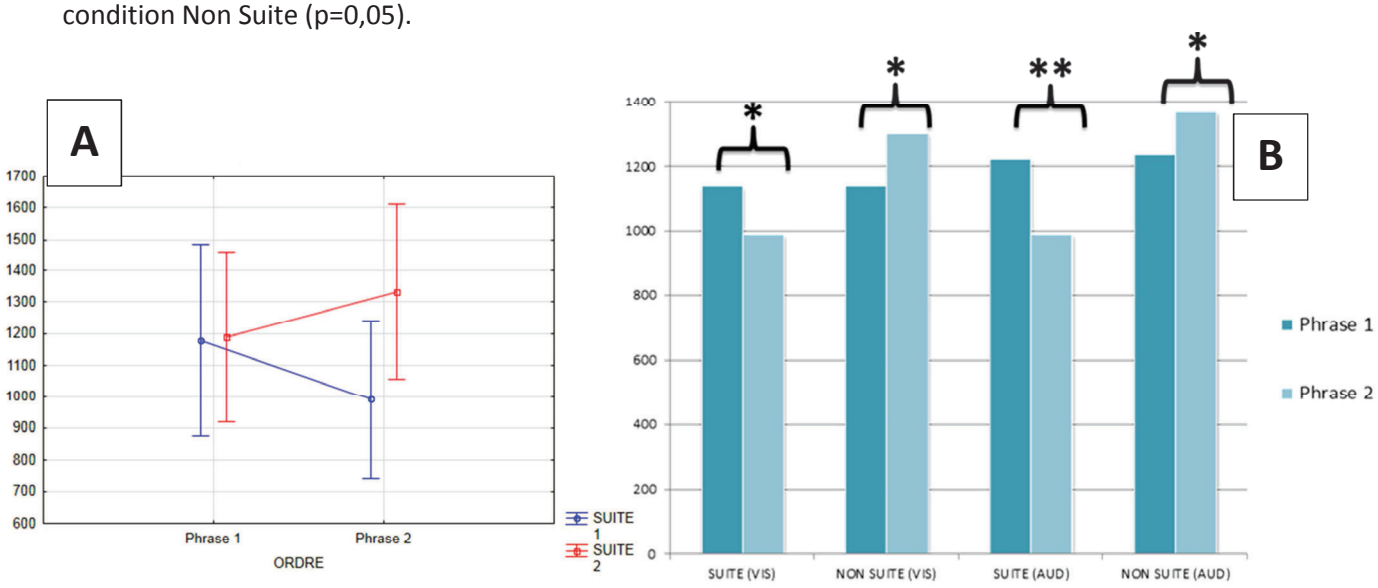


Fig 59) Illustrations de l'effet d'interaction Suite x Ordre sur le TR
La figure A représente l'interaction elle-même : les TR sont plus lents pour la phrase 2 (1334 ms) par rapport à la phrase 1 (1188 ms) dans la condition NonSuite et plus rapides pour la phrase 2 (990 ms) par rapport à la phrase 1 dans la condition Suite (1180 ms). Suite 1 : condition Suite ; Suite 2 : condition Non Suite
La figure B représente les moyennes du TR sur l'ensemble des sujets pour la phrase 1 et pour la phrase 2 (pour les conditions Suite/Non Suite et pour les deux modalités, auditive et visuelle). Les différences significatives sont représentées par une * (test de Bonferroni)

- *Interaction Suite x Groupe*

- COR : L'analyse des scores par une anova prenant en compte le facteur inter-sujet Groupe a révélé une interaction significative Suite x Groupe ([$F_{(2, 28)}$ = 4.92, p=0.01]). Un test Bonferroni post-hoc montre une différence significative des CORs dans la condition Suite entre G1 et G2 (p<0.0001), G1 et

G3 (p<0.0001), tandis pour la condition Non Suite les différences concernant le COR sont significatives (p<0.0001) entre tous les groupes (G1 et G2, G2 et G3, G3 et G1). (Voir Figure 60-A).

Pour les groupes « Mauvais représentateurs » (G1) et « Représentateurs Intermédiaires » (G2), les différences entre conditions Suite et Non Suite sont significatives (p=0.0006 et p=0.03). En revanche, il n'y a pas de différence entre ces 2 conditions pour le groupe des « Bons représentateurs » (G3) comme l'illustre la figure ci-dessous (60-B)

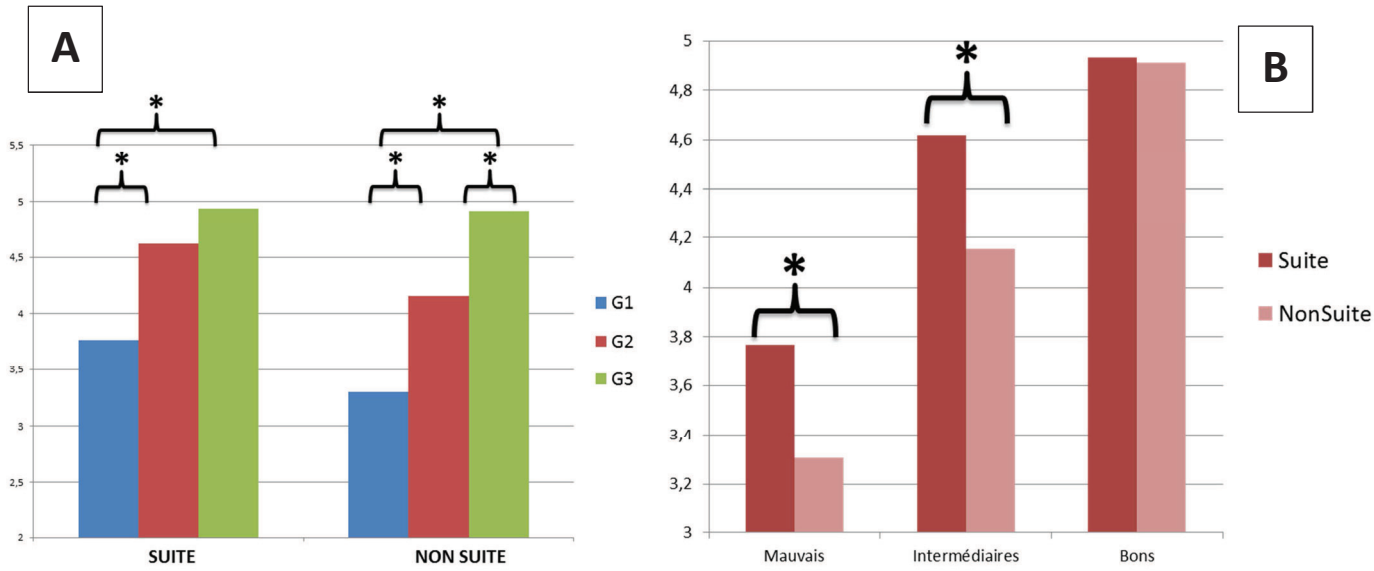


Fig 60) Illustrations de l'effet d'interaction Suite x Groupe sur le COR
La figure A représente les COR moyennés pour les 3 groupes de sujets dans les conditions Suite et Non Suite (G1=mauvais représentateurs, G2=représentateurs intermédiaires, G3=bons représentateurs).
La figure B représente la comparaison des COR moyennés entre la condition Suite et la condition Non Suite pour le G1, le G2 et pour le G3.
Les différences significatives sont représentées par une * (test de Bonferroni).

- TR : On retrouve la même interaction significative Suite x Groupe pour les TRs ([$F_{(2, 28)}$ = 4.59, p<0.01]). Le test post-hoc Bonferroni, effectué sur cette interaction, permet de montrer qu'il existe des différences significatives entre les groupes en fonction du facteur Suite (figure 61-A). En ce qui concerne la condition Suite, le G1 a des TR significativement plus lents (1382,02 ms) que ceux du G3 (638,58 ms, p=0.01) ; en revanche, il n'y a pas de différence entre les autres groupes. Pour la condition Non Suite, là aussi seuls les TR du G1 sont significativement plus lents (1471,96 ms) que ceux du G3 (772,65 ms, p=0.01). En ce qui concerne les différences entre les conditions Suite et Non Suite, seul le G2 (représentateurs intermédiaires) présente une différence significative entre les TR de ces deux conditions (1057,87ms vs 1377,22ms, p=0.0001). (Voir figure 61-B).

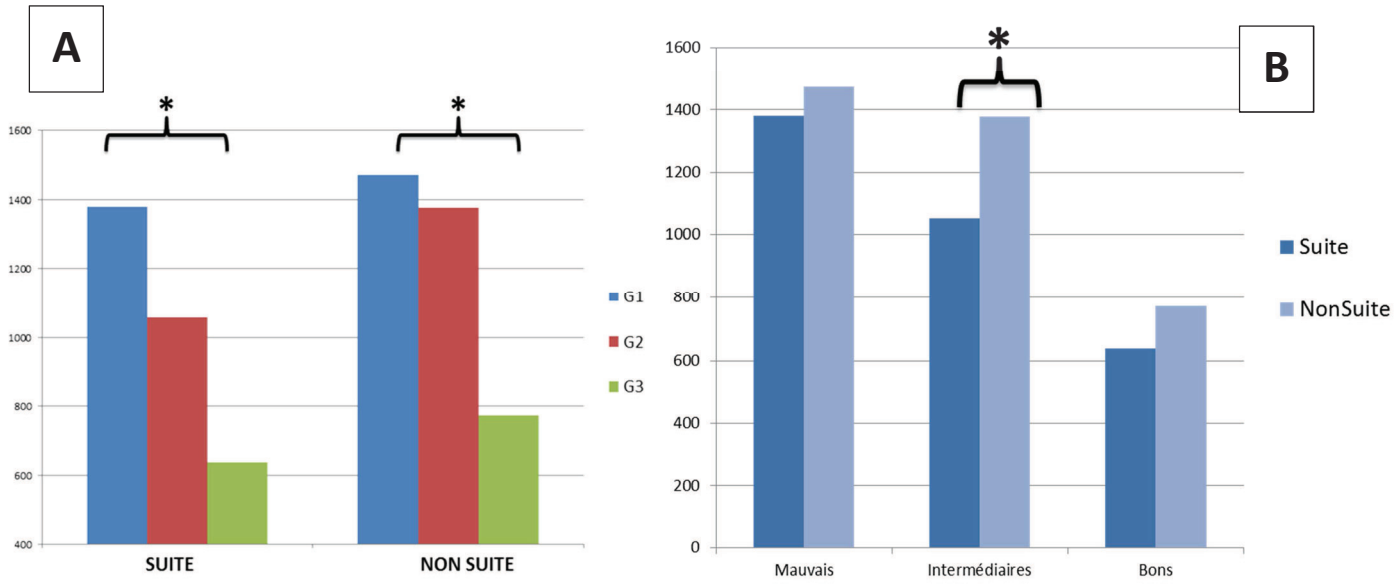


Fig 61) Illustrations de l'effet d'interaction Suite x Groupe sur le TR
La figure A représente les TRs moyennés pour les 3 groupes de sujets dans les conditions Suite et Non Suite (G1=mauvais représentateurs, G2=représentateurs intermédiaires, G3=bons représentateurs).
La figure B représente la comparaison des TRs moyennés entre la condition Suite et la condition Non Suite pour le G1, le G2 et pour le G3.
Les différences significatives sont représentées par une * (test de Bonferroni).

- *Interaction Suite x Ordre x Groupe*

- COR : En ce qui concerne le COR, l'interaction Suite x Ordre x Groupe est à la limite du seuil de significativité ([$F_{(2, 28)}$ = 3.14, p=0.05]) et un test post-hoc de Bonferroni permet de montrer qu'il existe un effet spécifique pour le groupe G2 se traduisant par une diminution significative des scores entre la Phrase 1 (4,49) et la Phrase 2 (3,81) pour la condition NonSuite (p=0.0167). Ce test post-hoc confirme également les différences qui existent entre les groupes de sujets avec en particulier des différences très marquées entre le G1 et le G3 quels que soient la condition ou l'ordre de présentation des phrases. (Voir Figure 62)

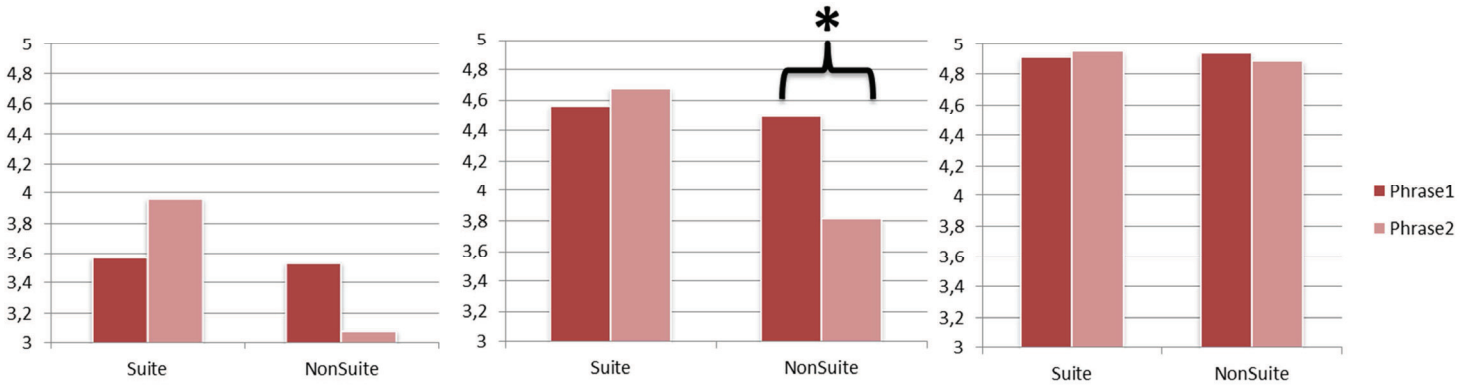


Fig 62) Illustration de l'effet d'interaction Suite x Ordre x Groupe sur le COR
Les différences significatives sont représentées par une * (test de Bonferroni)

- TR : Concernant les temps de réponses, l'interaction Suite x Ordre x Groupe est également significative ([$F_{(2, 28)}$ = 5.96, p<0.0069]). Un test Bonferroni post-hoc permet de préciser à quels facteurs est liée la significativité de cette interaction. Il existe une différence significative entre Phrase 1 et Phrase 2 pour la condition Suite chez les groupes G1 et G2 (p=0,01 et p=0.04), ainsi qu'une différence pour le G2 dans la condition Non Suite (p=0.002). En revanche, il n'y a aucune différence pour le groupe G3 quelle que soit la condition (Suite ou Non Suite) ou l'ordre de présentation de la phrase. (Voir Figure 63).

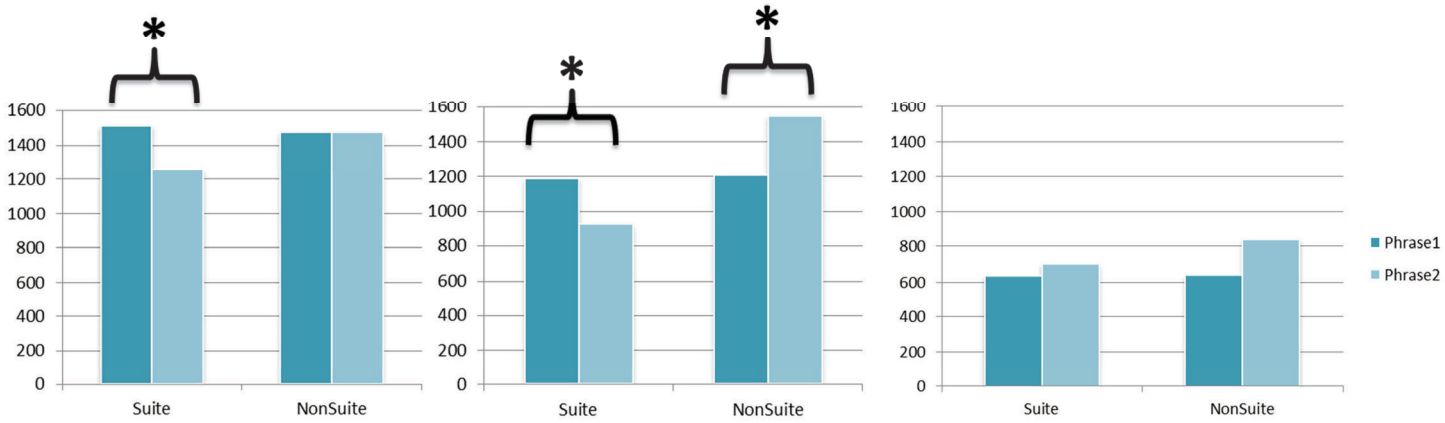


Fig 63) Illustration de l'effet d'interaction Suite x Ordre x Groupe sur le TR
Les différences significatives sont représentées par une * (test de Bonferroni)

- *Interaction Mot x Groupe (interaction spécifique aux TL)*

- TL : Pour les temps de lecture, l'analyse est un peu différente puisque nous avons considéré les stimuli provenant des phrases 1 et 2 comme un ensemble et non comme deux phrases séparées, c'est donc l'ordre des mots qui est important dans cette analyse. L'anova spécifiquement réalisée sur les TL montre qu'il y a un effet lié à l'ordre des mots ([F = 17.12, p p<0.0001]) et que l'interaction Mot x Groupe est significative ([F = 1.70, p=0.03]). Nous avons préféré aborder l'effet du facteur Mot dans le paragraphe consacré à l'effet d'interaction Mot x Groupe car les résultats de ces deux effets sont plus intéressants lorsqu'ils sont considérés conjointement.

L'effet Mot se traduit par une augmentation progressive de la rapidité du temps de lecture tout au long des mots qui composent la phrase 1 (mots 1 à 6), augmentation que l'on retrouve également pour les mots de la phrase 2 (mots 7 à 11). Mais cet effet est globalement le même chez tous les groupes. En effet, une analyse post-hoc plus poussée (test Scheffé) sur l'interaction Mot x Groupe nous a permis d'établir que les effets significatifs observés ne relevaient pas de différences majeures entre les différents groupes de sujets.

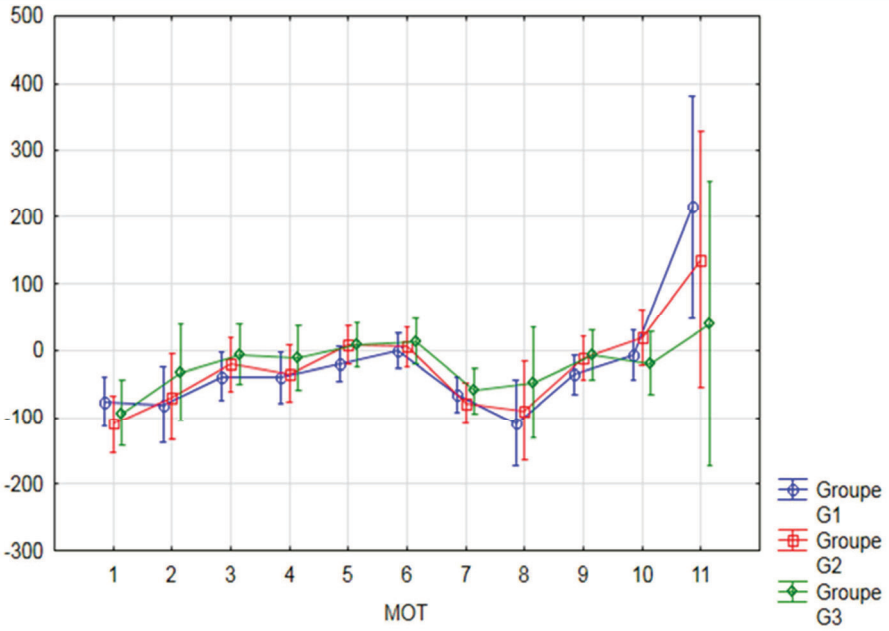


Fig 64) Illustration de l'effet d'interaction Mot x Groupe sur le TL
(G1=mauvais représentateurs, G2=représentateurs intermédiaires, G3=bons représentateurs)

L'interaction significative est principalement liée à des effets sur le dernier mot de la seconde phrase (mot 11) qui est globalement traité plus rapidement que tous les autres mots et surtout chez le G1. Pour tous les autres mots, l'analyse post-hoc révèle plusieurs éléments lorsqu'on prend en compte les groupes d'appartenance des sujets. Tout d'abord il n'y a pas de différence entre le temps de lecture des mots d'une phrase pour un même groupe de sujets (ex : MOT1 pour G1 n'est pas différent de MOT2 pour G1, MOT3 pour G1 etc…) mais surtout un même mot n'est pas traité différemment d'un groupe à un autre de sujets (MOT1 pour G1 = MOT1 pour G2 = MOT1 pour G3, etc.). A l'exception des effets mineurs observés pour le mot 11, on peut considérer qu'il n'y a aucune différence dans les temps de lecture en fonction du groupe d'appartenance.

- **<u>Synthèse des résultats</u>**

De ces analyses statistiques, plusieurs effets ressortent clairement quant aux paramètres comportementaux étudiés :

- Il n'y a aucun effet de la modalité (vision vs audition), que ce soit pour le COR ou pour les TR.

- Il y a un effet du facteur Suite à la fois sur les TR et le COR : pour la condition Non Suite, les scores sont globalement plus faibles et les TR plus lents que pour la condition Suite, indépendamment de la

modalité de présentation et de l'ordre des phrases. En revanche, le facteur Suite n'a aucune influence sur le TL.

Mais les résultats sont plus clairs lorsque l'on considère l'interaction Suite x Ordre (qui est elle aussi significative pour les scores et les TR). Quand on considère l'ensemble des résultats avec des tests post-hoc, on se rend compte qu'il y a :

- un effet de facilitation entre Phrase 1 et Phrase 2 quand les 2 stimuli font partie d'une suite logique (Condition Suite) qui se traduit par une diminution significative du TR et une augmentation significative de la note donnée pour la Phrase 2 par rapport à la Phrase 1.
- un effet d'inhibition entre Phrase 1 et Phrase 2 quand les 2 stimuli ne font pas partie d'une suite logique (Condition Non Suite), d'où une augmentation significative du TR et une diminution significative du score de représentabilité pour la Phrase 2 par rapport à la Phrase 1.

En résumé, indépendamment de la modalité de présentation, ces résultats montrent que : les sujets sont plus rapides pour répondre et qu'ils se représentent plus facilement la seconde phrase quand elle fait suite à la première en terme de cohérence sémantique, alors qu'ils sont plus lents et que cela « perturbe » leur représentation lorsque la seconde phrase manque de cohérence avec la première. Il existe donc une sorte de priming représentationnel où l'information issue du contexte serait importante pour la création et/ou la mise à jour d'une représentation.

Les résultats sont d'autant plus intéressants quand on les étudie en tenant compte de la classification des sujets en différents groupes selon leur profil de représentateurs (profils obtenus à partir du COR).

- On observe un effet global du facteur Groupe, à la fois sur les TR et le COR : les sujets G1 (mauvais représentateurs) ont globalement des scores plus bas et des TR plus lents que les sujets G2 (représentateurs intermédiaires) qui sont eux-mêmes moins bons que les sujets G3 (bons représentateurs). Les sujets G3 ont les scores les plus hauts et les TR les plus rapides. Que ce soit au niveau du COR ou des TR, l'existence de différents groupes dont les capacités de représentation ne sont pas les mêmes, semble se confirmer. En revanche, on ne retrouve pas d'effet du facteur Groupe associé aux temps de lecture.

- Par ailleurs, que ce soit pour le COR ou pour les TR, on trouve une interaction significative entre les facteurs Groupe et Suite, marquée par des différences significatives entre les conditions Suite et Non Suite pour G1 et G2 en ce qui concerne le COR (scores plus bas pour la condition Non Suite) mais seulement pour G2 en ce qui concerne les TR (TR plus lents pour la condition Non Suite).

Dans les deux cas (scores et TR), il n'y a aucune différence pour le G3. Là aussi, il n'y a aucune différence entre les groupes pour le TL des conditions Suite et Non Suite.

- On retrouve également l'interaction significative Suite x Ordre, à la fois pour les scores et les TR, interaction dont nous avons déjà discuté précédemment en termes d'effets facilitateur et inhibiteur. Mais ces effets sont modulés selon l'appartenance aux différents groupes. S'il existe un effet facilitateur sur les TR pour la condition Suite à la fois pour G1 et G2, on ne retrouve pas cet effet sur les scores. Le G2 (représentateurs intermédiaires) semble être le groupe dont les performances sont les plus influencées par l'effet contextuel puisque les sujets de ce groupe sont les seuls à présenter un effet perturbateur entre Phrase 1 et Phrase 2 pour la condition Non Suite, effet qui se répercute à la fois sur les scores et sur les TR. Le G3 ne présente aucun effet facilitateur ou perturbateur, quels que soient la condition ou l'ordre de présentation et ce, à la fois sur les scores et sur les TR.

Ce qui est particulièrement intéressant, c'est qu'aucun des effets observés sur le COR ou le TR ne peuvent être attribués aux temps de lecture puisque il n'y a pas de différence notable sur les TL pour les différents facteurs considérés, ni entre les groupes. La rapidité à répondre (TR) ou la facilité à noter la représentabilité d'une phrase ne semblent pas dépendantes des capacités en lecture des individus.

C) Analyse des réponses au questionnaire final

A la fin de cette expérimentation comportementale, un bref questionnaire a été adressé aux participants pour essayer de comprendre leurs réponses a posteriori (voir annexe 5). Il s'agissait d'un questionnaire assez libre, ce qui ne nous permet pas de quantifier précisément les réponses. Il apporte, cependant, quelques éléments de réflexion intéressants.

- **<u>Difficulté liée à la tâche</u>**

En ce qui concerne l'effet contextuel (induit par la présentation de 2 phrases successives sémantiquement cohérentes ou non), 55% des sujets ont spontanément déclaré qu'ils avaient eu l'impression d'être perturbés dans la création d'une nouvelle représentation lorsque la seconde phrase n'était pas cohérente avec la première ; 32 % des sujets n'ont pas signalé éprouver de difficulté particulière et 13% des sujets ont d'eux-mêmes spécifié n'avoir ressenti aucune différence selon que les 2 phrases présentées successivement étaient contextuellement cohérentes ou non (voir Figure 65). Il est assez amusant de constater que parmi les 13% de sujets qui ont indiqué ne

ressentir aucune différence entre Suite et Non Suite, les ¾ de ces sujets appartiennent au groupe des bons représentateurs.

Par ailleurs, alors que nos statistiques n'ont montré aucune différence concernant les TR et les scores en fonction de la modalité (auditive ou visuelle), la plupart des sujets semblent avoir une modalité de prédilection et les réponses sont assez partagées : 36 % des sujets disent éprouver plus de difficulté en modalité LECTURE contre 45 % en modalité AUDITIVE, 19 % ne signalent pas de différences spécifiques entre les 2 modalités (voir figure 65).

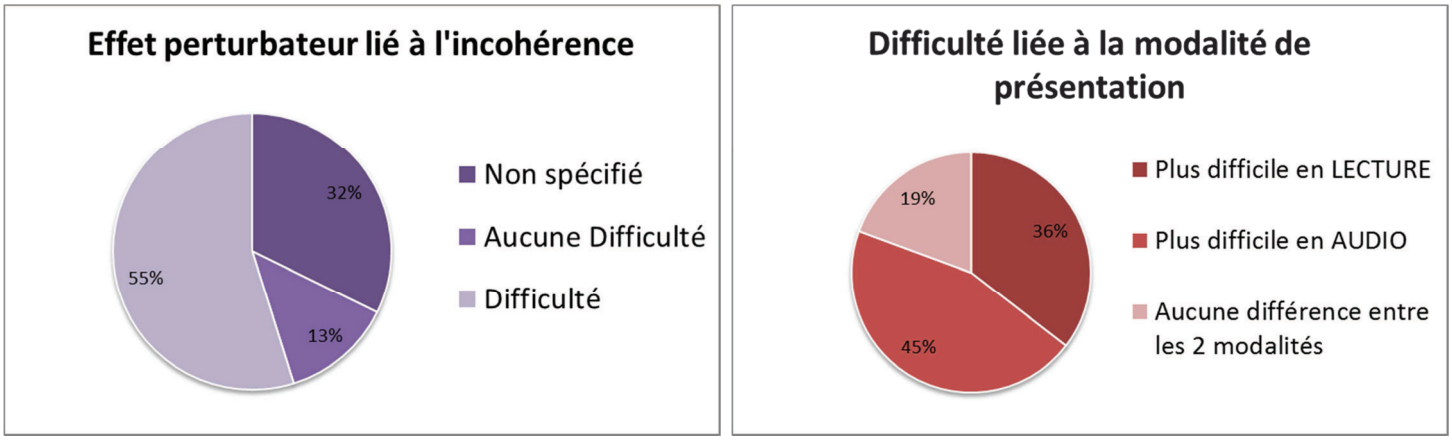


Fig 65) Répartition des sujets selon leur ressenti d'un effet perturbateur lié à l'incohérence (gauche)
Répartition des sujets selon leur difficulté éprouvée face à une modalité de présentation (droite)

Bien sûr, ces réponses sont assez subjectives et ne nous permettent pas de déterminer précisément comment se forment les représentations des sujets selon l'une ou l'autre des modaltiés. Cependant, nous avons voulu vérifier si les sujets étaient de bons « juges » de leurs performances en réalisant des analyses statistiques supplémentaires (anova) sur notre population de sujets, répartis selon leur difficulté éprouvée face à l'incohérence sémantique ou leur modalité de prédilection. Si l'on retrouve des résultats proches de ceux discutés précédemment (caractérisés par des effets significatifs du facteur Suite et du facteur Ordre), il n'y a pas d'effets spécifiques liés aux groupes. Donc, malgré la sensation des sujets d'être perturbés ou non par l'effet d'incohérence contextuelle ou leur impression d'avoir plus ou moins de difficulté dans l'une ou l'autre modalité, ces ressentis ne sont absolument pas validés statistiquement et ils n'ont aucune influence sur la rapidité de réponse des sujets (TR) ou sur leur facilité à se représenter une phrase (COR).

- **<u>Nature des représentations</u>**

Les réponses à certains items du questionnaire étant assez libres, la plupart des sujets ont tenu à préciser spontanément la nature de leurs représentations. Les patterns de réponses sont variés : certains sujets disent s'imaginer eux-mêmes faisant l'action ou imaginer des membres de la

famille ou des proches ; d'autres disent voir des scènes issues de films ou de publicités ; d'autres encore se représentent des personnes inconnues, voire même des silhouettes sans visage. Ce qui apparaît assez nettement, c'est qu'il y a une certaine cohérence dans les réponses des sujets (par exemple, ceux qui se représentent principalement des inconnus ne font pas référence à des représentations impliquant les membres de leur famille, et inversement). De plus, cette cohérence semble être liée au profil de représentateurs des sujets. Bien que ces données soient très difficiles à évaluer et à quantifier, nous avons tenu à illustrer les réponses des sujets en fonction des différents groupes COR (voir figure 66).

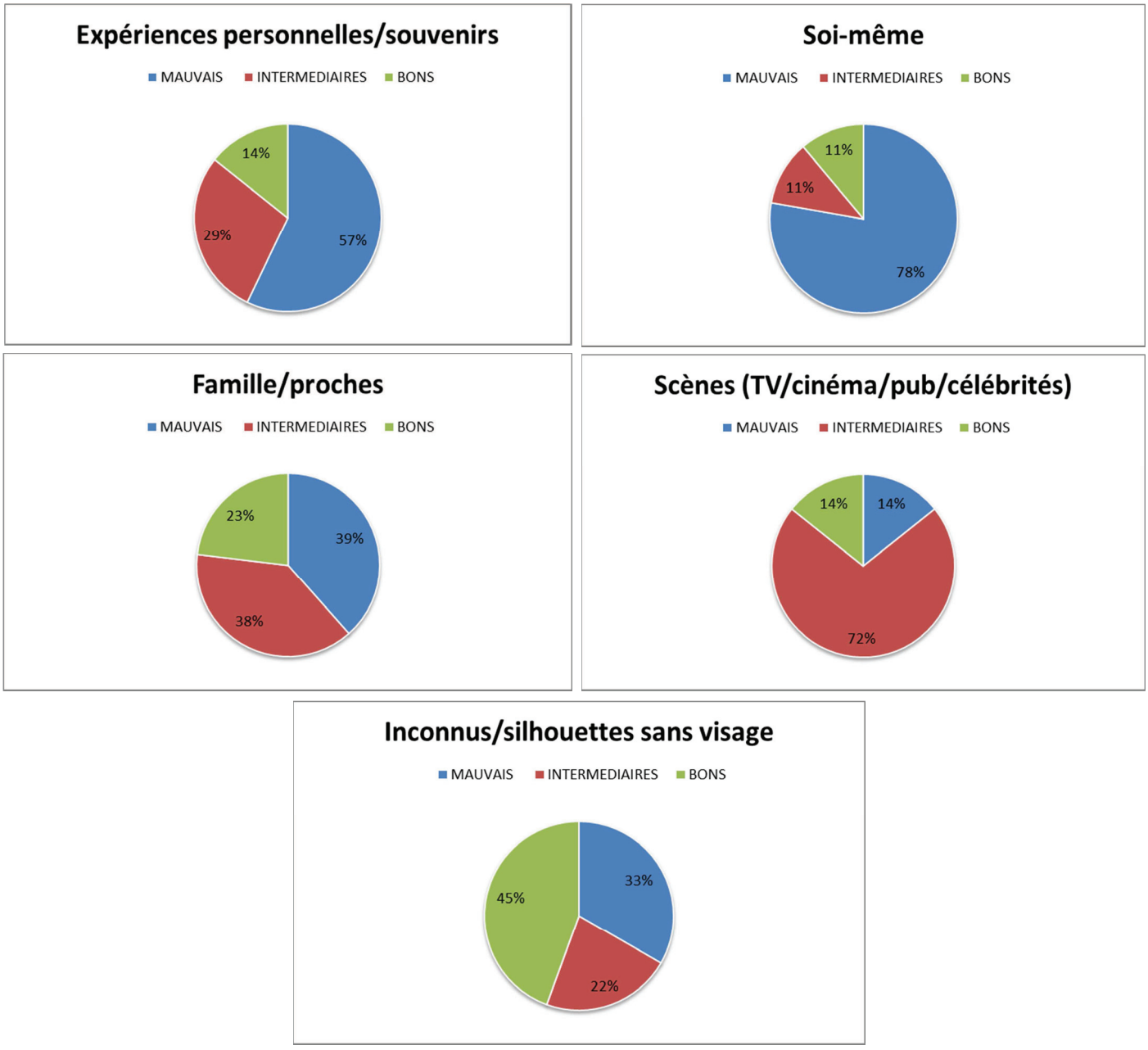


Fig 66) Répartition des types de représentations spontannément mentionnés par les sujets, en fonction du profil COR des sujets

(bleu= mauvais représentateurs, rouge= représentateurs intermédiaires, vert= bons représentateurs)

En effet, les « mauvais représentateurs » semblent plus généralement faire appel à un référentiel personnel puisqu'ils évoquent majoritairement des représentations d'eux-mêmes faisant l'action (78% par rapport à l'ensemble des sujets) ou de souvenirs qui leur sont propres (57 %). Ils rapportent aussi des représentations impliquant des proches (39%), dans la même mesure que les « représentateurs intermédiaires » (38%). Ces derniers semblent cependant avoir des représentations qui sont plus dépendantes d'un référentiel externe mais qui ne leur est pas personnel puisqu'ils évoquent majoritairement des représentations de scènes issues de publicités, films ou impliquant des célébrités (72%). Enfin, dans le groupe des « bons représentateurs », les représentations font intervenir principalement des inconnus ou des silhouettes (42% alors que pour toutes les autres catégories de représentations, les pourcentages de ce groupe se situent autour de 10-20%).

- **Influence de l'expertise en lecture sur les capacités de représentation**

En lien avec le niveau de lecture, un autre résultat intéressant est apparu lors de l'étude des réponses obtenues via le questionnaire post-protocole. En effet, il a été demandé aux sujets de préciser s'ils aimaient lire et combien d'ouvrages ils lisaient par an environ. Si les sujets qui appartiennent aux groupes des « bons représentateurs » et des « représentateurs intermédiaires » ne diffèrent pas vraiment, avec une moyenne de 10 livres lus par an (9,86), le groupe des « mauvais représentateurs » a un nombre de livres lus par an qui est bien inférieur (4,5) et significativement différent du nombre de livres lus par les autres groupes (test de Student p=0.02). Par ailleurs, parmi les gens qui ont précisé ne pas aimer la lecture et ne pas lire du tout (20% des sujets), 83,3% de ces personnes se situent dans le groupe des « mauvais représentateurs ». Si ces données sont très subjectives et difficilement quantifiables, il semblerait tout de même que la lecture ou plutôt le goût de la lecture pourrait être un bon prédicteur de la facilité à se représenter une phrase.

II.2.*6*) Conclusion et perspectives

Un premier aspect, qui ressort très fortement de nos différentes analyses, c'est qu'il n'y a pas d'influence de la modalité de présentation : que les phrases soient présentées visuellement ou auditivement n'influence pas les performances des sujets (scores et TR) ou leur répartition dans les différents groupes selon leur capacité de représentation et ce, malgré la sensation de certains sujets d'avoir une modalité où le traitement des phrases est facilité.

Une autre « absence de résultats », qui reste cependant un élément très intéressant à discuter, concerne les analyses des temps de lecture. Globalement nous n'avons pas trouvé d'effets

marquants, mais nous rediscuterons ce point dans les paragraphes concernant respectivement les effets de cohérence contextuelle et les effets liés aux groupes.

A propos de notre hypothèse concernant l'existence d'un priming représentationnel, elle est tout à fait confirmée puisqu'on trouve globalement des effets de facilitation pour la seconde phrase quand celle-ci est sémantiquement cohérente avec la première. Ces effets se traduisent par une plus grande facilité à se représenter la seconde phrase (score significativement plus grand) et des temps de réaction plus rapides. A l'inverse, lorsque la seconde phrase n'est pas cohérente avec la première, il y a un effet inverse qui se caractérise par des scores plus bas et des TRs plus lents pour la seconde phrase. Des résultats qui sont confirmés par les témoignages des participants puisque plus de la moitié d'entre eux a ressenti une gêne et dit avoir l'impression que la représentation d'une seconde phrase sémantiquement incohérente est « parasitée » par la représentation déjà construite à partir de la première phrase. Ce qui est intéressant, c'est que dans notre analyse sur les temps de lecture, nous avons montré qu'il n'y avait aucune différence pour la lecture des phrases des conditions Suite ou NonSuite alors qu'on aurait pu s'attendre à un ralentissement pour les mots de la seconde phrase lorsque celle-ci n'est pas cohérente avec la première (ralentissement qui aurait pu correspondre à la gêne décrite par les sujets pour se créer une représentation d'une seconde phrase lorsque celle-ci n'est pas sémantiquement cohérente avec la première). Les différences comportementales observées (TR et scores), entre condition Suite et Non Suite, sont donc distinctes de la vitesse de lecture puisque les mots sont lus à la même vitesse dans les deux conditions. Ces informations sont importantes pour mieux comprendre le timing de la création d'une représentation et il sera crucial de les mettre en lien avec les résultats de l'expérimentation similaire menée en EEG (voir Chapitre II.3).

Les résultats les plus pertinents de ce protocole comportemental sont sûrement ceux qui portent sur l'étude des groupes dont les capacités de représentation sont différentes. Il est tout d'abord intéressant de constater que cette différence que nous avions déjà observée (voir chapitre IRMf-DTI) est reproductible et statistiquement valable pour un nombre plus important d'individus : il est clair que les capacités de représentation ne sont pas les mêmes entre les individus. De plus, les profils de ces groupes, initialement caractérisés sur le COR, sont confortés par le fait que les différences sont aussi observées sur un autre marqueur comportemental, le temps de réaction. Contrairement à notre assertion d'origine qu'il existait 2 groupes dont les capacités de représentation sont distinctes (groupes basés sur un test T de Student), il semble qu'il existe en fait des profils plus fins puisque d'autres analyses statistiques (K-moyennes) ont révélé l'existence de 3 groupes (les bons représentateurs, les représentateurs intermédiaires et les mauvais représentateurs). Les bons représentateurs semblent se représenter plus facilement et plus

rapidement les phrases que les représentateurs intermédiaires et les mauvais représentateurs. Ce qui est intéressant, c'est que là aussi ces différences intergroupes ne semblent pas associées à des différences dans les temps de lecture, puisque les différences que nous avons trouvées sont minimes et ne semblent pas être des facteurs explicatifs convaincants des différences interindividuelles observées au niveau comportemental. Ces résultats montrent que, à compétences égales de lecture, il existe des différences entre les sujets en ce qui concerne les capacités de représentation.

En revanche, les différences entre les groupes se reflètent fortement sur les effets contextuels où des profils différents apparaissent. Que ce soit pour les scores ou les TR, les « bons représentateurs » ne sont pas perturbés par l'incohérence sémantique entre deux phrases puisque leurs données comportementales sont très bonnes (scores hauts et TRs rapides) et ne diffèrent pas significativement entre les deux conditions (Suite et Non Suite). Les « représentateurs intermédiaires » seraient le groupe qui s'appuie le plus sur les informations contextuelles. Chez ce groupe de sujets, dont les performances restent relativement bonnes (scores et TR plus proches du groupe des bons que des mauvais), on trouve des effets facilitateur et inhibiteur qui sont peu présents pour les autres groupes. Les capacités de représentation des mauvais représentateurs seraient nettement inférieures à celles des autres groupes (scores très bas et TR lents), de ce fait, elles ne bénéficieraient pas des informations contextuelles fournies par la 1ère phrase. En tout cas, l'existence de ces 3 groupes, dont l'utilisation des indices contextuels n'est pas la même, nous conforte dans l'idée que les capacités de représentation ne sont pas égales d'un individu à un autre. Nous allons maintenant essayer de comprendre à quoi peuvent être liées ces différences.

- **<u>Différences interindividuelles et capacités de représentation</u>**

Comme nous l'avons déjà précisé dans l'introduction de ce chapitre, la plupart des travaux qui portent sur les différences individuelles durant la lecture de phrases se sont concentrés sur les troubles de la compréhension et majoritairement chez des enfants. Cependant, capacité de représentation et compréhension seraient étroitement liées. Ainsi, pour Cain (2006), le résultat d'une bonne compréhension est une représentation cohérente de la signification d'un texte. Chez les jeunes enfants qui apprennent à lire, les capacités de compréhension sont fortement associées aux capacités en lecture : devenir un bon lecteur (c'est-à-dire un enfant dont la lecture est fluide) dépend à la fois de compétences phonologiques, mais aussi de la compréhension linguistique (Gough & Tunmer, 1986 ; Tunmer & Hoover, 1992 ; Nation & Snowling, 1998, 2004). Durant l'apprentissage de la lecture, les capacités de décodage et de compréhension semblent interdépendantes : les enfants qui développent des compétences en lecture normale ont à la fois de bonnes compétences en

compréhension et en décodage phonologique, à l'inverse des enfants mauvais lecteurs (« poor-readers ») qui présentent généralement des difficultés dans les deux domaines (Gough & Tunmer, 1986). Cependant, ce ne sont pas ces difficultés en lecture qui expliqueraient une mauvaise compréhension. Ainsi, on trouve des enfants ayant des compétences adaptées pour leur âge en ce qui concerne la lecture des mots, mais qui ont une mauvaise compréhension en lecture. Plusieurs études mettent ainsi en évidence l'existence de différences dans les capacités de compréhension pour des enfants d'âge appariés et dont les compétences verbales sont similaires (vocabulaire et niveau de lecture). Ce phénomène peut toucher jusqu'à 10% des écoliers âgés entre 7 et 11 ans (Yuill & Oakhill, 1991), donc des enfants qui ont déjà appris à lire (en anglais, ils sont désignés sous le terme de « poor comprehenders », Cain, 2006). En ce qui nous concerne, notre étude semble également montrer que, chez des adultes, les capacités de représentation ne sont pas dépendantes des capacités de lecture puisque nous n'avons pas mis en évidence de différence spécifique entre les différents groupes concernant le facteur « temps de lecture ». Par contre, comme le suggère le questionnaire post-protocole, la lecture pourrait avoir une influence sur le développement de ces capacités de représentation puisque nous avons observé que les personnes qui aiment le moins lire sont majoritairement des « mauvais représentateurs ». La lecture correspondrait à une sorte « d'entraînement » pour les capacités de représentation : les personnes qui aiment lire et qui lisent le plus, pourraient être celles qui auraient développé les meilleures capacités à se représenter une situation à partir d'une phrase. A moins que ce ne soit le contraire et que ces personnes aiment plus la lecture car elles ont de bonnes capacités de représentation à la base. Bien sûr, ces hypothèses sont purement spéculatives et il serait très intéressant de mener une étude développemental chez des jeunes enfants pour savoir quand et comment se développent les capacités de représentation.

Selon Cain (2006), la mémoire de travail (Baddeley, 1986) semble jouer un rôle important dans ces différences de compréhension. En effet les enfants ayant des capacités de compréhension moins bonnes semblent avoir des difficultés spécifiques par rapport à toute une gamme de compétences dépendantes de la mémoire, compétences susceptibles d'influencer la construction d'une représentation basée sur le sens (Cain, 2006). Ce sont spécifiquement des compétences tributaires de la mémoire de travail qui sont touchées, car ces compétences nécessitent qu'une information soit stockée en mémoire alors que d'autres sont traitées simultanément (par exemple, stocker en mémoire le sens de la phrase précédente alors qu'on lit la phrase suivante). Chez les enfants, un certain nombre de travaux mettent en évidence ces déficits spécifiques pour les « poor comprehenders » dans diverses tâches de mémoire de travail en lien avec le langage (tâche d'empan de mots, stockage et traitement simultanés d'informations linguistiques, etc…) alors même que leurs performances dans des tâches impliquant la mémoire à court terme sont intactes (Leather & Henry,

1994 ; Swanson & Berninger, 1995; Yuill, Oakhill, & Parkin, 1989 ; de Beni & Palladino, 2000; Oakhill, Yuill, & Parkin, 1986 ; Oakhill, Cain, & Yuill, 1998). Plus précisément par rapport aux tâches impliquant des compétences sémantiques, ils semblent moins bons pour produire des récits narratifs avec une structure causale cohérente (Cain, 2003), établir des liens entre des phrases individuelles pour établir une cohérence locale, intégrer de nouvelles informations à une représentation déjà existante pour établir une cohérence globale (Cain & Oakhill, 1999; Cain, Oakhill , Barnes, et Bryant, 2001; Oakhill, 1982, 1984) ou encore à conserver le sens du texte au fur et mesure qu'ils le lisent (Oakhill, Hartt, et Samols, 2005).

Chez les adultes aussi, les mesures de la capacité de mémoire de travail (que ce soit dans des tâches de lecture ou d'écoute) semblent significativement reliées au niveau de compréhension ainsi qu'à des compétences spécifiques qui favorisent la compréhension telles que la capacité à faire des inférences (Daneman & Merikle, 1996). Une des hypothèses est que les personnes qui ont des capacités de compréhension moins riches auraient des mécanismes d'inhibition[6] déficients qui les rendent moins aptes à éliminer de la mémoire de travail les informations qui ne sont pas ou plus pertinentes (de Beni, Palladino, Pazzaglia, et Cornoldi, 1998). Ces informations parasites peuvent nuire à l'intégrité de l'information importante à traiter, voire même interférer avec la capacité à traiter l'information, et des informations cruciales peuvent être altérées voire perdues si le système est surchargé. De cette façon, la faiblesse des capacités inhibitrices peut nuire à la compréhension d'un texte. Il existe effectivement un certain nombre de preuves que les adultes et les enfants dont la compréhension en lecture est plus pauvre ont des mécanismes inhibiteurs déficients. Par exemple, ils sont plus lents à supprimer les significations non pertinentes de mots ambigus (Barnes, Faulkner, Wilkinson, et Dennis, 2004 ; Gernsbacher & Faust, 1991). Adultes et enfants « poor comprehenders » sont plus susceptibles de retenir des informations qui ne sont pas pertinentes par rapport aux informations principales d'un texte qu'ils viennent d'étudier (de Beni & Palladino, 2000). Dans des tâches de mémoire, les adultes retiennent plus souvent des mots des essais précédents qui auraient dû être oubliés (de Beni et al., 1998) et chez les enfants, on trouve un plus grand nombre d'intrusions durant les phases de rappel de tâches de mémoire (de Beni & Palladino, 2000). Comme le souligne Cain (2006), ces résultats suggèrent une différence qualitative dans le fonctionnement de la mémoire de travail entre les personnes ayant des bonnes capacités de compréhension et celles dont ces capacités sont plus limitées. Si cette hypothèse est plus discutable chez les enfants dont les compétences linguistiques sont encore en développement, elle est très intéressante chez les adultes où en général « good » et « poor comprehenders » sont recrutés parmi les populations d'étudiants

[6] L'inhibition n'est pas un terme unitaire, il peut faire référence à différents processus cognitifs ou comportementaux ; mais dans cette hypothèse l'inhibition fait référence au processus de régulation de la mémoire de travail (Cain, 2006).

du premier cycle (de Beni et al., 1998), on peut donc supposer que les deux groupes de sujets ont un niveau raisonnable et équivalent en ce qui concerne les compétences linguistiques (ce qui est aussi le cas dans notre étude où les participants ont au minimum le baccalauréat).

Ces données semblent en adéquation avec les nôtres : de la même façon que pour la compréhension, des personnes qui ont des capacités de représentation moins bonnes auraient des difficultés à supprimer les informations linguistiques non pertinentes en raison de processus inhibiteurs moins adaptés qui conduisent à un moins bon fonctionnement de la mémoire de travail. C'est pour cette raison que les personnes ayant de moins bonnes capacités de représentation seraient plus influencées par l'information issue de la phrase précédente et que les effets perturbateurs pour la condition NonSuite seraient les plus marqués dans notre protocole (en tout cas pour les représentateurs intermédiaires). Les bons représentateurs pourraient avoir, quant à eux, de meilleures capacités en mémoire de travail, ce qui leur permettrait d'intégrer plus facilement le sens d'une nouvelle phrase, que celle-ci soit cohérente avec la première ou non. Initialement, nous n'avions pas envisagé cette hypothèse concernant la mémoire de travail.

- **Perspectives**

Dans une prochaine expérimentation, il serait pertinent de tester comment ces aspects mnésiques peuvent influencer les capacités de représentation, en ajoutant un test de la mémoire de travail à notre protocole. Par ailleurs, comme nous l'avons mentionné, la plupart des travaux cités concernent la compréhension, il serait aussi très intéressant de mesurer à la fois les capacités de compréhension et de représentation pour savoir jusqu'à quel point ces deux processus sont liés et si le résultat d'une bonne compréhension est effectivement une représentation cohérente de la signification (Cain, 2006). Il faudrait alors vérifier si on retrouve les mêmes différences interindividuelles : est-ce que les sujets qui ont les meilleures capacités de représentation sont aussi ceux qui ont les meilleures capacités de compréhension ? On pourrait envisager de mettre en place le même protocole avec une question qui concerne le degré de représentation (comme c'est le cas actuellement dans notre protocole) à laquelle on pourrait ajouter une question visant à tester la compréhension. Les phrases utilisées dans le protocole actuel étant très simples, nous pourrions utiliser des stimuli plus complexes (par exemple : « *Le marin qui monte au mât regarde le capitaine hurler après la mouette qui vient de lui voler son poisson.* »), on pourrait alors employer des questions visant à tester plus profondément la compréhension (question : « *est-ce le capitaine qui crie ? »*). Il serait d'ailleurs tout à fait intéressant de tester les capacités de représentation des sujets sur ce type de phrases : est-ce que les bons représentateurs ont toujours des capacités équivalentes

lorsque la difficulté des stimuli est croissante ou est-ce que les différences entre les groupes tendent à disparaître ? Est-ce qu'une meilleure représentation de ces phrases complexes permet d'aboutir à une meilleure compréhension ?

Bien sûr, pour étudier pleinement les liens entre capacités de représentation et de compréhension, il serait nécessaire de compléter cette réflexion par des protocoles en neuroimagerie. Nous pourrions utiliser le même type de protocole en IRMf soit en donnant aux sujets la consigne de se représenter la phrase (exactement comme pour ce protocole), soit en ne donnant aucune consigne particulière (comme dans notre premier protocole IRMf). Comparer directement ces deux conditions nous permettrait de savoir si les mêmes zones cérébrales sont impliquées dans la représentation et dans la compréhension et ainsi d'en savoir plus sur les liens entre ces deux processus. Des expériences en neuroimagerie seraient également cruciales pour mieux comprendre l'existence de différences individuelles dans les capacités de représentation.

Enfin un dernier sujet, qui mérite réflexion et donnera lieu à des expérimentations plus poussées, concerne l'accès aux représentations en mémoire. Nous ne parlons plus cette fois de mémoire de travail, mais cette nouvelle piste n'est pas incompatible avec le rôle supposé de la mémoire de travail dans les capacités de représentation. En effet, le questionnaire post-protocole a amené des éléments de réflexion intéressants sur la nature des représentations : les différences interindividuelles pourraient en fait correspondre à des « stratégies » représentationnelles différentes, faisant appel à des systèmes mnésiques distincts et qui se refléteraient dans les résultats comportementaux des sujets. Les « représentateurs intermédiaires » seraient les plus dépendants du contexte car ils auraient besoin de passer par une représentation mettant en scène des proches ou de faire appel à une scène particulière déjà construite (une pub passant à la télé, une scène d'un film). Ils auraient besoin de se remémorer une situation dans un contexte donné pour se créer une représentation, ce qui pourrait faire appel à la mémoire épisodique. Les « mauvais représentateurs » auraient, quant à eux, besoin de se projeter dans la scène et de se représenter en train de faire l'action ; des représentations qui feraient peut-être appel à la mémoire autobiographique. Mauvais représentateurs et représentateurs intermédiaires feraient appel à des souvenirs nécessitant un référentiel précis, qui leur est personnel (soit visualisation de proches, soit visualisation d'eux-mêmes). Ce qui pourrait expliquer les différences au niveau des données comportementales entre « représentateurs intermédaires » et « mauvais représentateurs », ce serait le besoin de ces derniers de se représenter faisant l'action : cette nécessité de projection du soi dans l'action pourrait être plus difficile (que de simplement se rappeler un proche faisant l'action ou une scène vue à la télé), notamment pour des actions qui leur sont moins familières et cela pourrait influencer à la fois les scores et les TR. Les « bons représentateurs » auraient une stratégie plus automatique basée sur

des connaissances plus généralisées, comme si le répertoire même des actions faisait partie d'une « sémantique de l'action » (l'action de faire la vaisselle, l'action d'aller au cinéma…). En tout cas, ils n'utiliseraient pas un référentiel spécifique, lié à la mémoire épisodique comme les autres groupes (une scène précise vue en vrai ou à la télé, impliquant des proches ou des personnages connus, ou encore une projection du soi dans l'action) puisqu'ils décrivent majoritairement des inconnus, des personnages sans visage ou des silhouettes. Cette compétence des « bons représentateurs » ferait plus appel à la mémoire sémantique, dans sa définition de système mnésique par lequel l'individu stocke des connaissances générales. En résumé, les bons représentateurs auraient des capacités de représentations plus automatisése, faisant appel à la mémoire sémantique ; les représentateurs intermédiaires utiliseraient plus leurs souvenirs et la mémoire épisodique tandis que les mauvais représentateurs feraient appel à leur mémoire autobiographique pour se créer une représentation. Cette hypothèse d'une différence d'utilisation des systèmes mnésiques nécessite d'être confirmée par des travaux plus précis en neuroimagerie. Mais les résultats IRMf entre bons et mauvais représentateurs (résultats que nous n'avons pas détaillés et qui sont consultables en annexe 3) apportent une piste intéressante. En effet, notre analyse statistique du pourcentage de changement du signal a permis de mettre en évidence deux aires où le fonctionnement entre bons et mauvais représentateurs n'est pas le même : le noyau caudé et le lobe temporal. Ces deux régions, ont été mises en cause dans l'hypermnésie, une pathologie très rare qui consiste en une exaltation de la mémoire épisodique (tous les souvenirs personnels sont exceptionnellement bien mémorisés) ; des études IRMf sur un patient chez qui cette pathologie est extrêmement exacerbée ont montré que le noyau caudé et le lobe temporal étaient plus volumineux chez ce patient, suggèrant un développement neurologique atypique. Chez nos sujets sains, ces régions ne présenteraient pas des différences morphologiques mais elles pourraient être impliquées différentiellement et expliquer, en partie, un recours différent à la mémoire épisodique/autobiographie/sémantique et les capacités de représentation variées des sujets.

Synthèse des résultats Partie Expérimentale 2 – Expérience comportementale

Avec ce protocole comportemental, nous souhaitions étudier les différences interindividuelles dans la capacité à se représenter des phrases, présentées en modalité auditive ou visuelle. Nos résultats montrent tout d'abord que la capacité à se représenter une phrase n'est pas dépendante du mode de présentation, ce qui suggère que cette capacité fait appel à un système sémantique amodal. Nous avons démontré qu'ils existaient bien des différences entre les individus et que ces différences se traduisaient par des effets sur des marqueurs comportementaux comme la facilité à se représenter une phrase (évaluée à travers le COR) et la vitesse de réponse (TR). Il existe différents profils de représentateurs parmi les sujets, mais ceux-ci ne sont pas liés aux capacités de lecture des individus (TL). Par ailleurs, nos résultats ont aussi permis de déterminer que l'effet contextuel était important : le contexte induit par une première phrase influence la capacité à se représenter une seconde phrase différemment selon que celle-ci est sémantiquement cohérente ou non avec la première. Les différents profils de représentateurs ne seraient pas sensibles de la même façon à cet effet contextuel.

II.3) Partie Expérimentale 3 : Utilisation de l'EEG pour mettre en évidence les processus temporels communs aux stimuli verbaux et non-verbaux ainsi que les effets liés aux informations contextuelles sur la création d'une représentation

II.3.*1*) Introduction

S'il est largement admis, aujourd'hui, que l'accès aux connaissances conceptuelles appelle l'activation d'informations en mémoire sémantique, certaines questions demeurent toujours ouvertes, notamment en ce qui concerne la temporalité de l'accessibilité aux représentations et leur stockage selon que l'on se réfère à une notion de système sémantique amodal ou modalité-spécifique (mots vs images).

Comme nous l'avons déjà mentionné, la théorie du double codage de Paivio (1986) postule en faveur de l'existence de routes de codage sémantique dissociées pour les images et les phrases et d'un traitement plus rapide pour la compréhension des images. C'est ce que soutiennent les modèles qui s'inscrivent dans une perspective modalité-spécifique du système conceptuel (Paivio, 1971, 1991 ; Shallice, 1988; Warrington & Shallice, 1984 ; Warrington & McCarthy, 1994). Bien qu'ils existent des processus spécifiques indéniables selon que les stimuli sont présentés sous la forme d'images ou de mots et que ceux-ci se traduisent par une implication de zones cérébrales différentes, mise en évidence par les travaux en neuroimagerie fonctionnelle (Vandenbherge, 1996 ; Moore & Price ; 1999, Devlin et al., 2002 ; Bright, Moss & Tyler ; 2004 ; Chee et al. 1999 ; Shinkareva et al. 2011, etc.), la plupart de ces mêmes travaux mettent également en évidence l'existence de zones cérébrales communes. Dans notre propre travail, nous avons démontré précédemment qu'un système commun était impliqué dans la compréhension de phrases et de scènes visuelles complexes ; ces résultats sont plus compatibles avec les modèles affirmant que toutes les voies de traitement convergeraient vers un seul ensemble de représentations conceptuelles, système qui serait commun aux deux modalités (Caramazza, Hillis, Rapp, et Romani, 1990 ; Lambon Ralph, Patterson, & Hodges, 1997; Lambon Ralph, Graham, Patterson, & Hodges, 1999; Rapp, Hillis, & Caramazza, 1993 ; Bright et al., 2004 ; Plaut, 2002). Ces résultats concorderaient aussi avec les résultats portant sur des données temporelles qui montrent que les activations sensorimotrices précoces soutiennent la compréhension (Jirak et al. 2010 ; Pulvermüller, 2008; Gallese, 2007, 2008 ; Glenberg 2012).

Les analyses EEG, en particulier la méthode des potentiels évoqués (PEs), permettent d'apporter des éléments de réponse supplémentaires sur la nature amodale ou unimodale du système sémantique grâce à leur excellente résolution temporelle : elles nous renseignent sur l'organisation des connaissances sémantiques via des données sur le timing de l'intégration

sémantique propres à chaque modalité, mais également sur l'existence de marqueurs spécifiques et/ou communs impliqués dans le traitement des stimuli verbaux et non verbaux. La technique des PE est l'un des meilleurs moyens de mettre en évidence des modifications minimes et transitoires de l'activité cérébrale en réponse à des stimulations sensorielles car ils permettent d'étudier le décours temporel de l'activité électrique cérébrale correspondant à différents types d'opérations et/ou de traitements cognitifs, ; de plus son usage est simple et non invasif.

- **La méthodologie EEG**

L'électro-encéphalographie (EEG) permet de mesurer l'activité électrique du cerveau par des électrodes placées sur le cuir chevelu. Les enregistrements électro-encéphalographiques représentent une méthodologie de choix en neurosciences pour étudier en temps réel les processus cognitifs liés à une tâche particulière. Cette technique provient de l'électroencéphalogramme (EEG), dont les premiers enregistrements ont été obtenus d'abord chez le chien (Pravdich-Neminsky, 1913) puis chez l'homme, en 1924, par le psychiatre allemand Hans Berger (Berger, 1929), mais des travaux encore plus anciens chez le singe, le lapin et le chien avaient déjà montré l'existence de phénomènes électriques dans le cerveau (Richard Caton, 1875 ; Adolf Beck, 1890). Il a fallu attendre les années 50 pour que l'EEG se développe vraiment et se répande dans la pratique médicale courante, en particulier avec les travaux sur l'épilepsie ; ce n'est pas avant les années 80 que l'EEG est devenue une méthode fréquemment utilisée pour l'étude des processus cognitifs (Kolb et Whishaw, 1986).

L'activité nerveuse est un processus électrochimique et bien que les potentiels électriques produits par un neurone individuel soient extrêmement faibles, l'activité synchrone de grandes populations de neurones génère des potentiels suffisamment grands pour qu'on puisse les enregistrer sur le scalp. Elle permet ainsi l'enregistrement de l'activité électrique cérébrale spontanée grâce à des électrodes posées à la surface du scalp car les tissus du cerveau, du crâne et du cuir chevelu conduisent passivement les courants électriques produits par l'activité synaptique. La technique d'électroencéphalographie mesure les variations de voltage résultant des courants ioniques à l'intérieur des cellules nerveuses. Les neurones sont électriquement chargés (ou polarisés) par des protéines de transport membranaires qui pompent les ions à travers les membranes. Les neurones communiquent entre eux principalement au moyen de neurotransmetteurs chimiques, captés par la membrane post-synaptique d'un neurone cible. L'excitation d'un neurone entraîne l'ouverture de canaux ioniques au niveau de sa membrane. La composition ionique étant différente à l'intérieur et à l'extérieur des cellules, l'ouverture des canaux engendre un mouvement de particules chargées dans le milieu intra et extracellulaire. Ainsi, les neurones échangent constamment des ions avec le milieu

extracellulaire, par exemple pour maintenir un potentiel de repos ou pour propager des potentiels d'action. Quand un potentiel post-synaptique excitateur (PPSE) parvient à un neurone cible, celui-ci est alors le siège d'une dépolarisation transitoire. A l'inverse, quand un potentiel post-synaptique inhibiteur (PPSI) parvient à un neurone cible, cela génère une hyperpolarisation transitoire. Ces phénomènes sont dus à des échanges d'ions, donc à des courants électriques, traversant la membrane cellulaire du neurone. Des ions de charges semblables se repoussent les uns les autres et forcent les ions à sortir des neurones ; quand ce phénomène se produit dans de nombreux neurones en même temps, cela provoque une vague de déplacements (volume conduction). L'amplitude de ces micro-courants décroît rapidement lorsqu'on s'éloigne de leur source, ce qui rend difficile l'acquisition de l'activité provenant de sources profondes, plus difficiles à détecter (Klein et al, 2006). Cependant, on peut généralement enregistrer ces micro-courants car ils peuvent former un dipôle unitaire parallèle à l'axe de la dendrite pour chaque neurone cortical actif. Lorsque cette onde d'ions atteint les électrodes situées au niveau du cuir chevelu, elle interagit avec le métal des électrodes qui est conducteur (en poussant ou en tirant les électrons du métal des électrodes) et on peut ainsi mesurer la différence de voltage entre deux électrodes grâce à un voltmètre. Cela permet d'obtenir l'activité électro-encéphalographique en enregistrant ces différences de voltage sur une période de temps et un nombre d'électrodes important. Le potentiel électrique généré par un seul neurone est trop petit pour être enregistré par l'EEG, mais l'EEG reflète toujours la sommation de l'activité synchrone des milliers ou des millions de neurones qui ont une orientation spatiale similaire (Nunez et al, 1981 ; Tatum et al. 2008). Les neurones pyramidaux du cortex sont connus pour produire un signal EEG important en raison de l'organisation en colonnes fonctionnelles, de leur alignement et parce qu'ils déchargent de façon synchrone et produisent ainsi un dipôle important, correspondant à la somme vectorielle des dipôles unitaires des neurones de la colonne. En effet, lorsque les dendrites d'un neurone sont organisées de façon radiale ou aléatoire autour du corps cellulaire, le dipôle unitaire résultant est souvent nul, alors que dans le cas des cellules pyramidales du cortex qui ont une arborisation dendritique orientée de façon perpendiculaire à la surface corticale, des dipôles unitaires équivalents non nuls peuvent être observés, générant les micro-courants (de l'ordre de la dizaine de microvolts) que l'on peut enregistrer à la surface du scalp. L'EEG de scalp permet de recueillir des oscillations de diverses fréquences, ces oscillations représentent l'activité synchronisée de réseaux de neurones. Certains patterns d'oscillations sont bien connus et peuvent être associés à différents processus (ex : rythmes du sommeil).

- *La technique des potentiels évoqués (PEs)*

En l'absence de tâche cognitive spécifique, on enregistre un EEG basal qui reflète l'ensemble des activités neuronales spontanées ; mais lorsqu'on présente un stimulus particulier, l'activité

cérébrale est modifiée transitoirement et une activité spécifique au traitement de ce stimulus vient s'ajouter au signal associé à l'activité spontanée. On peut alors enregistrer un potentiel évoqué ou réponse évoquée (PE, dit aussi ERP pour l'anglais Event-Related Potential) qui désigne la modification du potentiel électrique produite par le système nerveux en réponse à la présentation d'un stimulus particulier. Un potentiel évoqué est lié à la présentation d'un stimulus dans une modalité sensorielle (un son, une image, etc.) ; on parle alors de potentiel mais il peut aussi être lié à un événement interne, notamment une activité cognitive (attention, préparation motrice, etc.). En neurosciences, on fait parfois la distinction entre un potentiel évoqué « exogène » (evoked potential), qui désigne les modifications neuro-électriques consécutives à une stimulation externe, modifications observables aussi bien dans le système nerveux périphérique que dans le système nerveux central (voire même pouvant être associé aux mouvements des effecteurs dans le cas des potentiels évoqués moteurs) et un potentiel cognitif lié à un événement (ERP) ou potentiel évoqué « endogène », qui désigne les changements de potentiel électrique associés à un processus mental. Les courbes ERP reflètent non seulement l'intégration des caractéristiques objectives des stimuli (souvent dans les 200 premières millisecondes), mais aussi les influences endogènes liées aux réactions des sujets face à la tâche, aux stimuli et aux instructions des expérimentateurs (en général entre 200 et 1500 ms post-stimulus).

La technique des potentiels évoqués (PEs) consiste à obtenir et à moyenner les réponses évoquées sur des portions d'EEG synchronisées par rapport à la présentation d'un stimulus particulier ; ce qui permet ainsi de distinguer l'activité corticale liée à ce stimulus de l'activité corticale spontanée (Rugg & Coles, 1995). Ces signaux électriques des PEs étant en général très faibles (de l'ordre du micro-Volt), il est souvent nécessaire de répéter l'enregistrement un grand nombre de fois pour augmenter le rapport signal sur bruit en moyennant l'ensemble de ces mesures et obtenir ainsi une caractérisation fiable du potentiel évoqué. La réponse spécifique à la présentation d'un stimulus particulier se manifeste dans le PE par une succession de fluctuations négatives ou positives par rapport à un niveau basal ayant une distribution temporelle et spatiale propre. Ainsi, un certain nombre de « marqueurs évoqués » ont pu être caractérisés car un potentiel peut être décrit par différents paramètres comme sa polarité (ondes négatives N ou positives P), son amplitude, sa latence (N400 pour une onde négative ayant une latence de 400ms) ou d'après leur ordre d'apparition dans la succession des marqueurs de même polarité (N1/P2 si la deuxième onde positive suit le premier marqueur négatif). Certains marqueurs particuliers sont nommés en fonction du type de tâche dans lequel ils apparaissent (MMN pour Mismatch Negativity qui est observée lorsqu'un stimulus inattendu est inséré au sein d'une série régulière de stimuli identiques) ou en fonction de leur caractéristiques spatiales ou spatio-temprorelles (LAN pour left anterior negativity, une onde

négative de distribution spatiale antérieure gauche). Un ERP bien défini a toujours une latence et une forme caractéristique et est, en général, associé à une ou plusieurs fonctions cognitives spécifiques (par exemple, l'onde N170 observée en réponse à la présentation de l'image d'un visage), même si la grande variabilité des tâches dans lesquelles on peut trouver un potentiel évoqué rend parfois difficile la comparaison d'études, la similitude ou la différence entre différents effets étant parfois compliquée à établir.

- *Analyses Temps-fréquence*

Si les analyses ERPs représentent la façon la plus classique d'analyser l'évolution du signal liée à un événement en ne tenant compte que des aspects temporels, les données obtenues en EEG peuvent aussi être caractérisées différemment via des analyses Temps-fréquence, qui apportent une perspective différente et complémentaire pour caractériser les traitements temporels liés à un processus cognitif. Le signal électroencéphalographique (EEG) étant essentiellement caractérisé par son contenu fréquentiel, l'activité électrique cérébrale peut être décrite comme la somme des oscillations qui codent pour le transfert de l'information d'une population neuronale à une autre (le signal pouvant être représenté comme la somme des oscillations des différentes fréquences, représentées par différentes fonctions sinus et cosinus). Dans certaines situations expérimentales (mais aussi physiologiques ou pathologiques), les composantes fréquentielles peuvent subir des modifications relativement brèves et mesurables. Les oscillations sont des événements transitoires et imprévisibles et seule une analyse des variations fréquentielles permet de décrire ces événements. Les modifications des composantes fréquentielles sont assez bien décrites par les représentations temps-fréquences (RTF) car elles permettent une décomposition pertinente du signal. La méthode de décomposition spectrale permet de déterminer quelles fréquences interviennent dans un signal EEG. Mais surtout, les analyses temps-fréquence des enregistrements EEG permettent d'étudier simultanément l'évolution du signal au niveau temporel et fréquentiel : le but de ce type d'analyses est non seulement d'appréhender le contenu fréquentiel du signal mais aussi de déterminer à quel moment dans le temps la contribution de certaines fréquences est la plus importante. On peut ainsi déterminer la phase particulière d'une fréquence précise à un instant donné (phase locking), notamment dans un protocole expérimental, certains patterns d'oscillations pouvant être associés à différents processus cognitifs.

Observer un signal au cours du temps permet de déterminer son début, sa fin et d'éventuelles variations qualitatives, mais décomposer les périodicités (fréquences) d'un signal est moins évident. Pour effectuer les analyses fréquentielles du signal EEG, on utilise généralement la transformée de Fourier afin de décomposer un signal en une fréquence fondamentale et les différentes harmoniques

qui composent le spectre du signal, chaque fréquence correspondant à une fonction sinusoïdale dépendante du temps et apportant une information sur la régularité globale du signal. Mais la transformée de Fourier présente des limites pour l'analyse du signal EEG, car la transformée de Fourier d'un signal EEG ne permet pas une bonne définition de l'enchaînement des oscillations neuronales, elle fait simplement ressortir les différents types d'oscillations mises en jeu. D'autres modèles d'analyses sont donc appliqués pour mieux traiter les signaux non stationnaires (Fourier à fenêtre glissante, transformée en ondelettes de Morlet…) afin d'avoir une plus grande adéquation entre données fréquentielles et temporelles.

- *Avantages et Inconvénients des techniques d'analyse EEG*

Les différentes techniques d'analyse EEG (temps-fréquence et PEs) ont pour avantage principal de permettre une mesure directe de l'activité neuronale (contrairement à l'IRMf) et d'offrir ainsi une très bonne résolution temporelle. L'acquisition de données EEG pouvant s'effectuer de façon extrêmement rapide (de 500 à 1000 mesures par seconde pour les amplificateurs classiques), les analyses du signal EEG permettent donc d'obtenir des informations très fines sur les dynamiques de traitement, sur l'organisation temporelle et fréquentielle de différents processus engagés dans une tâche cognitive. De plus la méthode EEG est entièrement non invasive et moins contraignante que d'autres méthodes telles que l'IRMf (méthode plus bruyante, confinement dans la machine IRMf…). Mais la principale limite des analyses EEG réside dans leur faible résolution spatiale. Que ce soit l'observation d'un marqueur de surface par la méthode des PEs ou du spectre temps-fréquence d'une ou plusieurs électrodes, ces méthodes ne permettent pas de localiser précisément la source ou les sources de ces phénomènes. En effet, une activité de surface reflétant l'intégration de l'ensemble des dipôles unitaires actifs au niveau cortical, un effet de surface peut donc représenter aussi bien l'activation d'une population neuronale isolée que la somme d'activations dues à plusieurs populations séparées spatialement. De ce fait, déterminer la localisation intracorticale de la source d'un marqueur de surface (résolution du problème inverse) génère systématiquement une infinité de réponses, même si de nombreuses améliorations récentes dans l'acquisition et dans le traitement des données EEG sont prometteuses (nombre d'électrodes, meilleurs algorithmes de reconstruction de source). De plus, une autre limitation provient de la sensibilité même de l'EEG à l'activité neuronale, qui ne permet d'enregistrer efficacement que les populations de neurones organisées en colonnes et placées radialement ou tangentiellement par rapport à la surface du scalp et non l'activité des colonies orientées différemment. Ainsi certains processus pourraient échapper à la résolution temporelle de ces différentes techniques d'analyses EEG.

- **Potentiels évoqués et traitement sémantique des mots et des images**

L'application de la technique des PEs à l'étude du langage a émergé au début des années 70, mais elle s'intéressait alors essentiellement à la localisation électro-corticale des processus généraux de production et de perception du langage ainsi qu'aux phénomènes de latéralisation hémisphérique (Morrell & Huntington, 1971 ; McAdam & Whitaker, 1971). Brown et al. (Brown, Marsh & Smith, 1973 ; 1976) s'essayèrent à l'interprétation de PEs générés par la lecture de phrases, mais leurs résultats furent peu concluants car leur étude ne portait que sur un très petit nombre de phrases. Kutas et Hillyard (1980) furent finalement les premiers à identifier un marqueur électrophysiologique dont l'amplitude est modulée par la présence de mots sémantiquement discordants dans le contexte d'une phrase : la N400.

- *la N400*

Dans cette expérience princeps, Kutas et Hillyard présentaient des phrases mot à mot se terminant soit par un mot concordant prévisible (e.g. c'était son premier jour de *travail*), soit par un mot discordant qui engendre une violation sémantique (exemple, elle tartine son pain de *chaussettes*), les auteurs faisant l'hypothèse que la différence entre ces deux conditions se traduirait par une variation de l'amplitude de la composante P300. C'est en fait un marqueur négatif qui s'est révélé être évoqué par les mots discordants : l'onde N400, qui est une onde négative (N), faisant un pic vers 400ms après l'apparition d'un mot (fenêtre de 200–600 ms post stimulus) et ayant une topographie centro-pariétale. Les auteurs ont interprété cette composante comme le reflet du traitement supplémentaire nécessaire pour intégrer une information sémantique anormale dans un contexte linguistique. L'interprétation fonctionnelle de la N400 a longtemps été débattue, notamment pour la distinguer d'une autre composante négative plus précoce, la N200. Mais il a été établi que la N200 serait plus spécifiquement associée à des processus de catégorisations des stimuli et qu'elle était incontestablement distincte de la N400 (Deacon et al., 1991 ; Pritchard et al., 1991 ; Kutas & Van Petten 1988, 1994).

Dès le début de la caractérisation de la N400, il est devenu clair assez rapidement que c'était l'amplitude de la réponse qui était la plus sensible aux manipulations expérimentales (l'amplitude devenant plus petite lorsque les éléments sont attendus et donc plus faciles à traiter) et que c'était une mesure susceptible de varier en fonction des mêmes facteurs que ceux qui influencent les TR (Kutas & Federmeier 2009 ; 2011). Au niveau physiologique, ces réductions d'amplitude pourraient refléter de plus petits potentiels post-synaptiques des mêmes neurones ou l'activation de moins de neurones parmi une population, ou encore une synchronie temporelle moindre entre les neurones

qui génèrent la N400. La latence de la N400 est, en revanche, un facteur qui reste généralement stable quels que soient les stimuli ou la tâche.

La N400 a tout d'abord été majoritairement associée à des processus langagiers via des études utilisant du matériel verbal. Il a, par exemple, été montré que l'amplitude de la N400 était proportionnelle à la probabilité d'occurrence d'un mot dans un contexte précis mais aussi qu'elle reproduisait l'effet d'amorçage sémantique observé au niveau comportemental puisque l'amplitude de la N400 est plus importante pour des couples où l'amorce et la cible sont discordantes sémantiquement (tigre-banane) vs couple amorce-cible concordantes (chien-chat) (Kutas et al., 1984 ; Federmeier & Kutas, 1999 ; Bentin et al., 1985 ; Kutas & Van Petten 1988). Des travaux ultérieurs consacrés à l'étude de la N400 ont montré qu'elle pouvait également être évoquée, en modalité auditive et visuelle, avec d'autres types de stimuli qui n'étaient pas liés à des éléments langagiers : visages, objets, actions, couleurs, sons… (Holcomb et al., 2005 ; West & Holcomb, 2002 ; McPherson & Holcomb, 1999 ; Rebai et al., 1997 ; Olivares et al. 1999 ; Kovalenko et al. 2012 ; Ganis, Kutas, & Sereno, 1996; Nigam, Hoffman, & Simons, 1992; Sitnikova, Kuperberg, & Holcomb, 2003). De plus, un certain nombre d'études confirment également qu'il faut qu'il y ait un traitement d'ordre sémantique pour générer une N400. Ainsi, une discordance musicale n'évoque pas de N400 mais des composantes positives plus précoces ou plus tardives, la P300b ou la P600 (Besson & Macar, 1987 ; Kutas, 1999). De même, un nombre inattendu à la fin d'une suite logique de chiffres (Polich et al., 1985), un silence inattendu dans une série de stimulations langagières (Besson et al., 1997) ne provoquent pas de N400, pas plus que la violation des règles d'accord, de l'ordre des mots dans une phrase ou de règles grammaticales basiques (Kutas & Hillyard, 1983 ; Osterhout & Holcomb, 1992 ; Hagoort et al., 1993 ; Osterhout & Mobley, 1995 ; Kutas & Kluender, 1994).

Finalement, après 30 ans de recherche sur la N400, il s'avère que cette composante est avant tout un marqueur essentiel de la compréhension sémantique et non pas une signature liée à des effets de violation sémantique ou une réaction face à la présentation de n'importe quel stimulus arbitraire (Kutas et Federmeier, 2011). Pour Kutas & Federmeier, les premiers travaux sur la N400 suggéraient déjà que cette composante était liée à des aspects profonds de traitement du sens et d'accès à la mémoire sémantique, indépendamment de la modalité : la dimension essentielle associée à cette composante serait une différence sens/non sens et non pas linguistique/non linguistique (même si cela reste une mesure particulièrement pertinente pour l'étude du langage). Ainsi, la large sensibilité de la N400 à la pertinence des stimuli et aux manipulations sémantiques la rend particulièrement intéressante pour s'interroger sur la nature de la mémoire sémantique (en particulier sur l'existence d'un système amodal unique ou de plusieurs sous-systèmes) et sur la façon dont l'information liée au sens est stockée dans le cerveau (Kutas & Federmeier 2000, 2011 ; Kutas & Van Petten 1994). Pour

Kutas & Federmeier, la réponse fournie par les données concernant la N400 soulève de sérieuses questions quant à la validité du débat amodal vs modalité-spécifique, car même s'il y a des similitudes importantes (forme de la N400, évolution temporelle, sensibilité fonctionnelle…), il y a aussi des différences importantes (notamment en terme de topographie de scalps) entre les différents types de stimuli qui peuvent générer une N400 (visages, sons, mots…) ; par exemple les images suscitent une N400 similaire à celle des mots concrets (Ganis et al.1996) mais avec une distribution plus frontale alors que celle associée à des mots abstraits qui est plus centro-postérieure. Cependant, ces différences pourraient aussi résulter d'autres facteurs liés aux stimuli ou à la tâche, puisque dans une autre étude comparant images et mots, Nigram et al. (1992) n'ont pas trouvé de différence significative entre les paramètres de la N400 (latence, amplitude, distribution topographique).

Willems et al. (2008) ont aussi comparé directement les effets de la N400 liés au langage et aux images dans une tâche de détection de l'incongruence et ils ont mis en évidence une composante N400 centro-frontale similaire quant à sa localisation, sa latence et son amplitude. Ces résultats vont à l'encontre d'une vision où les informations linguistiques seraient traitées différemment (dans le temps ou fonctionnellement) des informations non-linguistiques issues de l'analyse des objets et des scènes (Kutas & Federmeier, 2011). Il reste encore un certain nombre de questions non résolues concernant le rôle de cette composante et son implication, en tant que marqueur de l'intégration sémantique dans un traitement amodal.

Pour Kutas & Federmeier (2011), ces différences topographiques pourraient refléter l'existence de différentes aires cérébrales impliquées dans la construction du sens et dont le recoupement anatomique serait seulement partiel. Ainsi, les effets N400 seraient sensibles à la modalité puisqu'on observe des différences topographiques entre les différents types de stimuli/modalités, mais ils ne seraient pas pour autant modalité-spécifiques puisque ces différents types de stimuli génèrent des effets N400 proches ; les effets N400 pourraient peut-être représenter l'interface entre un fonctionnement unimodal et amodal. Bien sûr, il est très difficile d'identifier les sources neurales responsables de ces effets N400, à savoir quelles aires sont engagées mais aussi quand et comment, d'autant plus que les mêmes opérations fonctionnelles peuvent être effectuées par différents substrats neuroanatomiques. De même, il est difficile de savoir s'il existe des types distincts de N400 (en fonction des modalités/stimuli) sur la base des potentiels de surface en raison des chevauchements temporels possibles avec d'autres réponses (nous discuterons plus amplement de la localisation mais aussi des nouvelles conceptions concernant le rôle de la N400 dans les parties Résultats et Conclusion de ce chapitre).

- *la N300*

Les premiers à mettre en évidence un effet d'amorçage sémantique spécifique aux images furent Barrett et Rugg (1990) : en utilisant des images d'objets réels dans une tâche d'association sémantique, ils obtinrent 2 composantes négatives, une N300 et une N450, dont l'amplitude était plus importante pour les images d'objets non liés sémantiquement que pour les images d'objets liés. La N300 serait distincte de la N400 car, en plus de la distinction au niveau temporel, elle aurait une distribution topographique plus frontale alors que la N450 que les auteurs ont identifiée aurait une topographie comparable à celle de la N400 classiquement identifiée.

Plusieurs autres études utilisant des stimuli images ont confirmé l'existence d'une N300, distincte topographiquement de la N400. Mais, comme la N400, l'amplitude de la N300 dépendrait du degré avec lequel une image d'objet est activée dans un réseau sémantique mais aussi de la prédictibilité d'un stimulus dans un contexte donné (McPherson et Holcomb, 1999 ; Kahlaoui et al., 2007). Pour ces auteurs, la N300 serait spécifiquement associée au traitement sémantique propre aux images, les travaux comparant images et mots n'ayant pas permis de retrouver cette composante associée à des stimuli langagiers (Nigam et al., 1992 ; Ganis et al., 1996 ; West et Holcomb, 2002 ; Hamm et al., 2002) ; elle pourrait refléter l'effort d'appariement avec une représentation lorsque des stimuli sont non concordants. Hamm et al. (2002) ont choisi des paires mots-images pour tester l'amorçage sémantique : les mots pouvant servir d'amorce faisaient soit référence à une catégorie (chien) ou à une sous-catégorie spécifique (colley), les images cibles étaient soit congruentes avec le mot (mot : chien, image : colley ou mot : colley, image : colley), soit incongruentes mais de la même catégorie (mot : colley, image : bulldog) ou appartenant à une catégorie différente (mot : colley, image : voiture). L'effet N400 ressort pour tous les types de mots utilisés comme amorces et pour tous les mismatchs. Par contraste, on ne rencontre la N300 frontale que lorsque le mot amorce appartient à une sous-catégorie spécifique et seulement pour les mismatchs avec une catégorie basique différente. Provoquer une composante N300 requiert la violation d'une attente sémantique fortement reliée à la catégorie ; ces résultats semblent indiquer que la N300 serait reliée à des processus de catégorisation initiale des objets plutôt basiques alors que la N400 refléterait des processus de catégorisation sémantique de plus haut niveau (Hamm et al.2002). Cependant, l'existence d'une composante spécifique aux images ne remet pas nécessairement en cause l'existence d'un système sémantique commun. Pour McPherson & Holcomb (1999), le chevauchement temporel des composantes N300 et N400, dans le cas des images, pourrait être responsable des distributions de scalps différentes entre mots et images : il y aurait un recoupement partiel entre la N400 centro-postérieure et la N300, qui est plus précoce et distribuée plus frontalement. Cette interférence entre les deux composantes négatives pourrait modifier l'effet de la

N400 liée aux images et masquer le fait que des sources neurales communes puissent intervenir pour les deux types de stimuli. La signification fonctionnelle précise de la N300 reste à élucider ; McPherson et Holcomb suggèrent qu'elle est liée au traitement de «l'information spécifique à un objet", ils ont constaté que la N300 était moins sensible que la N400 au degré de chevauchement sémantique entre deux éléments mais plus sensible que cette dernière à l'information structurelle des objets (différences entre des pseudo-objets qui ne respectent pas les principes structurels et les objets « scrambles » qui n'ont pas de signification mais qui sont « possibles »). Par ailleurs, d'autres études n'ont pas retrouvé cet effet de N300 spécifique aux images, c'est le cas de Kovalenko et al. (2012) qui ont comparé les effets associés à un amorçage sémantique de 800 paires d'objets sémantiquement reliés ou non et qui rapportent des effets de N400 pour ces images mais pas de N300.

- **Autres composantes impliquées dans le traitement des informations verbales et non verbales**

Sans parler de traitement sémantique, il y a un certain nombre de différences fondamentales entre mots et objets qui sont connues pour affecter la nature du traitement de l'information qui sous-tend l'identification de ces deux types de stimuli. Ces différences pourraient expliquer pourquoi la catégorisation des objets est plus rapide que celle des mots ainsi que l'existence de composantes ERP spécifiques (Yum et al., 2011). Par exemple, les parties d'un objet (la patte d'un animal) peuvent avoir du sens alors que les parties d'un mot (les lettres) n'en ont pas individuellement. De même, la forme globale d'un objet apporte des informations auxquelles la forme d'un mot ne permet pas d'accéder. Pour rendre plus comparable le traitement des mots et des images et annuler ces effets, Yum et al. (2011) se sont intéressées à la comparaison d'idéogrammes chinois, de mots et d'images dans une tâche de catégorisation sémantique. Pour ces auteurs, il existerait différents niveaux d'informations liés à des composantes spécifiques. Entre 150 et 200 ms post-stimulus, les effets seraient liés aux traitements spécifiques de l'aspect des stimuli (scripts logographiques vs scripts alphabétiques). Les stimuli qui ont un aspect logographique, tels que les idéogrammes, différeraient des mots et de leurs aspects alphabétiques non seulement parce qu'ils ont une ressemblance brute avec les dessins d'objets mais aussi parce qu'ils permettent une association directe avec des unités qui ont du sens (un idéogramme = un seul caractère = une représentation sémantique spécifique). En revanche, les mots reposent sur des symboles abstraits (les lettres de l'alphabet) et une association est nécessaire entre les caractéristiques visuelles des mots et leur sens. Des effets modalité-spécifiques, indépendants de l'aspect logographique ou alphabétique, interviendraient entre 200 et 300 ms et ils seraient censés refléter les différences d'activations de représentations abstraites

spécifiquement liées au traitement des images ou des mots. Les effets qui apparaissent vers 400 ms, seraient amodaux et reflèteraient des activations dépendantes du degré de signification des stimuli.

Ces résultats sont en contradiction avec ceux de Van Doren et al. (2012) qui se sont intéressés à la chronométrie précise de l'identification des mots et des images et qui montrent des effets précoces communs, indépendants du format de présentation. Ces auteurs ont employé une version EEG de leur protocole déjà utilisé en IRMf (Van Doren et al., 2010) où la présentation des stimuli est très brève, ce qui permet de mettre en évidence des processus très précoces qui ne sont pas attribuables à de la verbalisation ou à des jugements de type sémantique. Dans cette étude ERP, les auteurs ont montré que des effets amodaux très précoces, liés à l'identification des stimuli, se produisaient entre 180 et 200 ms : ils ont une polarité positive et une localisation centropostérieure (CPz). Ils sont suivis de l'apparition d'une autre composante positive centrofrontale et de composantes négatives bilatérales temporo-pariétales, entre 270 et 450 ms ; là aussi ces effets sont communs aux deux modalités. Pour Van Doren et al. (2012), ces résultats, associés à leurs données IRMf, sont cohérents avec l'hypothèse que les étapes initiales d'identification perceptuelle et d'accès à la mémoire visuelle à court terme sont supportées par des connexions longue-distance du gyrus frontal inférieur et du sillon occipito-temporal gauches ; ces connexions seraient le support anatomique permettant l'accès à la mémoire visuelle à court terme car elles favoriseraient la perception consciente et l'encodage amodal des stimuli (mots ou images).

D'autres composantes précoces (N200, P200, P300,…) pourraient également être impliquées dans la compréhension sémantique mais elles seraient plutôt associées à des processus attentionnels ou d'identification catégorielle et ne refléteraient pas nécessairement l'accès à la mémoire sémantique ou l'intégration du sens. Des composantes plus tardives, moins bien définies en terme de localisation de latence et de polarité, semblent également jouer un rôle dans l'intégration sémantique (Münte et al., 1998 ; Patel et al., 1998). Nous discuterons plus en détail de l'implication de certaines de ces composantes dans la partie Résultats de cette section.

- **<u>Traitement sémantique et information contextuelle</u>**

Nous avons vu que le traitement sémantique dépendait fortement du lien de congruence de présentation d'un stimulus dans un contexte donné et que loin de refléter uniquement des processus de violation sémantique, cette interaction était essentielle à la compréhension du sens. Elle est particulièrement intéressante dans le cadre du langage pour comprendre comment une représentation sémantique est comprise dans un contexte plus général (un mot au sein d'une

phrase, une phrase au sein d'un texte…), mais aussi dans des stimuli visuels plus complexes (succession d'images, vidéos…). Comment le cerveau fait-il pour mettre bout à bout les différents éléments disponibles (mots/phrases, objets et personnages/scène) et former une représentation d'ensemble cohérente ? Quels sont les marqueurs essentiels aux étapes de la compréhension : faut-il que l'intégration sémantique soit complète pour obtenir une représentation cohérente ou des processus précoces (notamment attentionnels et contextuels) peuvent-ils participer avant même qu'une représentation sémantique soit accessible ? Encore une fois, ces différentes questions sont essentielles pour comprendre comment fonctionne le système conceptuel et s'il est amodal ou modalité-spécifique.

- Information contextuelle et stimuli verbaux

Dans le cadre du langage, beaucoup de recherches se sont penchées sur cette question : comment les mots sont reconnus puis combinés en phrases, comment l'information au niveau du message général d'une phrase est intégrée dans un ensemble plus large de type texte ou discours (Schwanenflugel & LaCount, 1988; Van Berkum, 2008). Une approche qui a prévalu pendant longtemps, et qui reste encore répandue, suppose l'existence d'un flux de traitement hiérarchique : les mots sont reconnus individuellement et leur sens récupéré de façon stéréotypée et indépendamment du contexte dans lequel le mot apparaît. C'est seulement après l'achèvement de cette étape de reconnaissance individuelle des mots que les significations de chacun d'entre eux sont combinées pour créer des unités de sens d'ordre supérieur, telles que des phrases, et que l'information contextuelle devient accessible et capable d'influencer le traitement du langage (Forster, 1981 ; Swinney, 1979). A l'opposé, un certain nombre de travaux suggèrent que l'information contextuelle est disponible très tôt et qu'elle peut affecter le traitement des mots avant même que des effets sémantiques soient observés : l'information contextuelle disponible dans la phrase ou au niveau du discours pourrait avoir une influence sur le traitement sémantique au niveau des mots, ces effets étant caractérisables en ERP (Kamide et al., 2003; Van Berkum et al., 2003, 2008 ; Van Petten, 1993 ; 1995 ; Wlotko & Federmeier, 2012 ; Federmeier et al, 2003;. Kutas, 1993 ; Coulson et al., 2005 ; Federmeier, 2007). De même la congruence générale du message semble moduler les réponses ERP avant même que la reconnaissance des mots soit effectuée (Van Petten et al, 1999). Le traitement sémantique commencerait donc avant même que toutes les informations potentiellement pertinentes soient disponibles mais il serait modulé par toute information contextuelle déjà disponible ; la reconnaissance des mots ne serait pas l'étape finale essentielle à la compréhension mais la reconnaissance conceptuelle et l'accès aux informations sémantiques dépendraient d'une mise à jour continue de l'information provenant du stimulus mais aussi du contexte plus large dans lequel il est intégré (Wlotko & Federmeier, 2012). Les informations

temporelles obtenues grâce à des analyses ERPs sont essentielles pour comprendre si ce processus de traitement est effectivement réalisé « on line ».

Pour van Berkum et al. (2003), la N400 correspondrait au processus qui permet de relier l'information d'une phrase à un discours plus large. Dans leur expérience, ces auteurs ont montré des effets de la N400 liés à un mot anormal en fin de phrase par rapport au contexte fourni par un ensemble de phrases présentées précédemment, effets qui disparaissent quand ces phrases avec un mot « anormal » sont soumises de façon isolée. Pour Metusalem et al. (2012), les connaissances sur des événements du monde réel jouent un rôle important pour guider la compréhension on line du langage. Dans cette expérience, les ERPs ont été enregistrés pendant que les participants lisaient de courts scénarios décrivant des événements réels typiques dont la fin pouvait être cohérente ou non (scénario : jouer dans la neige, construire un bonhomme de neige/construire une serviette). Un mot anormal par rapport au contexte provoque une N400 qui disparaît lorsque le mot est contextuellement normal ou lorsque les phrases ne sont pas intégrées à un discours. Durant la compréhension, les personnes activeraient des connaissances générales sur les événements décrits qui participent à la création d'une représentation mentale de ces événements ; cette représentation accessible immédiatement aurait une influence sur le traitement linguistique. Federmeier & Laszlo (2009) suggèrent que la N400 refléterait l'existence d'une contrainte temporelle régissant la façon dont les informations issues du contexte peuvent être utilisées pour façonner les activations sémantiques initiales des mots. Comme nous l'avons déjà souligné précédemment, la N400 est générée par une variété de tâches impliquant des manipulations sémantiques qui vont moduler l'amplitude de la N400 mais pas sa latence (par exemple lorsque des mots sont associés à d'autres mots qui sont cohérents ou placés dans une phrase dont le contexte est congruent, l'amplitude de la N400 est réduite par rapport aux mêmes stimuli présentés hors-contexte ou dans un contexte incongru). La réduction d'amplitude pourrait refléter une réduction des exigences de traitement sémantique (Wlotko & Federmeier, 2012). Pour Federmeier & Laszlo, cette stabilité de la latence de la N400 (en opposition aux variations d'amplitude) suggère qu'il existerait un processus fonctionnel d'assemblage sous la dépendance d'un timing précis où les informations issues du contexte et des traitements bottom-up liés au stimulus sont regroupées, via cette synchronie temporelle, pour créer une représentation stable. Selon le point de vue de Federmeier et Laszlo, les différences dans l'amplitude de la N400 refléteraient l'influence conjointe des activations sémantiques (contraintes au niveau temporel) provenant de la partie bottom-up des traitements initiés par le stimulus et des états d'activation précédemment induits par l'information contextuelle de différents types (linguistique mais aussi basée sur des signaux non verbaux dans le cas d'une interaction réelle).

D'autres auteurs (Hagoort, 2005 ; Baggio & Hagoort, 2011) proposent une vision légèrement différente de ce processus fonctionnel d' « assemblage » (binding) dans le traitement du langage, basée sur des interactions oscillatoires entre les régions cérébrales frontales et temporales, interactions dont le rôle serait de permettre l'unification des informations bottom-up et top-down pour aboutir à la construction d'un message. Ce qui est important dans ces approches, c'est qu'elles impliquent que certains aspects de l'information contextuelle ne sont pas disponibles avant qu'une représentation sémantique initiale ne soit formée et qu'ils pourraient donc correspondre à des processus de réévaluation sémantique ultérieurs à la N400 (Wlotko & Federmeier, 2012).

Comme le mentionnent Wlotko & Federmeier (2012), de façon complémentaire à tous les travaux qui montrent que l'information contextuelle semble utilisée par le cerveau aussi tôt que possible pour permettre la compréhension (et ce, même avant toute information bottom-up lorsque les mots sont fortement prévisibles), de nombreuses études révèlent que dans le cas où des informations pertinentes sont disponibles plus tardivement au cours du traitement, la représentation initiale peut-être révisée, remodelée. C'est notamment ce que soulignent les travaux impliquant un langage figuratif, l'humour, l'ironie, des jeux de mots, des métaphores etc. où une réinterprétation du sens est souvent nécessaire à la compréhension car ce type de langage nécessite que le tri soit fait entre les multiples significations de l'information contextuelle pour que le sens littéral et humoristique émerge. C'est ainsi que plusieurs auteurs ont mis en évidence des processus plus tardifs dans le traitement du sens littéral (composante négative lente commençant vers 500 ms, avec une localisation fronto-antérieure) qu'ils associent à des mécanismes de réévaluation du sens et de « frame-shifting » (Coulson & Kutas, 2001 ; Coulson & Lovett, 2004; Coulson & Williams, 2005). Il semble qu'on retrouve cette composante fronto-négative gauche dans des protocoles basés sur des phrases normales plausibles mais qui sont modérément contraignantes, c'est-à-dire des phrases moins détaillées où le sens plus ambigu peut générer des représentations multiples mais dont le message est à peu près équivalent (Davenport & Coulson, 2011; Kutas, 1993). Ainsi, Wlotko et al. (2012) ont cherché à étudier les mécanismes multiples qui peuvent contribuer à la construction des représentations (au niveau du message général d'une phrase) car ils considèrent que certains de ces mécanismes pourraient être engagés après ou en même temps que la formation de la représentation initiale du sens et pourraient alors servir à réviser le sens initial. Dans cette étude, les auteurs ont enregistré les ERP pendant que les participants lisaient des phrases plausibles mais dont le sens général pouvait varier en fonction de la contrainte contextuelle imposée par le mot final, plus ou moins plausible dans le contexte de la phrase. Par exemple une phrase comme « The little girl refused to go to sleep until he told her a… » génère, dans 100% des cas, le mot « story » lors d'une tâche de complétion (contrainte forte) ; en revanche une phrase du type «Each night the campers

built a huge…", est completée seulement dans 79% des cas par le mot "fire" (contrainte modérée, d'autres alternatives étant possibles eg : « tent »). Leurs résultats confirment que l'amplitude de la N400 est fonction de la prédictibilité puisque plus la contrainte contextuelle est forte (mot final cohérent et attendu par rapport au reste de la phrase), plus l'amplitude de la N400 décroît. Mais leurs résultats ont aussi mis en évidence une composante négative latéralisée à gauche, large et lente, commençant vers 400-500 ms qui est plus importante pour les phrases dont la contrainte contextuelle est modérée ; au sein de cette catégorie de phrases, la négativité de la composante était plus grande pour les phrases qui ont peu de fins alternatives possibles que pour celles qui en ont beaucoup. Pour ces auteurs, cet effet serait lié à une réinterprétation de l'information contextuelle plus tardive, réinterprétation nécessaire lorsque les sujets ont déjà construit une représentation et d'autant plus difficile à effectuer lorsque les représentations sont proches. Ces résultats seraient cohérents avec ceux qui décrivent des effets de « frame shifting », processus psychologique dans lequel les éléments d'une représentation mentale sont ré-analysés à la lumière de nouvelles informations entrantes et ils suggèrent que malgré une analyse sémantico-contextuelle très précoce (entre 200 et 400 ms), la finalisation de la construction d'une représentation cohérente d'un événement peut nécessiter des processus plus tardifs de réévaluation sémantique.

- Information contextuelle et stimuli non-verbaux

Des questions similaires se sont posées concernant les stimuli visuels et les effets de violation contextuelle durant le traitement d'une scène. Dans la vie réelle, nous sommes habitués à voir certains objets dans des contextes spécifiques et récurrents, ce sont ces régularités contextuelles qui nous aident à effectuer des jugements rapides sur les objets qui composent une scène ; de nombreuses études comportementales ont confirmé que les objets, dont la probabilité d'apparition dans un contexte donné était faible, sont identifiés de façon plus lente et moins précise que les objets congruents avec le contexte, mais qu'ils ont aussi tendance à attirer plus l'attention (Mudrik et al., 2010). Ceci peut paraître paradoxal, comme le signalent Mudrik et al. (2010) : si les objets incongrus sont identifiés plus tardivement, comment peuvent-il attirer l'attention de façon précoce ? Selon le modèle de Loftus & Mackworth (1978) sur la perception des scènes, le point essentiel de la scène (le « gist ») pourrait être extrait rapidement, durant des étapes perceptuelles précoces où les objets qui composent la scène seraient partiellement reconnus. Cette reconnaissance partielle pourrait être suffisante pour déterminer ce qui choque dans la scène et qui nécessite une analyse ultérieure plus poussée, les informations extraites précocement entraînant la transmission rapide de l'information aux structures temporales médiales et au cortex frontal qui en retour guiderait l'analyse au sein du cortex temporal ventral en redirigeant l'attention sur les éléments perturbateurs

(Bar, 2004 ; Bar et al., 2006). L'analyse détaillée qui suit ces étapes précoces permettrait d'identifier précisément les incohérences entre les objets et le contexte.

De nombreux travaux comportementaux confirment que des observateurs humains sont capables de reconnaître une scène ou des objets isolés très rapidement, avec des durées d'exposition de l'ordre d'une vingtaine de ms, peut-être même avant que l'étape de traitement perceptuel soit achevée et avant tout traitement sémantique (Rieger, Braun, Bulthoff, & Gegenfurtner, 2005; Rousselet, Joubert, & Fabre-Thorpe, 2005 ; Hochstein & Ahissar, 2002 ; Oliva & Schyns, 2000; Oliva & Torralba, 2006 ; Delorme, Richard, & Fabre-Thorpe, 2000; Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001; Rousselet, Mace, & Fabre-Thorpe, 2003 ; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007).

Plusieurs modèles tentent d'expliquer ces phénomènes : le modèle du schéma perceptuel, le modèle de sélection d'objet ou modèle d'appariement, le modèle d'isolation fonctionnelle (voir Mudrik et al. 2010 pour plus de références). Les deux premiers modèles prédisent que les violations contextuelles ont des effets qui sont antérieurs à l'identification complète des objets. Selon le premier modèle, l'information contextuelle conduit à une activation rapide d'un schéma de scène qui, à son tour, active des processus d'attention sélective facilitant la détection ultérieure d'éléments perceptifs associés à des objets appartenant au schéma de scène. Les violations contextuelles entraveraient l'analyse des objets incongrus à des stades précoces car ils ne seraient pas cohérents par rapport au schéma de la scène. Selon le modèle de sélection d'objet, l'information contextuelle générerait également un schéma de la scène qui amorce l'identification de la cohérence des schémas d'objets et qui module l'attention de l'observateur en fonction de la quantité d'information perceptive nécessaire pour faire le lien avec une représentation spécifique d'un objet. Quand l'input visuel a été suffisamment analysé pour être comparé à des représentations stockées, l'activation contextuelle faciliterait la sélection de l'interprétation la plus probable. Le traitement de l'information contextuelle correspondrait plutôt à un processus top-down qui influencerait la décision à propos de l'identification des objets mais pas l'analyse perceptuelle de la scène. En revanche, le troisième modèle (le modèle d'isolation fonctionnelle) rejette toute influence contextuelle précoce et attentionnelle, l'information contextuelle n'interviendrait qu'après l'identification des objets pour influencer l'activation des connaissances sémantiques et des processus de décision. Ce modèle est en désaccord avec les résultats comportementaux qui montrent que des violations contextuelles perturbent les performances de sujets même lorsque le temps d'exposition est très court. Mais comme le précisent Mudrik et al. (2010), les résultats comportementaux ne nous renseignent pas sur le décours temporel du traitement des violations contextuelles et ne peuvent permettre d'affirmer que celui-ci a bien eu lieu, à la différence des analyses ERP.

Ganis & Kutas (2003) ont utilisé cette approche pour explorer le traitement des violations contextuelles au sein d'une scène visuelle : après avoir indiqué par une croix de fixation où l'objet d'intérêt devait apparaître, les auteurs présentaient une scène de la vie quotidienne (des joueurs de foot sur un terrain), au bout de 300 ms un objet apparaissait à l'endroit préalablement indiqué par la croix, qui pouvait être soit congruent avec la scène (un ballon de foot), soit incongruent (un rouleau de papier toilette). Les auteurs évoquent une composante N390 liée à l'effet de congruence (300-500 ms) qui ressemble fortement par sa morphologie et son décours temporel à la N400 que l'on rencontre lorsque des stimuli langagiers ou visuels ne sont pas cohérents avec un contexte sémantique, même si la N390 n'est pas aussi frontale que la N400 liée aux stimuli visuels (Federmeier & Kutas, 2001; Ganis et al., 1996). Pour Ganis & Kutas (2003), le sens d'une scène visuelle influencerait l'identification des objets qui en font partie : la ressemblance entre la N390 et la composante sémantique N400, mais aussi l'absence de différences plus précoces de l'activité cérébrale associée aux effets de congruence, font dire à ces auteurs que l'influence contextuelle serait postérieure aux étapes d'analyses sémantiques. Les effets de congruence illustrés par la N390 refléteraient l'activation d'un schéma visuel de la scène qui serait stocké dans le lobe temporal antérieur. Demiral et al. (2012) font référence à des résultats similaires.

Mais comme le soulignent Mudrik et al. (2010), l'étude de Ganis & Kutas (2003) présente des aspects méthodologiques qui peuvent nuancer l'interprétation de cette composante N400. Tout d'abord, le contexte et la localisation de l'objet sont connus avant que celui-ci n'apparaisse et les sujets ont donc des attentes très fortes concernant l'objet et lorsqu'un objet incongru apparaît, il y a non seulement une violation du contexte lui-même mais aussi des attentes préalablement formées des sujets. Par ailleurs, la tâche des sujets étant d'identifier l'objet d'intérêt, ceux-ci ont déjà aussi préparé une réponse associée, en plus d'attentes fortes, et lorsqu'un objet différent de celui-attendu apparaît, ils doivent changer leur réponse en fonction. Pour Mudrik et al. (2010), la composante N390 identifiée par Ganis et Kutas pourrait ne pas refléter des effets de violation contextuelle mais plutôt la discordance entre les attentes ou la réponse préalablement formées par les sujets et l'objet qui apparaît ou la réponse différente qui lui est associée. Ainsi, la N390 correspondrait bien à la N400 généralement déclenchée par une discordance entre les attentes du sujet et réponses et un stimulus dans un contexte donné ; ce protocole pourrait masquer les effets liés spécifiquement au traitement contextuel. C'est pourquoi, dans leur protocole, Mudrik et al. (2010) ont cherché à évaluer spécifiquement ces effets et à éviter toute influence des attentes préalables des sujets, en présentant une scène où un objet n'est pas cohérent sans que le contexte ou la localisation de l'objet aient été indiqués auparavant et sans tâche spécifique concernant l'identification de l'objet cible. Leurs résultats confirment que les effets de congruence contextuelle affectent le traitement d'une

scène relativement tôt puisqu'ils confirment l'existence de composantes négatives antérieures qui ressemblent aux composantes N300/N400 précédemment associées à l'intégration sémantique (similaire à la N390 décrite par Kutas et Ganis, 2003). Mais ces travaux mettent aussi en évidence une négativité plus distribuée et plus tardive (entre 650 et 850 ms) qui pourrait être reliée à des processus ultérieurs d'évaluation sémantique et de préparation de la réponse, suggérant que les informations contextuelles pourraient également concerner les processus de sélection d'un modèle de scène approprié et influencer les étapes ultérieures liées à l'activation des connaissances sémantiques et / ou à la prise de décision.

Les différents travaux sur des stimuli verbaux et non verbaux mettent en évidence des processus assez similaires dans l'analyse d'une scène qu'elle soit décrite dans une phrase ou représentée visuellement. Cependant, dans le cadre de la vision, la plupart de ces travaux se sont intéressés à des cas de violations sémantiques fortes (objet qui n'a rien à faire dans une scène comme un camion dans le ciel, Demiral et al. 2012) et pas à des cas où la situation finale est moins attendue mais cependant plausible, ce qui ne rend pas tout à fait comparables les travaux cités concernant des stimuli verbaux et non verbaux. Cependant, plusieurs travaux de Sitnikova et ses collègues (Sitnikova et al. 2008 ; Sitnikova et al. 2003) concernaient la construction des représentations d'événements visuels, avec l'enregistrement des ERPs de participants en train de visualiser de courtes vidéos représentant des actions de la vie de tous les jours (e.g. couper du pain). La fin de ces films pouvait être contextuellement appropriée ou inappropriée, soit par une situation violant totalement le but attendu d'une action (par exemple : « couper » du pain avec un fer à repasser) soit en présentant une action différente, qui n'était pas attendue, mais qui est possible (repasser le linge alors qu'on s'attendait à voir la fin de l'action « couper du pain »). La présentation d'une information contextuellement inappropriée à la fin du film évoque une activité négative antérieure similaire aux composantes N300/N400 (indépendamment du fait que la fin du film soit inattendue mais plausible ou totalement impossible). Mais les résultats montrent aussi une seconde composante plus tardive, postérieure et positive (1000-1500ms), intervenant spécifiquement quand les objets présentés à la fin des films violent le but attendu de l'action dans un contexte donné (fer à repasser dans le scénario couper le pain). Pour Sitnikova et al. (2008), ces résultats suggèrent que la compréhension visuelle du monde réel pourrait être réalisée par l'intermédiaire de deux mécanismes neurophysiologiques d'intégration sémantique différents, les différences morphologiques, temporelle et spatiale entre le complexe N300/N400 et l'effet de positivité postérieure plus tardif suggérant que les mécanismes cognitifs qui sous-tendent ces effets sont réalisés par des processus neuronaux anatomiquement distincts. Le premier mécanisme, qui se traduit par une négativité antérieure similaire à la N400, ferait le lien entre les informations entrantes et les concepts en

mémoire sémantique. Le second mécanisme, qui se traduit par une positivité postérieure plus tardive, évaluerait les informations entrantes au regard des exigences requises par les connaissances sur le monde réel, sur les objets et sur les actions. Ces deux mécanismes ne sont pas contradictoires et ils interviendraient dans la compréhension des événements de différentes manières : le premier mécanisme serait plus adapté à l'évaluation des situations familières alors que le second serait plus impliqué dans la compréhension des situations nouvelles.

II.3.*2*) Buts de l'expérience et hypothèses

L'objectif principal de l'expérience en EEG était d'utiliser une version assez proche du protocole IRMf (voir Chapitre II.1) afin d'étudier cette fois, non pas les corrélats neuroanatomiques d'un réseau sémantique commun mais les processus temporels qui peuvent intervenir dans l'intégration du sens, à la fois pour des images et pour des phrases. Nous avons choisi d'utiliser un protocole EEG pour répondre à cette question afin d'essayer de mieux comprendre quand et comment le système sémantique commun, mis en évidence par les expériences en neuroimagerie, pourrait intervenir dans la création des représentations sémantiques. Si on ne peut nier qu'images et phrases sont des stimuli différents, n'impliquant pas nécessairement les mêmes voies d'entrée et de traitements primaires, on peut supposer qu'à un certain moment des processus cognitifs doivent converger pour permettre d'aboutir à une compréhension assez similaire entre les deux types de stimuli.

En lien avec notre expérience comportementale, nous souhaitions également étudier les décours temporels associés à la construction et au maintien d'une représentation, lorsque celle-ci doit s'intégrer à un contexte plus complexe. Dans notre expérience comportementale (voir chapitre II.2), nous avons mis en évidence qu'une première représentation, associée à une phrase, facilite ou perturbe la représentation associée à un second stimulus selon que celui-ci est cohérent ou non avec le premier (ce qui se reflète sur des indices comportementaux tels que le COR ou le TR). Que se passe-t-il, au niveau cérébral, lorsque l'on perturbe une « représentation en cours » ? Peut-on associer ces différences comportementales à des marqueurs temporels qui seraient distincts/communs pour des stimuli verbaux et non verbaux ? C'est pourquoi, dans ce protocole EEG, nous avons choisi d'ajouter un deuxième stimulus (une deuxième phrase ou une deuxième image), stimulus qui n'était pas présent dans notre protocole IRMf ou seulement présent pour des phrases dans notre protocole comportemental puisque celui-ci ne comporte que du matériel verbal (voir la partie Matériel et Méthodes pour plus de détails).

Comme nous l'avons déjà mentionné, l'intégration du sens dans un contexte donné est un élément essentiel à la construction d'une représentation et de sa signification, c'est un facteur connu pour générer des effets de N400 pour des mots ou des images. Cependant, dans ce protocole, nous avons fait le choix de présenter les phrases dans leur intégralité et non mot à mot, même si nous savions que cela masquerait les composantes associées au traitement de chaque mot (voir partie résultats pour plus de détails). Les études qui se sont intéressées au traitement spécifique d'un « mot critique » dans une phrase, se sont souvent arrangées pour placer ce mot en fin de phrase et caler les ERP sur la fin de présentation du stimulus (Friederici et al., 1999; Gunter et al., 1997). Mais les phrases ainsi constituées sont peu naturelles et plusieurs études ont montré qu'il y avait des effets spécifiques reliés à cette présentation où le mot critique est placé en fin de phrases (Hagoort, Brown et Groothusen , 1993; Osterhout & Holcomb, 1992, 1993; Osterhout et Nicol, 1999 ; Hagoort, 2003) ; rendant difficile la possibilité de dissocier si les effets sont liés au mot lui-même, à la fin de présentation de la phrase, à l'intégration sémantique ou à des processus de prise de décision et de préparation de réponse. La présentation en un seul bloc permet d'étudier la façon dont les phrases sont traitées de façon plus naturelle et même si elle ne permet pas d'accéder à un décours temporel précis, elle peut être utilisée pour visualiser des différences entre deux conditions (Hahne et Friederici, 1999 ; Pannekamp et al., 2005 ; Metusalem et al., 2012 ; Federmeier et al., 2003). C'est dans ce but que nous avons volontairement choisi de présenter les phrases en un seul bloc : avoir une présentation plus naturelle, plus proche de la façon dont nous lisons des phrases dans la réalité, mais aussi savoir si l'on peut détecter des modulations entre deux conditions à travers ce mode de présentation (nous avions également des a priori sur les effets liés au verbe de la seconde phrase, mais nous en parlerons plus en détails dans la partie Résultats).

Pour répondre à ces différentes interrogations, nous avons choisi de réaliser à la fois des analyses ERPs et Temps-fréquence. Cependant les acquisitions ayant été achevées récemment, les analyses sont encore en cours, nous ne présenterons ici que les résultats préliminaires de nos analyses ERP. Dans la conclusion de ce sous-chapitre, nous ferons le point sur les analyses et améliorations qui pourraient venir compléter ce protocole.

Hypothèses

Avec ce protocole, nous voulions donc tester l'hypothèse qu'il y aurait, pour les images et les phrases, des processus précoces modalités-spécifiques et des processus communs, plus tardifs, qui joueraient un rôle dans l'intégration du sens. Notre hypothèse porte plus précisément sur la perturbation de l'intégration d'une « représentation en cours » à un contexte de compréhension plus global (un ensemble de deux phrases ou de deux images). Nous cherchons à déterminer si le fait de

présenter un second stimulus qui n'est pas cohérent avec le premier provoque des changements dans le décours temporel associé au traitement de ce stimulus. Ces processus devraient être révélés en comparant les ERPs liés à la présentation de stimuli successifs, images ou phrases, qui correspondent ou non à une suite sémantiquement logique du premier stimulus présenté. Pour répondre à ces questions, nous avons comparé l'activité dans les différentes conditions pour le stimulus 1 avec celle mise en jeu pour le stimulus 2.

Pour les images, la comparaison des ERPs liés à la présentation du second stimulus devrait faire apparaître des différences selon que le stimulus est une suite logique du premier ou non, différences qui ne devraient pas être observées pour le stimulus 1. On s'attend à trouver des effets au niveau du complexe N300/N400 ainsi que des différences plus précoces des composantes liées à des processus attentionnels et d'identification de la scène, mais aussi des effets plus tardifs différentiels lorsque l'intégration sémantique contextuelle est cohérente.

Pour les phrases, comme précisé précédemment, nous savons que les effets de marqueurs sémantiques tels que la N400 risquent d'être masqués ; nous nous attendons cependant à observer des décours temporels différents pour le stimulus 2 (Suite logique vs Non Suite) qui ne devraient pas être présents pour le stimulus 1.

Si les effets précoces sont différents pour le stimulus 2 selon que celui-ci est une phrase ou une image, des différences plus tardives communes aux deux modalités pourraient cependant indiquer l'existence de processus communs concernant l'intégration sémantique et/ou la réévaluation en fonction du contexte.

II.3.*3*) Matériels et Méthodes

- **Sujets**

Dix-neuf sujets sains, droitiers, de langue maternelle française et ayant une vision normale ou corrigée ont participé à cette étude (10 hommes, âge moyen : 25,5 ans ; SD : 4,11). Aucun participant n'avait d'antécédent de trouble médical psychiatrique ou neurologique majeur. Tous les participants ont donné leur consentement éclairé avant la session. Les acquisitions EEG ayant été effectuées à Lyon alors que celles en IRMf avaient été acquises à Paris, aucun sujet n'a participé aux deux protocoles. Parmi ces 19 acquisitions, les données d'un sujet ont été rejetées en raison d'un trop grand nombre d'artefacts (probablement liés à un problème de déplacement des électrodes durant l'acquisition).

- **Stimuli**

a) Sélection des stimuli

Comme l'intérêt premier était d'avoir un jeu de données assez similaires à celles utilisées en IRMf pour mettre en évidence un réseau sémantique, nous avons choisi des stimuli (scènes de la vie quotidienne) assez semblables à ceux de la première expérimentation (cf Chapitre II.1). Cependant, comme mentionné précédemment, nous souhaitions rendre ce protocole plus complexe et tester l'intégration sémantique contextuelle par l'ajout d'un second stimulus (action 1, puis action 2 qui peut être dans la continuité ou non de la 1ère action). Ce second stimulus implique toujours le ou les mêmes personnages que dans le premier stimulus, mais il peut être dans la continuité ou non du premier (ce qui correspond à deux conditions que nous avons appelées Suite vs Non Suite). Dans l'expérience en IRMf, les scènes représentaient une action qui se suffisait à elle-même (ex : un petit garçon jouant au football), dans cette seconde expérience nous avons essayé de privilégier des actions qui peuvent impliquer une suite pour créer une contrainte contextuelle plus forte (e.g. image 1 : un petit garçon s'apprêtant à tirer dans les buts, image 2 : le petit garçon a marqué un but). Nous avons choisi des stimuli qui nécessitent une compréhension plus approfondie de la scène, notamment par la compréhension des intentions des personnages (e.g. image 1 : une petite fille regarde un gâteau → elle a envie de le manger ? image 2 : effectivement, elle mange le gâteau), mais aussi avec la construction de scénarii logiques plus complets (e.g. image 1 : une femme regarde une robe en vitrine ; image 2 : elle rentre dans la boutique pour l'essayer). De ce fait, les stimuli utilisés sont plus complexes que dans l'expérience IRMf, ils intègrent notamment des actions plus compliquées, un nombre plus important de personnages interagissant entre eux et avec des objets (si un personnage fait l'action de « donner quelque chose », il faut quelqu'un pour recevoir). Bien sûr, lorsque les deux stimuli successifs représentent une Suite, ce n'est pas forcément la suite la plus logique mais c'est une suite possible se déroulant dans un contexte proche et cohérent (par exemple : se laver puis s'habiller). Lorsque les deux stimuli sont Non Suite, ce sont des événements qui paraissent difficilement liés parce qu'ils impliquent un contexte ou une action radicalement différents (par exemple se lever de son lit, puis boire une bière au bar) sans pour autant être des événements qui n'ont pas de sens (pas d'événements impossibles). Ainsi, entre stimulus 1 et stimulus 2, il y a un certain nombre d'éléments qui n'ont pas besoin d'être réinterprétés si les deux stimuli forment une suite d'actions (comprendre les rôles thématiques des différents personnages, comprendre l'action en cours….) Même si les deux stimuli ne représentent une suite logique d'actions, nous avons fait très attention en choisissant nos stimuli (et en les pre-testant) à ce que certains aspects ne deviennent pas ambigus entre le stimulus 1 et le stimulus 2, par exemple que des

personnages qui aient une relation apparente dans une image (ex : couple, mère/fille…) gardent la même relation dans une autre image.

Stimuli Images

Nous avons sélectionné 120 images sur la banque d'images (Getty images ©) ; cette base de données contient des photos des mêmes acteurs, c'est ce qui a rendu possible la sélection d'images représentant les mêmes personnages mais impliqués dans des situations différentes. Nous avons donc choisi 60 paires d'images : 30 paires « Suite » où la seconde image peut être considérée comme une suite logique de la première (mêmes personnages, même contexte, continuité de l'action) et 30 paires « Non Suite » où la seconde image implique les mêmes personnages mais dans une situation différente, ce qui laisse penser que ce n'est pas la suite de l'action précédente. (cf Figure 67).

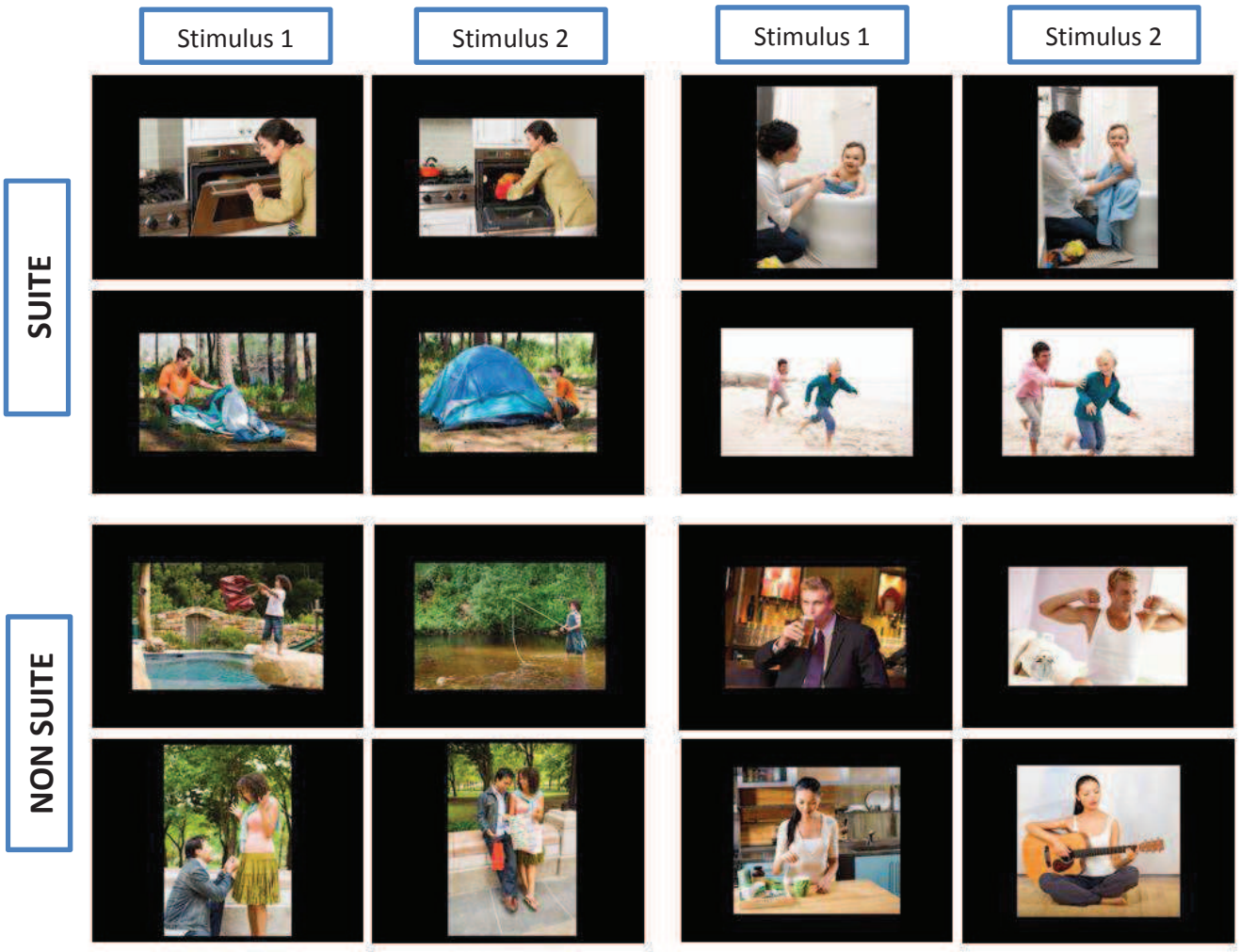


Fig 67) Exemples de couples de stimuli Images utilisés pour le protocole EEG
Les 4 couples d'images du haut représentent des stimuli de la condition Suite ; les 4 couples d'images du bas représentent des stimuli de la condition Non Suite.

Stimuli Phrases

Comme pour le protocole IRMf, nous avons créé nos stimuli « phrases » en nous inspirant de nos stimuli « images ». Cependant, nous n'avons pas pu appliquer tout à fait la même méthodologie que pour l'expérience IRMf où nous avions créé deux jeux de données (Jeu A et Jeu B) pour éviter qu'un même sujet voit une image et une phrase décrivant une même action. En raison de l'utilisation de deux stimuli consécutifs impliquant les mêmes personnages, il était difficile de trouver assez de situations proches mais avec des acteurs différents pour pouvoir créer 2 jeux de données. Dans une première version du protocole, les phrases étaient directement inspirées des images : nous avions décrit les mêmes actions dans les phrases que dans les images, mais en changeant les personnages (sexe, âge) et/ou le contexte… Cependant, une série de pilotes comportementaux nous a convaincus de changer de stratégie et d'utiliser des phrases plus distinctes des images choisies car plusieurs sujets pilotes nous ont rapporté avoir fait le lien entre les deux conditions malgré les changements effectués. Pour éviter d'avoir cet effet de « déjà vu », nous avons préféré créer un nouvel ensemble de phrases n'étant plus directement inspirées des images, mais restant proches toutefois, dans la mesure où phrases et images sont toujours des événements de la vie quotidienne (cf Figure 68 qui montre un exemple des stimuli utilisés).

De la même façon que les images peuvent impliquer les mêmes personnages dans des situations différentes à travers la présentation de deux stimuli consécutifs, les phrases décrivent un ou plusieurs personnages en train de faire une action qui peut se prolonger ou non dans la seconde phrase. Nous obtenons donc ainsi 30 couples de phrases « Suite » et 30 couples de phrases « Non Suite », soit 120 phrases au total. Pour reproduire un effet similaire à celui des images, à savoir qu'un certain nombre d'informations ne sont pas à ré-analyser complètement (identifier les personnages, définir leurs rôles ou la relation qui les unit…), nous avons choisi d'utiliser une phrase complète pour le stimulus 1 *(« La femme et l'homme regardent les affiches du cinéma »*) et seulement des pronoms pour se référer aux mêmes protagonistes dans le stimulus 2 (*« Ils achètent des places au guichet »*).

Fig 68) Exemples de couples de stimuli Phrases utilisés pour le protocole EEG
Les 4 couples de phrases du haut représentent des stimuli de la condition Suite ; les 4 couples de phrases du bas représentent des stimuli de la condition Non Suite.

b) Validation et répartition des stimuli

Comme pour nos précédents protocoles, tous les stimuli ont fait l'objet de pré-tests comportementaux (chez 10 sujets) pour rendre les stimuli les plus similaires possible entre matériel verbal et non verbal, mais aussi entre sessions expérimentales. Durant les pré-tests, plusieurs aspects ont été évalués afin de s'assurer de la validité des stimuli.

Validation des Stimuli Images

Concernant les images, il a été demandé aux sujets pilotes de : (1) Décrire la situation représentée sur la première image qui compose un essai (pour tester la compréhension/ambigüité des scènes) ; (2) Décrire la situation représentée sur la deuxième image qui compose l'essai (tester la compréhension/ambigüité des scènes) ; (3) Une fois vues les deux images, dire si ces deux images forment une suite cohérente ou non (afin de vérifier la répartition de nos stimuli en catégories Suite

ou Non Suite) ; (4) Signaler tout problème de compréhension (action des protagonistes peu claire, non reconnaissance des personnages entre la première et la deuxième image, image floue, gêne quelconque par rapport à l'une des images…). Les résultats des pré-tests nous ont permis de confirmer que les images avaient bien été sélectionnées et réparties entre les deux conditions puisque 100% des images ont été comprises sans aucune difficulté et 93% des stimuli ont bien été classifiés en Suite/NonSuite par plus de 90% des sujets. Tout stimulus posant problème (interprétation, reconnaissance des personnages…) a été supprimé pour être remplacé par un stimulus moins ambigu.

Validation des Stimuli Phrases

Concernant les phrases, la tâche était un peu différente ; il a été demandé aux sujets pilotes de : (1) Dire s'ils comprenaient bien l'action décrite dans la première phrase ; (2) Donner un score de représentabilité à la première phrase : arrivez-vous facilement à vous représenter cette phrase ? Sur une échelle de 1 à 5, de 1= très difficile, 2=difficile, 3=intermédiaire, 4=facile, 5= très facile ; (3) Dire s'ils comprenaient bien l'action décrite dans la deuxième phrase ; (4) Donner un score de représentabilité à la deuxième phrase : arrivez-vous facilement à vous représenter cette phrase ? Même échelle. (5) Une fois lues les deux phrases, dire si ces deux phrases forment une suite cohérente ou non (pour vérifier la répartition des stimuli en Suite/NonSuite) ; (6) Signaler tout autre problème qui gênerait la compréhension de la phrase (action décrite ambiguë, mots employés rares, mal connus ou mal utilisés…). Là encore les résultats montrent que la compréhension des phrases était très bonne (même si quelques stimuli ambigus ont été retirés ou modifiés par précaution) et 92,5% des stimuli ont bien été classifiés en Suite/NonSuite par plus de 90% des sujets. Globalement, les sujets n'avaient pas non plus de difficulté à se représenter nos phrases, puisqu'on obtient une moyenne de 4,53 pour le score de représentabilité sur l'ensemble des stimuli.

Sur la base de ces différents résultats, les stimuli ont été répartis afin que les différentes sessions et conditions soient équivalentes en terme : d'orientation portrait ou paysage (pour les images), genre, âge et nombre de personnages (pour images et phrases), nombre de lettres dans une phrase (nombre moyen de lettres pour la phrase1 et la phrase 2), constructions grammaticales (ex : verbes réflexifs ou proposition relative), fréquences films et fréquences livres de la base Lexique pour les phrases (voir Préambule et Annexe 6). Comme nous l'avons précisé dans le préambule, trouver une base de données pour étalonner les fréquences des mots pose toujours problème. Comme pour l'expérience IRMf, nous avons utilisé la base online Lexique mais cette fois, nous avons fait le choix de nous baser préférentiellement sur la fréquence Films pour répartir les stimuli entre les 2 sessions, tout en vérifiant cependant que les fréquences Livre étaient également équivalentes. Les fréquences

obtenues représentent la moyenne des fréquences des Open-Class Words utilisés dans la phrase (nous ne souhaitions pas intégrer la fréquence des articles et des prépositions, la fréquence de ces mots est très importante, elle aurait noyé les fréquences des autres mots de la phrase durant le moyennage).

- **Protocole expérimental**

a) Session expérimentale

Comme l'illustre la figure 69, les stimuli (120 phrases et 120 images) étaient répartis en 4 sessions (phrases et images n'étant jamais mélangées) :

- 2 sessions contenant des images
- 2 sessions contenant des phrases

Une session Images comporte : 30 couples d'images de la condition Suite (Image 1 Suite + Image 2 Suite) et 30 couples d'images de la condition Non Suite (Image 1 Non Suite + Image 2 Non Suite).

De même, une session Phrases comporte : 30 couples de phrases de la condition Suite (Phrase 1 Suite + Phrase 2 Suite) et 30 couples de phrases de la condition Non Suite (Phrase 1 Non Suite + Phrase 2 Non Suite).

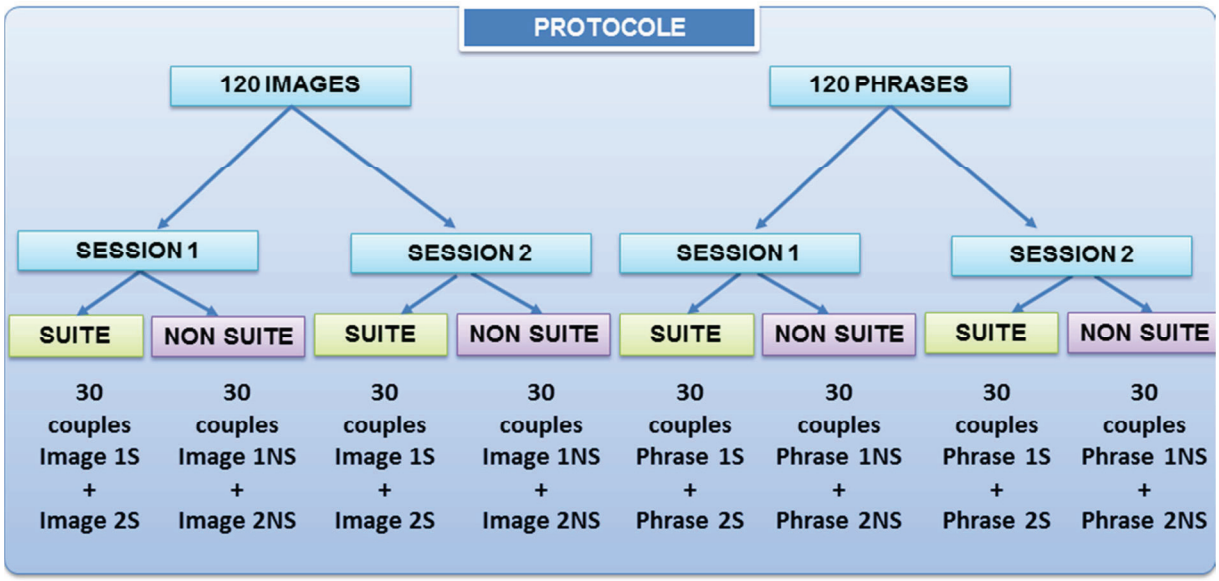


Fig 69) Protocole expérimental EEG
Les images sont réparties en deux sessions (session 1 et session 2), chaque session comporte 30 couples de stimuli de la condition Suite (Image 1S + Image 2S) et 30 couples de la condition Non Suite (Image 1NS + Image 2NS). Les phrases sont elles aussi réparties en deux sessions (session 1 et session 2), chaque session comporte 30 couples de stimuli de la condition Suite (Phrase 1S + Phrase 2S) et 30 couples de la condition Non Suite (Phrase 1NS + Phrase 2NS).

b) Procédure expérimentale

Pendant l'expérience, chaque session dure environ 12 minutes et les 4 sessions expérimentales sont présentées dans un ordre aléatoire ; au sein de chaque session, l'ordre des essais est lui aussi randomisé (la gestion des stimuli est assurée par le logiciel Presentation).

Le sujet est placé dans le noir, face à l'écran d'ordinateur et il utilise deux touches du clavier pour pouvoir effectuer la tâche comportementale (réponse oui ou non). Avant le début du protocole, une période d'entraînement supervisée par l'expérimentateur permet de s'assurer que le sujet a bien compris le déroulement du protocole et la tâche à effectuer.

Durant un essai, un cercle vert, qui indique le début d'un essai, apparaît au centre d'un écran noir pendant une durée aléatoire de 400 à 1000 ms (pour éviter des effets d'anticipation des stimuli). Il est suivi d'un premier stimulus phrase ou image présenté sur fond noir pendant 2 secondes, le premier stimulus est suivi d'un délai (point de fixation sur écran noir, durée aléatoire de 1000 à 1500 ms) ; un second stimulus phrase ou image apparaît pendant 2 secondes, également suivi d'un délai (1000 à 1500 ms). Les sujets ont pour consignes de rester attentifs à tous les stimuli car, dans 2/3 des essais, une question survient après la présentation du second stimulus.

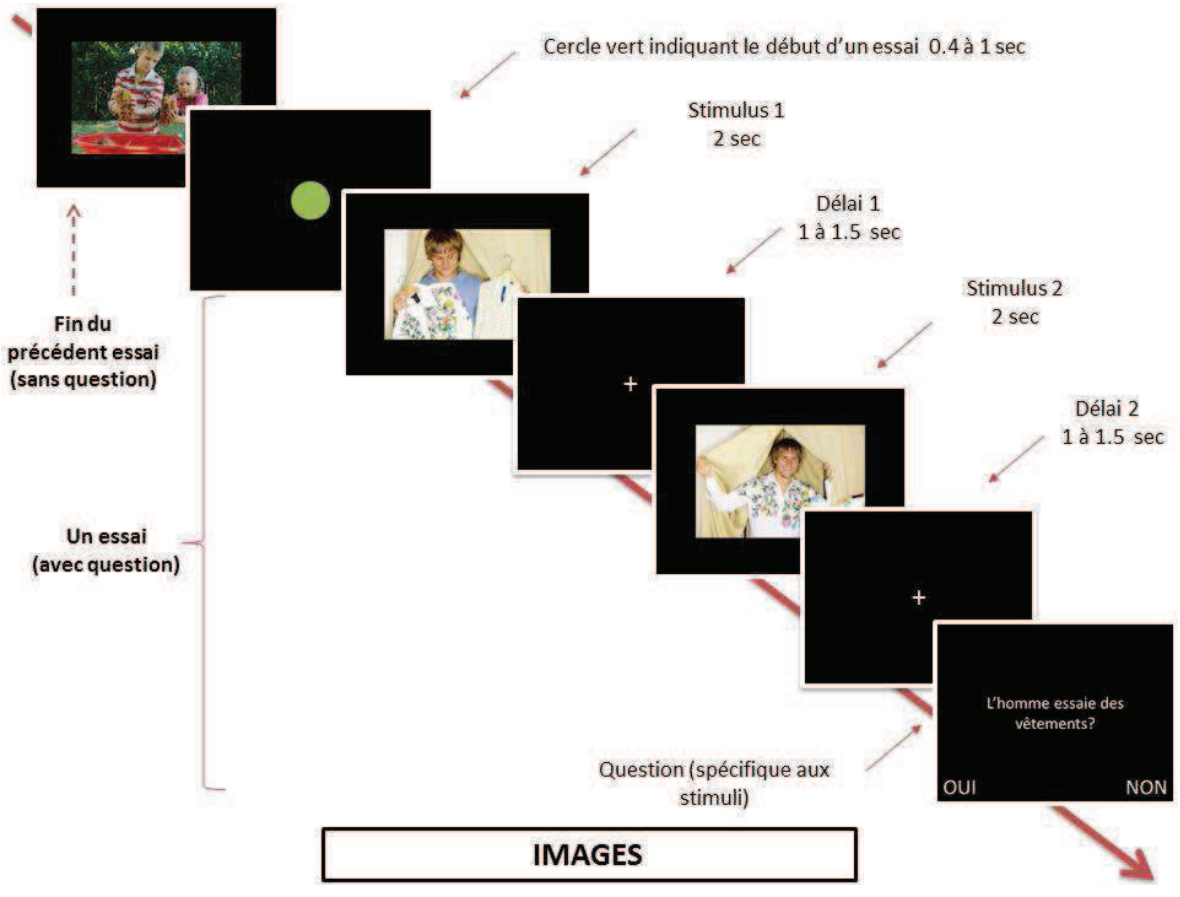


Fig 70) Déroulement d'un essai
Le déroulement d'un essai avec question est illustré ici avec des stimuli de la session Images.

Les questions sont relativement simples (voir figure 71) mais elles sont différentes à chaque fois et peuvent porter sur le premier et/ou le second stimulus (le sujet ne sait jamais à l'avance sur quel stimulus va porter la question et la réponse peut-être oui ou non). Les questions sont principalement là pour maintenir l'attention des sujets, mais pour pouvoir y répondre, cela suppose que les sujets aient une compréhension et une représentation assez précises de la scène. Un protocole EEG étant assez fatigant pour les sujets car ils doivent éviter de cligner des yeux pendant la présentation des stimuli, il leur a été indiqué avant l'expérience qu'ils pouvaient prendre tout leur temps pour répondre et pour se reposer au moment où la question apparaît. Nous avons donc choisi de ne pas effectuer d'analyses sur des marqueurs comportementaux tels que le temps de réaction pour ce protocole. Le nombre de bonnes réponses comportementales, qui dépasse les 95%, indique que les sujets étaient attentifs et que la tâche a été effectuée correctement.

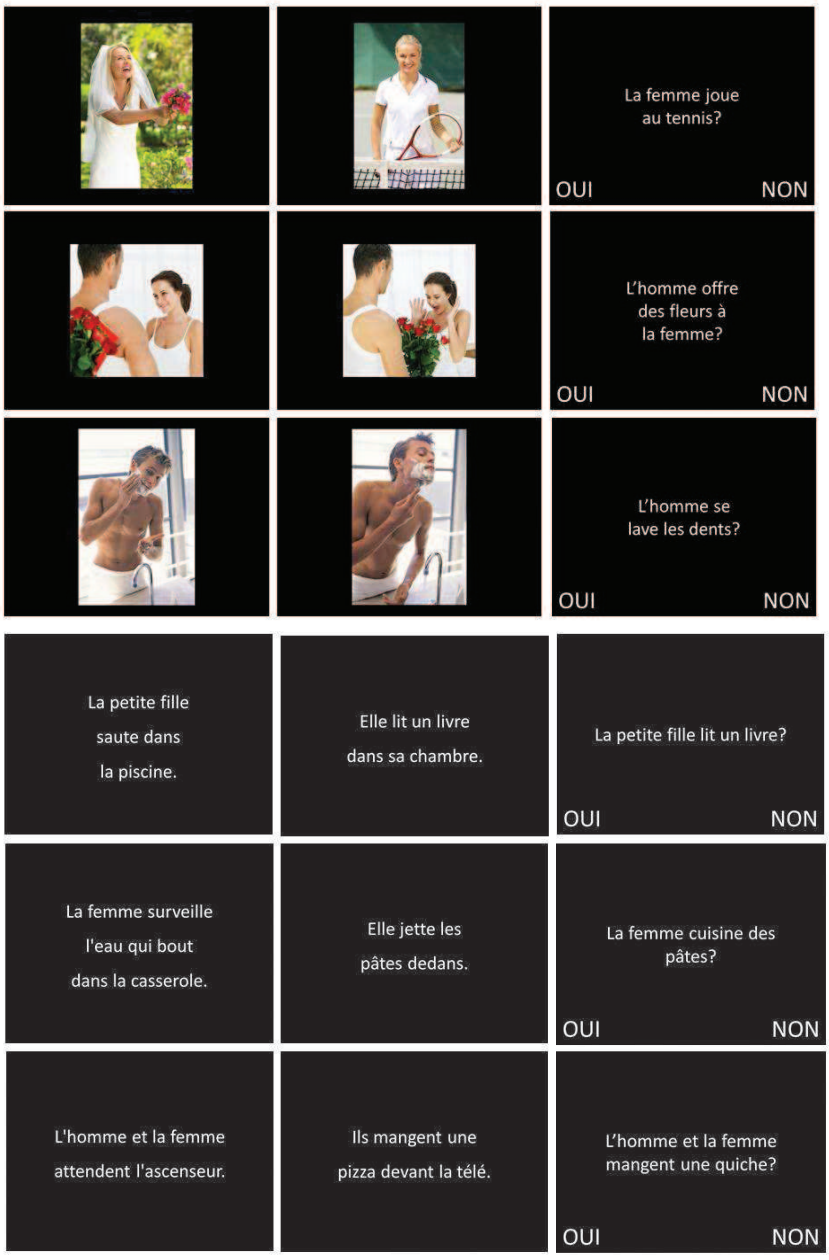


Fig 71) Quelques exemples de stimuli issus des essais avec question
(pour les stimuli Images et pour les stimuli Phrases)

c) Acquisition des données EEG

Les enregistrements EEG ont été effectués de façon continue, avec un taux d'échantillonnage à 2kHz (filtre passe-bas analogue, 400 Hz) à partir de 64 électrodes Ag-AgCl (Electro-Cap International, INC, Ohio, USA; Biosemi, ActiveTwo, version 5.36), référencées par rapport à une référence commune (CMS) et par rapport à une référence terre (DRL) directement intégrées au bonnet et positionnées selon le système international 10-20. Huit électrodes externes supplémentaires font également partie du montage de routine: 2 électrodes sur le nez et sur la joue, 4 autour des yeux (au dessus, à côté et en dessous) pour enregistrer les mouvements oculaires (électro-oculograme EOG) et 2 au niveau des mastoïdes bilatérales en vue de renforcer la détection des artefacts liés aux mouvements oculaires et permettre un re-référençage ultérieur potentiel.

II.3.*4*) Analyses des données EEG

- **Prétraitements**

Les prétraitements et les analyses des enregistrements EEG ont été réalisés en utilisant EEGLAB qui fonctionne sous environnement MATLAB (R2010a, The MathWorks, Inc., Natick, Massachusetts, USA). EEGLAB est un outil développé spécifiquement pour étudier les données éléctrophysiologiques (Delorme & Makeig, 2004). Les premières étapes de prétraitements (filtrage, définition des électrodes et epoching) ont été réalisées grâce à la boîte à outil FASTER qui simplifie les analyses en utilisant les fonctionnalités d'EEGLAB de façon automatique via un algorithme de rejets des artefacts dont les paramètres ont été optimisés (Nolan, Whelan & Reilly, 2010). Ces premières étapes ont été réalisées sur les données de chaque sujet indépendamment. Durant la phase de filtrage, un filtre passe-haut de 1Hz et un filtre passe-bas de 95 Hz ont été appliqués aux données. Un filtre « Notch on » a également été appliqué pour nettoyer le bruit autour de 50 Hz. Toutes les électrodes ont ensuite été re-référencées par rapport à une électrode de référence unique (Fz). Les époques correspondant aux différents événements d'intérêt (stimulus 1 Suite, stimulus 1 NonSuite, stimulus2 Suite, stimulus2 NonSuite) ont été définies et moyennées pour les différents types de stimuli (phrases, images, contrôles), de 200 ms avant la présentation du stimulus et pouvant aller jusqu'à 3000 ms après le stimulus pour pouvoir inclure la période minimale de délai qui suit un stimulus.

- **Analyse en composante indépendante**

Une baseline allant de -200 ms à 0ms a été soustraite à l'activité et une étape de rejet automatique des époques trop différentes des autres a été effectuée. Une analyse en composante indépendante (ICA) a ensuite été réalisée sur les 64 électrodes ainsi que sur les électrodes EOG pour chaque sujet. L'étape d'ICA a été faite en utilisant EEGLAB via un algorithme de séparation de source aveugle (BSS=

blind source separation), l'algorithme UW-Sobi (Belouchrani et al., 1997; Tang et al., 2004, 2005 ; Yeredor, 2000, 2002). Nous avons préféré privilégier l'algorithme UW-Sobi plutôt que l'algorithme INFOMAX qui est implémenté dans EEGLAB et qui est généralement utilisé pour effectuer la séparation de sources, car les algorithmes de type HOS (higher order statistics), comme INFOMAX, sont plus susceptibles d'identifier de fausses composantes que celles qui sont associées à des sources neurologiques réelles, ce qui altère fortement la qualité du signal récupéré et les analyses ultérieures de localisation de source. Les algorithmes SOS (second order statistics) tels que UW-Sobi sont moins sensibles à ce biais (Lio & Boulinguez, 2012). Les résultats de la décomposition issue de l'ICA ont été inspectés visuellement chez chaque sujet pour identifier et rejeter les composantes associées aux clignements des yeux, aux mouvements horizontaux et verticaux des yeux, aux artefacts cardiaques et respiratoires ou encore à des problèmes de bruit durant l'acquisition.

Analyses statistiques ERPs

Avant d'effectuer les analyses statistiques, les données ont été re-référencées par rapport aux 2 électrodes mastoïdes car cette référence est la plus adaptée pour obtenir des effets liés à la N400 (Kutas & Federmeier, 2011) et sur une période de temps correspondant à la présentation du stimulus (0-2000 ms), l'analyse des périodes de délai qui font suite au stimulus sera l'objet d'une analyse ultérieure.

Les analyses statistiques sur les ERPs ont été faites en utilisant le module STUDY de EEGLAB qui permet de définir un modèle (nombre de sujets, nombre de conditions) à partir des données individuelles nettoyées lors des pré-traitements. Ce module d'EEGLAB permet d'utiliser des statistiques paramétriques ou non paramétriques pour calculer et estimer la fiabilité des différences entre des conditions et/ou des groupes. Nous avons ainsi pu définir différents modèles (stimulus 1 vs stimulus 2, Suite vs Non Suite) et des analyses statistiques paramétriques ont été utilisées pour les tester. Pour comparer la différence entre deux conditions, la fonction EEGLAB réalise un test T apparié, à chaque temps de latence sur les ERPs moyennés des différents sujets. Lorsque plus de deux conditions sont comparées, les résultats sont soumis à une ANOVA. Les résultats significatifs (selon le seuil et la méthode d'analyse choisis) sont visualisables en grisé pour chaque électrode ou par une période noire pour le moyennage de plusieurs électrodes sélectionnées par l'utilisateur. En ce qui nous concerne, seuls les résultats significatifs sur plusieurs électrodes concomitantes et sur une période de temps non négligeable ont été pris en considération.

II.3.*5*) Résultats

Nous rappelons que notre but principal était de tester l'existence, pour les images et les phrases, de processus précoces modalité-spécifiques et de processus communs plus tardifs impliqués dans l'intégration sémantique et la construction d'une représentation cohérente en fonction du contexte. Ces processus devant être révélés par la comparaison des ERPs associés à la présentation de stimuli successifs sémantiquement reliés ou non, nous avons choisi d'examiner les effets liés à la présentation de 4 stimuli : Stimulus 1 Suite, Stimulus 1 NonSuite, Stimulus 2 Suite et Stimulus 2 Non Suite. Nous nous concentrerons particulièrement sur les comparaisons Stimulus 1 Suite vs NonSuite et Stimulus 2 Suite vs NonSuite, les sujets ne pouvant pas savoir à l'avance si le stimulus 1 fait partie d'une suite logique, il ne devrait pas y avoir de différence pour le stimulus 1 alors qu'elle devrait exister pour le stimulus 2. Ces comparaisons seront réalisées indépendamment pour les images et les phrases. Compte tenu de la différence inhérente entre stimuli verbaux et non-verbaux, une comparaison directe n'aurait pas été pertinente, mais les résultats seront discutés conjointement.

- **ERPs Images Suite vs NonSuite**

Les résultats présentés ici correspondent à des analyses paired t-test ou anova quand les 4 conditions sont présentées (p unc = 0.001) sur 18 sujets, pour une période de temps post-stimulus allant jusqu'à 2000ms, avec un lissage à 25Hz. Les figures les plus pertinentes ont été sélectionnées pour illustrer les résultats.

La différence Suite vs NonSuite pour Image 1 ne permet pas de trouver des différences notables significatives (voir le résultat pour les 64 électrodes, Annexes 7 et 8). Seul un léger effet, marqué par une négativité plus importante pour Image 1 Suite est notable vers 800 ms pour quelques électrodes pariéto-temporales gauches. Malgré toutes nos précautions expérimentales, il n'est pas impossible qu'il y ait une légère différence entre les stimuli des conditions Image 1 Suite et Image 1 NonSuite, cependant l'amplitude de cet effet est très faible (<2mv) et cet effet disparaît lorsque l'on moyenne les électrodes concernées.

En revanche, la différente Suite vs NonSuite pour l'Image 2 produit plusieurs effets présents sur de nombreuses électrodes centro-frontales. Ces effets se caractérisent par :

- une positivité plus importante qui commence vers 180 ms et se termine autour de 250, avec un pic plus marqué vers 200 ms pour la condition Image 2 Non Suite
- suivi d'un effet de positivité avec un pic vers 300ms où il n'y a pas de différence entre les deux conditions

- une négativité plus importante entre 400 et 550 ms, avec un pic vers 500ms, pour Image 2 Non Suite
- une positivité plus marquée et plus tardive, vers 900-1100 ms, pour Image 2 Non Suite

La figure ci-après illustre ces effets sur l'électrode FCz (meilleure électrode=électrodes où les effets sont les plus marqués).

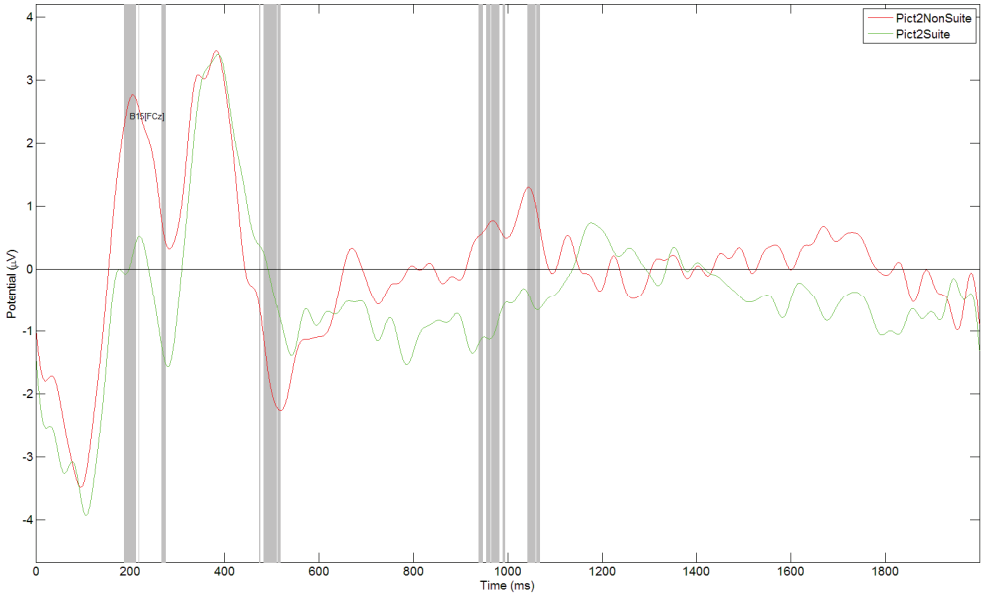


Fig 72) Comparaison des ERPs des conditions Suite et Non Suite pour le Stimulus 2 Image
Ces résultats illustrent la comparaison des ERPs, moyennés sur l'ensemble des sujets, de la condition Image 2 Suite (vert) et ceux de la condition Image 2 Non Suite (rouge) sur l'électrode FCz où les effets sont les plus marqués. Les effets significatifs apparaissent en grisé (p unc = 0.001, 18 sujets).

Ces effets sont présents sur la plupart des électrodes frontales, mais ils sont particulièrement significatifs sur les électrodes centro-frontales (F1, AFz, Fz, F2, FCz) pour la différence Image 2 Suite/NonSuite (alors que ces effets ne sont pas présents sur les mêmes électrodes pour la différence Suite/NonSuite Image 1). La figure 73 représente les différences entre Image 1 Suite/Non Suite et Image 2 Suite/Non Suite pour le regroupement des électrodes centro-frontales suivantes : F1, AFz, Fz, F2, FCz.

Lorsqu'on compare les 4 conditions images par une ANOVA (Image 1 Suite, Image 1 NonSuite, Image 2 Suite, Image 2 NonSuite), on note également plusieurs effets intéressants (voir Figure 73).

- Globalement, les stimuli 2 (Suite et Non Suite) génèrent un pic négatif vers 100 ms, plus marqué que pour les stimuli 1 (Suite et Non Suite)
- la composante positive vers 200 ms caractérise une activité différente pour Image 2 Suite, les courbes des autres conditions (Images 1 Suite & Non Suite, Image 2 Non Suite) étant équivalentes en termes d'amplitude et de latence. Elle est d'ailleurs suivie d'un pic négatif vers 300 ms qui lui aussi

semble caractériser une activité différente pour Image 2 Suite. En revanche, la composante positive entre 300-400 ms ne présente aucune différence significative entre ces différentes conditions.

- la composante négative dont le pic se situe vers 450 ms est nettement plus marquée pour la condition Image 2 Non Suite que pour les autres conditions.
- de même que pour les effets précoces situés entre 200-400 ms, les effets plus tardifs (entre 900-1100 ms) semblent particulièrement liés à un traitement différentiel de la condition Image 2 Suite

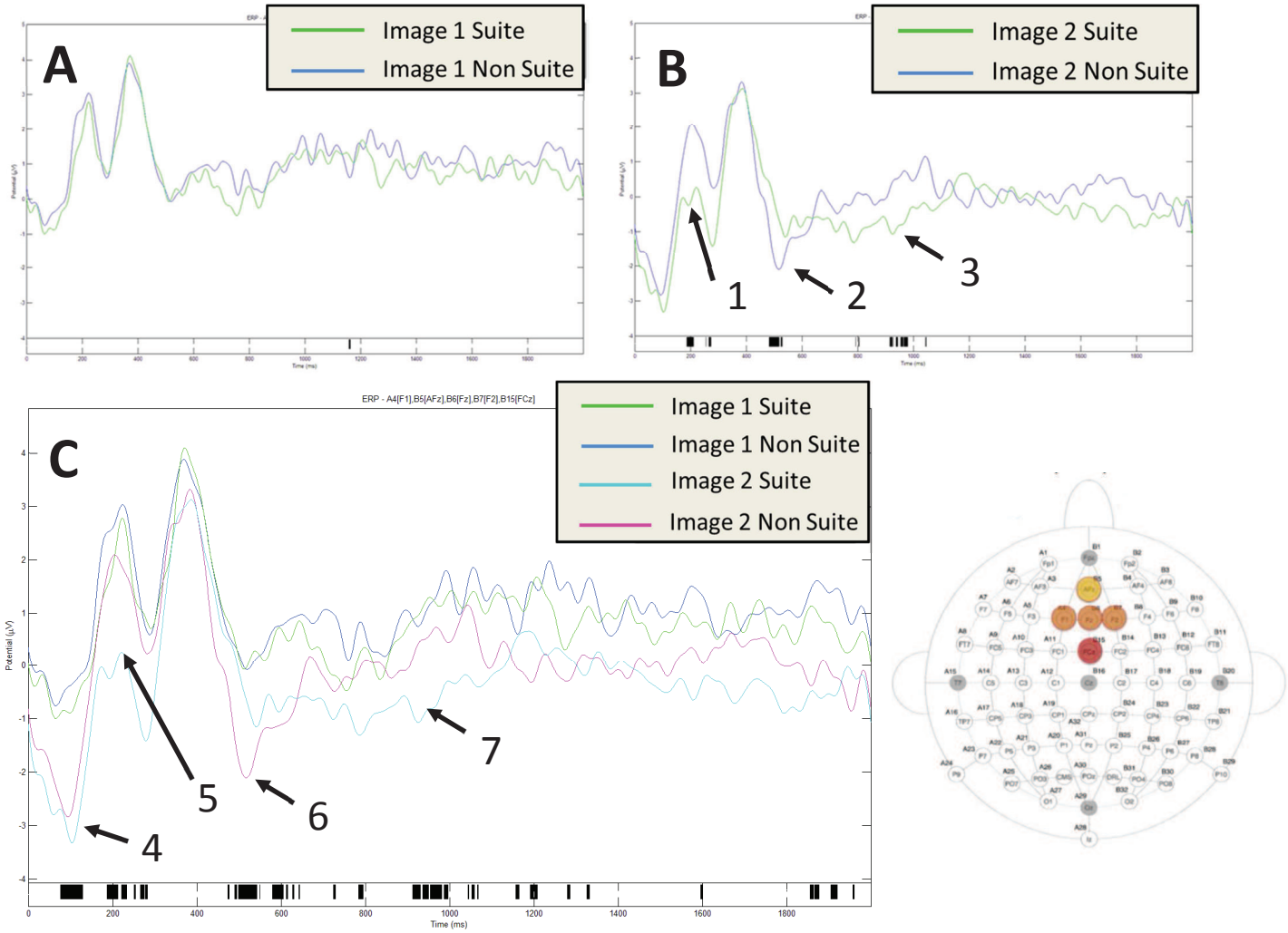


Fig 73) Synthèse des résultats pour les 4 conditions Images

La figure A représente les ERPs pour les conditions Image 1 Suite et Image 1 Non Suite, conditions pour lesquelles on n'observe essentiellement aucune différence. La figure B représente les ERPs pour les conditions Image 2 Suite et Image 2 Non Suite où l'on constate des différences. Le premier effet (1) apparaît vers 200ms où la condition Image 2 Non-Suite génère une positivité P200 plus importante. Vers 450ms, la condition Non Suite produit une négativité plus importante, caractéristique de la N400 (2). Cette négativité relative s'inverse ensuite, se transformant en un effet de positivité tardive où la condition Non-Suite est significativement plus grande que la condition Suite vers 900-1000 ms (3). La figure C représente la comparaison des ERPs pour les 4 conditions Images. En comparant les réponses pour Image 1 et Image 2, un nouvel effet apparaît : à 100 ms, l'amplitude de la N100 est significativement plus importante pour Image 2 (4). Vers 200-300ms, on constate que la condition Image 2 Suite génère un complexe P200-N300 marqué par une amplitude réduite par rapport à toutes les autres conditions (5). Vers 400-500ms, on observe deux effets (6) : la condition Image 2 Non Suite est plus négative que toutes les autres conditions ; il apparaît aussi que les deux stimuli Image 2 (conditions Suite et Non Suite) ont des valeurs plus négatives que les stimuli Image 1. Enfin, dans la période plus tardive (900-1000ms), la condition Image 2 Suite présente une positivité réduite suggérant un effort d'intégration moins important (7). Tous les résultats présentés correspondent à des ERPs moyennés, pour les 18 sujets, sur 5 électrodes centro-frontales (F1, AFz, Fz, F2, FCz) au seuil p unc = 0.001. Les périodes où les différences sont significatives sont indiquées par un marquage noir sur l'échelle du temps (différences entre deux conditions, Fig A et B, ou au moins pour une condition dans l'ANOVA, Fig C).

- **ERPs Phrases Suite vs NonSuite**

Les résultats présentés ici correspondent à des analyses paired t-test ou anova quand les 4 conditions sont présentées (p unc = 0.01) sur 18 sujets, pour une période de temps post-stimulus allant jusqu'à 2000ms, avec un lissage à 25Hz. Les figures les plus pertinentes ont été sélectionnées pour illustrer les résultats.

Comme attendu, le mode de présentation en bloc des phrases ne permet pas de caractériser des composantes clairement associées aux différents mots telles que la N400. Cependant, là encore la comparaison des différences entre stimulus 1 Suite/Non Suite et stimulus 2 Suite/Non Suite donnent des résultats significatifs pour le stimulus 2 alors qu'il n'y en a pas pour le stimulus 1.

En effet, la différence Suite vs NonSuite ne permet pas de trouver des différences significatives notables pour Phrase 1 (voir le résultat pour les 64 électrodes, Annexes 9 et 10). En revanche, la différente Suite vs NonSuite pour Phrase 2 produit plusieurs effets présents sur de nombreuses électrodes (voir figure 74 pour les effets sur la meilleure électrode FCz). Ces effets se caractérisent par :

- une positivité précoce plus importante entre 200 et 350 ms, avec un pic plus marqué vers 300 ms pour la condition Phrase 2 Non Suite
- une positivité plus marquée et plus tardive, vers 1000-1200 ms ainsi que des effets mineurs entre 1600-1800ms pour Phrase 2 Non Suite.

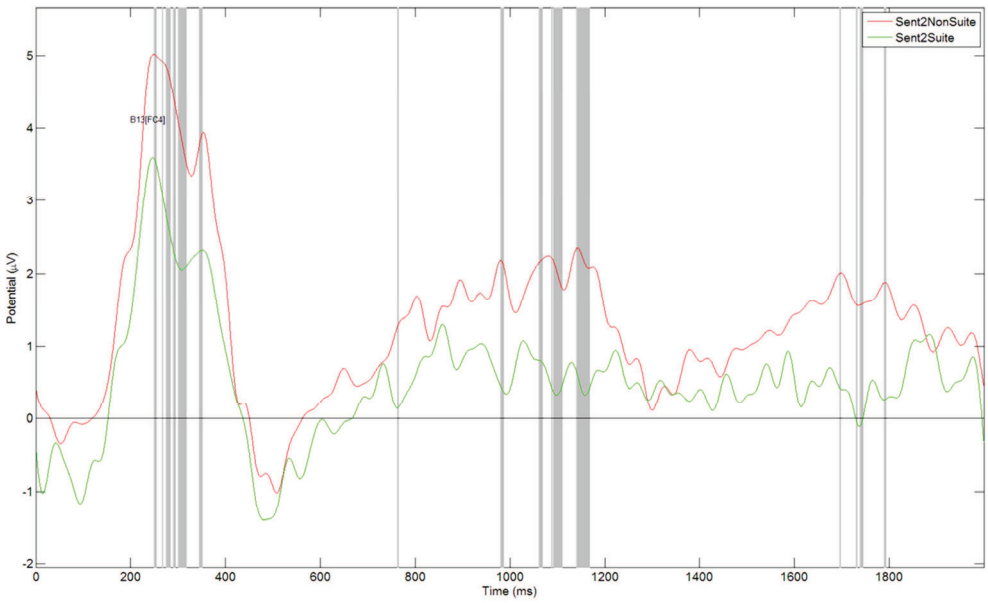


Fig 74) Comparaison des ERPs des conditions Suite et Non Suite pour le Stimulus 2 Phrase
Ces résultats illustrent la comparaison des ERPs, moyennés sur l'ensemble des sujets, de la condition Phrase 2 Suite (vert) et ceux de la condition Phrase 2 Non Suite (rouge) sur l'électrode FCz où les effets sont les plus marqués. Les effets significatifs apparaissent en grisé (p unc = 0.001, 18 sujets).

Si les effets concernant la composante positive située autour de 300 ms sont significatifs sur plusieurs groupements d'électrodes dont les électrodes fronto-centrales (Fpz, AFz, Fz, FCz), postéro-centrales droites (FC4, C2, C4, C6, CP4) mais aussi latérales droites (F8, FT8, T8, TP8), ce sont sur les électrodes fronto-centrales qu'ils sont les plus importants (Figure 75). De même si les effets plus tardifs (1000-1200 ms et 1600-1800 ms) sont présents sur un groupement frontal antérieur (Fp1, AF7, Fpz, FP2, AF8), ils semblent plutôt être liés à une localisation postérieure latérale droite, surtout temporale (FC4, C2, C4, C6, CP4 et F8, FT8, T8, TP8) car c'est sur ces regroupements d'électrodes que les effets sont les plus significatifs, mais aussi temporale gauche dans une moindre mesure (F7, FT7, T7, TP7). Ces effets ne sont pas présents, pour les mêmes groupements d'électrodes, pour la différence Suite/Non Suite phrase 1.

Par ailleurs, lorsqu'on compare les 4 conditions phrases par une ANOVA (Phrase 1 Suite, Phrase 1 NonSuite, Phrase 2 Suite, Phrase 2 NonSuite,), on note également plusieurs effets intéressants (voir Figure 75) :

- des effets différentiels précoces (<200 ms) sont présents à la fois sur les regroupements d'électrodes frontales mais aussi postéro-latérales, ils se caractérisent par une négativité plus marquée pour les Phrases 1 (Suite et Non Suite) par rapport aux Phrases 2 (Suite et Non Suite).
- Le phénomène de positivité précoce plus importante entre 200 et 350 ms, avec un pic plus marqué vers 300 ms semble caractériser un traitement différent pour la condition Phrase 2 NonSuite, puisqu'il est plus marqué dans cette condition que dans les autres conditions
- Il en est de même pour la positivité plus marquée et plus tardive, vers 1000-1200 ms ainsi que pour des effets mineurs entre 1600-1800ms qui sont différents pour Phrase 2 Non Suite par rapport aux autres conditions
- la comparaison des 4 conditions met aussi en évidence des effets différentiels entre les Phrases 1 (Suite et Non Suite) et Phrases 2 (Suite et Non Suite) entre 750 et 900 ms, marqués par une positivité plus importante pour les Phrases 2. Cependant, comme les stimuli des Phrases 1 et 2 ne sont pas tout à fait comparables (construction et longueur de phrases), ces effets sont difficilement interprétables car ils peuvent refléter des processus différents mais aussi le fait que les sujets ne soient pas à une étape équivalente du traitement de la phrase compte tenu des différences entre les phrases et d'un sujet à un autre.

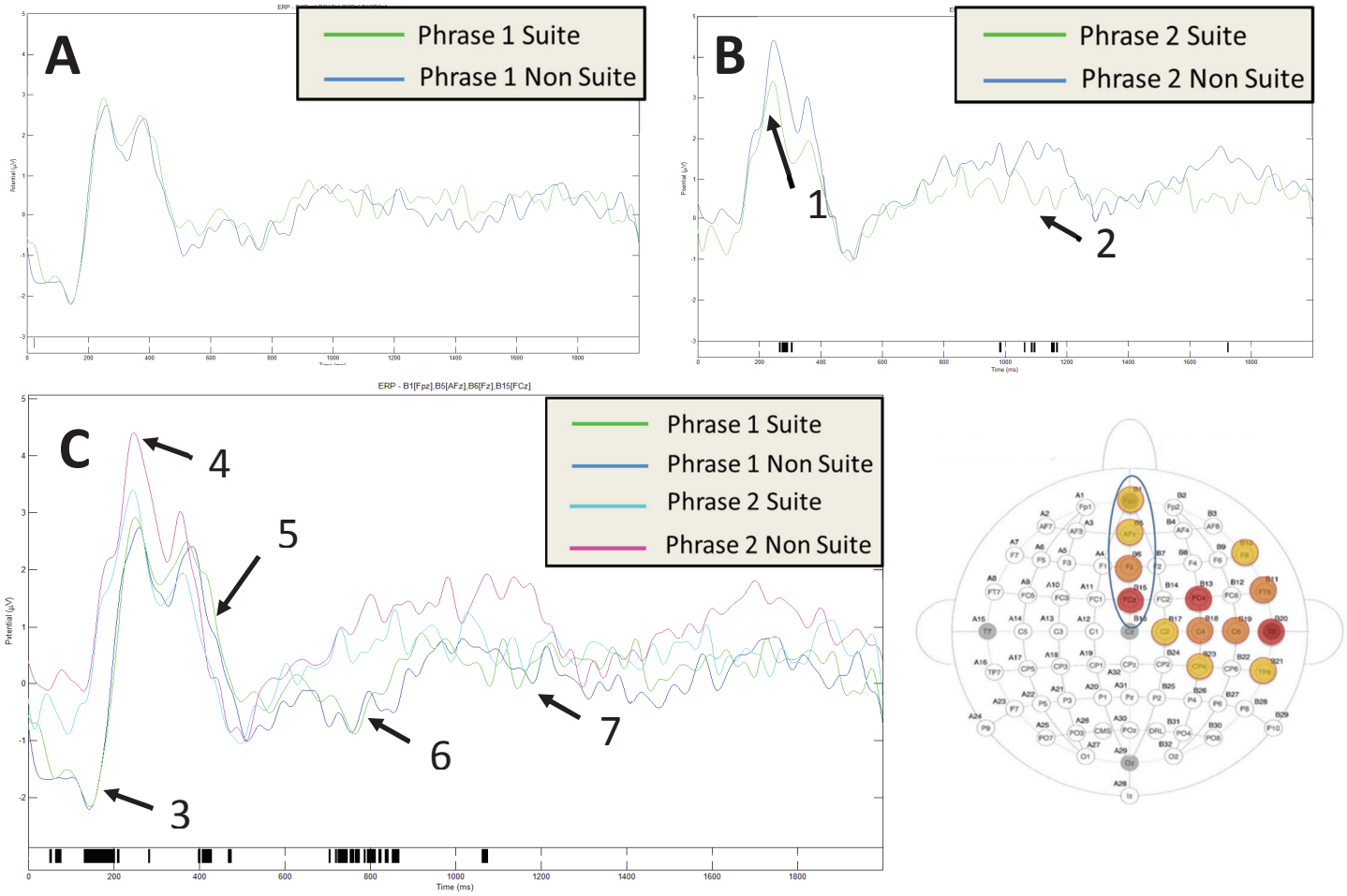


Fig 75) Synthèse des résultats pour les 4 conditions Phrases

La figure A représente les ERPs pour les conditions Phrase 1 Suite et Phrase 1 Non Suite, conditions pour lesquelles on n'observe essentiellement aucune différence. La figure B représente les ERPs pour les conditions Phrase 2 Suite et Phrase 2 Non Suite où l'on constate des différences. Le premier effet (1) apparaît vers 200ms où la condition Phrase 2 Non-Suite génère une positivité P200 plus importante. On observe ensuite une positivité plus importante et plus tardive pour la condition Phrase 2 Non Suite (2) qui devient significative vers 1000ms. La figure C représente la comparaison des 4 conditions Phrases. En comparant les réponses pour Phrase 1 et Phrase 2, un nouvel effet apparaît : vers 100-200 ms, l'amplitude de la N100 est significativement plus importante pour Phrase 1 Suite et Non Suite (3). A 200ms, on constate que la condition Phrase 2 Non Suite génère une composante P200 dont l'amplitude est plus importante par rapport aux trois autres conditions (4). Vers 400 ms, on note un léger effet entre les conditions Phrase 1 et Phrase 2 marqué par une négativité plus importante pour les stimuli Phrase 2, même si cet effet ne caractérise pas typiquement une N400 (5). A partir de 700ms, on observe un effet inverse qui se traduit par une positivité plus importante pour les stimuli Phrase 2 (6). Finalement, vers 1000ms, les stimuli de la condition Phrase 2 Non Suite génèrent une positivité tardive plus importante alors que les effets associés à la condition Phrase 2 Suite reviennent au niveau des autres conditions (7). Tous les résultats présentés correspondent à des ERPs moyennés, pour les 18 sujets, sur 4 électrodes centro-frontales (Fpz, AFz, Fz, FCz) au seuil p unc = 0.001. Les périodes où les différences sont significatives sont indiquées par un marquage noir sur l'échelle du temps (différences entre deux conditions, Fig A et B, ou au moins pour une condition dans l'ANOVA, Fig C).

- **Synthèse et discussion des résultats ERPs**

  - différences entre Stimulus1 et Stimulus 2

Tout d'abord, que ce soit pour les images ou pour les phrases, les résultats montrent des différences entre Stimulus 1 et Stimulus 2 en fonction du fait que le second stimulus est dans la continuité logique du premier : il n'y a pas de différence entre les conditions Suite et Non Suite pour le Stimulus 1 (ce qui était attendu puisqu'au moment où on leur présente le premier stimulus, les sujets ne savent pas s'il sera suivi d'un stimulus Suite ou Non Suite) alors qu'il y a des différences pour le Stimulus 2. Cette variabilité dans le décours temporel post-stimulus, à la fois pour les phrases et pour les images, montre qu'il existe un effet global lié à la cohérence contextuelle et à la différence de lien sémantique entre deux stimuli. Il se manifeste par un traitement différent selon qu'un second stimulus présenté a une continuité sémantique ou non avec un premier stimulus proposé. Même si pour les phrases, les effets sont plus minimes et que le mode de présentation en bloc rend les interprétations plus difficiles, le fait qu'il existe une variabilité du décours temporel spécifique pour le Stimulus 2 va dans le sens de l'existence d'un effet du contexte. Cet effet se traduit par des modulations du décours temporel, spécifiques pour le Stimulus 2, durant des périodes très précoces (entre 200 et 300 ms) et des périodes plus tardives (autour de 1000 ms) pour les images et les phrases, ainsi que par un effet de négativité autour de 400 ms spécifique aux images. Le fait qu'il y ait des effets très précoces associés à la contrainte contextuelle est cohérent avec l'idée que les informations contextuelles sont utilisées très précocement pour guider un traitement sémantique plus approfondi et que la reconnaissance/compréhension commence avant même la fin du traitement perceptivo-sémantique (Rieger et al., 2005; Rousselet et al., 2005 ; Joubert et al., 2007). Par ailleurs, les effets plus tardifs semblent montrer la nécessité d'un processus de réévaluation ultérieure ou d'intégration sémantique plus complexe, processus différent en fonction de l'information contextuelle préalablement disponible. En tout cas, ces résultats sont tout à fait cohérents avec l'idée que l'accès aux informations sémantiques dépend d'une mise à jour continue de l'information provenant du stimulus mais aussi du contexte plus large dans lequel il est intégré (Wlotko & Federmeier, 2012).

  - Stimuli Images

Traitements précoces attentionnels et contextuels

Concernant les images, on s'aperçoit tout d'abord qu'il y a un traitement différent très précoce entre Image 1 et Image 2 au niveau de la composante N100 dont l'amplitude est plus grande pour les Images 2 que pour les Images 1 (en dehors de leur appartenance à une condition Suite vs

Non Suite). Cet effet témoigne probablement de l'engagement attentionnel des sujets dans la tâche qui est plus grand pour le second stimulus que pour le premier car les sujets savent qu'il y a un second stimulus à traiter et qu'il risque d'être différent du premier. Etant donné qu'il n'y a pas de différence Suite/Non Suite, ce phénomène correspondrait plutôt à une réaction attentionnelle de bas niveau, sans influence due au traitement de l'information contextuelle au niveau sémantique. Ce rôle attentionnel de la N100 est cohérent avec des travaux qui ont montré son implication dans l'attention sélective, son rôle de filtrage perceptif et le fait que l'amplitude de la N100 était plus grande lorsque l'attention était focalisée ou lorsque des stimuli étaient attendus (Parasuraman, 1980 ; Woods, 1990 ; Coull, 1998).

En revanche, le complexe d'ondes positives-négatives entre 200-300ms révèle déjà des différences liées à une analyse contextuelle, avec une amplitude réduite de la P200 et un pic plus négatif vers 300ms, qui suit la P200 pour la condition Image 2 Suite alors que les trois autres conditions (Image 2 Non Suite, Images 1 Suite et NonSuite) sont traitées indistinctement. La première composante, de par son amplitude, sa latence et sa localisation centro-frontale, semble correspondre à la description de la P200, une composante généralement générée par la présentation de stimuli visuels et dont le rôle reste mal connu car elle semble modulée par des tâches cognitives variées. Toutefois, il semble y avoir un consensus sur le fait que cette composante est liée à des traitements perceptifs d'ordre supérieur modulés par l'attention. Elle a notamment été étudiée dans des processus de recherche visuelle et d'attention et du traitement de l'information contextuelle ; elle serait impliquée plus généralement dans des aspects mnésiques car elle pourrait participer au système qui permet de comparer les informations sensorielles entrantes avec les informations en mémoire (Freunberger et al., 2007 ; Luck et al., 1994 ; Evans & Federmeier, 2008; Phillips & Takeda, 2009; Dunn et al., 1998; Lefebvre et al., 2005). La P200 a été traditionnellement étudiée dans le cadre de la perception visuelle, en particulier par rapport à la manière dont les stimuli sont évalués, notamment dans le cadre de paradigmes d'amorçage visuel où l'on cherche à comprendre comment une information préalable peut façonner une réponse subséquente. Dans ce type de paradigmes expérimentaux, on présente brièvement aux participants une image ou un mot, suivi d'un délai, puis d'un second stimulus sur lequel les sujets doivent faire une tâche de classification ; il a été montré que les stimuli cibles généraient une amplitude plus importante de la P200 par rapport à des stimuli contrôles (Luck et al., 1994). Ces résultats ne sont pas vraiment cohérents avec nos propres résultats, qui sont plus en accord avec les travaux montrant que l'attention accrue des sujets tend à faire diminuer l'amplitude de la P200 car une attention plus importante restreint la taille de l'espace de recherche ou le nombre d'associations qui doivent être faites, ce qui peut alors faciliter la recherche visuelle durant des stades très précoces du traitement perceptif (Phillips & Takeda, 2009). Un plus

grand nombre de cibles probables diminue également l'amplitude de la P200 car celle-ci est sensible au nombre de distracteurs lors de la recherche visuelle. C'est ce qui pourrait se produire avec nos stimuli : en plus d'une attention augmentée par le but de la tâche et la détection précoce de la congruence entre les deux images, lorsque la seconde image est cohérente avec la première, il y a moins de détails à ré-analyser car un certain nombre d'éléments sont communs aux deux situations (personnages, objets, contexte…). Cela pourrait expliquer l'amplitude plus basse de la P200 pour les stimuli Image 2 Suite. Ce que semblent aussi confirmer les différences plus tardives entre cette condition et les autres conditions car on observe un traitement différentiel de cette condition, tout au long de la présentation du stimulus. Ce traitement différentiel est marqué par une négativité générale plus importante tout au long du décours temporel de cette condition (avec des différences significatives vers 900-1000 ms mais aussi au-delà) alors que les autres conditions sont toutes les trois traitées de façon équivalente, comme des stimuli nouveaux. Une autre interprétation de la modulation de l'amplitude de cette composante est fournie par les travaux qui étudient la théorie de l'esprit. En effet, Wang et al. (2010) ont montré une amplitude réduite de la P200 lorsque les sujets regardaient des dessins représentant des scènes sans personnage (un paysage de neige, un salon…) par rapport à des scènes avec des personnages où une compréhension poussée des états mentaux et des intentions des personnes est nécessaire. Peut-être est-ce le cas avec nos stimuli où lorsque la seconde image est contextuellement cohérente avec la première, il y aurait une analyse moins approfondie des intentions de la personne et des buts de l'action puisque ce sont plus ou moins les mêmes que dans la première image. Alors que, lorsque la seconde image n'est pas cohérente, il faut à nouveau effectuer cette analyse des intentions et des buts des personnages, ce qui pourrait là-aussi expliquer les résultats équivalents en ce qui concerne l'amplitude de la P200 pour Image 2 Non Suite et Image 1. Par ailleurs, Wang et al (2010) ont également montré l'existence des effets plus tardifs (vers 1000-1300 ms) associés à une composante positive tardive dont l'amplitude est plus positive lorsque le sujet doit utiliser la théorie de l'esprit pour comprendre la situation que lorsqu'il doit juste percevoir le personnage ou la scène. Ces effets plus tardifs sont aussi concordants avec ceux mis en évidence dans notre étude, ce qui pourrait laisser penser que la P200 a effectivement un rôle associée à la TOM.

Si la composante P200 est clairement dissociable dans notre étude, le pic négatif qui lui fait suite est plus difficilement interprétable. De par sa latence et sa localisation, il ressemble à la N300 décrite précédemment comme une composante spécifiquement associée au traitement sémantique des images (Barrett & Rugg, 1990 ; McPherson et Holcomb, 1999 ; Nigram et al., 1992). Cependant, ces précédents travaux ont plutôt montré une modulation de l'amplitude de cette composante, marquée par une négativité importante lorsque des stimuli ne sont pas cohérents les uns par rapport

aux autres, négativité supposée refléter un effort d'appariement supplémentaire lorsque les stimuli sont non concordants (images d'objets non liés vs objets liés sémantiquement, prédictibilité d'un stimulus dans un contexte donné…). Ce n'est pas le cas dans notre étude où l'amplitude de cette composante est au contraire réduite lorsque des stimuli contextuellement cohérents (Image 2 Suite) sont présentés par rapport aux autres conditions. Une autre interprétation pourrait être que cette composante n'est pas une N300 mais une N200 tardive, une composante qui a souvent une distribution fronto-centrale et dont l'amplitude mais aussi la latence (entre 200 et 350 ms) peut varier en fonction d'un grand nombre de stimuli. Si la composante N200 a été initialement considérée comme un « détecteur » des défauts d'appariements (mismatch), on pense aujourd'hui qu'elle pourrait refléter des fonctions de contrôle exécutif et elle a notamment été largement étudiée dans le cadre du langage (Folstein & Van Petten, 2008; Schmitt, Münte, et Kutas, 2000). Ce qui peut laisser penser que cette composante serait en fait une N200 tardive, c'est le fait que l'on trouve juste ensuite une composante P300. En effet, la N200 et la P300 souvent étudiées ensemble sont considérées comme un complexe de composantes (N200, P300 divisée en P3a et P3b) réagissant conjointement, la N200 ayant une réponse fonctionnelle très proche de la P3b notamment parce que la probabilité d'un stimulus affecte de façon similaire l'amplitude de ces deux composantes. La N200 et la P300 seraient sensibles aux manipulations expérimentales similaires et représenteraient une connexion de mécanismes mentaux qui travaillent ensemble pour interpréter les changements contextuels (Patel et al., 2005 ; Azizian et al., 2006). Cependant, là aussi, cette interprétation semble contraire avec nos résultats puisque l'amplitude de la N200 tardive (ou N300) n'est pas modulée par les stimuli non cohérents mais au contraire par les secondes images qui sont dans la continuité sémantique des premières et que, de plus, l'amplitude de la P300 ne montre aucune différence significative entre les différentes conditions. Quoiqu'il en soit, que l'on considère que la composante négative soit une N300 spécifique aux images ou une N200 tardive faisant partie d'un complexe N200/P3b, nos résultats ne permettent pas vraiment de donner une interprétation claire de la fonction de cette composante**.** La baisse d'amplitude de cette composante est-elle liée à une augmentation/baisse générale de l'attention pendant la compréhension de ces stimuli ? Doit-on associer un nouveau rôle fonctionnel à cette composante ? Des travaux complémentaires sont nécessaires pour mieux comprendre le rôle de cette composante dans l'intégration sémantique (nous y reviendrons ultérieurement).

L'étude de la composante P300 est aussi très intéressante dans le cadre de ce travail. Les premières observations de la P300 (plus spécifiquement la composante qui serait appelée plus tard la P3b ; Squires et al., 1975) ont été faites dans le milieu des années 60 avec les travaux de Chapman & Bragdon (1964) qui mirent en évidence des réponses ERP spécifiques pour des nombres en

comparaison à des flashs de lumière ; leurs résultats furent interprétés comme une différence liée au fait que les stimuli aient une signification ou pas. Par la suite, la P300 a été subdivisée en deux composantes : la P3a qui est une composante plus précoce (250-280 ms) associée à des effets attentionnels et au traitement de la nouveauté (Comerchero & Polich, 1999 ; Polich, 2003 ; 2007) et la P3b qui correspond à la P300 décrite initialement et qui a été particulièrement étudiée dans le cadre du traitement de l'information. De façon générale, il a été montré que les événements improbables généraient une P3b et que moins ils sont probables, plus la P3b est importante (Donchin, 1981 ; Hillyard & Kutas 1983). Cependant ces effets seraient fortement dépendants de la tâche demandée et des réponses du sujet qui lui sont associées, l'amplitude de la P3b étant une bonne mesure de la charge cognitive requise par une tâche et de la pertinence du traitement d'un stimulus spécifique pour réaliser une tâche (Donchin, 1981 ; Courchesne et al., 1975; Debener et al., 2005; Dien et al., 2004; Goldstein et al., 2002; Johnson, 1988; Polich & Comerchero, 2003; Spencer et al., 1999). Dans notre étude, il n'y aurait pas d'effet sur l'amplitude de la P300 car tous les événements ont une signification, ils ne sont pas impossibles même si certains sont moins probables dans un contexte donné. Par ailleurs, la tâche à effectuer par les sujets n'étant pas une tâche de détection de l'incongruence, le traitement des stimuli 1 et 2 (Suite et NonSuite) sont tous les deux pertinents pour la tâche (car il est probable qu'une question soit posée sur l'un ou l'autre de ces événements), ce qui pourrait également expliquer que la P300 ne soit pas différente dans les 4 conditions. De nombreuses études ont montré des liens forts entre l'amplitude de la P3b et la confirmation d'une attente, mais ces études ont surtout porté sur des tâches de classification catégorielle et sur des stimuli ne nécessitant pas un traitement sémantique, tels que des figures ou des chaînes de lettres (voir Folstein & Van Petten, 2011 ; Van Petten & Luka, 2012). Le rôle plus général de la P300 (qui n'est pas associé à un traitement sémantique) serait plutôt un rôle de mise à jour de la mémoire de travail, concernant les informations contextuelles, ce qui permettrait leur stockage ultérieur en mémoire (Donchin & Coles, 1988 ; Polich, 2007 ; Van Petten, 2012). Ce serait plutôt cette fonction de la P300 que l'on retrouverait dans notre étude.

Traitements sémantiques et contextuels

Si cette composante N300 ou N200 tardive (faisant partie d'un complexe N200/P300) est difficile à catégoriser et que son rôle fonctionnel manque de clarté dans notre étude, nos résultats montrent pourtant clairement que cette dernière est dissociée temporellement et fonctionnellement de la N400, en accord avec les travaux de la littérature (Hamm et al. 2002 ; McPherson et Holcomb, 1999). Nos résultats révèlent en effet l'existence d'un effet avec un pic négatif vers 490 ms dont la latence et la localisation centro-frontale sont concordants avec les travaux qui ont mis en évidence cette composante, même si la latence de la composante que nous observons semble un peu plus

tardive que celle observée généralement pour les stimuli images (Barrett & Rugg, 1990). Nos résultats montrent que cette composante n'est pas seulement liée au traitement des violations sémantiques fortes (objets incongrus, actions impossibles…), elle est dépendante de l'information contextuelle disponible. En effet, le profil de cette composante, observé dans notre étude, semble lié à un effort d'intégration sémantique contextuelle plus important lorsque les stimuli ne sont pas cohérents puisque cette composante montre un effet spécifique (amplitude plus négative) pour la condition Image 2 Suite par rapport aux autres conditions Image. Nos résultats confirment (si tant est que cela soit nécessaire) que la N400 a avant tout un rôle de marqueur reflétant des processus sémantiques et qu'elle n'est pas seulement associée au traitement d'informations linguistiques (Kutas & Federmeier, 2011) ; ils sont tout à fait cohérents avec le rôle attribué actuellement à la N400 d'intégration des informations dans un contexte sémantique plus large en fonction des connaissances sur les événements réels (Van Berkum, 2003 ; Metusalem et al., 2012). Les différences observées sur cette composante semblent par ailleurs confirmer son rôle dans un traitement sémantique de haut niveau, alors que les effets liés aux composantes plus précoces (notamment la possible N300) refléteraient des processus de catégorisation initiale plus basiques et de nature pré-sémantique, en accord avec les résultats qui montrent que la N400 est beaucoup plus sensible aux aspects sémantiques reliant différents éléments alors que la N300 est plus concernée par l'identification des propriétés structurelles des éléments présents dans un stimulus visuel (Hamm et al.2002 ; McPherson et Holcomb, 1999).

- Stimuli Phrases

En ce qui concerne les phrases, il est évident que le mode de présentation n'est pas le plus adapté pour déterminer les composantes spécifiquement impliquées dans le traitement des différents mots qui composent une phrase et les rôles qui peuvent leur être associés. Cependant, les résultats mettent en avant des différences intéressantes qui peuvent laisser supposer l'existence de processus communs à l'intégration sémantique, quelle que soit la modalité, mais aussi la participation d'un certain nombre d'effets précoces différents, révélant des processus et des stratégies propres à chaque modalité.

Traitements précoces attentionnels et traitements tardifs contextuels

Si l'analyse des stimuli images avait mis en évidence une composante N100 plus négative pour les Images 2, l'effet pour les phrases est inverse : on trouve une amplitude négative plus importante pour les Phrases 1 (Suite et Non Suite) que pour les Phrases 2 (Suite et Non Suite).

Pourtant la tâche est tout à fait la même dans les deux conditions et on aurait pu penser que l'attention serait dirigée de la même façon pour les images et pour les phrases, avec une attention accrue lors de la présentation du Stimulus 2. Cet effet et cette différence entre les conditions sont difficiles à interpréter (car le mode de présentation des phrases en bloc perturbe le timing des composantes précoces de type N1-P2) ; il semble que les stratégies attentionnelles initiales soient différentes pour les conditions images et phrases. Peut-être que la simulation générée par les phrases suscite une « remise à niveau attentionnelle » afin de traiter au mieux la seconde phrase, alors que pour les images, les processus attentionnels sont amorcés par des éléments visuels (objets, personnages…) présents dans la scène et qui attirent l'attention de façon automatique.

Par ailleurs, un effet particulièrement intéressant c'est la différence entre Phrase 2 Suite et Phrase 2 Non Suite tout au long de la présentation des stimuli. Elle est marquée par une positivité plus importante entre 200-400 ms avec un pic à 300 ms, ainsi que des effets plus tardifs vers 1000-1200 ms pour la condition Phrase 2 Non Suite.

Nos ERPs étant alignés sur le début de présentation de la phrase, ces composantes ne sont pas aussi nettement dissociables qu'avec une présentation mot à mot. Il faut préciser qu'initialement, nous souhaitions combiner des données Eye Tracker à ces analyses ERPs avec l'hypothèse que, dans la seconde phrase, le verbe est un élément essentiel de la simulation (cf. travaux comportementaux et de neuroimagerie qui ont montré des modulations spécifiques liés au verbe, voir Chapitre I). Le verbe apporte des éléments essentiels et permet très rapidement de savoir si une seconde phrase s'inscrit dans une cohérence contextuelle par rapport à une phrase présentée précédemment. Notre but était donc de recaler le début des ERPs sur la lecture du verbe pour chaque sujet. Nous avons volontairement construit nos stimuli dans cette optique, avec des secondes phrases où le verbe ne représente pas une violation sémantique en terme d'action possible ou non, mais où, dès la lecture du verbe de la seconde phrase, on pressent d'emblée que la phrase va être cohérente ou non avec la première. Ex : *« La femme sort une clé de son sac à main. Elle **ouvre** la porte de l'appartement. »* ; *« L'homme sort le linge de la machine à laver. Il **étend** le linge dans le jardin. » ; « L'homme et la femme se disputent dans la cuisine. Ils **dansent** sur la piste. » ; « La jeune femme achète un ticket de métro au guichet. Elle **sèche** ses cheveux avec une serviette. »* C'est aussi pour faciliter le traitement du verbe dans la seconde phrase que nous avons utilisé des phrases plus courtes avec seulement un pronom placé avant le verbe. Malheureusement, les acquisitions Eye tracker n'ayant pas fonctionné pour la totalité des sujets et des stimuli, nous n'avons que des données partielles qui ne nous permettent pas pour le moment de recaler précisément les ERPs sur la lecture du verbe. Cependant, nous ne rejetons pas l'idée que les effets observés dans nos données (effets précoces et plus tardifs) pour la Phrase 2 pourraient être liés au traitement du verbe. En effet,

dans la lecture normale (c'est-à-dire avec une présentation qui ne soit pas en bloc), nous sommes capables d'identifier dix à douze lettres par saccade et une présentation de 50 millisecondes est suffisante à la lecture. Par ailleurs, les petits mots grammaticaux comme les auxiliaires, les pronoms ou les articles ne sont parfois pas traités ou alors comme faisant partie d'un même groupe (par exemple : pronom + nom) (Dehaene, 2007). Dans notre protocole expérimental (voir Chapitre II.2), le pronom qui précède le verbe est traité très rapidement (environ 300 ms) ; de plus dans ce protocole, le traitement se fait mot à mot et les sujets doivent appuyer sur un bouton, ce qui rallonge le temps à cause de la production d'une réponse motrice. Bien que nous ne puissions l'affirmer, les effets que nous observons pourraient être liés spécifiquement à l'information issue du traitement du verbe. Si les effets précoces vers 300 ms font penser à une composante P3b (localisation et timing) qui correspondrait à des processus attentionnels impliqués dans le traitement de la nouveauté, les effets plus tardifs pourraient refléter des processus plus complexes d'intégration et de réévaluation sémantiques. Des effets tardifs ont été observés durant le traitement de phrases en modalités auditive ou visuelle ; en général ces effets sont associés avec une intégration complexe, à la fois sémantique et syntaxique. Münte et al. (1998) ont notamment montré des effets de ce type : des violations sémantiques, syntaxiques et orthographiques génèrent des déviations positives du signal ERP dans des fenêtres temporelles tardives (600-800, 800-1000 et 1000-1200 ms). Dans cette étude, les effets de positivité associés aux violations sémantiques étaient localisés principalement au niveau pariétal (ce qui pourrait correspondre à la localisation pariéto-temporale que nous observons). Patel et al. (1998) ont eux aussi observé des effets de positivité tardifs, vers 800-1100 ms, associés au traitement des incongruités structurelles dans des phrases plus ou moins complexes. Ces effets ont une large distribution frontale, centrale et pariétale. Ces données donnent une perspective d'interprétation de nos observations où l'on note des effets de positivité tardifs (800-1000 ms) lorsque le second stimulus est sémantiquement incongru. Cependant, nos résultats semblent montrer que la non-congruence est détectée assez tôt puisqu'on note une activité différente pour Phrase 2 Non Suite, entre 200 et 400 ms, avec un pic vers 300 ms qui serait donc un marqueur physiologique précoce de cette détection. Cette composante ressemble à une P300 et, comme nous l'avons déjà mentionné, les événements les moins probables génèrent les P3b les plus importantes (Donchin, 1981 ; Hillyard & Kutas 1983), ce qui pourrait être cohérent si les sujets ont déjà traité le verbe à ce moment.

Les effets plus tardifs de réévaluation sémantique pourraient aussi être spécifiquement reliés au traitement du verbe qui permet d'évaluer très rapidement la cohérence contextuelle. Ils sont

cohérents avec l'hypothèse et les résultats de Toshio Inui[7] (2010, 2012) qui montrent qu'en comparaison avec des pseudo-verbes, des verbes d'actions présentés pendant 300 ms génèrent des effets spécifiques vers 700-900 ms, caractérisés par une phase de synchronisation des ondes Theta (3.5 Hz) des électrodes frontales et pariétales (voir figure 76). Selon T.Inui, cette synchronisation correspondrait à la signature neurale de l'unification sémantique. Ici, les processus tardifs refléteraient plutôt une intégration définitive plus complexe lorsque le second stimulus n'est pas cohérent (information issue du traitement du verbe) ou des processus de réévaluation, de type « frame-shifting » (Coulson & Kutas, 2001; Coulson & Lovett, 2004; Coulson & Williams, 2005). Cependant, la version actuelle de ce protocole ne nous permet pas d'affirmer à quoi sont liés ces effets (traitement du verbe, traitement d'un autre mot, phénomène attentionnel, catégorisation de bas niveau).

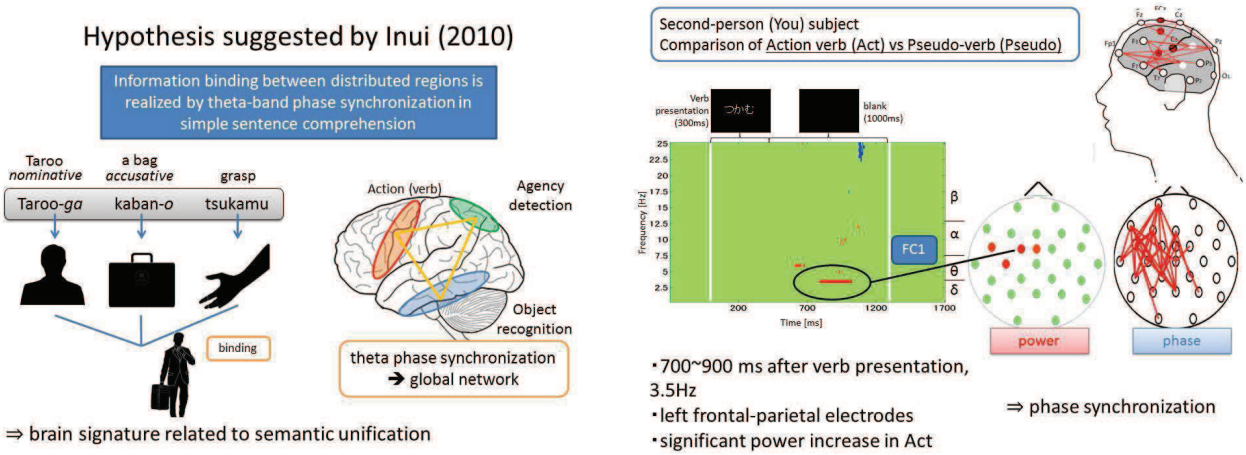


Fig 76) Hypothèse de T.Inui (2010) sur le traitement des verbes
Les verbes d'actions présentés génèreraient des effets spécifiques vers 700-900 ms, caractérisés par une phase de synchronisation des ondes Theta (3.5 Hz) entre les électrodes frontales et pariétales. C'est cette synchronisation qui serait responsable du regroupement de diverses informations (telles que la reconnaissance des objets ou détection de l'agentivité) et d'aboutir à une unification sémantique permettant la compréhension d'une phrase.

- Processus communs et processus modalité-spécifiques

On retrouve ces effets de positivité tardive dans une même fenêtre temporelle, à la fois pour des images ou des phrases. La source fronto-centrale de ces effets pourrait également être en partie commune (même si les phrases semblent aussi engager des composantes plus fortement latéralisées à droite, notamment avec une localisation temporo-pariétale). Rappelons que les seconds stimuli ont été construits de telle sorte que, tout seul, le second stimulus ne soit pas problématique en soi et les effets observés sont donc réellement liés à des traitements différentiels lorsque les stimuli ne sont pas cohérents avec un contexte sémantique créé par le premier stimulus. Ces effets tardifs présents

[7] Je remercie le Dr Inui de m'avoir autorisée à utiliser ses figures pour illustrer mon propos.

dans les deux modalités (voir Figure 77) suggèrent un mécanisme commun pour intervenir dans la construction/réévaluation sémantiques, indépendamment de la modalité. Il serait éventuellement susceptible d'engager le système fronto-temporo-pariétal mis en évidence en IRMf.

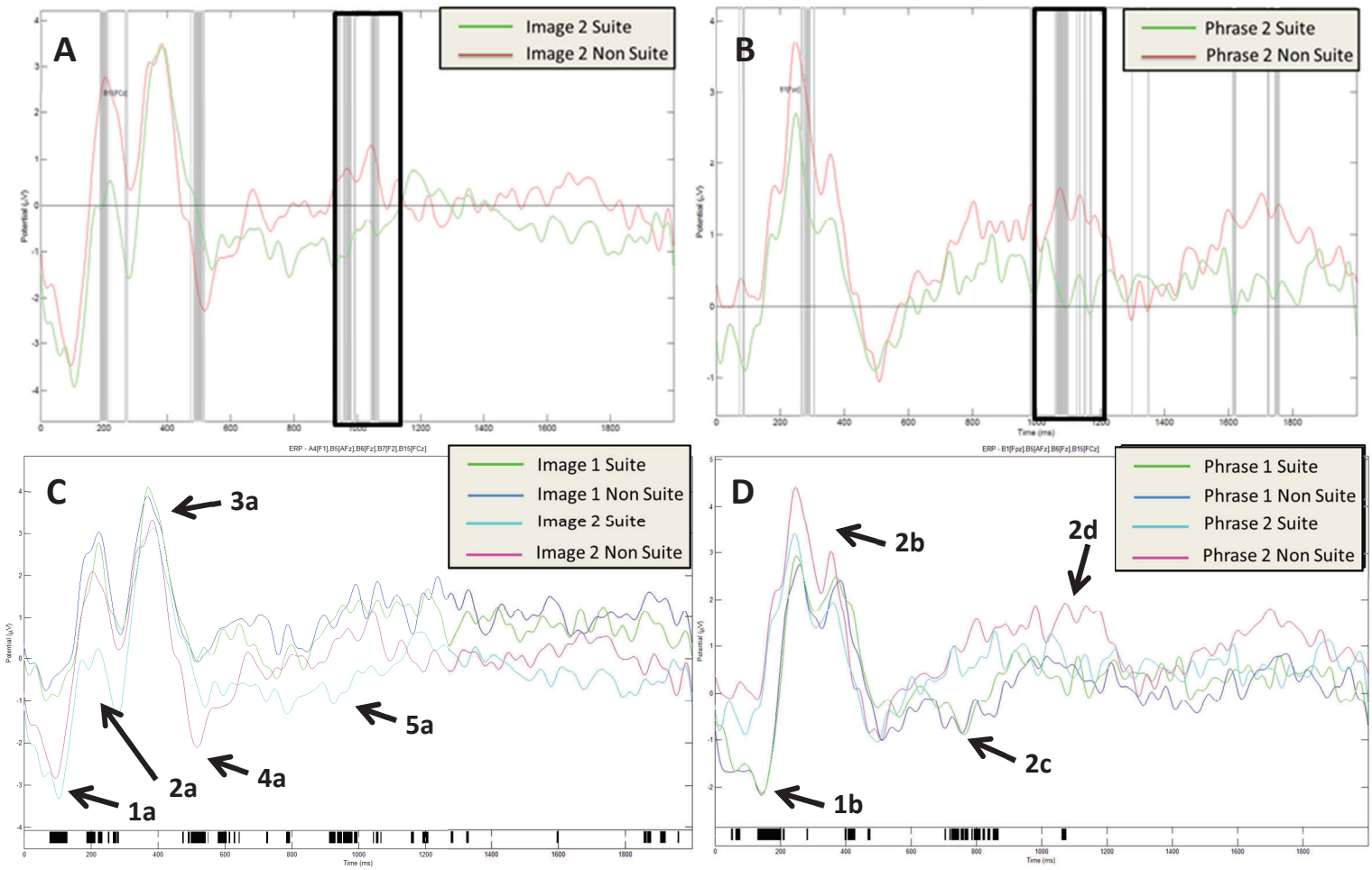


Fig 77) Synthèse des principaux résultats ERPs pour les conditions images et phrases

Les 2 figures du haut représentent la comparaison des ERPs pour les stimuli 2, pour images (A) et phrases (B) : on note des effets tardifs similaires entre les deux types de stimuli, caractérisés par une positivité tardive plus importante (vers 1000ms) pour les stimuli Non Suite qui pourraient traduire des effets d'intégration sémantique communs aux deux modalités.

Les 2 figures du bas représentent les principales composantes générées par les conditions Images (C) et Phrases (D). Pour les images : l'effet (1a) représente un traitement précoce différentiel entre Image 1 et Image 2, qui se caractérise par une composante N100 dont l'amplitude est plus grande pour les Images 2. Il pourrait correspondre à un engagement attentionnel plus important pour le second stimulus. L'effet (2a) se traduit par une amplitude réduite de la P200, suivi d'un un pic plus négatif vers 300ms, spécifiques pour la condition Image 2 Suite. Là encore, cela pourrait correspondre à des effets attentionnels pré-sémantiques puisque l'attention accrue des sujets tend à faire diminuer l'amplitude de la P200. L'effet (3a) correspondrait à un processus de mise à jour des informations contextuelles en mémoire de travail, permettrant leur stockage ultérieur de ces informations. L'effet (4a) représente une N400 qui est connue pour être impliquée dans les processus sémantiques et dans l'intégration de l'information contextuelle. Quant à l'effet (5a), il correspondrait, comme mentionné précédemment, à des effets plus tardifs d'intégration sémantique.

Pour les phrases : l'effet (1b) représenterait des stratégies attentionnelles liées à la composante N100. L'effet (2b) correspondrait à une P300 (P3b) et ferait plus spécifiquement appel à des processus attentionnels impliqués dans le traitement de la nouveauté. Les effets (2c) et (2d) représenteraient des processus plus tardifs d'intégration sémantique, avec un traitement spécifique pour les stimuli de la condition Phrase 2 Non Suite vers 1000ms.

Cependant, malgré ces effets plus tardifs qui pourraient être les marqueurs de processus communs sous-jacents, on note un certain nombre de différences précoces modalité-spécifiques qui font que même si ces processus plus tardifs ont une origine commune, ils reposent sur des informations issues de stratégies d'analyse perceptive différentes. Si certaines différences paraissent tout à fait normales et semblent reliées à des phénomènes perceptivo-attentionnels propres à chaque modalité, certaines différences de traitement notamment lorsque l'on compare les Stimuli 1 et les Stimuli 2 sont plus difficiles à appréhender, d'autant plus qu'il semble y avoir des effets inversés entre les modalités. Dans les images, c'est tout d'abord la condition Image 2 Suite qui est traitée différemment (entre 200-400 ms, mais aussi plus tardivement) alors que les 3 autres conditions semblent équivalentes pour les mêmes périodes, et pourtant l'effet N400 est spécifique à la condition Image 2 Non Suite car l'amplitude de la N400 est moins importante pour la condition Image 2 Suite (jusqu'ici traitée différemment des autres conditions), ce qui est cohérent avec l'interprétation selon laquelle cette réduction d'amplitude peut refléter une réduction des exigences de traitement sémantique (Wlotko & Federmeier, 2012). Lorsque l'on compare Image 2 Suite et Non Suite, la condition Non Suite semble marquée par une positivité tardive plus importante et on s'aperçoit en fait que cette positivité est comparable à celle de l'image 1 (Suite et Non Suite). En revanche, dans les phrases, c'est bien la condition Phrase 2 Non Suite qui a, la plupart du temps, un décours temporel différent par rapport à la condition Phrase 2 Suite mais aussi par rapport à Phrase 1 (Suite et Non Suite). On ne peut donc pas rejeter l'idée qu'il y aurait des stratégies perceptives, attentionnelles et/ou sémantiques différentes pour traiter les aspects de non-congruence contextuelle (Phrase 2 Non Suite, effet N400 pour image 2 Non Suite) mais aussi lorsqu'il y a un lien sémantique entre deux conditions (Image 2 Suite).

- <u>Cohérence contextuelle et traitement sémantique</u>

Contrairement à certains résultats qui montrent une absence de différences précoces de l'activité cérébrale associée aux effets de congruence (Ganis & Kutas, 2003), nous pensons que l'influence contextuelle n'est pas postérieure aux étapes d'analyses sémantiques. Si les effets précoces sont modalités-spécifiques et donc moins reliés à un mécanisme commun, comment les différences contextuelles sont-elles traitées de façon aussi précoce ? Il y a probablement des effets «top-down», où le système sémantique commun génère des prédictions sur ce qui doit apparaître et où les systèmes modalités-spécifiques réagissent dans les 200-300 premières ms sur la base de ces prédictions mettant en jeu des stratégies d'analyse assez différentes durant la compréhension contextuelle. Malgré toutes nos précautions dans la construction de cette expérience et la volonté d'utiliser une tâche impliquant un traitement sémantique assez profond (questions différentes à chaque fois pouvant porter sur le stimulus 1 et/ou 2), nous ne pouvons pas tout à fait rejeter l'idée

que les différences que nous observons dans notre protocole et que nous interprétons comme des stratégies attentionnelles différentes sous un contrôle top-down, pourraient en fait refléter un niveau de traitement différent entre phrases et images. Si dans la tâche de lecture on observe des effets réellement liés au traitement sémantique et à une intégration contextuelle plus compliquée lorsque la seconde phrase n'est pas cohérente ,dans les images il s'agit peut-être plus d'effets de catégorisation de bas niveau, basée sur une analyse perceptive des différences de la scène (personnages qui n'ont pas les mêmes vêtements donc couleurs différentes, décor différent, objets différents...). Cela ne veut pas dire qu'il n'y a pas de traitement sémantique durant l'analyse des images (ce que confirment les résultats comportementaux de la tâche qui montrent que les sujets ont analysé sémantiquement la scène puisqu'ils sont capables de répondre à des questions pointues, ainsi que la présence d'une N400 pour le traitement des images, composante spécifiquement associée au traitement sémantique) ; cela tendrait seulement à démontrer que les différences précoces ne seraient pas dues à des différences de traitement sémantique et d'intégration contextuelle. Cependant, même si ces effets ne sont pas liés à l'analyse sémantique elle-même, ils pourraient quand même refléter des étapes nécessaires à l'intégration sémantico-contextuelle. Comment le cerveau peut-il détecter la non-congruence aussi précocement et construire une représentation cohérente en conséquence ? Comment les connaissances sur les événements réels peuvent-elles influencer la compréhension des phrases et des scènes visuelles (Metusalem et al., 2012) ? Un certain nombre de questions restent en suspens et le timing ainsi que le rôle précis des composantes impliquées dans ce traitement contextuel restent à déterminer. Là encore, des analyses ou des tâches supplémentaires pour essayer de mieux comprendre le rôle fonctionnel de ces composantes précoces dans le traitement sémantique seront nécessaires.

Ces résultats ERPs, même s'ils ouvrent des pistes intéressantes, sont principalement exploratoires et s'inscrivent dans un projet plus vaste d'analyses, qu'ils demandent à être confirmés par des statistiques plus précises sur des fenêtres temporelles et des clusters, ainsi que par des analyses complémentaires (ERPs/Eye tracker et Temps fréquence). Nous espérons que ces différentes analyses nous éclaireront sur certains de ces effets liés à l'intégration contextuelle, notamment sur le rôle de la N400 comme marqueur de l'intégration sémantique mais aussi sur l'effet de P300 pour les phrases : une distinction contextuelle peut-elle être aussi précoce dans le cadre d'une phrase ? Si oui, est-elle le marqueur d'un traitement sémantique ou un effet de catégorisation perceptive de plus bas niveau ? A quel élément de la phrase est-elle liée (verbe) ? Cet effet est-il équivalent, différent ou complémentaire, pour traiter la cohérence contextuelle, de l'effet de N400 démontré ici pour les images et que l'on trouve généralement dans le langage ?

II.3.*6*) Conclusion et perspectives

Si nos données mettent en évidence des résultats préliminaires assez intéressants et qui nous permettent d'affirmer que la compréhension implique des processus différents (attentionnels mais aussi d'évaluation sémantique) selon que les stimuli sont intégrés dans un contexte plus vaste de compréhension, un certain nombre d'analyses complémentaires et d'améliorations sont à apporter pour avoir une vue plus large des processus temporels de la compréhension sémantique, notamment en ce qui concerne le langage où des données supplémentaires sont nécessaires pour pouvoir identifier les composantes impliquées dans la compréhension sémantique du langage et les comparer à celles qui sont recrutées pour le traitement des scènes visuelles.

- **Analyses FRPs (fixation related potentials)**

Comme nous l'avons déjà mentionné, concernant les stimuli langagiers, notre but initial était de combiner des données Eye tracker à nos données EEG afin de recaler précisément les ERPs sur la lecture de chaque mot d'une phrase et de pouvoir ainsi obtenir les composantes spécifiquement associées à chaque mot (en particulier au verbe de la seconde phrase). A nos yeux, corréler le décours temporel tout au long de la lecture des phrases avec des données Eye tracker reste une approche essentielle pour comprendre quand et comment est réalisée l'intégration sémantique au sein d'une phrase.

La lecture correspond à une alternance de fixations (d'une durée de 150 à 300 ms) et de brèves saccades (30 ms), le traitement de l'information étant majoritairement limité aux périodes de fixations (Kliegl et al., 2006). La lecture est contrainte par une zone appelée empan de perception visuelle des lettres, liée à l'étroitesse de la région fovéale qui permet de capter les lettres avec suffisamment de précision pour permettre de les reconnaître. Nous n'identifions que dix à douze lettres par fixation et de façon asymétrique avec un biais en faveur de ce qui se trouve à la droite du centre du regard (3 ou 4 lettres à gauche et 7 à 8 lettres à droite). Au-delà de cette zone de perception visuelle, nous ne sommes plus sensibles à l'identité des lettres, mais seulement à la présence des espaces qui délimitent le mot suivant car ils nous fournissent des indices de longueur et nous permettent de préparer la prochaine saccade afin que le regard atterrisse à proximité du centre du prochain mot (Dehaene, 2007). Par ailleurs, il suffit de 50 millisecondes de présentation pour que la lecture progresse quasiment normalement, un bon lecteur lisant environ 250 à 500 mots par minute. Dans la lecture normale, il y a souvent des mots qui sont ignorés ou traités plus rapidement (auxiliaires, pronoms ou articles). Mais la durée de fixation d'un mot est aussi dépendante des informations disponibles, elle est par exemple réduite en fonction de la fréquence des mots ou de leur prévisibilité issue des mots précédemment présentés au sein d'une phrase. Cependant, les

processus perceptifs et attentionnels qui guident les mouvements des yeux tout au long de la lecture d'une phrase restent mal connus : dans quelle mesure le comportement du mouvement oculaire est affecté par des facteurs oculomoteurs de bas niveau par rapport aux processus cognitifs d'ordre supérieur ? Quelle est la quantité d'information qui est extraite de la partie située à droite du centre du regard ? Les lecteurs traitent-ils les informations issues de plusieurs mots en même temps de façon parallèle ? (Kliegl et al., 2006 ; Starr & Rayner, 2001).

Depuis les années 80, un grand nombre de travaux basés sur des données oculomotrices ont tenté de répondre à ces différentes interrogations. Deux conceptions dominaient alors le débat (Just & Carpenter, 1980), conceptions selon lesquelles un traitement distribué de mots, à travers les fixations des différents mots d'une phrase, est impossible. D'après « the immediacy-of-processing assumption », l'interprétation se fait mot à mot, sans attendre que les ambiguïtés visuelles, lexicales ou sémantiques soient résolues, au risque de choisir un sens faux. Selon « the eye-mind assumption», le lecteur maintient sa fixation sur le mot jusqu'à ce que celui-ci soit traité entièrement, la durée de fixation du regard étant un élément révélateur du temps de traitement cognitif nécessaire et qui correspondrait à la somme de toutes les fixations sur un mot lors de la première rencontre avec celui-ci. La complexité inhérente au processus de lecture a rapidement révélé de sérieuses limites à ces hypothèses et d'autres modèles ont vu le jour, tels que « the process-monitoring assumption » (Morrison, 1984 ; Reichle, Pollatsek, & Rayner, 2003) qui restreint « the eye-mind assumption » à un processus permettant l'accès lexical et la programmation des saccades oculaires. L'idée de base de cette hypothèse est que, lors de la lecture, nous ne traitons qu'un mot à la fois dans notre lexique mental, ce processus d'accès au lexique mental étant régi par l'esprit qui programme les mouvements oculaires suivants permettant de compléter de l'intégration sémantique à travers la lecture des mots suivants. Même si cette hypothèse suppose l'existence d'un lien fondamental entre processus oculomoteurs et langagiers, elle conserve une vision strictement séquentielle du traitement des mots où les changements attentionnels sont contingents à l'accès lexical. Ce modèle permet toutefois une forme limitée de traitement distribué de mots à travers les fixations oculaires car il suppose que la programmation des saccades est déclenchée par une étape initiale d'accès lexical et que, par conséquent, les yeux peuvent se déplacer avant ou après que l'étape d'accès au lexique soit complète. Mais le traitement distribué dans cette approche reste limité car le début des saccades est lié à un traitement lexical sériel, cette hypothèse impliquant notamment que la durée de fixation peut être influencée par la difficulté inhérente au traitement du mot précédent et non par la difficulté du mot suivant (Kliegl et al., 2006).

D'autres modèles (Kolers, 1976 ; Bouma & DeVoogd, 1974) présentent une vision radicalement différente. Selon ces modèles, la durée de fixation (entre 150 et 300 ms) est trop courte pour que la

compréhension soit effectuée uniquement durant cette période. Selon Kolers et ses collègues, l'irrégularité des mouvements oculaires suggère que les yeux se déplacent en grande partie à leur propre rythme et que les processus de contrôle sur les mouvements oculaires n'interviennent que lorsque c'est nécessaire car l'esprit serait capable de traiter l'ordre sériel des mots ultérieurement et indépendamment de l'ordre réel dans lequel les mots ont été fixés par le regard. Si la proposition de Kolers est intéressante car elle met en évidence le fait que les processus de compréhension ne peuvent s'effectuer uniquement durant le temps très court de fixation des mots, elle mérite d'être réévaluée sérieusement (Kliegl et al. 2006). En effet, le traitement d'un mot semble plutôt se prolonger durant les deux ou trois fixations suivantes, ce qui serait plus en accord avec une vision distribuée du traitement lexical, formalisée par « the cognitive lag hypothesis » (Rayner, 1977, 1978), hypothèse selon laquelle le traitement linguistique d'un mot continue à s'effectuer, même lorsque les yeux se sont déplacés sur le mot suivant. De plus, les durées de fixation sont influencées à la fois par différents niveaux de traitement lexical, visuel et oculomoteur. Pour Kliegl et al. (2006), l'analyse de nombreux travaux sur le temps de lecture, ainsi que leurs propres résultats, montrent que la plupart du temps, plusieurs mots sont traités en parallèle à différents niveaux cognitifs et perceptuels grâce à des effets dynamiques de perception et de modulation attentionnelle qui permettraient de guider le regard humain durant la perception et l'interprétation en temps réel de la scène décrite par les mots. Tout d'abord, si le traitement fovéal est essentiel, on a aussi démontré l'existence de prétraitements parafovéaux susceptibles d'influencer la durée et la position d'une fixation actuelle et l'empan perceptif définirait une région d'accès qui va largement au-delà des limites d'un mot car elle concernerait le traitement des mots en vision fovéale mais aussi les prétraitements parafovéaux. La récupération en mémoire amorcée par le contexte d'une phrase pourrait aussi retarder les saccades oculaires. Par ailleurs, le traitement perceptivo-lexical et le contrôle oculomoteur ne se synchroniseraient qu'occasionnellement pour ne pas perdre le fil de la compréhension ou pour corriger une erreur oculomotrice (mot sauté ou mal lu). Ces différents aspects suggèrent que, durant la lecture, le cerveau a un temps d'avance sur le mouvement des yeux et qu'il est capable de diriger la prochaine saccade vers une cible, via des traitements parafovéaux qui permettent de faire des prédictions sur le mot à venir en utilisant l'information contextuelle disponible dans la phrase en tant qu'indice de récupération. Au même moment, le cerveau agirait sur le contrôle du mouvement des yeux pour achever la reconnaissance des mots, l'analyse syntaxique et établir la cohérence sémantique. Tous ces éléments constitueraient des preuves de la nature distribuée des traitements lexical, visuel et oculomoteur qui interviennent sur les fixations oculaires pendant la lecture. Pour Kliegl et al. (2006), le challenge futur sera d'utiliser le potentiel de ces informations sur les mouvements oculaires comme principaux indicateurs des dynamiques comportementales et cognitives impliquées dans la compréhension d'une phrase.

C'est précisément ce que nous aimerions faire en combinant données Eye Tracker et ERPs durant une tâche de lecture on line, car si les protocoles de lecture mot à mot (SVP, serial visual presentation : un seul mot au centre d'un écran pendant plusieurs centaines de ms, suivi d'un délai avant la présentation du mot suivant) permettent d'avoir des composantes claires et délimitées, ils représentent une simplification de la façon dont nous lisons et ne permettent pas d'accéder réellement à une compréhension du fonctionnement cérébral lors de la lecture naturelle d'une phrase. Contrairement à une présentation SVP, la lecture normale a un timing beaucoup plus rapide, auquel s'ajoute le prétraitement des mots suivants en vision parafovéale. De plus, les mots ne sont pas toujours fixés centralement (comme c'est le cas en SVP), mais sont souvent traités à partir de positions de visualisation non-optimales à la limite des mots (Nuthmann, Engbert, & Kliegl, 2005).

Si relier des données EEG à des données oculaires n'est pas une méthode nouvelle puisqu'on utilise depuis longtemps des potentiels moyennés sur les onsets de fixations ou sur les saccades (fixation related potentials FRPs et saccade-related potentials) à partir des données obtenues via les électrodes enregistrant le signal oculaire (EOG), ces données manquent de précision sans un enregistrement Eye-tracker concomitant (voir Dimigen et al. 2011, pour plus de références). En combinant Eye tracker et EEG, nous aimerions savoir si, lors de la lecture continue d'une phrase, tous les mots génèrent des composantes spécifiques similaires (notamment la N400). Dans notre expérience comportementale (chapitre II.2), nous avons pu constater que les sujets mettaient environ 1800 ms pour lire une phrase du type des phrases utilisées pour le stimulus 2 (phrases de 4 à 8 mots) et 391,65 ms en moyenne pour lire un mot (avec bien sûr des différences en fonction du nombre de lettres, du type de mot et de son positionnement dans la phrase). Compte tenu du fait que notre protocole comportemental avait un mode de présentation mot par mot et que les sujets devaient appuyer sur un bouton pour faire défiler chaque mot, il est difficilement envisageable que chaque mot soit complètement traité durant ce court laps de temps grâce à l'intervention des composantes sémantiques de type N400 et autres composantes plus tardives. Nos résultats comportementaux sont plus en accord avec l'existence d'un traitement sémantique distribué où plusieurs mots seraient traités en parallèle pour permettre de faire émerger une représentation cohérente. Cette vision n'est pas incompatible, au contraire, avec l'idée que le verbe dans la seconde phrase pourrait avoir un rôle essentiel dans la compréhension et l'intégration sémantique car il permet d'apporter des éléments forts qui lèvent l'ambiguïté concernant la cohérence contextuelle. Cependant, un certain nombre de travaux ont montré que le traitement du verbe, supposé apporter des informations sémantiques fortes, ne module pas nécessairement l'amplitude de la N400 (Kuperberg, 2007). Il est donc nécessaire de retravailler nos données pour pouvoir démontrer plus précisément comment peuvent interagir les différentes composantes cérébrales, marqueurs de

l'intégration sémantique, dans la construction d'une représentation (notamment celles qui sont associées au traitement du verbe). Par ailleurs, nous pourrions obtenir également des éléments de réponse plus pertinents sur l'implication d'un système sémantique amodal dans l'intégration contextuelle et sur le rôle de la composante N400. En effet, si une seule N400 associée à une scène visuelle permet d'identifier un lien sémantico-contextuel entre deux stimuli, on pourrait penser qu'il y a un indéniable avantage temporel pour le traitement des images : la perte d'ambigüité du contexte et la construction d'une représentation sémantique cohérente seraient alors plus rapides pour les images que pour le langage où chacun des mots d'une phrase devrait subir des étapes de traitement équivalentes avant de pouvoir être intégrés en une représentation sémantique cohérente. Seule l'idée que plusieurs mots sont traités en parallèle permet de réconcilier le fait que la construction d'une représentation soit aussi rapide dans le cadre d'une phrase que dans le cas d'une image et pourrait impliquer un système sémantique fronto-pariétal équivalent. Le fait de supposer l'existence d'un traitement parallèle dans la lecture et la compréhension des phrases soulève cependant un certain nombre de questions : les sources de ce traitement sont-elles multiples ? Ce traitement recrute-t-il des aires différentes et/ou supplémentaires ? Les composantes associées au traitement de chaque mot ne risquent-elles pas de se confondre ou de s'interpoler ? Certains mots sont-ils particulièrement importants lors de ce traitement distribué ?

Si plusieurs auteurs reconnaissent que combiner données oculomotrices et EEG représente une méthodologie avantageuse pour étudier le langage (Sereno et al, 1998 ; Sereno & Rayner, 2003 ; Dimigen et al., 2011), à notre connaissance peu d'études se sont penchées sur une association eyetracker et ERPs pour mieux caractériser les processus temporels impliqués dans la compréhension sémantique d'une phrase. Sereno et al. (1998) ont par exemple observé que des aspects lexicaux (mots vs pseudomots), de fréquence (haute vs basse) et de régularités (correspondances orthographe-son exceptionnelles vs régulières) généraient des différences très précoces dans le décours temporel associé au mot traité, différences précoces cohérentes avec les données oculomotrices qui montrent des durées de fixation très courtes et avec l'idée que l'accès aux connaissances lexicales doit être initié dans les 100-150 premières ms. Mais cette étude porte sur des mots présentés de façon isolée et non dans le contexte d'une phrase. Camblin et al. (2007) ont montré que les violations de la congruence du discours (induites par un mot spécifique) ont des effets précoces et persistants à la fois sur les ERPs et les mesures d'Eye tracker durant la lecture d'une phrase, ce qui indique que les représentations discursives ont des effets rapides qui influencent le traitement lexico-sémantique. Cependant, les auteurs ont réalisé les acquisitions Eye tracker et EEG séparément et n'ont pas cherché à savoir comment ces deux mesures pourraient s'influencer mutuellement. A notre connaissance, l'étude de Dimigen et al. (2011) est, pour le

moment, la seule à avoir utilisé directement l'enregistrement de mouvements des yeux par Eye-Tracker combinés à des données EEG pour obtenir des FRPs (fixation-brain related potentials) durant la lecture naturelle de phrases. Avec cette expérience, les auteurs ont répliqué des effets solides liés à la prédictibilité des mots concernant la N400 mais ils ont aussi mis en évidence des effets similaires sur les mouvements des yeux, suggérant que ces mesures sont toutes deux influencées par un processus commun sous-jacent. Leur protocole démontre l'intérêt et la faisabilité de combiner EyeTracker et données EEG, même si les auteurs soulignent un certain nombre de contraintes expérimentales dont il faut tenir compte pour analyser ce type de données (voir Dimigen et al., 2011).

Nous aimerions reproduire les mêmes effets que Dimigen et al. (2011) et pouvoir les comparer à ceux engagés dans le traitement des scènes visuelles. En effet, là aussi, la combinaison de l'Eye tracker et de l'EEG pourrait apporter des informations considérables. Les stimuli visuels ont largement démontré qu'ils pouvaient être utilisés pour générer des ERPs dont les différentes composantes peuvent être reliées à des processus attentionnels et/ou sémantiques (ce que nous avons nous-mêmes observé dans nos résultats). L'Eye tracker a aussi beaucoup été utilisé pour évaluer quantitativement ce que les sujets analysent dans une scène et faire des hypothèses sur les mécanismes sous-jacents, notamment l'interaction des processus top-down et bottom-up (Smilek et al., 2006 ; Riby & Hancock, 2008, 2009 ; Carmi & Itti, 2006 ; Malcom & Henderson, 2010 ; Geisler, 2008). A notre connaissance, là-aussi, peu d'études ont cherché à combiner les deux méthodologies, étant donné que la présentation des stimuli visuels génère très facilement des composantes ERPs distinctes. Il serait pourtant intéressant d'étudier de façon plus approfondie le timing de cette analyse visuelle pour voir si certaines composantes perceptivo-attentionnelles et/ou sémantique sont liées plus spécifiquement à l'identification visuelle de certains éléments de la scène. Comparer ces données combinées pour des images et pour des phrases devrait nous apporter des éléments de réponse plus précis sur l'implication d'un système sémantique commun dans le traitement de ces deux types de stimuli.

Qui plus est, les données Eye tracker elles-mêmes pourraient déboucher sur des pistes intéressantes via des analyses quantitatives et qualitatives des mouvements oculaires, à la fois pour les images et les phrases. Est-ce que les sujets ont des mouvements oculaires différents sur l'ensemble de la phrase (ou sur certains mots) lorsque celle-ci n'est pas contextuellement cohérente ? De même, pour les images, il serait intéressant d'analyser ce que les sujets regardent en premier (les personnages, les objets, le décor…) et s'il y a un traitement différentiel lorsque deux stimuli sont contextuellement reliés ou non. Est-ce que les sujets passent moins de temps à explorer la scène de la seconde image lorsqu'elle est cohérente car elle partage un certain nombre d'invariants avec la première, ce qui par

exemple pourrait expliquer le traitement attentionnel différent au niveau de la P200 pour Image 2 Suite. En plaçant des régions d'intérêt a priori sur les éléments qui composent la scène (personnages, des objets et du reste de la scène), nous pourrions savoir s'il y a des différences liées aux conditions Suite et Non Suite dans l'exploration de ces éléments, différences qui pourraient être mises en relation avec nos analyses ERPs, notamment les différences attentionnelles et sémantiques que nous avons pu observer mais dont les rôles spécifiques restent incertains. Ces données comportementales semblent essentielles pour mieux comprendre les processus d'intégration et de réévaluation sémantiques dans chaque modalité. Une analyse des sujets pour lesquels les données Eye tracker ont été enregistrées correctement devrait déjà nous donner des éléments de réponses, que viendront compléter des acquisitions complémentaires.

- **Analyses temps fréquences : mise en évidence des processus temporels amodaux**

Rappelons que notre hypothèse première, associée à ce protocole, concernait la similarité des processus temporels impliqués dans la compréhension des phrases et des images : nous souhaitions mettre en évidence des marqueurs temporels qui pourraient présenter des similarités, témoignant ainsi de processus sous-jacents similaires et soutenant l'hypothèse qu'un même réseau fronto-temporo-pariétal serait impliqué de façon amodal.

Si la méthodologie ERPs était particulièrement intéressante pour analyser des différences liées à l'intégration contextuelle entre stimuli verbaux et non-verbaux, elle est moins appropriée pour établir les processus communs qui pourraient être engagés tout au long du traitement de ces deux types de stimuli. Dans cette optique, nous envisageons de compléter nos résultats par des analyses temps-fréquence grâce auxquelles nous espérons mettre en évidence des marqueurs fréquentiels similaires, impliqués dans la compréhension des phrases et des images. Elles devraient permettre de mettre en évidence des marqueurs communs tout au long de l'analyse d'un stimulus mais aussi des moments clés dans le décours temporel où l'intégration sémantique se ferait de façon similaire pour les deux conditions, pré-supposant l'implication d'un système sémantique commun (nous détaillerons davantage les analyses envisagées dans la conclusion de ce chapitre).

De nombreux travaux ont suggéré l'implication des bandes de fréquence gamma (30-60 Hz) dans la compréhension sémantique (Bastiaansen et Hagoort, 2006 ; Weiss et al., 2003 ; Van Berkum et al., 2003 ; Braeutigam et al. 2001). On trouve une augmentation de l'activité gamma associée à de nombreux processus tels que l'organisation perceptuelle (Fries et al., 1997 ; Tallon-Baudry et al., 1996, 1997; Keil et al., 1999), l'attention (Fries et al., 2001 ; Gruber et al., 1999 ; Taylor et al., 2005) et des processus mnésiques de reconnaissance, de mémoire associative et de mémoire de travail (Fell et al., 2001 ; Miltner et al., 1999; Gruber et al., 2008, Tallon-Baudry et al., 1998; Lutzenberger et al.,

2002; Kaiser et al., 2003). Une interprétation proposée stipule que cette activité pourrait marquer l'intervention de processus de « binding », permettant l'assemblage de différentes informations sensorimotrices et la création d'une représentation cohérente (Gray et al., 1989).

Un certain nombre de données montrent que les bandes de fréquence Theta, autour de 3-7 Hz, pourraient être également impliquées dans le traitement sémantique (Hald et al., 2006 ; Bastiaansen et al. 2005 ; Bastiaansen et al. 2002, 2008). Ainsi, Hald et al., 2006 ont observé que, dans une tâche de violation sémantique, il y avait une augmentation de l'activité Theta durant un intervalle de 300-800 ms après la présentation d'un mot cible alors que l'activité gamma était plus fortement associée au traitement des phrases sémantiquement correctes. Pour Bastiaansen et al. (2005), l'augmentation de l'activité théta temporale gauche refléterait l'activation d'un réseau impliqué dans la récupération des propriétés lexico-sémantiques des mots de type « open class words » (noms, verbes, …). Pour Inui et al. (2010, 2012), la synchronisation des oscillations theta issues des aires frontales et temporo-pariétales serait le marqueur de l'intégration sémantique durant la compréhension d'une phrase.

L'activité dynamique oscillatoire et interactive de ces différentes composantes serait un élément déterminant dans la compréhension du langage. Pour Bastiaansen et Hagoort (2006), les opérations de récupération en mémoire seraient plutôt associées à une augmentation de la synchronisation neuronale dans les bandes de fréquence alpha et theta alors que les opérations d'unification, notamment sémantique mais aussi syntaxique, impliqueraient plus fortement une synchronisation des betas hautes fréquences (12-30 Hz) et des gammas moyens (autour de 30 Hz) ; ces derniers seraient particulièrement reliés à l'unification sémantique. Pour Wang et al. (2012), l'augmentation dans le spectre du gamma serait à mettre en relation avec la prévisibilité d'un mot à venir en fonction du contexte de la phrase plutôt qu'avec l'intégration sémantique de ce nouveau mot dans le contexte précédent ; par ailleurs la synchronisation dans les bandes théta serait associée aux détections des erreurs dans le langage.

D'autres travaux ont aussi montré l'implication de ces composantes pour la compréhension de stimuli visuels (de Borst et al., 2012 ; Tallon-Baudry et al., 1998; 1997 ; Tallon-Baudry & Bertrand, 1999). Pour Tallon-Baudry & Bertrand (1999), un type particulier d'activité gamma (les gammas induits qui se caractérisent par une latence plus importante) pourrait constituer la signature neurologique de l'identification des objets. En effet, nous traitons les objets dans leur ensemble, comme un tout. Cependant, certaines propriétés des objets sont codées et traitées dans différentes régions du cerveau. Comment et quand peuvent alors émerger les représentations cohérentes ? Ce serait la synchronisation rythmique des décharges neuronales associées aux bandes gamma (autour

de 40 Hz) qui pourrait fournir le lien spatial et temporel nécessaire à l'unification des traitements issus de différentes zones du cerveau pour construire une représentation cohérente, et ce, que l'objet soit perçu, imaginé ou rappelé à partir de la mémoire. De Borst et al., (2012) ont combiné IRMf et données temps-fréquence et pour ces auteurs les régions frontales seraient responsables de l'intégration des informations « what » et « where » lorsque l'on doit se représenter une scène. La synchronisation précoce des oscillations theta et beta serait associée aux activations en IRMf des cortex frontaux et refléterait des processus de récupération et d'intégration de l'information. Les régions frontales seraient le « chef d'orchestre » des régions occpito-temporales encodant les propriétés détaillées des représentations associées à des objets alors que les régions pariétales seraient responsables des informations « où » c'est-à-dire de l'encodage de l'organisation spatiale et de la formation d'une représentation mentale cohérente de la situation.

Par ailleurs, pour faire plus amplement le lien avec les travaux sur la cognition incarnée et la simulation sensorimotrice provoquée par l'observation d'actions et la lecture de phrases d'actions, nous nous intéresserons également aux rythmes Mu (fréquence alpha 9-13 Hz). Une des toutes premières validations de l'existence d'un système de neurones miroirs chez l'humain a été démontrée dans les années 50 par Gastaut et ses collègues (Cohen-Seat et al. 1954, Gastaut & Bert 1954), qui ont observé en EEG une désynchronisation des rythmes Mu à la fois lorsque les sujets produisent des mouvements et quand ils observent des mouvements effectués par les autres. Cette découverte a été confirmée par plusieurs expériences en EEG et en MEG qui ont aussi montré que la désynchronisation observée semble provenir du cortex situé au niveau du sillon central et refléter une modulation des neurones sensorimoteurs par les neurones miroirs du cortex prémoteur (Cochin et al. 1998, 1999 ; Altschuler et al. 1998, 2000 ; Hari et al. 1998 ; Muthukumaraswamy et al. 2004 ; Oberman et al., 2005; Pineda, 2005). La désynchronisation des rythmes Mu est aussi rencontrée dans des tâches qui requièrent une analyse d'aspects sociaux tels que l'évaluation des émotions, ce qui laisse supposer que ces rythmes seraient aussi des marqueurs de processus cognitifs de haut niveau impliqués dans la théorie de l'esprit (Pineda & Hecht, 2008 ; Altschuler et al., 2000). S'il est maintenant reconnu que les rythmes mu auraient un rôle fonctionnel important qui relie perception et action, spécifiquement dans la transformation de ce qui est vu ou entendu en termes d'actions (transformation of ''seeing'' and ''hearing'' into ''doing.'', Pineda, 2005), peu de travaux se sont intéressés à l'activité de ces rythmes Mu dans le cadre du langage (van Elk et al., 2010 ; Bastiaansen & Hagoort, 2006) et encore moins à la comparaison directe d'images et de phrases d'actions. Nous essaierons donc de caractériser l'activité des rythmes Mu, à la fois pour les images et pour les phrases, et de comprendre comment ce marqueur peut intervenir avec les autres marqueurs

fréquentiels dans la compréhension, les différentes données existantes ne permettant pas d'avoir une vision unifiée du rôle de ces différentes composantes et de leurs interactions.

Comme pour nos analyses ERPs, nous nous intéresserons aux différences de patterns fréquentiels liés à la cohérence et à la contrainte sémantique contextuelle (Stimuli Suite/ Non Suite). Nous avons déjà obtenu quelques résultats préliminaires qui montrent une activité gamma induite plus importante pour les Stimuli 2 Non Suite, images ou phrases, témoignant peut-être d'une intégration sémantique plus difficile lorsque le second stimulus n'est pas cohérent avec le premier (des analyses plus spécifiques restent à effectuer). Mais contrairement aux analyses ERPs où nous cherchions principalement à mettre en évidence des différences entre les conditions, grâce aux analyses temps-fréquence, nous nous intéresserons particulièrement aux processus de construction des représentations qui seraient communs aux images et aux phrases, en comparant l'activité durant la présentation du stimulus 1 pour ces deux types de stimuli. Comme la catégorie Stimulus 1 contient 120 images et 120 phrases, nous espérons que ce nombre important de stimuli sera suffisant pour faire apparaître ces processus. Une comparaison des résultats ERPs et temps-fréquence serait également pertinente, plusieurs travaux ayant montré qu'il pouvait exister une relation inverse entre ces deux types de marqueurs de l'activité cérébrale, notamment dans la compréhension du langage où des violations sémantiques produisent des effets N400 inversement corrélés avec l'amplitude de l'activité oscillatoire des bandes de fréquence thêta ou encore une relation seulement partielle entre les marqueurs sémantiques ERPs et l'activité gamma (Davidson & Indefrey, 2007 ; Wang et al., 2012).

- **Traitements sémantiques vs non sémantiques**

Bien qu'ils n'aient pas été décrits dans la partie Matériel et Méthodes pour ne pas alourdir la description du protocole, nous avons également utilisé des stimuli contrôles dans cette expérience. Assez similaires à ceux utilisés en IRMf, les contrôles images sont des scrambles issus des images d'origine et pour les contrôles phrases, nous avons utilisé des suites de voyelles ou de consonnes. Analyser ces contrôles et comparer les activités (Temps-fréquences et ERPs) qui leur sont associées à ceux des stimuli qui ont du sens, nous aidera probablement à mieux caractériser le rôle de certaines composantes (notamment les composantes précoces de type P200, N200/N300 et P300) et cela nous permettra de distinguer celles qui sont liées à des processus sémantiques ou à des processus perceptifs et attentionnels de plus bas niveau. Un premier résultat issu des analyses préliminaires de ces contrôles semble montrer que la variabilité, associée à des composantes plus tardives (autour et au-delà de 1000 ms) pour phrases et images, n'est pas présente pour les stimuli contrôles, ce qui laisse penser que ces effets seraient bien liés à un traitement sémantique.

- **Analyses EEG et profils de représentateurs**

Il serait également très intéressant d'analyser nos résultats (ERP, temps-fréquence, Eye tracker) en fonction des profils de représentateurs des sujets, profils que nous avons réunis pour la plupart des sujets car ils ont aussi passé l'expérience comportementale (facteur COR = coefficient de représentabilité). L'analyse comportementale (Chapitre II.2) a montré que le profil des sujets, bons mauvais ou représentateurs intermédiaires, ne s'expliquait pas par des différences dans les temps de lecture. Les données IRMf (Chapitre II.1) n'ont pas permis de révéler clairement des régions distinctes ou supplémentaires qui seraient activées chez l'un ou l'autre groupe de sujets. C'est donc peut-être au niveau temporel que nous pourrions mettre en évidence des différences liées à la capacité à se représenter les phrases, les ERPs étant une méthode particulièrement appropriée pour tester les différences entre deux populations, en terme d'amplitude et de latence. Federmeier et al. (2002) ont par exemple montré une amplitude de la N400 significativement réduite pour les personnes âgées par rapport à des adultes plus jeunes, dans une tâche où des mots inattendus violent fortement les attentes induites par le contexte d'une phrase : il semble que les personnes âgées utilisent moins l'information fournie par le contexte prédictif. Dans une autre tâche sémantique, Ferdermeier et ses collègues (Federmeier et al. 2003) ont relevé plusieurs effets sur la latence de différentes composantes en comparant des sujets jeunes et des sujets plus âgés, confirmant que l'âge peut produire des changements sur des processus langagiers de haut niveau. Si l'âge peut être un facteur, il est tout à fait envisageable que nous puissions mettre en évidence des différences liéés aux capacités de représentabilité propres aux sujets qui, comme nous l'avons vu précédemment, pourraient être un élément important de l'organisation et de l'accès aux représentations langagières. Le COR significativement plus important chez les bons représentateurs pourrait être le reflet d'une meilleure synchronisation entre les aires cérébrales impliquées dans le traitement sémantique et/ou d'une communication plus rapide. Il serait donc tout aussi intéressant d'étudier les réponses face aux stimuli images en tenant compte du facteur COR, car si nous avons mis en évidence le COR grâce à des stimuli langagiers, on peut imaginer retrouver des effets similaires dans les images : peut-être que les personnes qui ont un meilleur COR ont aussi des facilités d'accès au sens à partir d'une scène visuelle ou au contraire, peut-être qu'il y a une dissociation en fonction de la modalité et que les personnes avec un bon COR pour le langage auront plus de difficultés pour la modalité image.

- **EEG intra-cortical**

Enfin, ce protocole a été légèrement modifié (même principe, mêmes stimuli mais tâche contrôle différente + ajout d'une condition en modalité auditive) pour être utilisé en EEG intra-cortical en vue d'une collaboration avec JP Lachaux et l'hôpital neurologique de Grenoble. Lachaux et ses collègues (Vidal et al., 2012) ont été parmi les premiers à réaliser des enregistrements directes pendant la lecture naturelle de phrases et à montrer des modulations des bandes gamma haute fréquence (40-150 Hz) associées. Ces auteurs ont utilisé une tâche simple visant à contraster processus sémantiques (lecture d'une phrase avec du sens « Le roi a quitté son château »), phonologiques (groupes de lettres formant un mot à prononcer mentalement « PPP HHH OOO NNN EEE ») et purement visuels (caractères alpha-numériques « HDF3Z NGTD SZY2G »). L'analyse de ces conditions a révélé différents résultats concernant le réseau cérébral impliqué dans la lecture, notamment que toute lecture de phrase se caractérise par une augmentation des fréquences gamma et que les réponses neurales locales induites par la lecture sont corrélées au sein de ce réseau via des connexions cortico-corticales longue distance qui sont différentes selon la nature du traitement (sémantique, phonologique ou visuel). Les analyses ont montré, en particulier, que dans le cas du traitement sémantique, les fluctuations de l'activité neuronale étaient corrélées entre le gyrus frontal inférieur (GFi) et le gyrus temporal moyen (MTG). Grâce à notre protocole adapté pour cette étude en intra-cortical, nous souhaitons tester si l'on retrouve le même type d'activité dans les bandes gamma et les mêmes corrélations fronto-temporales, impliquées à la fois dans le traitement sémantique des phrases et des images. Nous espérons que ces acquisitions chez des patients épileptiques en bilan pré-opératoire, sélectionnés sur la localisation majoritairement frontale et temporo-pariétale des électrodes implantées, apporteront des éléments de réponse essentiels pour mieux caractériser les processus communs impliqués dans la compréhension des phrases et des images, ainsi que des informations sur les localisations plus précises associées à ces processus. Ces résultats devraient contribuer à combler l'actuel manque de connaissances sur les interactions corticales longue distance sous-jacentes au réseau impliqué dans la compréhension sémantique (Price, 2010 ; Vidal et al., 2012) ainsi que sur les sources potentielles de marqueurs caractéristiques de l'intégration sémantique tels que la N400. Des études basées sur des enregistrements intracraniaux ont déjà permis de déterminer que les sources de la N400 semblaient localisées au niveau de la partie antérieure du lobe temporal médian, mais aussi dans les aires temporales moyennes, inférieures et supérieures ainsi que dans les aires préfrontales (Kutas & Federmeier, 2011).

Malgré tout l'intérêt qu'il peut y avoir à comparer des stimuli en modalités verbale et non verbale, mener à bien cette comparaison reste peut-être illusoire : on ne peut nier qu'il y a un certain nombre de différences entre les deux types de stimuli, susceptibles d'engager des traitements différents. Il y a, dans les images, une immédiateté de l'information que l'on ne retrouve pas dans les phrases où plusieurs interprétations possibles peuvent rentrer en conflit pour une même phrase. Nous nous sommes aperçus en faisant passer notre protocole que l'information contextuelle déjà disponible, supposée faciliter le traitement sémantique d'un second stimulus, pouvait au contraire la perturber et ce, particulièrement dans le cadre du langage. Par exemple, dans une phrase telle *que « La petite-fille glisse le long du toboggan »*, plusieurs personnes nous ont rapporté s'imaginer la petite fille sur un toboggan finissant dans un bac à sable. La suite *« Elle atterrit dans la piscine »* lève l'ambigüité contextuelle mais rentre en conflit avec les attentes et la représentation initiale que le sujet avait construites. Il est difficile de déterminer si nous mesurons réellement un processus correspondant à une intégration sémantique facilitée lorsqu'il y a un lien sémantique entre deux phrases, tant la variabilité des représentations d'un sujet à un autre est grande : des phrases que nous considérions comme sémantiquement cohérentes sont sûrement traitées comme des « violations sémantiques » par certains sujets. Malgré toutes les précautions expérimentales prises et toutes les améliorations que nous pourrions encore apporter à ce protocole, il y aura toujours une ambigüité propre au langage, car le matériel est créé sur les propres représentations de l'expérimentateur qui le conçoit et que rien ne nous permet d'affirmer que les sujets ont les mêmes représentations… Les stimuli visuels présentent aussi une certaine ambigüité : nous devons nous appuyer sur des indices minimes, des détails infiniment complexes, pour être capables de nous représenter des choses aussi fines que les intentions, les émotions, l'état d'esprit de la personne à partir d'une simple image, là où, au contraire, une simple phrase permettrait de clarifier les choses. D'où toute la difficulté à utiliser des stimuli aussi différents que des phrases et des images et la partialité des réponses concernant le débat sur la nature amodale du système sémantique. Dans ce type d'expérimentation utilisant des phrases ou des images, la dimension subjective dans le choix des stimuli par l'expérimentateur et dans les réponses données par les sujets est importante, mettant en jeu des composantes très personnelles telles que des souvenirs, des émotions…ce qui rend plus difficile l'analyse des résultats obtenus. Toutefois, malgré la grande variabilité des représentations inhérentes au langage et les détails infimes sur lesquels se base l'analyse visuelle, notre compréhension ne semble pas perturbée dans un cas comme dans l'autre. Il est évident que nous sommes capables de nous adapter perpétuellement lorsque nous interagissons avec autrui, que cette interaction fasse appel au langage ou non. C'est aussi là tout ce qui fait la « magie » du cerveau et tout l'intérêt de cette thématique de recherche.

Synthèse des résultats Partie Expérimentale 3 – EEG

Nous souhaitions tout d'abord analyser, au niveau temporel, les processus communs susceptibles d'intervenir dans la compréhension sémantique d'événements verbaux et non-verbaux (Images/Phrases). Par ailleurs, nous voulions évaluer l'impact d'un contexte sémantique cohérent ou incohérent sur la construction d'une représentation (en comparant les ERPs associés à la présentation des stimuli 1 ou 2 pour les conditions Suite et Non Suite).

Nos analyses mettent en évidence une forte influence de l'information contextuelle qui se traduit par des modulations du signal EEG spécifiques à la présentation d'un second stimulus lorsque celui-ci est incohérent : on observe des effets très précoces (entre 200 et 300 ms) pour les mots et les images ainsi qu'une négativité autour de 400 ms, spécifique aux images et caractéristique de la N400. Ces effets précoces, propres à chaque modalité, sont dus à des mécanismes attentionnels et ils traduisent la difficulté à intégrer une information sémantique non cohérente avec le contexte. Les informations contextuelles seraient utilisées très précocement pour guider un traitement sémantique plus approfondi. Ce second traitement pourrait correspondre aux effets plus tardifs observés dans notre étude (vers 1000 ms) qui sont similaires entre les deux modalités et qui pourraient correspondre à un processus d'intégration sémantique plus complexe faisant appel au système sémantique amodal.

Conclusion

III.1) Synthèse des principaux résultats

Ce travail de thèse avait pour but d'affiner certaines connaissances sur le système sémantique, à travers trois protocoles expérimentaux mettant en œuvre des méthodologies variées (IRMf/DTI, EEG et expérimentation comportementale).

Nos protocoles ont tout d'abord permis de confirmer l'existence d'un système sémantique amodal, impliqué dans la compréhension de stimuli verbaux et non-verbaux. Ils ont aussi mis en évidence l'importance de l'information contextuelle pour créer une représentation cohérente d'une situation, ainsi que l'existence de différences interindividuelles dans les capacités représentationnelles. Si nos différents protocoles donnent des informations complémentaires sur la manière dont est organisé le système sémantique, c'est sans doute le protocole en IRMf/DTI qui nous a permis d'apporter les éléments de réponse les plus convaincants car il répond aux principales interrogations liées à la thématique de ce travail (les deux autres protocoles nécessitant des développements ou des analyses supplémentaires).

III.3.*1*) Système sémantique commun à la compréhension des phrases et des scènes visuelles

Les résultats de nos différents protocoles convergent avec les données de la littérature qui montrent que le système conceptuel est accessible quelle que soit la modalité d'entrée. En effet, nous avons montré que les mêmes régions cérébrales étaient activées en IRMf lors de la présentation de stimuli complexes verbaux ou non verbaux. Bien que notre protocole comportemental ne fournisse pas d'information sur la localisation cérébrale, il a permis de déterminer que les phrases présentées en modalité visuelle n'étaient pas traitées différemment de celles présentées en modalité auditive, ce qui suggère également l'utilisation d'un même système sémantique. Concernent notre protocole EEG, s'il met en évidence des processus précoces modalité-spécifiques (probablement attentionnels), il révèle aussi des effets tardifs d'intégration sémantique communs aux images et aux phrases (même timing, vers 1000 ms et mêmes électrodes centrofrontales impliquées).

De façon cohérente avec les théories de la cognition incarnée, nos résultats en IRMf démontrent l'implication d'une importante composante motrice qui correspond au MNS puisqu'on retrouve plusieurs des régions décrites comme contenant des neurones miroirs chez l'homme (le cortex prémoteur BA6, l'aire de Broca GFi pars triangularis BA45, le cortex pariétal sensorimoteur et le LPI BA 37, 39 et 40). Mais nos résultats suggèrent aussi que la composante motrice ne représente qu'un

aspect du réseau sémantique qui s'étend plus vastement au niveau cortico-striatal, en impliquant particulièrement les aires préfrontales et temporo-pariétales.

Ce qui ressort principalement de nos résultats IRMf-DTI, c'est que les régions qui composent ce réseau sont des régions clés de par leur implication dans des processus cognitifs variés mais aussi de par leur connectivité anatomique. Ces régions correspondent à des « hubs » intégratifs et on ne peut que constater la grande similitude (anatomique mais aussi fonctionnelle) avec les zones de convergence telles que définies dans les modèles de Damasio ou Barsalou (Damasio, 1989-Barsalou, 1999).

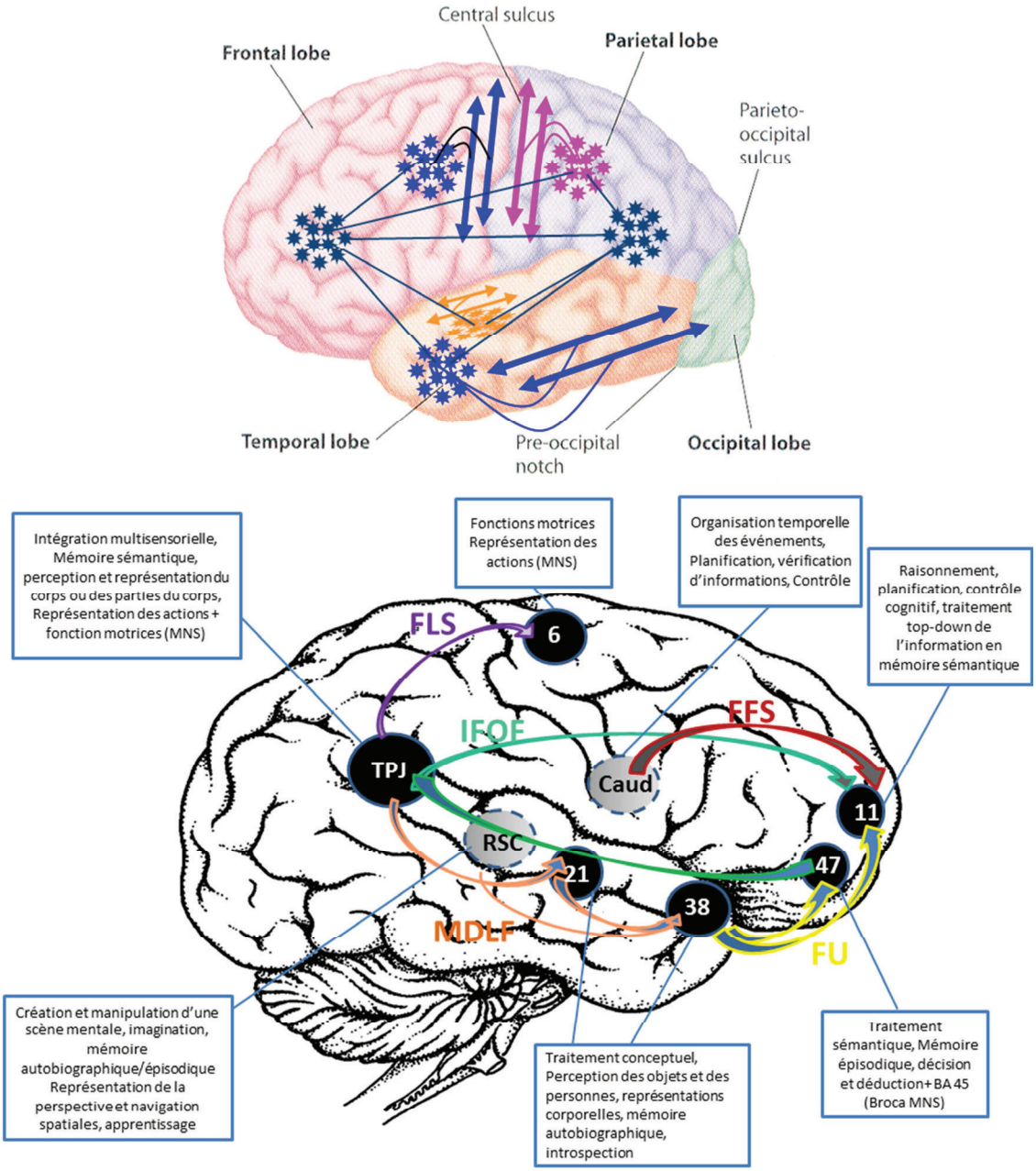


Fig 78) Zones de convergence et connectivité du réseau sémantique commun

La figure du haut représente le réseau de zones de convergence impliquées dans la compréhension sémantique du langage, telles qu'elles sont décrites dans les modèles de Damasio (1989) et Barsalou (1999). On ne peut que constater la proximité avec les aires cérébrales mises en évidence dans notre propre travail (figure du bas), qui sont impliquées dans la compréhension sémantique d'événements complexes verbaux et non verbaux mais aussi dans des processus cognitifs multiples, ce qui réaffirme leur rôle en tant que zones de convergence. Le réseau sémantique commun est soutenu par une architecture spécifique de fibres de matière blanche, qui assure la communication entre ces zones de convergence.

En effet, le TPJ (gyrus angulaire) est une région fortement associée à la mémoire sémantique mais aussi à l'intégration multi-sensorielle (Binder & Desai, 2011) où les informations issues de plusieurs modalités convergeraient pour former une représentation cohérente. Le pôle temporal antérieur est aussi considéré comme un « hub » intégratif en raison de son large rôle dans le traitement conceptuel (Simmons & Martin, 2009 ; Visser et al., 2010 ; Kiefer, & Pulvermüller, 2012) mais ces régions ne stockeraient pas les connaissances conceptuelles proprement dites, elles ne feraient qu'influencer la simulation via l'activation des zones sensorimotrices. Les régions situées entre la partie plus antérieure du lobe temporal et le TPJ ont des propriétés associées à la perception et à la représentation des corps et des objets ; mais ces régions ont aussi des fonctions liées à la mémoire autobiographique ou épisodique ainsi qu'à l'introspection, la plupart des régions observées dans notre étude recoupant le réseau du mode par défaut (lobe temporal médian, le précuneus, le cortex cingulaire postérieur et notamment le cortex rétrosplénial). Parmi ces régions, le RSC serait également une zone de convergence cruciale puisque cette région combine des propriétés associées à l'imagination, à la mémoire autobiographique/épisodique, ainsi qu'à la représentation de la perspective et à la navigation spatiale permettant la création et manipulation d'une scène mentale. Enfin, la simulation sensorimotrice impliquant les aires temporo-pariétales serait sous la dépendance des aires frontales et striatales (noyau caudé) en raison de leur rôle dans la planification structurelle et le contrôle cognitif. Au niveau des régions frontales, le cortex préfrontal médial, tout particulièrement, influencerait spécifiquement la simulation au niveau des aires sensorimotrices en raison de son rôle dans le contrôle top-down de l'activité en mémoire sémantique (Lallée et al 2010, Madden et al 2010).

Les données issues de nos analyses DTI vont dans le même sens puisqu'elles montrent que le réseau sémantique commun est soutenu par une architecture impliquant des voies de matière blanche spécifiques. Le TPJ a notamment une situation privilégiée en tant que hub, de par sa localisation anatomique à l'intersection des pôles pariétal et temporal et sa connectivité via un réseau à trois branches qui relient les régions temporales (MdLF), frontales (IFOF) et prémotrices (FLS). La connectivité de la partie antérieure du lobe temporal semble particulièrement importante puisque les régions temporales (GTs BA21 et GTm BA38) sont également reliées aux régions frontales médiale et latérale (BA11 et BA47) par l'intermédiaire du faisceau unciné. La connectivité du réseau commun correspond à la voie ventrale sémantique (FOF, l'UF, l'ILF ou MdLF), généralement mise en évidence à travers les études en DTI (Mandonnet et al., 2007 ; Catani et al., 2002 ; Saur et al., 2008 ; Duffau, 2008 ; Menjot de Champfleur et al., 2012).

Le fait que les régions de notre étude chevauchent fortement celles décrites dans le mode par défaut a une implication théorique importante qui montre bien toute la complexité fonctionnelle des zones

de convergence : les régions qui soutiennent la simulation sensorimotrice responsable de la compréhension sémantique sont les mêmes que celles impliquées dans des tâches de haut niveau encourageant le sujet à se tourner vers une mentalisation interne, à faire appel à la TOM pour comprendre les intentions d'autrui, à se remémorer des souvenirs ou à imaginer le futur. Ce serait avant tout l'implication de ce réseau qui permettrait à un sujet d'avoir la capacité de se représenter soi-même et de construire la représentation d'une scène en se projetant mentalement au moment présent dans une simulation d'un autre temps, lieu ou perspective (Buckner & Carroll, 2007 ; Hassabis & Maguire, 2007 ; Spreng, 2009), ce qui expliquerait en particulier les liens entre ce réseau et la mémoire autobiographique.

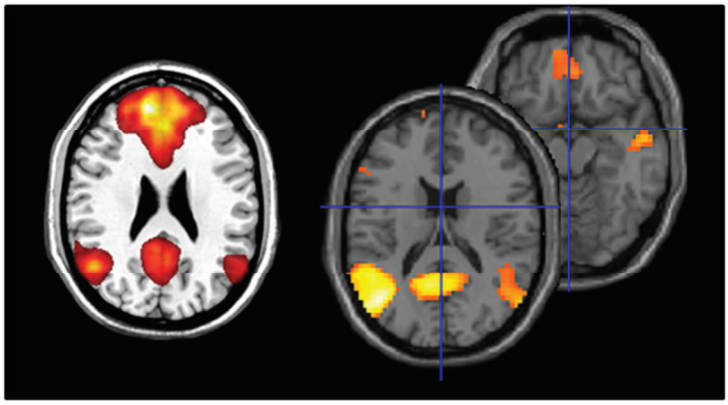


Fig 79) Default Mode Network

A gauche, les régions cérébrales activées en IRMf par des tâches spécifiquement utilisées pour impliquer le mode par défaut (Harrison et al., 2008). A droite, les activations issues de notre propre analyse de conjonction et qui correspondent aux principales régions du mode par défaut : le TPJ bilatéralement, le RSC et le cortex préfrontal.

Différentes régions cérébrales seraient associées à des fonctions plus ou moins spécifiques, à travers des traitements de plus bas niveaux sous la dépendance de sous-régions (perception du corps et des objets, mécanismes attentionnels…) et on peut considérer que les connaissances conceptuelles seraient « incarnées » et représentées dans les zones sensorielles et motrices, dont l'activation serait facilitée et influencée par les régions temporales et frontales. Mais c'est avant tout le fonctionnement en réseau de l'ensemble de ces régions qui permettrait de faire le lien entre les éléments perçus et les informations conceptuelles en mémoire sémantique pour aboutir à la représentation d'une situation à partir d'une image ou d'une phrase, ce fonctionnement en réseau étant soutenu par les faisceaux de la voie sémantique ventrale au niveau de la connectivité anatomique. Par ailleurs, l'important recoupement anatomique entre le réseau sémantique amodal mis en évidence dans notre étude et le mode par défaut, fortement associé à la mémoire autobiographique, permet d'envisager comment la simulation sensorimotrice nécessaire à la compréhension sémantique fait appel à nos expériences personnelles pour créer une représentation.

Si notre étude IRMf/DTI est celle qui apporte le plus de renseignements sur la localisation des régions cérébrales qui forment le système sémantique, des analyses supplémentaires sur des aspects temporels plus précis pourront nous permettre de comprendre plus précisément comment et à quel moment ces régions interagissent pour aboutir à une représentation cohérente. Une piste intéressante serait de chercher à savoir si ces régions interagissent séquentiellement ou conjointement. L'une des hypothèses que nous avons pour continuer à explorer cette thématique est que les connexions courtes (BA21-BA38, caudé-préfrontal, cortex temporal antérieur-cortex préfrontal) doivent soutenir des mécanismes locaux, potentiellement séquentiels, qui font le lien entre les processus perceptifs de plus bas niveau (perception des objets, du corps, détection de la perspective…) et la simulation sensorimotrice, alors que les connexions plus longues (comme celle qui relie le TPJ au préfrontal ou le TPJ-cortex temporal antérieur) supposent des mécanismes de synchronisation longue distance. Ces mécanismes de synchronisation impliquant les principales zones de convergence pourraient correspondre à l'émergence d'une représentation cohérente et soutenir des mécanismes d'analyses de plus haut niveau (TOM, capacité à imaginer le futur ou le passé, etc…).

Les analyses temps-fréquence de nos données EEG devraient nous fournir de plus amples renseignements sur ces possibles mécanismes de synchronisation mais elles devraient aussi nous renseigner sur la mise en jeu, au niveau temporel, de processus d'intégration sémantique communs à la compréhension des stimuli verbaux et non-verbaux. Nous envisageons également de compléter les analyses effectuées sur les données IRMf par des analyses de type DCM qui devraient contribuer à une meilleure compréhension des mécanismes de connectivité fonctionnelle et de synchronisation pour l'une ou l'autre des modalités. Enfin, l'utilisation de notre protocole EEG pour une expérimentation en intra-cortical devrait nous permettre de faire le lien entre les mécanismes au niveau temporel et des données de localisation beaucoup plus précises (les électrodes étant implantées dans plusieurs zones de convergence qui nous intéressent à la fois pariétales, temporales et frontales) pour mieux comprendre comment s'effectuent les synchronisations mais aussi quand et où la compréhension sémantique converge vers un processus commun entre les deux modalités.

III.3.2) Importance de l'information contextuelle pour la création d'une représentation

Notre protocole EEG, mais également notre expérience comportementale, nous ont permis de mettre en évidence le rôle crucial de l'information contextuelle dans la formation d'une représentation cohérente. En effet, nous avons montré que l'accessibilité à une représentation était facilitée lorsque celle-ci était initiée par une première représentation contenant un lien sémantique.

En revanche, lorsque deux stimuli ne présentent pas de lien sémantique, la création d'une seconde représentation est perturbée. Ces mécanismes se traduisent par des effets au niveau de marqueurs comportementaux (TR, COR), mais aussi par une influence sur le signal EEG qui correspond à des modulations du décours temporel durant des périodes très précoces (entre 200 et 300 ms). Ces effets sont spécifiques au second stimulus présenté, que ce soit pour des images ou des phrases.

Le fait qu'il y ait des effets très précoces associés à la contrainte contextuelle est cohérent avec l'idée que les informations contextuelles sont utilisées très précocement pour guider un traitement sémantique plus approfondi qui serait caractérisé par des processus plus tardifs d'intégration sémantique, processus communs aux deux modalités. En revanche l'utilisation de l'information contextuelle disponible pourrait reposer sur des processus précoces (notamment attentionnels) spécifiques à chaque modalité.

Nos données nécessitent des analyses complémentaires pour mieux comprendre ces mécanismes, elles suggèrent cependant qu'une représentation conceptuelle nécessite une construction pour être intégrée à un contexte plus vaste d'informations et que le traitement sémantique diffère selon l'information contextuelle disponible.

III.3.*3*) Différences interindividuelles dans les capacités de représentation à partir d'une phrase

Notre protocole comportemental est celui qui ouvre les pistes de réflexion les plus prometteuses car les différences interindividuelles dans les capacités représentationnelles sont peu étudiées et elles ouvrent des perspectives de recherche particulièrement intéressantes pour mieux comprendre la nature des représentations conceptuelles dans le cadre du langage. En effet, cette expérience nous a permis de considérer la nature des représentations initiées par la lecture ou l'écoute d'une phrase. Si dans les deux modalités, les phrases permettent de déclencher une représentation visuelle (ce qui suggère l'utilisation d'un processus sémantique commun partagé par la vision et le langage) nous nous sommes aperçus que ces représentations visuelles déclenchées par un événement décrit verbalement n'étaient pas homogènes et qu'il y avait plusieurs catégories d'individus dont les capacités représentationnelles ne sont pas similaires.

Le fait que ces différences interindividuelles soient mesurables au niveau comportemental est intéressant et nos résultats relancent un débat que nous avions abordé dans l'introduction : comment envisager les représentations sémantiques de façon globale, sur une population d'individus, alors qu'elles sont par ailleurs uniques puisqu'elles sont le fruit de nos expériences propres et donc de nos souvenirs personnels stockés en mémoire autobiographique ou épisodique ?

Ces différences interindividuelles pourraient provenir d'une utilisation différente de nos souvenirs pour se construire une représentation, traduisant peut-être des différences d'utilisation du système sémantique commun et/ou une communication différente entre les systèmes mnésiques sémantique et autobiographique/épisodique. Nos données comportementales sont corroborées par le fait que nous avons trouvé une corrélation entre le nombre de fibres d'un faisceau appartenant à la voie sémantique, le MdLF, et les capacités représentationnelles des sujets. N'ayant pas mis en évidence de différence notable dans l'utilisation de zones cérébrales spécifiques (IRMf) entre les différents groupes de sujets, on peut aussi envisager que ces différences soient liées à une rapidité de traitement différente. Une hypothèse que nous aimerions étudier à travers l'étude de nos données EEG mais aussi à travers des analyses plus spécifiques de la conductivité du MdLF (analyse du FA).

L'une des limites de ce protocole comportemental réside dans le fait que nous ne pouvons affirmer que cette capacité représentationnelle supporte la compréhension ou qu'elle correspond à la simulation sensorimotrice décrite dans les théories de la cognition incarnée (puisque la consigne était de « se représenter » une phrase). Des expériences complémentaires seront nécessaires pour mieux appréhender si la compréhension d'une phrase réside en la facilité à se la représenter et si elle peut être liée aux compétences mnésiques des sujets et à des particularités cérébrales (rôle du MdLF, rapidité de traitement…).

III.2) Discussion

L'implication de zones de convergence et de l'architecture spécifique d'un réseau sémantique amodal a largement été abordée précédemment : nous avons discuté les interprétations théoriques de nos résultats dans les chapitres correspondants. Nous avons décrit comment un même réseau cérébral pouvait être à la fois impliqué dans la réalisation et la reconnaissance des actions, mais aussi dans leur compréhension au niveau sémantique. Il semble que l'approche de certains chercheurs, basée sur une perspective liée à l'évolution peut apporter un éclairage intéressant et non négligeable aux problématiques que nous avons abordées et ouvrir de nouvelles pistes de réflexion, notamment sur le rôle du système fronto-temporo-pariétal (et des zones de convergences qui le composent) dans des aspects de cognition sociale de haut niveau (empathie, TOM, etc).

C'est à ces aspects que nous nous intéresserons plus en détail dans cette discussion ; plus particulièrement nous nous focaliserons sur pourquoi et comment ce système pourrait jouer ce rôle dans la mesure où l'être humain est avant tout un être social, fait pour interagir avec ses congénères. Aujourd'hui, un certain nombre d'auteurs considèrent que l'évolution de notre cerveau aurait pu être façonnée pour fournir cette capacité à interagir socialement, à travers la mise en place de capacités spécifiques à l'être humain telles que le langage, la TOM, l'empathie (Ramachandran, 2000 ; Oberman et al., 2007 ; Gallese, 2008). Réinterpréter les connaissances neuroscientifiques selon cette perspective évolutive paraît essentiel, en effet elle pourrait fournir le lien entre action et langage et permettre de comprendre comment un même système peut être impliqué dans des fonctions cognitives aussi diverses.

L'étude neuroscientifique de la cognition sociale ouvre de nouvelles perspectives pour une meilleure compréhension de ce qu' « être humain » signifie (Gallese, 2008). Les interactions sociales représentent une part essentielle de l'existence humaine, certains pensent même que nous sommes nés avec un désir et une capacité innés pour ces interactions sociales (Meltzoff & Moore, 1997 ; Meltzoff & Decety, 2003). Pour certains, ce besoin intrinsèque d'interactions sociales serait un facteur ayant motivé l'évolution des capacités de cognition sociale qui sont spécifiquement humaines (Ramachandran, 2000 ; Oberman et al., 2007). Actuellement, on considère de plus en plus que le MNS ou, selon les auteurs, un réseau plus large d'aires fronto-temporo-pariétales[8] serait le système en charge de ces aspects liés à la cognition sociale (Gallese, 2008 ; Oberman et al., 2007 ; Arbib, 2005). Ce système et le mécanisme de simulation incarnée qu'il sous-tend (c'est-à-dire sa capacité à faire correspondre notre perception de l'environnement à des représentations sensori-motrices

[8] Dans la suite de ce paragraphe, nous parlerons surtout du MNS car la plupart des hypothèses développées dans la littérature concerne ce système. Mais bien sûr, ces hypothèses restent les mêmes pour un réseau fronto-temporo-pariétal ne se limitant pas au MNS.

internes) serait capable de jouer un rôle clé dans des aspects multiples, depuis la simple perception d'une action à l'empathie pour autrui (Oberman et al., 2007 ; Gallese, 2008).

III.2.*1*) Pertinence du système fronto-temporo-pariétal dans l'évolution de la cognition sociale

La présence de neurones miroirs chez différentes espèces de primates (comme les macaques et les humains) et les capacités qui leur sont associées chez les différentes espèces, semblent favoriser une vision continuiste de l'évolution de la cognition sociale. Cependant, cette même preuve doit aussi être réconciliée à la fois avec le caractère unique de la cognition sociale humaine et la grande variété de processus qu'elle recouvre (Gallese, 2008).

Il est probable que les singes utilisent leur MNS pour optimiser leurs interactions sociales, en donnant du sens aux comportements de leurs congénères et que, même chez les primates non humains, le MNS serait suffisamment sophistiqué pour soutenir une facilitation sociale. Bien que les capacités d'imitation purement motrice des singes n'aient été que rarement mises en évidence (Ferrari et al., 2006 ; Kumashiro et al. 2003), il a été montré que les singes étaient capables d'imitation cognitive (Subiaul, Cantlon, Holloway, et Terrace, 2004 ; Paukner, Anderson, Borelli, Visalberghi, & Ferrari, 2005), mais aussi d'inférer sur les connaissances de leurs congénères à propos du monde en se basant sur des indices comportementaux, tels que la direction du regard (Hare et al., 2000; Tomasello, Carpenter, Call, Behne, & Moll, 2005). Il a été notamment mis en évidence que les singes pouvaient établir un lien cognitif entre « voir » et « savoir », plusieurs études ayant montré que des singes rhésus choisissaient systématiquement de dérober de la nourriture au moment où l'expérimentateur humain ne pouvait les surprendre (Flombaum et Santos, 2005 ; Santos, Nissen, & Ferrugia, 2006).

Les chercheurs supposent que le MNS, chez les primates inférieurs (lémuriens, loris), aurait essentiellement pour fonction de faciliter la compréhension de l'action (Rizzolatti et al., 2001) tandis que chez les primates supérieurs non-humains (macaques, chimpanzés, etc.) ce même système aurait évolué pour soutenir les processus plus élaborés que nous avons mentionnés via l'activation des propriétés motrices des neurones miroirs (Rizzolatti et Arbib, 1998). Cependant, malgré ces capacités assez élaborées, on peut supposer que les singes n'attribuent pas explicitement des états mentaux à leurs congénères pour donner une explication causale aux comportements de ces derniers (Gallese, 2008). Chez l'homme, l'évolution du MNS serait encore plus avancée, ce qui leur permettrait de se représenter non seulement les aspects physiques d'une action mais aussi les intentions sous-jacentes, les pensées et les sentiments qui ont motivé cette action, éventuellement

par le biais de connexions réciproques entre d'autres régions du cerveau comme le système limbique ou le cortex préfrontal médial (Oberman, 2007). C'est cette évolution qui aurait « jeté les bases » pour le développement des compétences sociales uniques de l'homme, telles que l'imitation (Rizzolatti et al, 2001), la compréhension des intentions associées à une action (Rizzolatti, Fogassi, et Gallese, 2006) la TOM et la capacité à décoder les états mentaux d'autrui (Gallese, 2007; Gallese & Goldman, 1998), l'empathie (de Vignemont & Singer, 2006; Gallese, 2003 ; Sommerville & Decety, 2006) et surtout le langage (Gallese, 2001 ; Arbib, 2005 ; Gallese et Lakoff, 2005; Rizzolatti & Arbib, 1998).

La pertinence, pour la psychologie sociale, de la découverte du MNS semble évidente : l'interaction sociale, la coordination interpersonnelle, l'attention partagée autour d'actions conjointes exigent la capacité de comprendre et d'anticiper le comportement de nos partenaires. Le MNS semble être le substrat neurofonctionnel de toutes ces capacités (Gallese, 2008). Un intéressant travail en EEG de Oberman et al. (2007) confirme pleinement l'importance du MNS dans la cognition sociale. En effet, dans cet article, les auteurs montrent que le rythme des ondes Mu est modulé par le degré d'interaction sociale, durant une tâche où les sujets regardaient des vidéos montrant des individus plus ou moins engagés dans des interactions sociales. La suppression des rythmes Mu a souvent été utilisée comme un indice de l'activité des neurones miroirs (Gastaut & Bert, 1954; Cochin et al., 1998; Babiloni et al., 2002; Pineda et al., 2000 ; Altschuler et al., 1997; Oberman et al., 2005 ; Muthukumaraswamy et al., 2004). L'utilisation de cet indice étant, par ailleurs, validée par des preuves anatomiques et physiologiques de connexions cortico-corticales fortes, chez les primates humains et non humains, entre le cortex prémoteur ventral contenant les neurones miroirs et le cortex sensorimoteur primaire où les ondes Mu sont générées et enregistrées (Muakkassa et Strick, 1979; Godschalk et al, 1984;. Matelli et al, 1986 ; Nishitani et Hari, 2000 ; Tokuno et Nambu, 2000; Dum et Strick , 2002; Shimazu et al, 2004). Dans le travail d'Oberman et al., la condition de non-interaction est celle qui provoque le moins de suppression des rythmes Mu et inversement pour les conditions où l'interaction sociale est importante. Pour ces auteurs, ces données suggèrent que le MNS de l'homme est spécialisé non seulement dans le traitement des actions, mais plus particulièrement dans le traitement des stimuli qui ont une pertinence sociale.

III.2.*2*) Evolution et langage

Parmi les différentes compétences à forte connotation sociale, le langage est non seulement un pivot essentiel de la cognition humaine, mais c'est aussi celle qui est la plus unique et la plus spécifique de notre espèce. Jusqu'à présent, aucun équivalent au langage humain n'a été trouvé chez

d'autres espèces animales, y compris les grands singes et les oiseaux chanteurs (Berwick, 2011). Même si certaines espèces sont capables d'apprentissage vocal (éléphants, phoques, orques, certains oiseaux…) et si certaines des capacités et systèmes requis pour le langage ont des équivalents chez d'autres espèces (comme les oiseaux chanteurs qui ont une capacité d'apprentissage de l'imitation vocale et une organisation neuronale sous-jacente similaires à l'homme, Berwick, 2011), l'homme est le seul parmi les primates à présenter ces capacités (Jarvis, 2006). Le langage humain semble être un développement évolutif récent : les preuves archéologiques suggèrent qu'il est apparu dans les cent mille dernières années et qu'il n'a pas évolué de manière significative depuis que nos ancêtres humains ont quitté l'Afrique il y a environ 50 000-80 000 années (Tattersall, 2010). Pour Berwick et al. (2013), le fait que le langage soit unique pour les êtres humains provient d'une évolution spécifique de notre espèce permettant d'aboutir à cette capacité langagière dont des corrélats cérébraux sont identifiables.

- **L'hypothèse d'exploitation neuronale**

Pour Gallese (Gallese, 2008), puisque le MNS semble jouer un rôle important dans la cognition sociale, c'est aussi un parfait candidat pour expliquer la nature incarnée du langage : dans le langage, la signification serait intimement liée à l'expérience sociale que représente l'action. Pour formaliser cette hypothèse, Gallese parle d' « hypothèse d'exploitation neuronale » pour expliquer comment les principaux aspects de la cognition sociale humaine peuvent être soutenus par des mécanismes neuronaux originellement dédiés à l'intégration sensori-motrice. Ces mécanismes auraient progressivement changé, avec l'augmentation du pouvoir computationnel liée à l'évolution du cerveau humain, constituant ainsi la base d'une nouvelle architecture neurofonctionnelle pour la pensée et le langage, tout en conservant aussi leurs fonctions d'origine (Gallese & Lakoff, 2005; Gallese, 2003, 2007). Bien sûr, comme le précise Gallese, l'hypothèse de l'exploitation neuronale est à prendre avec prudence car elle postule que l'amélioration quantitative d'une architecture neurofonctionnelle préexistante peut engendrer une évolution qualitative de différentes aptitudes cognitives sociales. Cependant, cette hypothèse a le mérite d'inscrire la cognition sociale, et en particulier le langage, dans une perspective incarnée : pour attribuer un sens aux mots, une fusion est nécessaire entre les sons articulés des mots et une situation, que l'on expérimente souvent dans l'action ; c'est exactement ce que permet la simulation incarnée.

- **La théorie gestuelle de l'évolution du langage**

Certains auteurs postulent également que le MNS aurait été une plateforme initiale parfaite pour une évolution ultérieure du langage mais ils fournissent une explication plus détaillée sur l'origine de cette évolution : le langage aurait évolué à partir des gestes manuels, sorte de

« pantomime » correspondant à un système de communication primitif (Corballis, 2010 ; Gentilucci & Corballis, 2006). À un certain moment de l'évolution des hominidés, le système miroir aurait assumé un rôle supplémentaire dans le langage, tout en devenant latéralisé, mais en gardant aussi ses fonctions initiales liées aux activités manuelles. Plusieurs arguments semblent étayer cette théorie. Tout d'abord, le langage lui-même n'implique pas nécessairement de vocalisations, on sait aujourd'hui que le langage des signes, utilisé par les personnes sourdes, présente toute la complexité linguistique des langues parlées (Emmorey, 2002 ; Neidle, Kegl, MacLaughlin, Bahan, & Lee, 2000) et l'aire de Broca est activée aussi bien par la parole que par le fait de signer (Horwitz et al., 2003). De plus, les circuits neuronaux impliqués dans la génération des appels vocaux chez les primates non-humains (cortex cingulaire, diencéphale, structures du tronc cérébral) sont radicalement différents de ceux qui sont impliqués dans la parole humaine (Jürgens, 2002 ; Rizzolatti & Sinigaglia, 2008). Ces observations sont donc plus cohérentes avec l'hypothèse que le langage aurait évolué à partir des gestes manuels plutôt qu'à partir des vocalisations (Arbib, 2005 ; Armstrong & Wilcox, 2007; Corballis 2002, 2004 ; Fogassi & Ferrari, 2007; Rizzolatti & Arbib, 1998 ; Rizzolatti & Sinigaglia, 2008).

Bien que la communication vocale soit très répandue dans le règne animal, les vocalisations sont pour la plupart génétiquement structurées et l'intentionnalité est souvent limitée. Comme nous l'avons mentionné, peu d'espèces semblent capables d'apprentissage vocal, ce qui est bien sûr essentiel à la langue parlée (Corballis, 2010). En revanche, les gestes de communication utilisés par les grands singes (gorilles, chimpanzés, bonobos…) sont soumis à l'apprentissage social, sensibles à l'état attentionnel du destinataire et peuvent même être détachés du contexte chez les chimpanzés et les bonobos (Pika, Liebal, & Tomasello, 2003 ; 2005 ; Liebal, Call, & Tomasello, 2004 ; Pollick & de Waal, 2007), tous ces éléments étant des conditions préalables au langage.

Notre patrimoine, hérité de nos ancêtres hominidés, semble donc beaucoup mieux pré-adapté à une communication intentionnelle basée sur les mouvements du corps, et en particulier les mouvements des mains, qu'à une communication basée sur des vocalisations (Corballis, 2010). Et il semble donc raisonnable de supposer que les éléments précurseurs du langage humain soient fondés sur la mimesis, le corps et les mains étant libres de se déplacer avec beaucoup de degrés de liberté et d'imiter ainsi l'activité du monde extérieur (une évolution qui aurait été possible avec l'apparition de la bipédie chez l'homo erectus, il y environ 2 millions d'années, Donal, 1991). Avec la mimesis pure, la compréhension reposant sur des actions mimées serait plus ou moins directement associée aux actions elle-mêmes, ce qui expliquerait l'implication initiale du MNS. Au cours de l'évolution humaine, le langage serait devenu plus conventionnalisé et plus symbolique, perdant son aspect mimétique, et les vocalisations auraient remplacé les actes manuels comme principal moyen de communication. Cette transition, probablement progressive et naturelle, aurait permis l'intégration

des vocalisations dans le MNS (Corballis, 2010). On comprend mieux cette idée de transition entre un système gestuel initial et un système langagier basé principalement sur des vocalisations si l'on considère la parole elle-même non pas comme un système acoustique mais plutôt comme un système gestuel (basée sur des mouvements de la bouche et du visage), une idée capturée par la théorie motrice de la perception de la parole et les travaux sur la phonologie articulatoire (Gentilucci & Corballis, 2006). Les études sur le développement de l'enfant semblent confirmer l'existence d'un lien encore prégnant entre activité gestuelle et capacités langagières puisque les gestes manuels, antérieurs à l'apparition du langage chez les jeunes enfants, joueraient un rôle important dans le développement de la parole normale et seraient de bons prédicteurs des capacités langagières ultérieures (Bates & Dick, 2002 ; Iverson et Goldin-Meadow, 2005). Par exemple, le babillage canonique chez les enfants âgés de 6 à 8 mois est accompagné par des mouvements rythmiques de la main (Masataka, 2001) ; la compréhension de la parole chez les enfants entre 8 et 10 mois et les productions de mots entre 11 et 13 mois sont accompagnées par des gestes de pointage et des gestes indiquant la reconnaissance des mots (Bates & Snyder, 1987).

Corballis (2010) ajoute un élément intéressant à sa réflexion pour expliquer l'apparition de la grammaire, dans un second temps, permettant d'aboutir au système langagier tel qu'on le connaît aujourd'hui. Ce serait l'évolution de la mémoire épisodique et la capacité à « voyager dans le temps mentalement » qui aurait poussé le langage à se « grammaticaliser », impliquant une augmentation nécessaire du vocabulaire et des formes langagières pour pouvoir faire référence à des épisodes distincts, en temps et en lieu, du moment actuel. Cette évolution se serait probablement produite avec l'apparition du genre Homo au Pléistocène et aurait permis à l'être humain d'acquérir cette capacité unique à reconstituer des événements passés et à imaginer des événements futurs ou fictifs. Le langage aurait ainsi évolué pour permettre à nos « pérégrinations mentales » d'être partagées, au bénéfice du groupe mais aussi de l'individu (Corballis, 2010). Pour Corballis (2010), la capacité au « voyage dans le temps mental » et le langage auraient évolué conjointement pour aboutir à cette structure particulière qu'est l'esprit humain.

Bien sûr, encore une fois, ces hypothèses sont purement spéculatives et requièrent des éclaircissements, notamment concernant la datation de l'introduction du contrôle vocal par rapport au système des neurones miroirs. Cependant, un élément de réponse pourrait provenir de la génétique, avec la mise en évidence du gène FOXP2 (pour « forkhead box P2 ») grâce aux travaux sur une famille anglaise, la famille KE, dont plusieurs membres souffraient d'altérations spécifiques du langage suite à une mutation de ce gène. En effet, on trouve chez plusieurs membres de cette famille affectés par la mutation de ce gène des déficits très sévères de l'articulation vocale (Watkins, Dronkers, et Vargha-Khadem, 2002) et un défaut d'activation de l'aire de Broca dans des tâches

langagières (Liégeois et al., 2003). Ces éléments laissent penser que ce gène jouerait un rôle dans le recrutement de cette aire pour le contrôle vocal, lors de la mise en place des voies neuronales durant le développement embryonnaire (Corballis, 2004). Bien qu'extrêmement bien conservé phylogénétiquement chez les mammifères (la protéine qu'il produit est presque identique chez la souris et les primates qui sont séparés par environ 130 millions d'années d'évolution), le gène FOXP2 aurait subi deux mutations depuis la scission entre les hominidés et les chimpanzés, mutations qui se traduisent par 2 ou 3 acides aminés différents. Ce seraient ces mutations infimes qui font vraisemblablement la différence entre les animaux incapables de parler et l'être parlant que nous sommes… On estime que ces mutations auraient eu lieu il y a 200 000-100 000 ans, ce qui pourrait correspondre à l'apparition de l'Homo Sapiens (il y a environ 170.000 ans) et à la période d'émergence du langage articulé chez l'humain (Enard et al., 2002).

III.2.*3*) Multiples rôles du système fronto-temporo-pariétal

Pour conclure, nous reprendrons la réflexion de V.Gallese (2008), qui précise que la pertinence du rôle du MNS (ou d'un système fronto-temporo-pariétal plus étendu) dans tant d'aspects de la cognition sociale ne provient pas d'un « don » spécifique des neurones miroirs comme si ces neurones avaient des propriétés « magiques » permettant de répondre à toutes nos interrogations sur la nature humaine et la cognition sociale. Les neurones miroirs tirent leurs propriétés des connexions étroites qu'ils entretiennent avec d'autres populations de neurones dans le cerveau et si le MNS est impliqué dans autant d'aspects de la cognition sociale, c'est grâce à l'activation de circuits cortico-cortical multiples et parallèles qui génèrent les propriétés miroirs. Comme le mentionne Gallese (Gallese, 2008), on ne peut pas nécessairement attribuer toute la richesse de la cognition sociale humaine uniquement au MNS. Il est vrai que la découverte des neurones miroirs qui fut très médiatisée au départ (pour certains, la découverte des neurones miroirs serait équivalente, en psychologie, à la découverte de l'ADN pour la biologie ; Ramachandran 2000) fut par la suite banalisée et parfois vivement critiquée, provoquant même une certaine irritation dans certains domaines des sciences cognitives, ce qui tend à décrédibiliser les hypothèses spéculatives liées au rôle du MNS.

Il paraît presque inutile de rappeler tout l'intérêt qu'auront également ces connaissances sur le rôle des aires sensorimotrices dans la cognition sociale pour une meilleure compréhension de certaines pathologies congénitales, en particulier des troubles du spectre autistique. En effet, un certain nombre de travaux dans ce domaine semblent corroborer l'importance du MNS dans les compétences sociales puisqu'un dysfonctionnement de ce système pourrait fournir une explication

des déficits neuropsychologiques dans des pathologies telles que l'autisme où l'on sait que les compétences sociales et l'interprétation des émotions et états mentaux d'autrui sont des processus particulièrement touchés (Oberman et al., 2005, 2008 ; Altschuler et al, 2000 ; Nishitani et al, 2004; Dapretto et al, 2006; Oberman et al, 2005; Théorêt et al, 2005 ; Villalobos et al, 2005 ; Baron-Cohen, 1995 ; Iacoboni & Mazziotta, 2007).

Il semble bien que ce système, façonné par des millions d'années d'évolution, soit responsable de la plupart des compétences qui caractérisent l'être humain. C'est peut-être parce que nous sommes « faits » pour les interactions sociales que ce système hautement adaptatif permet de surmonter l'incertitude inhérente au langage et les différences interindividuelles pour que nous puissions nous comprendre et interagir. Malgré des capacités de représentation (compréhension ?) différentes, ce système nous permettrait ainsi d'accéder à une représentation plus ou moins partagée du monde, d'un individu à l'autre.

Quelle que soit la position soutenue concernant les liens entre compétences sociales et systèmes senrorimoteurs, ces hypothèses basées sur une perspective évolutive restent intéressantes et apportent un éclairage nouveau sur le rôle des régions sensorimotrices et sur la nature des représentations. Des travaux portant sur les capacités et l'anatomie comparées chez le singe et chez l'homme viendront sûrement compléter ces hypothèses dans les années à venir, travaux qui nous permettront de mieux comprendre qui nous fûmes mais surtout qui nous sommes…

III.3) Perspectives : Intérêt de la Cognition Incarnée pour la recherche en Robotique

Ayant largement abordé les interrogations persistantes liées à nos thématiques de recherche et les améliorations à apporter à nos protocoles précédemment, nous préférons ouvrir le débat dans ce chapitre « Perspectives » en évoquant l'intérêt des connaissances dans le domaine de la cognition incarnée pour le développement de robots intelligents. Fruit de la réflexion développée dans notre équipe où Neurosciences et Robotique sont étroitement liées, cette thématique peut paraître à première vue éloignée de nos préoccupations neurophysiologiques. Si la découverte des neurones miroirs fut un élément important de la recherche des années 90 ; la robotique est une thématique actuelle majeure et ce, probablement encore pour un certain nombre d'années à venir. On parle même de « robolution»[9]. Aujourd'hui, les développements robotiques s'appuient sur les connaissances neuroscientifiques : tout ce qui concerne le MNS fonto-temporo-pariétal et son rôle dans la cognition sociale, les simulations sensorimotrices et la compréhension dans le cadre de la cognition incarnée sont des éléments majeurs de la recherche en robotique. Et finalement, les deux thématiques ont peut-être plus de points d'ancrage communs qu'il n'y paraît à première vue…

Depuis les mythes fondateurs jusqu'aux romans et films de science-fiction actuels, l'homme a toujours rêvé de créer un être artificiel à son image, capable d'interagir avec lui de façon intelligente et autonome et, depuis une quinzaine d'années, on a assisté à une émergence extraordinaire de la robotique personnelle avec le concept d'Intelligence Artificielle. Concrétisation du rêve de passionnés de robotique, fruit de la recherche expérimentale dans ce domaine ou résultat d'études de marché menées par des entreprises à objectifs commerciaux, les robots actuels ont des compétences variées et des fonctions précises grâce aux progrès technologiques. Les premiers robots étaient principalement ludiques ou avaient une fonction domotique[10], puis sont apparus des robots capables de nous libérer de certaines tâches ménagères[11] ou de tâches fastidieuses, voire dangereuses[12] et même des robots (souvent humanoïdes) susceptibles de remplacer l'homme dans certaines activités à caractère social[13] ou à des fins thérapeutiques[14].

[9] Terme employé par Bruno Bonnell dans son ouvrage « Viva la robolution » pour définir l'intérêt grandissant pour la robotique et l'apparition progressive de robots dans notre société.

[10] Robots jouets de type Aïbo , Pleo, Genibo, WowWee… Robots domotiques : robot aspirateur, robot piscine, robot tondeuse à gazon…

[11] nettoyer, vider la poubelle, faire les poussières, arroser les plantes, ramasser des objets ou des vêtements, regrouper la vaisselle et la mettre dans l'évier ou le lave-vaisselle, tondre la pelouse, laver la piscine ou garder la maison.

[12] Robot travaillant sous la pluie et se déplaçant sur des terrains en pente ou sablonneux, robot vigile dans des immeubles ou robot pompier, robot soldat …

[13] Robots baby-sitter, hôtesses d'accueil, enseignants

[14] Robots compagnons auprès d'enfants malades ou assistants de vie pour des personnes âgées ou handicapées. Signalons, dans l'actualité, que Panasonic vient juste d'annoncer la commercialisation de ses robots dédiés au monde de l'hôpital.

Enfin, peu à peu, a germé l'idée que l'on pouvait concevoir des robots dotés d'intelligence et capables d'autonomie et d'adaptation, capables d'interagir avec l'homme. Certains chercheurs comme Cynthia Breazeal du MIT (Media Lab) affirme : *« un robot n'est pas un outil, mais un partenaire. Les robots doivent être dotés d'une personnalité et être capables de comprendre nos intentions et ce que nous sommes pour devenir de véritables partenaires, utiles, capables de collaborer avec nous et de s'intégrer socialement à notre environnement. »*

Créer des robots intelligents… le défi paraît énorme, on attend de ces robots qu'ils puissent parler, réagir de façon adaptée aux comportements et aux émotions d'une personne, imiter et comprendre les actions humaine, avoir une interaction avec leur environnement. S'il est vrai qu'aujourd'hui certains robots peuvent avoir des comportements assez variés et des interactions simples avec un utilisateur humain (se déplacer, réagir à la voix et au toucher, exprimer des émotions simples, percevoir leur entourage, parler…) malgré leur degré de sophistication technologique, la plupart de ces robots ont des comportements préenregistrés et en termes de capacités sociales, ils ne sont pas adaptés aux hautes compétences que l'on attend d'eux car ils manquent d'autonomie, de compréhension et sont donc assez limités en termes de réelle interaction. Pour créer cette nouvelle génération de robots « intelligents » de nombreux chercheurs ont eu l'idée de s'appuyer sur les travaux des neuroscientifiques, des spécialistes du comportement humain qui, pour leur part, ont aussi mesuré ce que pouvait leur apporter cette collaboration dans la connaissance de l'être humain et la prise en charge de certaines pathologies, les robots pouvant constituer un outil permettant de tester des hypothèses neuroscientifiques. Pour les roboticiens d'aujourd'hui, l'intelligence des robots ne dépend plus uniquement des algorithmes implémentés dans la machine et de leur puissance de calcul : l'idée que plus on introduirait d'algorithmes et plus les robots seraient performants n'est plus aussi importante. Trois axes majeurs, fortement reliés à la cognition incarnée, semblent cruciaux pour doter les robots de capacités d'interaction : leur donner un corps, développer leurs compétences d'apprentissage en interaction avec l'environnement (notamment les capacités de compréhension des actions humaines) et développer leurs compétences sociales.

- **<u>Robots : des agents « embodied », un corps adapté pour affûter l'intelligence motrice</u>**

Certes ce sont des algorithmes très complexes qui rendent le robot le plus autonome possible mais c'est aussi son corps : il est important qu'il puisse se déplacer rapidement sur deux jambes (pour être adapté aux monde des humains), changer de direction, éviter un obstacle... La dynamique de son mouvement, sa capacité d'équilibre (qui ne dépend pas uniquement des algorithmes mais également de la forme de son corps) sont indispensables à l'autonomie du robot et à son interaction avec l'homme. Des chercheurs comme Hans Moravec (Université de Carnegie

Mellon, USA), Rodney Brooks (MIT, USA) et Rolf Pfeifer (Université de Zurich, Suisse) pensent que l'intelligence du robot doit être conçue en termes de « sensibilité » physique plutôt que de « pensée symbolique ». Pierre-Yves Oudeyer précise l'idée développée par Rolf Pfeifer en 1999 : « *le corps lui-même, par ses propriétés physico-dynamiques, peut réaliser un calcul implicite qui participe tout autant au comportement que le calcul explicite réalisé par le cerveau.* »

L'enjeu pour la robotique actuelle est donc d'envisager le robot non pas comme un cerveau agissant sur le corps mais comme une unité cerveau-corps. Ainsi le LAAS,[15] s'intéresse au mouvement (locomotion, décision et interaction) et tout particulièrement à la compréhension et la modélisation de la démarche humaine afin de donner aux robots humanoïdes une autonomie d'action en utilisant la planification du mouvement de manière automatique. Un autre axe de recherche porte sur la chaîne perception/décision/action avec des travaux sur le suivi de personnes en mouvement, leur localisation (essentielle pour une interaction avec elles) avec des travaux sur la reconnaissance des mouvements humains, et surtout sur la signification des gestes en relation avec un ordre vocal (Sreenivasa et al., 2012 ; Taïx et al., 2012 ; Stasse et al., 2008). Reconstruire le mouvement de l'homme doit aussi servir à l'apprentissage de gestes que le robot pourra effectuer en toute autonomie, autonomie que les chercheurs améliorent grâce à de nouveaux algorithmes de déplacement mais aussi grâce à des applications neurorobotiques où la décision de mouvement viendrait uniquement de la perception visuelle et inertielle du robot.

Certains de ces travaux s'inspirent de l'anthropologie évolutionniste sur les robots cognitifs (notamment les travaux du Max Planck Institute) et visent à mettre en œuvre l'attention conjointe entre homme et robot pour permettre au robot de se mettre à la place de l'homme et d'imaginer ce qu'il peut percevoir. Simuler des processus mentaux n'est pas un domaine nouveau, les cogniticiens utilisent des ordinateurs pour cela depuis les années 50, mais beaucoup de ces simulations se sont avant tout focalisées sur les aspects cognitifs de haut-niveau (résolution de problème) ou sur la simulation de neurones artificiels. Il manquait un élément essentiel à ces types de simulation : le corps, l'expérience de cette corporalité étant nécessaire au développement cognitif puisque les traitements sensori-moteurs et les capacités motrices jouent un rôle clé dans le développement des capacités cognitives de plus haut niveau, y compris le langage et les comportements sociaux. Tout comme l'humain, c'est par ce corps que le robot sera capable d'interagir avec son environnement et avec nous. C'est aussi ce corps qui pourrait conditionner les « représentations » du robot et ses apprentissages. Il paraît évident qu'il faut au robot un système équivalent au système fronto-temporo-pariétal humain pour lui permettre de reconnaître les actions effectuées par un être

[15] laboratoire basé à Toulouse et utilisant le robot HRP-2 (Humanoid Robotic Platform 2 ou Promet, Kawada Industries-AIST), qui a de grandes capacités de mouvements.

humain, de les faire correspondre à ses propres représentations « sensorimotrices » et ainsi de leur donner un sens.

- **Apprendre de façon naturelle en interaction avec l'environnement**

Il existe de nombreuses formes d'apprentissage artificiel, basées sur des algorithmes ou des réseaux de neurones complexes (apprentissage par exploration, par optimisation, par approximation et interpolation, etc.). Plutôt que de nous égarer dans la présentation de données techniques, essayons d'illustrer l'apprentissage par quelques exemples concrets (Oudeyer & Kaplan, 2006 ; 2007 ; 2008 ; Kaplan & Oudeyer, 2007, 2008). Aujourd'hui un robot peut, par exemple, apprendre par ce que l'on appelle des prédictions : on lui présente un nombre limité de différents objets (balle, chaise) et il essaiera ensuite de classifier un nouvel objet dans l'une ou l'autre des catégories sur la base d'une « reconstruction » créée à partir des différents exemplaires rencontrés. A partir d'un nombre limité d'exemples, le robot essaie de trouver une correspondance à partir de « points communs » qu'il peut trouver entre la nouvelle situation qui lui est présentée et les situations précédemment rencontrées. Des algorithmes encore plus élaborés peuvent permettre à un robot d'apprendre à anticiper les conséquences perceptives de ses actions (visuelles, sonores, tactiles, proprioceptives, etc.) et il peut ainsi prédire les conséquences d'un nouvel acte à partir d'expériences passées. Une autre forme d'apprentissage est basée sur l'apprentissage par prototypes : la machine garde en mémoire tous les exemples précédents et quand elle doit faire une prédiction, elle recherche les exemples les plus proches déjà mémorisés.

Ces différentes méthodes permettent de fournir un embryon de « mémoire sémantique » au robot, dans la mesure où il peut conserver une trace des situations/objets qu'il a rencontrés. Cependant, ces méthodes d'apprentissage ne sont pas très naturelles : l'apprentissage par prédictions manque parfois de précision et d'efficacité (il y a toujours une approximation dans la reconstruction) tandis que l'apprentissage par prototypes risque de générer un manque de souplesse. C'est pourquoi, actuellement, on essaie de plus en plus de reproduire un apprentissage naturel en laissant les robots apprendre de leurs interactions avec leur environnement. Inspirée de la psychologie du développement, la robotique développementale est une méthodologie qui cherche à développer pour les robots des systèmes cognitifs dynamiques, évolués et capables de contrôler leur corps de façon autonome. La grande idée qui préside cette démarche est que le développement cognitif d'un système robotique complexe peut se faire étape par étape, à la manière dont un enfant se développe cognitivement grâce à des interactions avec le monde. Les chercheurs se sont inspirés d'une idée émise par Alan Turing (dans un article de 1950 « Computing machinery and intelligence ») et énonçant les bases de la «robotique expérimentale». « *Au lieu de simuler un esprit adulte, pourquoi*

ne pas essayer de simuler celui d'un enfant ? Si on le faisait, avec de l'apprentissage et de l'éducation, on pourrait obtenir un esprit adulte. »

Il existe plusieurs « bébés robots » dans différents laboratoires et c'est dans cette optique de robotique développementale que s'inscrivent les travaux réalisés, dans notre équipe, sur le robot iCub[16], dont « le « cerveau » est équipé de quelques capacités innées proches de celles observées chez les nourrissons, telles que reconnaître un visage humain ou détecter des objets dans un environnement. En dehors de ces quelques capacités, tout est à apprendre. Le cerveau de l'iCub correspond à une architecture logicielle à trois niveaux. Le premier niveau regroupe des informations sur ce que l'iCub voit et sent. Le second système est appelé « système de modulation » et est basé sur un ensemble de motivations basiques simulant la curiosité d'un jeune enfant pour les stimulations nouvelles et sa tendance à s'engager dans des interactions sociales. Il régit la circulation des informations provenant du premier niveau et les compare aux actions et aux informations sensorielles auxquelles l'iCub a déjà été confronté auparavant (reproduisant grossièrement les fonctions de l'hippocampe, des ganglions de la base et de l'amygdale dans le cerveau humain). Le troisième niveau utilise les expériences passées pour faire des prédictions (« que se passera-t-il si je bouge la balle avec telle force et dans telle direction ? ») puis renvoie ces informations au niveau intermédiaire pour l'aider à déterminer la prochaine action du robot.

L'objectif des différents laboratoires travaillant sur l'iCub est de rendre l'iCub aussi humain que possible dans ses déplacements, ses comportements, son langage. Actuellement, il est en phase d'apprentissage et l'objectif est de produire une machine capable d'adapter son comportement à différentes circonstances, d'agir, de communiquer et de coopérer avec l'humain ; ceci grâce à l'ensemble de capteurs et de moteurs qui le composent et qui lui permettent d'avoir une perception active mais surtout grâce aux apprentissages qui lui sont proposés pour lui apprendre à parler, à comprendre, à reproduire des actions, à explorer son environnement et à s'y adapter, en étant capable d'interpréter correctement les objets qui l'entourent, tout cela à partir de modèles d'activités humaines et sans programmation préalable de ces fonctions. Des chercheurs[17] tentent par exemple d'implémenter dans l'iCub une modélisation du système cérébelleux, le cervelet étant, chez l'homme, un élément crucial pour le mouvement dans la mesure où il gère le timing et le rythme de nos mouvements. D'autres chercheurs cherchent à développer chez l'iCub des représentations internes précises des objets et de leurs rôles afin de pouvoir les utiliser en coopération avec les humains ou encore, partant de l'idée que la manipulation et la communication co-évoluent durant

[16] Créé par l'Institut Italien des Technologies, IIT), l'iCub (i-Cognitive Universal Body, cub signifiant aussi « petit de mammifère ») est un robot humanoïde mis au point par le consortium RobotCub.
[17] Institute of Audiovisual Studies, Université Pompeu Fabra -Barcelone

les 3 premières années chez l'enfant, les chercheurs[18] tentent de faire émerger progressivement le langage humain chez l'iCub sans qu'aucune règle langagière ne soit implémentée au préalable dans l'architecture logicielle du robot. Par exemple, ils nomment un objet tout en le montrant à l'iCub ou en le laissant le manipuler, afin que celui-ci puisse faire des associations entre les sons qu'il entend et les données qui parviennent à son système visuo-moteur et acquérir peu à peu un langage naturel à la grammaire basique mais correcte. Pour apprendre une action, l'iCub apprend par observation et manipulation : on lui montre une action (ex : déplacer un objet), s'il ne sait pas ce que c'est on lui « explique » à quoi cette action correspond par éléments séquentiels (ex : saisir l'objet, déplacer sa main de gauche à droite, lâcher l'objet…) afin qu'il puisse acquérir une représentation interne de cette action, la conceptualiser et ensuite la reproduire tout en liant cela avec le langage afin qu'il puisse verbaliser les actions qu'il observe ou effectue. Ces apprentissages, tant au niveau des actions que du langage, reposent sur l'observation et l'imitation.

L'iCub de notre équipe est déjà capable d'intervenir dans une interaction entre des êtres humains qui collaborent, prenant le rôle de l'un ou l'autre des humains dans une tâche simple. A terme il pourra comprendre les actions réalisées par des personnes, anticiper les objectifs d'une action et pouvoir aider les personnes en interagissant avec elles, y compris verbalement (Dominey et al., 2008 ; Lallée et al., 2011). Les dernières avancées de notre équipe concernent la possibilité de donner une mémoire autobiographique à l'iCub afin qu'il puisse faire appel à des souvenirs qui lui sont propres pour analyser une situation présente.

- **Développer les compétences sociales des robots**

Un certain nombre de chercheurs (notamment Cynthia Breazeal et les équipes de robotique du Masachuchets Institute of Technology) travaillent plus particulièrement au développement des compétences sociales des robots, dans une démarche très proche de celle utilisée pour la reconnaissance des actions et la compréhension (chez des robots comme l'iCub) puisqu'il s'agit également d'une perspective développementale. Les deux approches diffèrent légèrement car Breazeal et ses collègues postulent que la création de robots « socialement intelligents » est l'aspect primordial à développer pour rendre les robots plus aptes à interagir avec les humains (Breazeal, 2000, 2003, 2004, 2007, 2011 ; Bar-Cohen et al., 2003 ; Breazeal, Buchsbaum et al., 2005). C'est pourquoi la plupart des robots développés au MIT sont des robots à fortes compétences émotionnelles. Les travaux menés par C. Breazeal portent essentiellement sur la compréhension et

[18] Recherches du Centre for Robotics and Neural Systems, University of Plymouth et de l'Institute of Cognitive Sciences and Technologies de Rome, dans le cadre du projet ITALK (Integration and Transfer of Action and Language Knowledge in Robots)

l'expression des émotions et sur la construction de robots capables d'apprendre à imiter les expressions faciales d'un humain et très nettement inspirés de mécanismes biologiques (Breazeal et al., 2005).

Cette approche est fortement influencée par la façon dont les enfants apprennent à communiquer avec leur entourage et à comprendre les actions des autres, avec en particulier pour base l'hypothèse que chez les enfants humains, les interactions imitatives (qui commencent par l'imitation faciale) sont une étape importante de leur développement cognitif. Ces comportements imitatifs (Meltzoff & Gopnik, 1993 ; Meltzoff & Moore, 1983) sont présents très tôt chez l'enfant, quasiment à partir de la naissance et ils se développent rapidement et dans de nombreux domaines (vocalisations, gestuelles, actions reliées à des objets…). Il semble que cette capacité précoce d'imitation soit déjà une forme puissante et flexible d'apprentissage social, précurseur des compétences sociales ultérieures et en particulier de la TOM, mécanisme qui permet de comprendre et d'interpréter les actions, les émotions et les états mentaux d'autrui, (Theory of Mind ou ToM, Premack & Woodruff, 1978). Soumis à des réajustements incessants et à l'intégration permanente de nouvelles composantes, les premiers échanges partagés bi-directionnels, non verbaux, les jeux entre les enfants et les adultes seraient la base de l'apprentissage de la communication et des habiletés de langage. Les réponses imitatives des adultes guident et modèlent l'apprentissage des enfants et peu à peu l'enfant apprend à anticiper et à gérer volontairement certaines séquences prédictives d'événements. Selon Meltzoff, ces jeux d'imitations réciproques seraient les premières bases de la compréhension d'autrui dans la mesure où ils permettent à l'enfant d'explorer, d'élaborer et de consolider une connaissance d'eux-mêmes, des autres et de ce qui relie ces deux entités. Peu à peu, les enfants ne percevront plus ces imitations comme une simple correspondance corporelle avec leurs propres expressions, mais comme des actions à part entière. Comprendre comment la capacité d'imitation peut contribuer au développement des mécanismes qui sous-tendent les compétences sociales est important pour développer des robots pouvant exprimer des capacités sociales. La théorie de la Simulation (domaine d'étude de la TOM) semble être l'une des hypothèses les plus plausibles expliquant le rôle de l'imitation dans le développement d'une théorie de l'esprit. La théorie de la Simulation postule que, en simulant les actions d'une autre personne et les stimulations qu'elle perçoit, grâce à nos propres mécanismes comportementaux et de traitements des stimuli, les humains peuvent faire des prédictions à propos de comportements et des états mentaux des autres. Il s'agit de penser comme si nous étions l'autre pour comprendre l'autre. Chez les enfants, comme nous l'avons vu précédemment, la façon d'apprendre à « simuler » l'autre dépendrait des interactions imitatives précoces ; imiter l'expression ou le mouvement des autres serait déjà une forme de simulation de leurs comportements. Toutes les théories développementales de l'imitation

ont, bien sûr, trouvé une validation à travers les nombreux travaux sur les neurones miroirs. En ce qui concerne le robot l'imitation pourrait être utile à deux niveaux : elle lui permettrait de développer des habiletés dans le domaine de la communication et elle permettrait d'apprendre au robot de nouvelles capacités motrices et actions. Des robots possédant des facultés d'imitation ont déjà été développés auparavant, mais en général, il s'agissait de systèmes capables d'imiter certaines actions précises et les chercheurs se sont focalisés avant tout sur la performance du robot plutôt que sur une interaction imitative globale.

Pour C. Breazeal, utiliser des processus d'imitation devrait principalement permettre d'aider le robot à apprendre davantage sur les autres autour de lui. Plus un moyen de développer l'intelligence sociale du robot qu'une façon de lui enseigner rapidement de nouvelles capacités, l'imitation devrait permettre d'amorcer le développement de compétences sociales chez le robot. Comme pour les enfants, avec les robots, il lui semble pertinent de commencer par l'imitation faciale et c'est ce qui a été testé avec le robot Leonardo, un petit robot humanoïde, conçu pour communiquer et effectuer des gestes et qui peut aussi bien interagir avec des personnes que manipuler des objets. Il est capable d'expressions faciales proches de celles des humains grâce à un système de vision binoculaire actif. Pour définir les modules à implémenter dans leur robot, les auteurs se sont basés sur un modèle de Meltzoff et Moore : le modèle AIM (active intermodal mapping), qui suggère qu'une combinaison de connaissances innées et de mécanismes spécialisés d'apprentissage sont à la base de la capacité des enfants à imiter dans un but particulier et d'une manière plurimodale. Ce modèle AIM comporte 3 éléments clés : l'identification des parties du corps impliquées dans le mouvement ou l'expression à imiter ; l'espace intermodal qui permet d'associer les expressions et les mouvements en fonction de leur source, sorte de format universel pour représenter les gestes et les postures réalisés par l'enfant lui-même ou ceux que l'enfant perçoit chez l'adulte ; et la phase de « babillage » moteur, phase d'exploration de son espace moteur et d'apprentissage en vue d' utiliser avec un contrôle maximal son corps et son visage. Imiter les expressions faciales humaines et détecter une adéquation entre les mouvements des autres et ses propres mouvements a permis au robot Leonardo de s'engager dans des interactions imitatives bi-directionnelles. A partir de son propre système inspiré du modèle AIM (le poséographe), Leonardo semble capable de créer une représentation de l'expression de l'humain quand il observe les expressions faciales d'un humain et au fur et à mesure, il peut réaliser grâce à son propre visage des approximations de plus en plus précises de l'expression du participant humain. Les chercheurs pensent avoir fait un premier pas dans l'implémentation d'un système de type neurones miroirs chez un robot. Cette technique utilisée pour les expressions faciales pourrait être généralisée à tout le corps, on peut ainsi rêver qu'un jour les robots soient capables de comprendre nos actions mais aussi nos intentions.

D'autres robots créés au MIT[19] dans l'optique de développer des robots sociaux sur la base de théories issues de la psychologie du développement ont permis d'étudier les aspects émotionnels des interactions avec l'entourage (Breazeal et al., 2008 ; Lee & Breazeal, 2010). Capables de percevoir chez l'humain une grande variété d'indices sociaux via les canaux auditifs et visuels, ces robots peuvent délivrer, en retour, des signaux sociaux à l'humain via la direction de leur regard, leur expression faciale, leur posture corporelle et du babillage vocal. Ainsi ils expriment une large palette d'émotions telles que la peur, la joie, la surprise, le dégoût, la colère… L'objectif du Personal-Robot-Group du MIT est de développer des robots sociaux capables d'apprendre avec succès ce qui est important pour un individu à partir d'interactions du type de celles que les humains développent, sur le long-terme, dans la vie de tous les jours. Mais pour rendre ces robots plus performants, Breazeal et ses collègues pensent qu'il est également nécessaire de mieux comprendre les facteurs humains qui interviendraient dans un enseignement de l'homme vers le robot (patience limitée si le robot ne comprend pas, données humaines ambigües pour le robot, quel type de modèle mental les gens se créent du robot en situation d'apprentissage…).

D'autres chercheurs, comme Kerstin Dautenhahn, professeur et chercheur en robotique et intelligence artificielle à l'université de Hertfordshire, voient dans l'intelligence sociale l'élément clé du développement intellectuel des robots, mais leurs théories se réfèrent plutôt à la phylogénèse qu'à l'ontogénèse pour expliquer la nature de l'intelligence sociale et s'appuient sur l'hypothèse de l'intelligence social (ou hypothèse du cerveau social, Dunbar 1993). D'après R Dunbar, chez les primates l'épouillage maintiendrait la cohésion sociale. Cette interaction sociale basique aurait été remplacée chez les groupes d'humains modernes (avec l'évolution du cerveau) par une forme plus d'interaction efficace : le langage. Pour Kirsten Dauhtenhahn, « appliquer l'hypothèse de l'intelligence sociale à l'AI implique que l'intelligence sociale soit un ingrédient clé de l'intelligence humaine et, en tant que tel, un candidat préalable à tout robot artificiellement intelligent » (Fong et al., 2003 ; Dautenhahn, 1995, 2003, 2004, 2012, ; Dautenhahn et al., 2002, 2007, 2011 ; Billard et Dautenhahn, 1998, 1999 ; Robins et al., 2012. Cela représente un challenge excitant qui requiert des points de vue interdisciplinaires que l'on rencontre dans le champ de l'interaction homme robot (HRI human-robot interaction). Ce qui compte ici pour rendre le robot intelligent ce n'est plus que le robot soit capable d'imiter, mais avant tout d'interagir.

[19] Kismet (Breazeal, 2000) et Nexi qui est un robot qualifié de MDS (mobile, dexterous, social = mobile, habile et social).

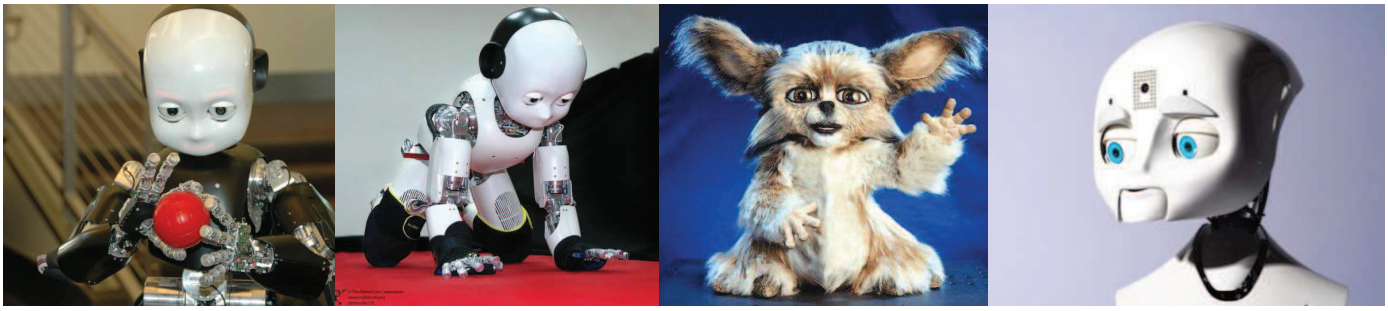


Fig 80) Quelques robots dont l'intelligence s'inspire de la cognition incarnée
Le robot iCub en train de manipuler un objet, l'iCub qui se déplace à quatre pattes (projet RobotCub) ; le robot Leonardo et le robot Nexi (MIT)

- **<u>Robots et Cognition Incarnée</u>**

Tout au long de cette réflexion sur pourquoi et comment développer des robots qui puissent être nos partenaires sociaux, un certain nombre de mots clés apparaissent : corporéité, imitation, perception des états mentaux d'autrui, décodage des intentions et des émotions, interactions, coopération, compétences sociales, compréhension sémantique et bien sûr langage… Ces thématiques, que nous avons abordées tout au long de ce travail, sont reliées par un fil conducteur : le fait qu'il existe dans le cerveau humain des mêmes zones de convergence qui pourraient être impliquées dans tous ces processus (avec un rôle particulier du MNS ou d'un réseau fronto-temporo-pariétal plus large).

Il paraît donc évident que c'est ce système qu'il faut reproduire pour doter les robots de toutes ces capacités et leur permettre d'acquérir une vaste palette de « processus » cognitifs et comportementaux. Il faut concevoir le robot comme une entité où corps et esprit sont aussi importants l'un que l'autre et interagissent conjointement. Les chercheurs en robotique ne s'y sont pas trompés et de plus en plus de travaux s'inspirent des théories de la « cognition incarnée » pour créer des robots intelligents. Dans le passé, les connaissances dans ce domaine ont déjà eu une influence importante sur les travaux en intelligence artificielle et en robotique. Cependant, ces notions étaient le plus souvent appliquées à des robots très simples, pour soutenir des mécanismes de bas niveau tels que la motricité dynamique et la navigation ou la résolution de problèmes, la plupart des modèles de calcul qui sous-tendent les plates-formes robotiques ne prenant pas en considération un fait essentiel : la nature incarnée des robots (Hoffman, 2012). Avec le développement des connaissances sur la cognition incarnée et l'amélioration technique des robots, on a assisté à un changement sur la conception de l'intelligence artificielle et sur son implémentation dans des systèmes autonomes, où apparaît comme cruciale la nécessité d'intégrer à la fois des capacités perceptives, cognitives, communicatives et de motricité pour créer des robots humanoïdes

vraiment autonomes capables d'agir conjointement avec les humains (Knoll et al., 2001). Avec cette approche prenant en compte la cognition incarnée, on espère aboutir à des robots plus proches de nous ; d'ailleurs la terminologie a également changé et on parle aujourd'hui plus volontiers d' « humains artificiels » ou d' « agents conversationnels incarnés » (Vogeley et Bent, 2010). En effet, on développe des robots qui ressemblent à des humains, qui sont interactifs et qui ont des capacités sociales complexes, avec une attention particulière portée sur le langage et les représentations, la communication non verbale et l'apprentissage social.

Certains chercheurs en robotique ont commencé à intégrer ces principes, inspirés de la cognition incarnée, dans leurs travaux (Breazeal, Gray & Berlin, 2009 ; Breazeal, 2002 ; Riek, Paul, et Robinson, 2010 ; Pfeifer & Scheier, 1997 ; Dominey et al., 2005 ; Gorniak & Roy, 2007; Lungarella, Metta, Pfeifer, & Sandini, 2003; Mavridis & Roy, 2006; Pfeifer & Gómez, 2005; Roy, Hsiao, & Mavridis, 2004; Vernon et al., 2007). C'est le cas, en particulier dans le cadre du langage et des représentations sémantiques où plusieurs travaux, dont les résultats sont prometteurs, confirment l'intérêt que présente le fait de s'appuyer sur l'approche de la cognition incarnée pour implémenter des compétences langagières (Cangelosi et al., 2010 ; Dominey et Boucher, 2005 ; Lallée et al., 2010 ; Marocco et al., 2010 ; Madden et al., 2010).

- *Robots et coopération*

Dans le domaine de la coopération, un certain nombre de travaux montrent déjà l'intérêt qu'il y a à utiliser les connaissances issues de la cognition incarnée, pour tous les niveaux que peut nécessiter une interaction coopérative (compréhension des intentions, imitation, langage…). Equipés de senseurs et de capteurs qui leur permettent d'exister dans le monde physique des humains et d'interagir avec eux , le corps et l'« esprit » des robots sont affutés pour en faire des êtres doués d'habiletés sociales mais il leur manque une spécificité de la cognition humaine qui rend les humains uniques par rapport aux autres animaux, c'est la capacité et la motivation à partager des états intentionnels avec les autres.

Michael Tomasello est chercheur dans le domaine de la cognition sociale, de l'apprentissage social et de la communication, champs qu'il a étudiés à partir de perspectives comparatives, développementales et culturelles. Il a beaucoup travaillé sur la comparaison de la cognition humaine avec celle des grands singes. Tomasello et al. (2005) ont avancé l'hypothèse que la faculté humaine à partager des émotions serait le résultat de deux capacités distinctes : la première serait la capacité à se représenter et à concevoir les autres comme des agents dirigés par un but intentionnel, à déterminer leurs intentions à travers l'observation de leurs comportements ; la seconde concerne la motivation à partager des intentions avec les autres. D'après un certain nombre de travaux en

cognition animale et en psychologie développementale, les primates non-humains et humains maîtrisent la première capacité, mais seuls les humains possèderaient la deuxième capacité : la motivation à coopérer. Et cette capacité à coopérer serait présente très tôt, chez les jeunes enfants. Par exemple, dans une étude (Warneken, Chen & Tomasello, 2005) qui met en scène des jeunes enfants de 18-24 mois ou des jeunes chimpanzés, impliqués dans des tâches orientés vers un but et des jeux sociaux faisant appel à la coopération avec un adulte humain, seuls les enfants participent aux deux types de tâches de façon enthousiaste, cherchant à réengager l'adulte dans la tâche ou à l'aider quand celui-ci arrête d'agir. Les chimpanzés ne sont intéressés que par les tâches quand elles sont dirigées vers un but (de la nourriture) et ne recherchent pas du tout l'interaction lorsqu'ils doivent coopérer. En effet, ils sont tout à fait capables de réaliser une tâche quand elle a pour but la nourriture mais jugent une tâche inutile si elle débouche seulement sur une interaction. D'autres études (Warneken & Tomasello, 2006) montrent que des jeunes enfants aident l'adulte spontanément dans des situations diverses, démontrant ainsi une motivation altruiste précoce à aider et une capacité à comprendre et se représenter les buts et les intentions des autres. De jeunes enfants sont aussi tout à fait capables d'effectuer une épreuve d'imitation inversée (Carpenter, Tomasello, & Striano, 2005) : dans une tâche où les rôles sont alternés, l'enfant peut très bien passer d'un rôle à l'autre, montrant qu'il est capable de maintenir une représentation de l'alternance des rôles des participants.

C'est sur cette idée de coopération que les chercheurs (Dominey et Warneken, 2011) ont décidé d'implémenter ces capacités dans un système robotique (« the Cooperator », un simple bras robotisé) et de les tester dans le cadre d'interactions homme-robot. Dans diverses tâches expérimentales, robots et humains coopéraient en déplaçant des objets physiques à différentes positions dans un espace de travail partagé. Le modèle implémenté semble plutôt bien fonctionner puisque le système robotique est capable d'accéder à des représentations basiques c'est à dire qu'il peut stocker et récupérer des actions dans un ordre séquentiel ; il est aussi capable d'attribuer chaque action à un agent. Il accède aussi à l'état physique du monde en temps réel grâce à une fonctionnalité du système (World Model) lui permettant d'encoder les localisations physiques des objets et de mettre à jour les informations à ce sujet grâce à la vision et à la proprioception. Les changements observés (objets déplacés) lui permettent de détecter des actions qui correspondent aux mouvements de ces objets et les actions sont représentées en termes d'agent, d'objet et de but de l'action (ex : X a bougé le cheval à côté du chien). Ces représentations peuvent être utilisées pour ordonner, reconnaître ou décrire une action et ensuite pour imiter et échanger à propos d'une action. Grâce à ce système, le robot est capable de réaliser des tâches d'imitation et d'imitation inversée ainsi que des jeux coopératifs plus ou moins complexes. Les capacités à inférer les

intentions et à coopérer avec autrui apparaissent donc comme des éléments clés pour développer des robots intelligents, la coopération intervenant dans le développement de nombreux aspects cognitifs (motricité, langage, etc..). Peu importe la tâche, on peut imaginer appliquer ce système à différentes situations (notamment au langage), l'essentiel étant vraiment la capacité à élaborer des représentations internes partagées avec autrui. De même, une plateforme robotique plus élaborée permettrait aussi plus de liberté : d'ailleurs l'iCub apprend déjà grâce à ce type d'interactions coopératives. Il peut ainsi participer à des jeux coopératifs ou aider un humain à construire une table en observant deux individus réaliser la tâche une première fois et ensuite, non seulement il peut prendre la place de l'un ou l'autre des humains, mais il est également capable d'anticiper l'action suivante à réaliser afin d'aider au mieux l'humain qui construit la table. Il sera bientôt doté d'un système lui permettant d'avoir des représentations internes fines sur lui-même et sur la personne avec qui il interagit lors d'une tâche coopérative. Bien sûr, pour le moment, ces formes de coopération ne sont pas généralisées et le robot ne peut intervenir que sur des tâches spécifiques, mais on voit combien cela pourrait être pertinent dans la vie de tous les jours.

- *Robots et langage : représentations conceptuelles et simulation*

Initialement, les modèles computationnels étudiant la communication du langage dans des systèmes robotiques ou dans des agents simulés étaient plutôt centrés sur les caractéristiques internes de l'agent qui permettent au lexique d'être construit sur la base d'un système symbolique autoréférentiel. Les robots ne possédaient qu'une série de symboles abstraits, utilisés pour la communication et pour la représentation de significations (Kirby, 2001). Mais ces modèles étaient sujets au problème de l'ancrage des symboles (Harnad, 1990) : les symboles sont des entités autoréférentielles qui requièrent l'interprétation d'un expérimentateur externe pour identifier la signification référentielle des éléments lexicaux. La recherche traditionnelle en robotique avait tendance à utiliser plutôt les théories amodales de la connaissance qui affirment que l'information est issue des stimuli traités perceptivement et convertis en symboles non perceptifs qui pourront ensuite être utilisés pour la recherche d'information, la prise de décision et la production d'action. Des conceptions contestées, comme nous l'avons déjà vu, par les découvertes récentes et les nombreux travaux dans le domaine de la cognition incarnée qui suggèrent plutôt un modèle où les concepts sont ancrés dans des représentations modales, utilisant les mêmes mécanismes que ceux de la perception et de l'action (Barsalou, 1999 ; Kosslyn, 1995). Les chercheurs en robotique sont de plus en plus convaincus que les représentations chez les agents robotiques ne peuvent être élaborées dans l'abstrait mais à partir d'un contexte concret d'interactions entre l'agent et son environnement.

Selon les approches inspirées de la cognition incarnée, les compétences linguistiques se développent grâce à l'interaction directe entre les robots et l'environnement avec lequel ils interagissent (humains ou autres robots faisant partie de cet environnement). Dans ces modèles, le monde extérieur joue un rôle essentiel dans le « façonnement » du langage utilisé par les robots et le langage est fondé sur le développement des connaissances cognitives et sensorimotrices des agents (Steels, 2003 ; Marroco, 2010). Concevoir le langage chez les robots dans une perspective incarnée nécessite que le lexique de base de l'agent soit lui aussi incarné et cela suppose la capacité de lier les représentions perceptuelles (et internes) à des éléments symboliques (Cangelosi et Riga, 2006). En robotique, la notion de représentation perceptuelle peut être traduite en modèles cognitifs computationnels de multiples façons. Comme nous l'avons vu, on peut implémenter des modèles théoriques de la simulation dans un système robotique pour déduire les croyances, les intentions, les émotions et les buts d'une action humaine (Breazeal, Gray, et Berlin, 2009 ; Gray, Breazeal, Berlin, Brooks, et Lieberman, 2005 ; Dominey et al, 2011, Lallée et al., 2010, 2012, 2013). Le robot utilise les mêmes systèmes perceptifs pour générer des comportements et pour la simulation de ces différents processus.

Gorniak et Roy (2007) ont ainsi présenté une théorie computationnelle de la compréhension du langage en fonction de l'environnement physique dans lequel se trouve le robot. Steels et ses collaborateurs ont beaucoup travaillé sur le langage, même si les premiers robots utilisés n'étaient pas « incarnés » à proprement parler mais seulement des « têtes parlantes » (Steels et al., 2002; Steels, 1998, 2001, 2003) ; ces premiers travaux ont permis de montrer qu'un lexique commun pour décrire le monde pouvait émerger à partir d'une interaction entre deux robots ou entre robots et humains (Steels et Kaplan, 2000). L'approche de Steels se caractérise par l'accent mis sur l'importance des mécanismes sociaux dans l'émergence du langage. D'autres modèles se sont concentrés sur les facteurs de développement qui favorisent l'acquisition du langage en cherchant à déterminer le rôle des motivations internes et d'un comportement exploratoire actif (Oudeyer et Kaplan, 2006 ; Oudeyer et al. 2007). Un certain nombre de modèles se sont intéressés à la capacité des robots à acquérir un lexique à travers des interactions avec des utilisateurs humains (Roy et al., 2003 ; Sugita et Tani, 2005 ; Yu, 2005). Une étude de Dominey et al. (2005) donne un aperçu du développement et de l'évolution du langage : comment des phrases simples permettent d'aboutir progressivement à des constructions grammaticales plus abstraites.

Si les premières expériences consistaient surtout à faire apprendre au robot des noms (mots pour décrire des entités physiques ou abstraites) et des verbes (mots représentant une action réalisée par le robot ou par l'expérimentateur), les protocoles expérimentaux et les compétences acquises par les robots deviennent de plus en plus élaborés. Une expérience réalisée avec le robot iCub dans notre

équipe (Lallée, 2010) montre comme la compréhension d'un langage « incarnée » permet au système d'élaborer des représentations pour traiter des actions dirigées vers un but (prendre, couvrir, donner…) et de comprendre les relations entre l'état initial et le résultat d'une action. Par ailleurs, cette étude montre également comment des éléments grammaticaux (catégories grammaticales et connecteurs causaux : parce que, si, alors…) peuvent structurer et enrichir les représentations de l'état du monde que le robot se crée. Marocco et al. (2010) ont présenté un modèle robotique (réseau de neurones artificiels fonctionnant par un algorithme de rétro-propagation à travers le temps) pour l'étude de la représentation incarnée de mots d'action. Ils ont montré comment le robot humanoïde iCub pouvait apprendre la signification de ces mots d'action en interagissant physiquement avec l'environnement et relier les effets de ses propres actions à un comportement observé sur des objets, avant et après l'action.

Hoffman (2012) propose un autre modèle computationnel très intéressant : il est lui aussi basé sur l'utilisation de représentations perceptives modales mais plus dans l'esprit des notions de zones de convergence (Damasio, 1989 ; Simmons & Barsalou, 2003) avec un système sous la forme de « flux de modalité », construit à partir de nœuds interconnectés. De plus, les capacités de ce système ne se limitent pas au système conceptuel puisque ce modèle permettrait aussi de développer les systèmes de mémoire et de prise de décision. Dans ce modèle, les nœuds peuvent correspondre à des entrées sensorielles brutes (une image), une caractéristique (une couleur dominante), une propriété (la vitesse d'un objet) ou à un niveau conceptuel supérieur, regroupement statistique des différentes caractéristiques et propriétés. Ces flux modalitaires sont reliés à un réseau de l'action constitué de nœuds d'action qui sont activés de la même manière que les nœuds de processus perceptif. Un nœud d'action, à son tour, permet la réalisation d'une action motrice. Les flux d'activations vont dans les deux sens : un sens afférent (depuis les systèmes sensoriels aux concepts et actions) et un sens efférent (dans la direction opposée), en lien avec le principe d'intégration perception/action. Pour Hoffman (2012), dans ces modèles, il est nécessaire de prendre en compte aussi le fait que développements moteur et cognitif ne sont pas disparates mais, au contraire, très interdépendants. Il souligne qu'en général l'exécution motrice est traditionnellement considérée, en robotique, comme quelque chose de passif, une fin de chaîne de traitements qui reçoit les flux et les données et qui est sous la dépendance d'un processus décisionnel central. Hoffman (2012), quant à lui, pense qu'il est possible de modéliser cette influence réciproque action-perception chez des robots interactifs, sous forme de deux réseaux d'activation symétriques perception-action. Dans ces réseaux, la perception exerce une influence sur les associations de niveau supérieur, menant à la sélection d'une action potentielle mais peut aussi inversement être biaisée par l'activité motrice. De plus, il faut tenir compte que le traitement perceptif n'est pas une analyse strictement bottom-up

des données disponibles initialement, comme il est souvent modalisé dans les systèmes robotiques et en IA. Au contraire, les simulations de processus perceptifs affectent l'acquisition de nouvelles données perceptives et la connaissance motrice est utilisée dans le traitement sensoriel, où les intentions, les objectifs et les attentes jouent un rôle dans la capacité à analyser le monde. En d'autres termes, les systèmes sensori-perceptifs sont pénétrables à la fois par des processus cognitifs et par des processus d'action. La notion d'influences top-down a déjà été explorée pour certaines tâches visuelles traitées par des systèmes informatiques (par exemple, Bregler, 1997; Hamdan, Heitz, et Thoraval, 1999), mais Hoffman (2012) pense que l'utilisation du traitement top-down a un fort potentiel pour des applications à des robots interagissant physiquement avec des humains où ce qui est simulé cognitivement (les actions, les concepts et les prédictions) pourrait influencer les modules perceptuels et sensoriels de niveau inférieur. Pour Hoffman (2012), ce sont ces trois niveaux (représentations perceptuelles modales, intégration action/perception et action/cognition, influence top-down de la simulation sur la perception) qu'il faut implémenter dans nos robots interactifs pour permettre une coordination plus fluide entre humains et robots, aboutissant à des interactions structurées socialement. Les résultats futurs, liés à ce modèle inspiré de la notion de zones de convergence, seront à suivre avec intérêt.

- *Robots « Incarnés » : le début d'une réflexion philosophico-scientifique*

Comme les robots agissent toujours dans un environnement physique, il est important de regarder leur présence physique non pas comme un filtre bruyant parasitant le traitement d'autres informations mais au contraire comme un atout pour développer de réels systèmes intelligents et autonomes. Il est nécessaire que les robots utilisent leur corporéité, en particulier dans les interactions avec les humains, de manière similaire à celle dont le cerveau humain gère sa propre situation physique (Hoffman, 2012). Ce sont ces réflexions qui ont permis à la cognition incarnée de trouver une place de choix dans la robotique et qu'un tournant a été amorcé par rapport aux préoccupations plus traditionnelles de l'IA (la prise de décision, l'apprentissage, la planification…) qui fait que l'on cherche aujourd'hui à représenter des processus cognitifs plus complexes en se basant sur les théories de la cognition incarnée. La recherche dans le domaine de la cognition incarnée ne cessant d'évoluer, la robotique de demain pourra sûrement bénéficier des nouvelles découvertes de ce champ et, grâce aux principes de la cognition incarnée, d'autres processus pourront être modélisés avec succès tels que la communication non verbale (Todorov et al., 2006), apprendre à se déplacer en interagissant avec son environnement (Li et al., 2013), le raisonnement (Havasi, Speer, Pustejovsky, & Lieberman, 2009; Liu & Singh, 2004), etc.

L'application des principes de la cognition incarnée à n'importe quel système informatique reste en soi un défi, car tout est calcul et algorithme : du plus bas niveau jusqu'à des aspects symboliques, abstraits et modulaires (Hoffman, 2012). De nombreux points techniques restent à résoudre pour pouvoir implémenter concrètement un système inspiré de la cognition incarnée dans des agents robotiques. Cela dit, on peut voir l'intelligence artificielle et la robotique interactive, d'un point de vue plus large, comme une opportunité de tester les hypothèses psychologiques et neuroscientifiques de la cognition incarnée en les implémentant dans des systèmes robotiques. Cette approche se traduit par l'expression « comprendre le cerveau en créant le cerveau » qui a été inventée pour expliquer comment les robots humanoïdes et les neurosciences computationnelles pourraient contribuer à accroître les connaissances dans la psychologie humaine et la neurophysiologie sous-jacente (Asada et al, 2001 ; Brooks, 1997 ; Cheng et al. , 2007; Kawato, 2008). Encore une fois, l'association des travaux en robotique et neurosciences nous permettra de mieux comprendre le fonctionnement du cerveau, mais aussi qui nous sommes et comment nous interagissons les uns avec les autres, une grande partie des travaux actuels en robotique/neurosciences étant tournés vers la cognition sociale (Chaminade & Cheng, 2009). Cette réflexion nous permettra aussi d'aboutir à une meilleure compréhension des interactions humains-robots et ainsi d'accroître leur acceptation sociale, ce qui semble nécessaire s'ils se révèlent être nos compagnons dans un futur proche…

Enfin, comme le mentionne Knoll et al. (2001), il faudra peut-être revoir nos ambitions à la baisse et accepter que les robots ne soient peut-être jamais une réplique exacte de l'intelligence humaine, malgré tous nos efforts pour modéliser les processus qui caractérisent l'être humain. Comment évolueront les robots de demain ? Auront-ils une « intelligence » qui leur soit propre mais à laquelle nous ne pouvons pas accéder ? Seront-ils comme certains animaux à qui nous attribuons un degré d'intelligence et de communication certain (baleines, oiseaux…), mais que nous ne pouvons comprendre sans correspondance avec nos propres représentations et émotions ? Ou les robots humanoïdes partageront-ils un terrain identitaire avec nous ? Avec la robotique, tout un champ de nouveaux questionnements s'offre à nous : de quelle nature seront les « représentations » de ces robots, agents artificiels incarnés ? Y-aura-t-il des différences inter individuelles entre les robots ? Sachant que chaque robot aura un corps qui lui est personnel, mais aussi une expérience propre, comment se développeront leurs systèmes mnésiques sémantique, épisodique et autobiographique ? Pour répondre à ces questions un effort concerté entre roboticiens, neuroscientifiques, psychologues et philosophes sera probablement nécessaire dans le futur. Ces réflexions nous apporteront sûrement un éclairage nouveau sur les différentes théories de la relation corps-esprit, développées et débattues depuis plus de 2000 ans…

Références bibliographiques

Abrams, R. A., & Balota, D. A. (1991). Mental chronometry: Beyond reaction time. *Psychological Science*, *2*(3), 153-157.

Adamo, M., & Ferber, S. (2009). A picture says more than a thousand words: Behavioural and ERP evidence for attentional enhancements due to action affordances. *Neuropsychologia*, *47*(6), 1600-1608.

Afifi, A. K., & Bergman, R. A. (1998). *Functional neuroanatomy: text and atlas*(pp. 303-329). New York: McGraw-Hill.

Alario, F. X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments, & Computers*, *31*(3), 531-552.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in cognitive sciences*, *4*(7), 267-278.

Allport, D. A. (1985). Distributed memory, modular subsystems and dysphasia.*Current perspectives in dysphasia*, 32-60.

Altschuler, E. L., Vankov, A., Wang, V., Ramachandran, V. S., & Pineda, J. A. (1998). Person see, person do: Human cortical electrophysiological correlates of monkey see monkey do cells?. In *JOURNAL OF COGNITIVE NEUROSCIENCE* (Vol. 10, pp. 91-91).

Altschuler, E.L., Vankov, A., Hubbard, E.M., Roberts, E., Ramachandran, V.S., Pineda, J.A. (2000). Mu wave blocking by observer of movement and its possible use as a tool to study theory of other minds. *Annual Meeting of the Society for Neuroscience*

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268-277.

Andersen, R. A., Snyder, L. H., Bradley, D. C., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual review of neuroscience*, *20*(1), 303-330.

Anderson, J. R.(1983). The architecture of cognition. *Cambridge, Mass.: Halyard University Pres*.

Andrews-Hanna, J. R., Reidler, J. S., Huang, C., & Buckner, R. L. (2010). Evidence for the default network's role in spontaneous cognition. *Journal of Neurophysiology*, *104*(1), 322-335.

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron*,*65*(4), 550-562.

Anwander, A., Tittgemeyer, M., von Cramon, D. Y., Friederici, A. D., & Knösche, T. R. (2007). Connectivity-based parcellation of Broca's area.*Cerebral Cortex*, *17*(4), 816-825.

Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and brain sciences*,*28*(02), 105-124.

Arbib, M. A. (2010). Mirror system activity for action and language is embedded in the integration of dorsal and ventral pathways. *Brain and language*, *112*(1), 12-24.

Armstrong, D. F., & Wilcox, S. (2007). *The gestural origin of language*. Oxford: Oxford University Press.

Asada, M., MacDorman, K. F., Ishiguro, H., & Kuniyoshi, Y. (2001). Cognitive developmental robotics as a new paradigm for the design of humanoid robots.*Robotics and Autonomous Systems*, *37*(2), 185-193.

Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nature neuroscience*, *7*(5), 542-548.

Azizian, Allen, et al. "Beware misleading cues: Perceptual similarity modulates the N2/P3 complex." *Psychophysiology* 43.3 (2006): 253-260.

Aziz-Zadeh, L., & Damasio, A. (2008). Embodied semantics for actions: Findings from functional brain imaging. *Journal of Physiology-Paris*, *102*(1), 35-39.

Aziz-Zadeh, L., Fiebach, C. J., Naranayan, S., Feldman, J., Dodge, E., & Ivry, R. B. (2008). Modulation of the FFA and PPA by language related to faces and places. *Social neuroscience*, *3*(3-4), 229-238.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, *16*(18), 1818-1823.

Babiloni, C., Babiloni, F., Carducci, F., Cincotti, F., Cocozza, G., Del Percio, C., ... & Rossini, P. M. (2002). Human cortical electroencephalography (EEG) rhythms during the observation of simple aimless movements: a high-resolution EEG study. *Neuroimage*, *17*(2), 559-572.

Baddeley A. (1986). Working Memory. Oxford: Clarendon

Baggio, G., & Hagoort, P. (2011). The balance between memory and unification in semantics: A dynamic account of the N400. *Language and Cognitive Processes*, *26*(9), 1338-1367.

Bajo M. T. (1988). Semantic facilitation with pictures and words. *Journal of experimental psychology. Learning, memory, and cognition 14*, 579-89.

Bajo, M. T. (1988). Semantic facilitation with pictures and words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(4), 579.

Bak, T. H., & Hodges, J. R. (2003). Kissing and dancing—a test to distinguish the lexical and conceptual contributions to noun/verb and action/object dissociation. Preliminary results in patients with frontotemporal dementia.*Journal of Neurolinguistics*, *16*(2), 169-181.

Bak, T. H., O'Donovan, D. G., Xuereb, J. H., Boniface, S., & Hodges, J. R. (2001). Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the motor neurone disease–dementia–aphasia syndrome. *Brain*, *124*(1), 103-120.

Bak, T. H., Yancopoulou, D., Nestor, P. J., Xuereb, J. H., Spillantini, M. G., Pulvermüller, F., & Hodges, J. R. (2006). Clinical, imaging and pathological correlates of a hereditary deficit in verb and action processing. *Brain*, *129*(2), 321-332.

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., ... & Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage*, *30*(3), 917-926.

Bar M (2004) Visual objects in context. *Nat Rev Neurosci* 5(8): 617–629

Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1235-1243.

Barbas, H. (2000). Connections underlying the synthesis of cognition, memory, and emotion in primate prefrontal cortices. *Brain research bulletin*, *52*(5), 319-330.

Bar-Cohen, Y., & Breazeal, C. (2003). Biologically inspired intelligent robots. In *Smart Structures and Materials* (pp. 14-20). International Society for Optics and Photonics.

Barnes, M. A., Faulkner, H., Wilkinson, M., & Dennis, M. (2004). Meaning construction and integration in children with hydrocephalus. *Brain and Language*, *89*(1), 47-56.

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. MIT press.

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition*, *14*(2), 201-212.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and brain sciences*, *22*(04), 577-660.

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.*, *59*, 617-645.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction.*Philosophical Transactions of the Royal Society B: Biological Sciences*,*364*(1521), 1281-1289.

Barsalou, L. W., Kyle Simmons, W., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in cognitive sciences*, *7*(2), 84-91.

Bartels, A., & Zeki, S. (2004). Functional brain mapping during free viewing of natural scenes. *Human brain mapping*, *21*(2), 75-85.

Bartoshuk, L. M. (2000). Comparing sensory experiences across individuals: recent psychophysical advances illuminate genetic variation in taste perception.*Chemical Senses*, *25*(4), 447-460.

Bartoshuk, L. M., Duffy, V. B., Chapo, A. K., Fast, K., Yiee, J. H., Hoffman, H. J., ... & Snyder, D. J. (2004). From psychophysics to the clinic: missteps and advances. *Food Quality and Preference*, *15*(7), 617-632.

Bastiaansen, M. C., Van Der Linden, M., Ter Keurs, M., Dijkstra, T., & Hagoort, P. (2005). Theta responses are involved in lexical—Semantic retrieval during language processing. *Journal of Cognitive Neuroscience*, *17*(3), 530-541.

Bastiaansen, M., & Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Progress in brain research*, *159*, 179-196.

Bastiaansen, M., Oostenveld, R., Jensen, O., & Hagoort, P. (2008). I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. *Brain and Language*, *106*(1), 15-28.

Bastiaansen, M., Posthuma, D., Groot, P. F., & de Geus, E. J. (2002). Event-related alpha and theta responses in a visuo-spatial working memory task.*Clinical neurophysiology*, *113*(12), 1882-1893.

Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain.*Developmental psychobiology*, *40*(3), 293-310.

Bates, E., & Snyder, L. S. (1987). The cognitive hypothesis in language development. In E. Ina, C. Uzgiris, & E. J. McVicker Hunt (Eds.), Infant performance and experience (pp. 168–204). Urbana, IL: University of Illinois Press.

Beauchamp, M. S. (2005). See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Current opinion in neurobiology*,*15*(2), 145-153.

Beauchamp, M. S., & Martin, A. (2007). Grounding object concepts in perception and action: evidence from fMRI studies of tools. *Cortex*, *43*(3), 461-468.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron*, *34*(1), 149-159.

Beauvois M-F, Saillant B, Mhninger V, Llermitte F. (1978). Bilateral tactile aphasia: a tacto-verbal dysfunction. *Brain* 101:381–401

Beauvois M-F. (1982). Optic aphasia: a process of interaction between vision and language. *Philos. Trans. R. Soc. Lond*. B 298:35–47

Beck, A. (1890). Die Bestimmung der Localisation der Gehirn und Rückenmarkfunctionen vermittelst der elektrischen Erscheinungen. *Cent. Physiol*, *4*, 473-476.

Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., & Saxe, R. (2008). Concepts are more than percepts: The case of action verbs. *Journal of Neuroscience*, 28(44), 11347–11353.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and clinical Neurophysiology*, *60*(4), 343-355.

Bergen, B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, *31*(5), 733-764.

Bergen, B., & Wheeler, K. (2010). Grammatical aspect and mental simulation.*Brain and Language*, *112*(3), 150-158.

Berger H., (1929). Ueber das Elektroenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 1929, Berlin, 87: 527-570.

Berl, M. M., Balsamo, L. M., Xu, B., Moore, E. N., Weinstein, S. L., Conry, J. A., ... & Gaillard, W. D. (2005). Seizure focus affects regional language networks assessed by fMRI. *Neurology*, *65*(10), 1604-1611.

Berthoz, A. ; Andrieu, B. (2011). Le corps en acte – Centenaire Maurice Merleau Ponty. *Presses Universitaires de Nancy.*

Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in cognitive sciences*,*17*(2), 89-98.

Berwick, R. C., Okanoya, K., Beckers, G. J., & Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in cognitive sciences*, *15*(3), 113-121.

Besson, M., & Macar, F. (1987). An Event-Related Potential Analysis of Incongruity in Music and Other Non-Linguistic Contexts. *Psychophysiology*,*24*(1), 14-25.

Billard, A., & Dautenhahn, K. (1998). Grounding communication in autonomous robots: an experimental study. *Robotics and Autonomous Systems*, *24*(1), 71-79.

Billard, A., & Dautenhahn, K. (1999). Experiments in learning by imitation-grounding and use of communication in robotic agents. *Adaptive Behavior*, *7*(3-4), 415-438.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory.*Trends in cognitive sciences*, *15*(11), 527-536.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767-2796.

Binkofski, F., Buccino, G., Posse, S., Seitz, R. J., Rizzolatti, G., & Freund, H. J. (1999). A fronto-parietal circuit for object manipulation in man: evidence from an fMRI-study. *European Journal of Neuroscience*, *11*(9), 3276-3286.

Blanke, O., Landis, T., Safran, A. B., & Seeck, M. (2002). Direction-specific motion blindness induced by focal stimulation of human extrastriate cortex.*European Journal of Neuroscience*, *15*(12), 2043-2048.

Boecker, H., Dagher, A., Ceballos-Baumann, A. O., Passingham, R. E., Samuel, M., Friston, K. J., ... & Brooks, D. J. (1998). Role of the human rostral supplementary motor area and the basal ganglia in motor sequence control: investigations with H2 15O PET. *Journal of neurophysiology*, *79*(2), 1070-1080.

Boecker, H., Jankowski, J., Ditter, P., & Scheef, L. (2008). A role of the basal ganglia and midbrain nuclei for initiation of motor sequences. *Neuroimage*,*39*(3), 1356-1369.

Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *The Journal of Neuroscience*, *16*(11), 3737-3744.

Bookheimer, S. Y., Zeffiro, T. A., Blaxton, T., Gaillard, W., & Theodore, W. (1995). Regional cerebral blood flow during object naming and word reading.*Human Brain Mapping*, *3*(2), 93-106.

Borghi, A. M., & Cimatti, F. (2010). Embodied cognition and beyond: acting and sensing the body. *Neuropsychologia*, *48*(3), 763-773.

Borghi, A. M., & Scorolli, C. (2009). Language comprehension and dominant hand motion simulation. *Human Movement Science*, *28*(1), 12-27.

Borghi, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition*, *32*(6), 863-873.

Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT.*Annu. Rev. Neurosci.*, *28*, 157-189.

Boronat, C. B., Buxbaum, L. J., Coslett, H., Tang, K., Saffran, E. M., Kimberg, D. Y., & Detre, J. A. (2005). Distinctions between manipulation and function knowledge of objects: evidence from functional magnetic resonance imaging.*Cognitive Brain Research*, *23*(2), 361-373.

Borreggine, K. L., & Kaschak, M. P. (2006). The action–sentence compatibility effect: It's all in the timing. *Cognitive Science*, *30*(6), 1097-1112.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., et al. (1994). The role of the right hemisphere in the interpretation of Wgu- rative aspects of language. A positron emission tomography activation study. *Brain*, 117(6), 1241–1253.

Botzung, A., Denkova, E., Ciuciu, P., Scheiber, C., Manning, L., (2008). The neural bases of the constructive nature of autobiographical memories studied with a self-paced fMRI design. *Memory* 16, 351–363.

Botzung, A., Denkova, E., Manning, L., (2008). Experiencing past and future personal events: functional neuroimaging evidence on the neural bases of mental time trav- el. *Brain Cogn*. 66, 202–212.

Boulenger, V., Mechtouff, L., Thobois, S., Broussolle, E., Jeannerod, M., & Nazir, T. A. (2008). Word processing in Parkinson's disease is impaired for action verbs but not for concrete nouns. *Neuropsychologia*, *46*(2), 743-756.

Boulenger, V., Roy, A. C., Paulignan, Y., Deprez, V., Jeannerod, M., & Nazir, T. A. (2006). Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *Journal of cognitive neuroscience*, *18*(10), 1607-1615.

Boulenger, V., Shtyrov, Y., & Pulvermüller, F. (2012). When do you grasp the idea? MEG evidence for instantaneous idiom understanding. *Neuroimage*, *59*(4), 3502-3513.

Bouma, H., & De Voogd, A. H. (1974). On the control of eye saccades in reading. *Vision Research*, *14*(4), 273-284.

Bowers, J. S. (2009). On the biological plausibility of grandmother cells: implications for neural network theories in psychology and neuroscience.*Psychological review*, *116*(1), 220.

Braeutigam, S., Bailey, A. J., & Swithenby, S. J. (2001). Phase-locked gamma band responses to semantic violation stimuli. *Cognitive brain research*, *10*(3), 365-377.

Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, *59*(1), 119-155.

Breazeal, C. (2003). Toward sociable robots. *Robotics and Autonomous Systems*, *42*(3), 167-175.

Breazeal, C. (2004). Social interactions in HRI: the robot view. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*,*34*(2), 181-186.

Breazeal, C. (2011). Social robots for health applications. In*Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE* (pp. 5368-5371). IEEE.

Breazeal, C. L. (2000). *Sociable machines: expressive social exchange between humans and robots* (Doctoral dissertation, Massachusetts Institute of Technology).

Breazeal, C., Buchsbaum, D., Gray, J., Gatenby, D., & Blumberg, B. (2005). Learning from and about others: Towards using imitation to bootstrap the social understanding of others by robots. *Artificial Life*, *11*(1-2), 31-62.

Breazeal, C., Gray, J., & Berlin, M. (2009). An embodied cognition approach to mindreading skills for socially intelligent robots. *The International Journal of Robotics Research*, *28*(5), 656-680.

Breazeal, C., Siegel, M., Berlin, M., Gray, J., Grupen, R., Deegan, P., ... & McBean, J. (2008). Mobile, dexterous, social robots for mobile manipulation and human-robot interaction. In *ACM SIGGRAPH 2008 new tech demos* (p. 27). ACM.

Bright, P., Moss, H., & Tyler, L. K. (2004). Unitary vs multiple semantics: PET studies of word and picture processing. *Brain and language*, *89*(3), 417-432.

Brooks, R.A., (1997). The cog project. Advanced robotics. J. Robot. Soc. Jpn. 15, 968– 970.

Brown, W. S., Marsh, J. T., & Smith, J. C. (1973). Contextual meaning effects on speech-evoked potentials. *Behavioral Biology*, 9(6), 755-761.

Brown, W. S., Marsh, J. T., & Smith, J. C. (1976). Evoked potential waveform differences produced by the perception of different meanings of an ambiguous phrase. *Electroencephalography and clinical Neurophysiology*, *41*(2), 113-123.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., ... & Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European journal of neuroscience*,*13*(2), 400-404.

Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., ... & Rizzolatti, G. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics: An fMRI study. *Journal of Cognitive Neuroscience*, *16*(1), 114-126.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Cognitive Brain Research*,*24*(3), 355-363.

Büchel, C., & Friston, K. (2000). Assessing interactions among neuronal systems using functional neuroimaging. *Neural Networks*, *13*(8), 871-882.

Büchel, C., & Friston, K. J. (1997). Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cerebral Cortex*, *7*(8), 768-778.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in cognitive sciences*, *11*(2), 49-57.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences*, *1124*(1), 1-38.

Buckner, R. L., Raichle, M. E., Miezin, F. M., & Petersen, S. E. (1996). Functional anatomic studies of memory retrieval for auditory words and visual pictures. *The Journal of neuroscience*, *16*(19), 6219-6235.

Burgess, N., Becker, S., King, J. A., & O'Keefe, J. (2001). Memory for events and their spatial context: models and experiments. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1413), 1493-1503.

Burianova, H., McIntosh, A. R., & Grady, C. L. (2010). A common functional brain network for autobiographical, episodic, and semantic memory retrieval.*Neuroimage*, *49*(1), 865-874.

Byrne, P., Becker, S., & Burgess, N. (2007). Remembering the past and imagining the future: a neural model of spatial memory and imagery.*Psychological review*, *114*(2), 340.

C. Breazeal (2007). "Sociable Robots." *Journal of the Robotics Society of Japan,* 24(5), 591-593.

Cabeza, R., & St Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends in cognitive sciences*, *11*(5), 219-227.

Cain, K. (2003). Text comprehension and its relation to coherence and cohesion in children's fictional narratives. *British Journal of Developmental Psychology*, *21*(3), 335-351.

Cain, K. (2006). Individual differences in children's memory and reading comprehension: An investigation of semantic and inhibitory deficits. *Memory*,*14*(5), 553-569.

Cain, K., & Oakhill, J. (2004). Reading comprehension difficulties. In *Handbook of children's literacy* (pp. 313-338). Springer Netherlands.

Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology*, *76*(4), 683-696.

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and writing*, *11*(5-6), 489-503.

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, *29*(6), 850-859.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of educational psychology*, *96*(1), 31.

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: the influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of educational psychology*, *96*(4), 671.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., ... & David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science*, *276*(5312), 593-596.

Calvo-Merino, B., Ehrenberg, S., Leung, D., & Haggard, P. (2010). Experts see it all: Configural effects in action observation. *Psychological Research PRPF*,*74*(4), 400-406.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: an FMRI study with expert dancers. *Cerebral cortex*, *15*(8), 1243-1249.

Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology*, *16*(19), 1905-1910.

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory and Language*, *56*(1), 103-128.

Canessa, N., Borgo, F., Cappa, S. F., Perani, D., Falini, A., Buccino, G., ... & Shallice, T. (2008). The different neural correlates of action and functional knowledge in semantic memory: an FMRI study. *Cerebral Cortex*, *18*(4), 740-751.

Cangelosi, A., & Riga, T. (2006). An embodied model for sensorimotor grounding and grounding transfer: Experiments with epigenetic robots. *Cognitive science*, *30*(4), 673-689.

Capitani, E., Laiacona, M., Mahon, B., & Caramazza, A. (2003). What are the facts of semantic category-specific deficits? A critical review of the clinical evidence. *Cognitive Neuropsychology*, *20*(3-6), 213-261.

Cappa, S. F. (2008). Imaging studies of semantic memory. *Current opinion in neurology*, *21*(6), 669-675.

Caramazza, A. (1996). Pictures, words and the brain. Nature 383: 216–217.

Caramazza, A., & Mahon, B. Z. (2003). The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends in cognitive sciences*, *7*(8), 354-361.

Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *Journal of cognitive neuroscience*,*10*(1), 1-34.

Caramazza, A., Hillis, A. E., Rapp, B. C., & Romani, C. (1990). The multiple semantics hypothesis: Multiple confusions?. *Cognitive Neuropsychology*, *7*(3), 161-189.

Carey, D.P., Perrett, D.I. & Oram, M.W. (1997) Recognizing, Understanding, and Producing Action, in Handbook of Neuropsychology: Action and Cognition Vol. 11 (Jeannerod, M. & Grafman, J., Eds.) Amsterdam: Elsevier, pp.111–130

Carmi, R., & Itti, L. (2006). Visual causes versus correlates of attentional selection in dynamic scenes. *Vision research*, *46*(26), 4333-4345.

Carmichael, S. T., & Price, J. L. (1995). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *Journal of Comparative Neurology*, *363*(4), 615-641.

Carpenter, M., Tomasello, M., & Striano, T. (2005). Role reversal imitation and language in typically developing infants and children with autism. *Infancy*, *8*(3), 253-278.

Carter, E. J., Hodgins, J. K., & Rakison, D. H. (2011). Exploring the neural correlates of goal-directed action and intention understanding. *Neuroimage*,*54*(2), 1634-1642.

Carter, R. (2010). *The human brain book*. Penguin.

Casey, B. J., Epstein, J., Buhle, J., Liston, C., Davidson, M., Tonev, S., ... & Glover, G. (2007). Frontostriatal connectivity and its role in cognitive control in parent-child dyads with ADHD. *American Journal of Psychiatry*, *164*(11), 1729-1736.

Castriota-Scanderbeg, A., Hagberg, G. E., Cerasa, A., Committeri, G., Galati, G., Patria, F., ... & Frackowiak, R. (2005). The appreciation of wine by sommeliers: a functional magnetic resonance study of sensory integration.*Neuroimage*, *25*(2), 570-578.

Catani, M., & Jones, D. K. (2005). Perisylvian language networks of the human brain. *Annals of neurology*, *57*(1), 8-16.

Catani, M., & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*, *44*(8), 1105-1132.

Catani, M., Howard, R. J., Pajevic, S., & Jones, D. K. (2002). Virtual in Vivo Interactive Dissection of White Matter Fasciculi in the Human Brain.*Neuroimage*, *17*(1), 77-94.

Caton, R. (1875). Electrical currents of the brain. *The Journal of Nervous and Mental Disease*, *2*(4), 610.

Cattaneo, L., Fasanelli, M., Andreatta, O., Bonifati, D. M., Barchiesi, G., & Caruana, F. (2012). Your actions in my cerebellum: subclinical deficits in action observation in patients with unilateral chronic cerebellar stroke. *The Cerebellum*, *11*(1), 264-271.

Chaminade, T., & Cheng, G. (2009). Social cognitive neuroscience and humanoid robotics. *Journal of Physiology-Paris*, *103*(3), 286-295.

Chaminade, T., & Decety, J. (2002). Leader or follower? Involvement of the inferior parietal lobule in agency. *Neuroreport*, *13*(15), 1975-1978.

Changeux, J.-P. (1983). L'homme neuronal. *Paris, Fayard.*

Chao, L. L., & Martin, A. (1999). Cortical regions associated with perceiving, naming, and knowing about colors. *Journal of Cognitive Neuroscience*, *11*(1), 25-35.

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage*, *12*(4), 478-484.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature neuroscience*, *2*(10), 913-919.

Chapman, R. M., & Bragdon, H. R. (1964). Evoked responses to numerical and non-numerical visual stimuli while problem solving.

Chatterjee, A., Southwood, M. H., & Basilico, D. (1999). Verbs, events and spatial representations. *Neuropsychologia*, *37*(4), 395-402.

Chee, M. W., O'Craven, K. M., Bergida, R., Rosen, B. R., & Savoy, R. L. (1999). Auditory and visual word processing studied with fMRI. *Human Brain Mapping*, *7*(1), 15-28.

Cheng, G., Hyon, S. H., Morimoto, J., Ude, A., Hale, J. G., Colvin, G., ... & Jacobsen, S. C. (2007). CB: A humanoid research platform for exploring neuroscience. *Advanced Robotics*, *21*(10), 1097-1114.

Chertkow, H., Bub, D., Deaudon, C., & Whitehead, V. (1997). On the status of object concepts in aphasia. *Brain and language*, *58*(2), 203-232.

Childers, J. B., & Tomasello, M. (2006). Are Nouns Easier to Learn Than Verbs? Three Experimental Studies. *Action meets word: How children learn verbs*, 311.

Chong, T. T. J., Cunnington, R., Williams, M. A., Kanwisher, N., & Mattingley, J. B. (2008). fMRI adaptation reveals mirror neurons in human inferior parietal cortex. *Current biology*, *18*(20), 1576-1580.

Cochin, S., Barthelemy, C., Lejeune, B., Roux, S., & Martineau, J. (1998). Perception of motion and qEEG activity in human adults.*Electroencephalography and Clinical Neurophysiology*, *107*(4), 287-295.

Cochin, S., Barthelemy, C., Roux, S., & Martineau, J. (1999). Observation and execution of movement: similarities demonstrated by quantified electroencephalography. *European Journal of Neuroscience*, *11*(5), 1839-1842.

Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. Neuroimage, 40(1), 353-366.

Cohen-Seat, G., Gastaut, H., Faure, J., & Heuyer, G. (1954). Etudes expérimentales de l'activité nerveuse pendant la projection cinématographique.*Rev. Int. Filmologie*, *5*, 7-64.

Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex.*Annual review of neuroscience*, *22*(1), 319-349.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review*, *82*(6), 407.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory.*Journal of verbal learning and verbal behavior*, *8*(2), 240-247.

Coltheart, M. (1978). Lexical access in simple reading tasks.

Coltheart, M. (2005). Modeling Reading: The Dual-Route Approach.

Comerchero, M. D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology*, *110*(1), 24-30.

Corballis, M. C. (2002). *From hand to mouth: The origins of language*. Princeton University Press.

Corballis, M. C. (2004). The origins of modernity: was autonomous speech the critical factor?. *Psychological Review*, *111*(2), 543.

Corballis, M. C. (2010). Mirror neurons and the evolution of language. *Brain and language*, *112*(1), 25-35.

Coslett, H. (1998). Evidence for a disturbance of the body schema in neglect.*Brain and cognition*, *37*(3), 527-544.

Cotelli, M., Borroni, B., Manenti, R., Alberici, A., Calabria, M., Agosti, C., ... & Cappa, S. F. (2006). Action and object naming in frontotemporal dementia, progressive supranuclear palsy, and corticobasal degeneration.*Neuropsychology*, *20*(5), 558.

Coull, J. T. (1998). Neural correlates of attention and arousal: insights from electrophysiology, functional neuroimaging and psychopharmacology. *Progress in neurobiology*, *55*(4), 343-361.

Coulson ,S., &Williams, R. F. (2005). Hemispheric asymmetries and joke comprehension. *Neuropsychologia*, *43*(1), 128-141.

Coulson, S., & Kutas, M. (2001). Getting it: Human event-related brain response to jokes in good and poor comprehenders. *Neuroscience letters*,*316*(2), 71-74.

Coulson, S., & Lovett, C. (2004). Handedness, hemispheric asymmetries, and joke comprehension. *Cognitive Brain Research*, *19*(3), 275-288.

Coulson, S., Federmeier, K. D., Van Petten, C., & Kutas, M. (2005). Right hemisphere sensitivity to word-and sentence-level context: evidence from event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 129.

Courchesne, E., Hillyard, S. A., & Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography and clinical neurophysiology*, *39*(2), 131-143.

Coynel, D., Marrelec, G., Perlbarg, V., Pélégrini-Issac, M., Van de Moortele, P. F., Ugurbil, K., ... & Lehéricy, S. (2010). Dynamics of motor-related functional integration during motor sequence learning. *Neuroimage*, *49*(1), 759-766.

Creem-Regehr, S. H., & Lee, J. N. (2005). Neural representations of graspable objects: are tools special?. *Cognitive Brain Research*, *22*(3), 457-469.

Cross, E. S., Hamilton, A. F. D. C., & Grafton, S. T. (2006). Building a motor simulation de novo: observation of dance by dancers. *Neuroimage*, *31*(3), 1257-1267.

Dale, A.M., Buckner, R.L. (1997). Selective averaging of rapidly presented individual trials using fMRI. Human *Brain Mapping* ; 5:329 -- 40.

Damasio, A. R. (1989). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation*, *1*(1), 123-132.

Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*,*33*(1), 25-62.

Damasio, A. R., & Damasio, H. (1994). Cortical systems for retrieval of concrete knowledge: The convergence zone framework. *Large-scale neuronal theories of the brain*, 61-74.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422-433.

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2005). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature neuroscience*, *9*(1), 28-30.

D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology*,*19*(5), 381-385.

Dautenhahn, K. (1995). Getting to know each other—artificial social intelligence for autonomous robots. *Robotics and autonomous systems*, *16*(2), 333-356.

Dautenhahn, K. (2001). The narrative intelligence hypothesis: In search of the transactional format of narratives in humans and other social animals. In*Cognitive Technology: Instruments of Mind* (pp. 248-266). Springer Berlin Heidelberg.

Dautenhahn, K. (2003). Stories of lemurs and robots: The social origin of story-telling. *Narrative intelligence*, *46*, 63.

Dautenhahn, K. (2007). Socially intelligent robots: dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 679-704.

Dautenhahn, K. (2012). Social Robots as Assistants. *Sponsoring Institutions*, 4.

Dautenhahn, K., & Nehaniv, C. L. (2002). *The agent-based perspective on imitation* (pp. 1-40). MIT Press.

Dautenhahn, K., & Saunders, J. (2011). *New frontiers in human-robot interaction* (Vol. 2). John Benjamins.

Dautenhahn, K., & Werry, I. (2004). Towards interactive robots in autism therapy: Background, motivation and challenges. Pragmatics & Cognition, 12(1), 1-35.

Dautenhahn, K., Ogden, B., & Quick, T. (2002). From embodied to socially embedded agents–implications for interaction-aware robots. *Cognitive Systems Research*, *3*(3), 397-428.

Davenport, T., & Coulson, S. (2011). Predictability and novelty in literal language comprehension: An ERP study. *Brain research*, *1418*, 70-82.

David, N., Bewernick, B. H., Cohen, M. X., Newen, A., Lux, S., Fink, G. R., ... & Vogeley, K. (2006). Neural representations of self versus other: visual-spatial perspective taking and agency in a virtual ball-tossing game. *Journal of Cognitive Neuroscience*, *18*(6), 898-910.

David, N., Cohen, M. X., Newen, A., Bewernick, B. H., Shah, N. J., Fink, G. R., & Vogeley, K. (2007). The extrastriate cortex distinguishes between the consequences of one's own and others' behavior. *Neuroimage*, *36*(3), 1004-1014.

Davidson, D. J., & Indefrey, P. (2007). An inverse relation between event-related and time–frequency violation responses in sentence processing. *Brain research*, *1158*, 81-92.

Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *The Journal of Neuroscience*, *23*(8), 3423-3431.

Davis, M. H., Meunier, F., & Marslen-Wilson, W. D. (2004). Neural responses to morphological, syntactic, and semantic properties of single words: an fMRI study. *Brain and language*, *89*(3), 439-449.

De Beni Paola Palladino Francesca Pazzaglia Cesare Cornoldi, R. (1998). Increases in intrusion errors and working memory deficit of poor comprehenders. *The Quarterly Journal of Experimental Psychology: Section A*,*51*(2), 305-320.

De Beni, Palladino, Pazzaglia, Cornoldi, R. (1998). Increases in intrusion errors and working memory deficit of poor comprehenders. *The Quarterly Journal of Experimental Psychology: Section A*,*51*(2), 305-320.

De Beni, R., & Palladino, P. (2000). Intrusion errors in working memory tasks: Are they related to reading comprehension ability?. *Learning and Individual Differences*, *12*(2), 131-143.

De Borst, A. W., Sack, A. T., Jansma, B. M., Esposito, F., De Martino, F., Valente, G., ... & Formisano, E. (2012). Integration of “what” and “where” in frontal cortex during visual imagery of scenes. *Neuroimage*, *60*(1), 47-58.

De Jong, P. F., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading*, *6*(1), 51-77.

De Vega, M. (2008). Levels of embodied meaning: From pointing to counterfactuals. *Symbols and embodiment. Debates on meaning and cognition*, 285-308.

De Vignemont, F., & Singer, T. (2006). The empathic brain: how, when and why?. *Trends in cognitive sciences*, *10*(10), 435-441.

De Witt Hamer, P. C., Moritz-Gasser, S., Gatignol, P., & Duffau, H. (2011). Is the human left middle longitudinal fascicle essential for language? A brain electrostimulation study. *Human brain mapping*, *32*(6), 962-973.

Deacon, D., Breton, F., Ritter, W., & Vaughan, H. G. (1991). The relationship between N2 and N400: Scalp distribution, stimulus probability, and task relevance. *Psychophysiology*, *28*(2), 185-200.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., Von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *The Journal of Neuroscience*, *25*(50), 11730-11737.

Decety, J., & Grèzes, J. (2006). The power of simulation: imagining one's own and other's behavior. *Brain research*, *1079*(1), 4-14.

Decety, J., & Lindgren, M. (1991). Sensation of effort and duration of mentally executed actions. *Scandinavian journal of psychology*, *32*(2), 97-104.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., ... & Fazio, F. (1997). Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain*, *120*(10), 1763-1777.

Decety, J., Jeannerod, M., Durozard, D., & Baverel, G. (1993). Central activation of autonomic effectors during mental simulation of motor actions in man. *The Journal of Physiology*, *461*(1), 549-563.

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., ... & Fazio, F. (1994). Mapping motor representations with positron emission tomography.

Dehaene, S. (2007). *Les neurones de la lecture*. Odile Jacob.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in cognitive sciences*, *15*(6), 254-262.

Dehaene-Lambertz, G., Dehaene, S., Hertz-Pannier, L. (2002) Functional neuroimaging of speech perception in infants. *Sciences* ;298: 2013-5

DeLong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends in neurosciences*, *13*(7), 281-285.

Delorme, A., Richard, G., & Fabre-Thorpe, M. (2000). Ultra-rapid categorisation of natural scenes does not rely on colour cues: a study in monkeys and humans. *Vision research*, *40*(16), 2187-2200.

Demb, J. B., Desmond, J. E., Wagner, A. D., Vaidya, C. J., Glover, G. H., & Gabrieli, J. D. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: a functional MRI study of task difficulty and process specificity. *Journal of Neuroscience*, *15*(9), 5870-5878.

Demiral, Ş. B., Malcolm, G. L., & Henderson, J. M. (2012). ERP correlates of spatially incongruent object identification during scene viewing: Contextual expectancy versus simultaneous processing. *Neuropsychologia*, *50*(7), 1271-1285.

Démonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., ... & Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, *115*(6), 1753-1768.

Démonet, J. F., Thierry, G., & Cardebat, D. (2005). Renewal of the neurophysiology of language: functional neuroimaging. *Physiological reviews*,*85*(1), 49-95.

Denis, M., & Golonelli, C. (1976). Mémorisation de stimulus dessins ou noms en fonction de la nature du matériel utilisé pour l'épreuve de reconnaissance.*L'année psychologique*, *76*(1), 39-53.

Denis, M., & Kosslyn, S. M. (1999). Scanning visual mental images: A window on the mind. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*.

Desai, R. H., Binder, J. R., Conant, L. L., & Seidenberg, M. S. (2010). Activation of sensory–motor areas in sentence comprehension. *Cerebral Cortex*, *20*(2), 468-478.

D'Esposito, M., Postle, B. R., & Rypma, B. (2000). Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies.*Experimental Brain Research*, *133*(1), 3-11.

D'Esposito, M., Postle, B. R., Ballard, D., & Lease, J. (1999). Maintenance versus manipulation of information held in working memory: an event-related fMRI study. *Brain and cognition*, *41*(1), 66-86.

Devlin, J. T., & Watkins, K. E. (2007). Stimulating language: insights from TMS.*Brain*, *130*(3), 610-622.

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Moss, H. E., Fadili, M. J., & Tyler, L. K. (2002). Is there an anatomical basis for category-specificity? Semantic memory studies in PET and fMRI. *Neuropsychologia*, *40*(1), 54-75.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental brain research*, *91*(1), 176-180.

Dien, J., Spencer, K. M., & Donchin, E. (2004). Parsing the late positive complex: mental chronometry and the ERP components that inhabit the neighborhood of the P300. *Psychophysiology*, *41*(5), 665-678.

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: analyses and review. *Journal of Experimental Psychology: General*, *140*(4), 552.

Dominey, P. F. (2005). Emergence of grammatical constructions: evidence from simulation and grounded agent experiments. *Connection Science*, *17*(3-4), 289-306.

Dominey, P. F., & Boucher, J. D. (2005). Developmental stages of perception and language acquisition in a perceptually grounded robot. *Cognitive Systems Research*, *6*(3), 243-259.

Dominey, P. F., & Warneken, F. (2011). The basis of shared intentions in human and robot cognition. *New Ideas in Psychology*, *29*(3), 260-274.

Dominey, P. F., Hoen, M., & Inui, T. (2006). A neurolinguistic model of grammatical construction processing. *Journal of Cognitive Neuroscience*,*18*(12), 2088-2107.

Dominey, P. F., Metta, G., Nori, F., & Natale, L. (2008). Anticipation and initiative in human-humanoid interaction. In *Humanoid Robots, 2008. Humanoids 2008. 8th IEEE-RAS International Conference on* (pp. 693-699). IEEE.

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Harvard University Press.

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating?. *Behavioral and brain sciences*, *11*(03), 357-374.

Donchin, E. (1981). Surprise!… surprise?. *Psychophysiology*, *18*(5), 493-513.

Downing, P. E., Chan, A. Y., Peelen, M. V., Dodds, C. M., & Kanwisher, N. (2006). Domain specificity in visual cortex. *Cerebral cortex*, *16*(10), 1453-1461.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*(5539), 2470-2473.

Duffau, H. (2005). Lessons from brain mapping in surgery for low-grade glioma: insights into associations between tumour and brain plasticity. *The Lancet Neurology*, *4*(8), 476-486.

Duffau, H., Capelle, L., Sichez, N., Denvil, D., Lopes, M., Sichez, J. P., ... & Fohanno, D. (2002). Intraoperative mapping of the subcortical language pathways using direct stimulations An anatomo-functional study. *Brain*, *125*(1), 199-214.

Dum, R. P., & Strick, P. L. (2002). Motor areas in the frontal lobe of the primate. *Physiology & behavior*, *77*(4), 677-682.

Dumoulin, S. O., Bittar, R. G., Kabani, N. J., Baker, C. L., Le Goualher, G., Pike, G. B., & Evans, A. C. (2000). A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning.*Cerebral Cortex*, *10*(5), 454-463.

Dunn, B. R., Dunn, D. A., Languis, M., & Andrews, D. (1998). The relation of ERP components to complex memory processing. *Brain and cognition*, *36*(3), 355-376.

Dupont, P., Orban, G. A., De Bruyn, B., Verbruggen, A., & Mortelmans, L. (1994). Many areas in the human brain respond to visual motion. *Journal of Neurophysiology*, *72*(3), 1420-1424.

Ehrsson, H. H., Fagergren, A., Jonsson, T., Westling, G., Johansson, R. S., & Forssberg, H. (2000). Cortical activity in precision-versus power-grip tasks: an fMRI study. *Journal of Neurophysiology*, *83*(1), 528-536.

Ellmore, T. M., Beauchamp, M. S., Breier, J. I., Slater, J. D., Kalamangalam, G. P., O'Neill, T. J., ... & Tandon, N. (2010). Temporal lobe white matter asymmetry and language laterality in epilepsy patients. *Neuroimage*, *49*(3), 2033-2044.

Emmorey, K. (2001). *Language, cognition, and the brain: Insights from sign language research*. Psychology Press.

Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S., Wiebe, V., Kitano, T., ... & Pääbo, S. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*, *418*(6900), 869-872.

Enrici, I., Adenzato, M., Cappa, S., Bara, B. G., & Tettamanti, M. (2011). Intention processing in communication: a common brain network for language and gestures. *Journal of Cognitive Neuroscience*, *23*(9), 2415-2431.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598-601.

Evans, K., and. Federmeier, K (2008). "Uncovering the functional significance of the P2 repetition effect." *PSYCHOPHYSIOLOGY*. Vol. 45.

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes.*Journal of cognitive neuroscience*, *13*(2), 171-180.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study.*European Journal of Neuroscience*, *15*(2), 399-402.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of neurophysiology*, *73*(6), 2608-2611.

Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: modality specificity and emergent category specificity.*Journal of Experimental Psychology: General*, *120*(4), 339.

Farrer, C., & Frith, C. D. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency.*Neuroimage*, *15*(3), 596-603.

Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A. C., Pozzo, T., ... & Fadiga, L. (2009). Encoding of human action in Broca's area. *Brain*, *132*(7), 1980-1988.

Federmeier, K. D., & Kutas, M. (1999). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research*, *8*(3), 373-392.

Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, *42*(2), 133-141.

Federmeier, K. D., & Laszlo, S. (2009). Time for meaning: Electrophysiology provides insights into the dynamics of representation and processing in semantic memory. *Psychology of learning and motivation*, *51*, 1-44.

Federmeier, K. D., McLennan, D. B., Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study.*Psychophysiology*, *39*(2), 133-146.

Federmeier, K. D., Van Petten, C., Schwartz, T. J., & Kutas, M. (2003). Sounds, words, sentences: age-related changes across levels of language processing. *Psychology and aging*, *18*(4), 858.

Fell, J., Klaver, P., Lehnertz, K., Grunwald, T., Schaller, C., Elger, C. E., & Fernández, G. (2001). Human memory formation is accompanied by rhinal–hippocampal coupling and decoupling. *Nature neuroscience*, *4*(12), 1259-1264.

Ferrari, P. F., Visalberghi, E., Paukner, A., Fogassi, L., Ruggiero, A., & Suomi, S. J. (2006). Neonatal imitation in rhesus macaques. *PLoS biology*, *4*(9), e302.

Ferreira, C. T., Giusiano, B., Ceccaldi, M., & Poncet, M. (1997). Optic aphasia: Evidence of the contribution of different neural systems to object and action naming. *Cortex*, *33*(3), 499-513.

Ferretti, T. R., Kutas, M., & McRae, K. (2007). Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(1), 182.

Fiez, J. A., Raichle, M. E., Balota, D. A., Tallal, P., & Petersen, S. E. (1996). PET activation of posterior temporal regions during auditory word presentation and verb generation. *Cerebral cortex*, *6*(1), 1-10.

Fiez, J. A., Raichle, M. E., Miezin, F. M., Petersen, S. E., Tallal, P., & Katz, W. F. (1995). PET studies of auditory and phonological processing: Effects of stimulus characteristics and task demands. *Journal of Cognitive Neuroscience*,*7*(3), 357-375.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: a review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology*, *61*(6), 825-850.

Fisher, J. E., Cortes, C. R., Griego, J. A., & Tagamets, M. A. (2012). Repetition of letter strings leads to activation of and connectivity with word-related regions.*Neuroimage*, *59*(3), 2839-2849.

Flöel, A., de Vries, M. H., Scholz, J., Breitenstein, C., & Johansen-Berg, H. (2009). White matter integrity in the vicinity of Broca's area predicts grammar learning success. *Neuroimage*, *47*(4), 1974-1981.

Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology*, *15*(5), 447-452.

Fodor, J. A. (1975). *The language and thought*. Harvard University Press.

Fogassi, L., & Ferrari, P. F. (2007). Mirror neurons and the evolution of embodied language. *Current Directions in Psychological Science*, *16*(3), 136-141.

Folstein, J. R., & Van Petten, C. (2011). After the P3: Late executive processes in stimulus categorization. *Psychophysiology*, *48*(6), 825-841.

Folstein, J., and Van Petten, V. (2008): "Influence of cognitive control and mismatch on the N2 component of the ERP: a review." *Psychophysiology* 45.1 152-170.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and autonomous systems*, *42*(3), 143-166.

Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *The Quarterly Journal of Experimental Psychology*, *33*(4), 465-495.

Freud S. 1891 (1953). On Aphasia. Transl. E Stengel. New York: Int. Univ. Press

Freunberger, R., Klimesch, W., Doppelmayr, M., & Höller, Y. (2007). Visual P2 component is related to theta phase-locking. *Neuroscience letters*, *426*(3), 181-186.

Frey, S. H., & Gerry, V. E. (2006). Modulation of neural activity during observational learning of actions and their sequential orders. *The Journal of neuroscience*, *26*(51), 13194-13201.

Frey, S., Campbell, J. S., Pike, G. B., & Petrides, M. (2008). Dissociating the human language pathways with high angular resolution diffusion fiber tractography. *The Journal of Neuroscience*, *28*(45), 11435-11444.

Friederici, A. D. (2006). What's in control of language?. *Nature neuroscience*,*9*(8), 991-992.

Friederici, A. D., & Gierhan, S. M. (2012). The language network. *Current opinion in neurobiology*.

Friederici, A. D., Steinhauer, K., & Frisch, S. (1999). Lexical integration: Sequential effects of syntactic and semantic information. *Memory & Cognition*,*27*(3), 438-453.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, *9*(10), 474-480.

Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annual review of neuroscience*, *32*, 209-224.

Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*,*291*(5508), 1560-1563.

Fries, P., Roelfsema, P. R., Engel, A. K., König, P., & Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proceedings of the National Academy of Sciences*, *94*(23), 12699-12704.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling.*Neuroimage*, *19*(4), 1273-1302.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 459-473.

Fuster, J. (2008). *The prefrontal cortex*. Access Online via Elsevier.

Gaillard, W. D., Balsamo, L., Xu, B., McKinney, C., Papero, P. H., Weinstein, S., ... & Theodore, W. H. (2004). fMRI language task panel improves determination of language dominance. *Neurology*, *63*(8), 1403-1408.

Galaburda, A. M., Corsiglia, J., Rosen, G. D., & Sherman, G. F. (1987). Planum temporale asymmetry, reappraisal since Geschwind and Levitsky.*Neuropsychologia*, *25*(6), 853-868.

Gall, F. J. (1825). Sur les fonctions du cerveau et sur celles de chacune de ses parties,... *J.-B. Baillière.*

Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of 'theory of mind'in verbal and nonverbal tasks. *Neuropsychologia*, *38*(1), 11-21.

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in cognitive sciences*, *4*(1), 14-21.

Gallese, V. (2003). A neuroscientific grasp of concepts: from control to representation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1435), 1231-1240.

Gallese, V. (2003). The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology*, *36*(4), 171-180.

Gallese, V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 659-669.

Gallese, V. (2008). Mirror neurons and the social nature of language: the neural exploitation hypothesis. *Social neuroscience*, *3*(3-4), 317–333.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in cognitive sciences*, *2*(12), 493-501.

Gallese, V., & Lakoff, G. (2005). The brain's concept's: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 21*, 455–479.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593-609.

Gangitano, M., Mottaghy, F. M., & Pascual-Leone, A. (2001). Phase-specific modulation of cortical motor output during movement observation. *Neuroreport*,*12*(7), 1489-1492.

Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Cognitive Brain Research*, *16*(2), 123-144.

Ganis, G., Keenan, J. P., Kosslyn, S. M., & Pascual-Leone, A. (2000). Transcranial magnetic stimulation of primary motor cortex affects mental rotation. *Cerebral Cortex*, *10*(2), 175-180.

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for "common sense": An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience*, *8*(2), 89-106.

Gastaut, H. J., & Bert, J. (1954). EEG changes during cinematographic presentation (Moving picture activation of the EEG). *Electroencephalography and clinical neurophysiology*, *6*, 433-444.

Gates, L., & Yoon, M. G. (2005). Distinct and shared cortical regions of the human brain activated by pictorial depictions versus verbal descriptions: an fMRI study. *Neuroimage*, *24*(2), 473-486.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2000). Neurosciences cognitives: la biologie de l'esprit. *De Boeck Supérieur.*

Gazzaniga, M. S., Ivry, R. B., Mangun, G. R., & Steven, M. S. (1998).*Cognitive neuroscience: The biology of the mind* (pp. 44-50). New York: WW Norton.

Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cerebral Cortex*, *19*(6), 1239-1255.

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annu. Rev. Psychol.*, *59*, 167-192.

Gennari, S. P., Mac Donald, M. C., Postle, B. R., & Seidenberg, M. S. (2007). Context-dependent interpretation of words: Evidence for interactive neural processes. *Neuroimage*, 35, 1278–1286.

Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience & Biobehavioral Reviews*, *30*(7), 949-960.

Gerardin, E., Sirigu, A., Lehéricy, S., Poline, J. B., Gaymard, B., Marsault, C., ... & Le Bihan, D. (2000). Partially overlapping neural networks for real and imagined hand movements. *Cerebral Cortex*, *10*(11), 1093-1104.

Geschwind, N., & Levitsky, W. (1968). Human brain: left-right asymmetries in temporal speech region. *Science*, 161, 186-187.

Geyer, S., Matelli, M., Luppino, G., & Zilles, K. (2000). Functional neuroanatomy of the primate isocortical motor system. *Anatomy and embryology*, *202*(6), 443-474.

Gianelli, C., Farnè, A., Salemme, R., Jeannerod, M., & Roy, A. C. (2011). The agent is right: when motor embodied cognition is space-dependent. *PloS one*,*6*(9), e25036.

Giraud, A. L., Kell, C., Thierfelder, C., Sterzer, P., Russ, M. O., Preibisch, C., & Kleinschmidt, A. (2004). Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. *Cerebral Cortex*, *14*(3), 247-255.

Gitelman, D. R., Nobre, A. C., Sonty, S., Parrish, T. B., & Mesulam, M. (2005). Language network specializations: an analysis with parallel task designs and functional magnetic resonance imaging. *Neuroimage*, *26*(4), 975-985.

Glasser, M. F., & Rilling, J. K. (2008). DTI tractography of the human brain's language pathways. *Cerebral Cortex*, *18*(11), 2471-2482.

Glenberg, A. M. (1997). What memory is for: Creating meaning in the service of action. *Behavioral and brain sciences*, *20*(01), 41-50.

Glenberg, A. M. (2007). Language and action: creating sensible combinations of ideas. *The Oxford handbook of psycholinguistics*, 361-370.

Glenberg, A. M., & Gallese, V. (2012). Action-based language: a theory of language acquisition, comprehension, and production. *Cortex*, *48*(7), 905-922.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action.*Psychonomic bulletin & review*, *9*(3), 558-565.

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, *28*(1), 1-26.

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *The Quarterly Journal of Experimental Psychology*, *61*(6), 905-919.

Godschalk, M., Lemon, R. N., Kuypers, H. G. J. M., & Ronday, H. K. (1984). Cortical afferents and efferents of monkey postarcuate area: an anatomical and electrophysiological study. *Experimental Brain Research*, *56*(3), 410-424.

Gold, B. T., Balota, D. A., Kirchhoff, B. A., & Buckner, R. L. (2005). Common and dissociable activation patterns associated with controlled semantic and phonological processing: evidence from fMRI adaptation. *Cerebral Cortex*, *15*(9), 1438-1450.

Goldenberg, G. (2001). Imitation and matching of hand and finger postures.*Neuroimage*, *14*(1), S132-S136.

Goldstein, A., Spencer, K. M., & Donchin, E. (2002). The influence of stimulus deviance and novelty on the P300 and novelty P3. *Psychophysiology*, *39*(6), 781-790.

Goode, J. (2007). Wine and the Brain. *Questions of taste: the philosophy of wine*, 79-98.

Gorniak, P., & Roy, D. (2007). Situated language understanding as filtering perceived affordances. *Cognitive Science*, *31*(2), 197-231.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and special education*, *7*(1), 6-10.

Gough, P. B., Hoover, W. A., & Peterson, C. L. (1996). Some observations on a simple view of reading. *Reading comprehension difficulties: Processes and intervention*, 1-13.

Gough, P. M., Nobre, A. C., & Devlin, J. T. (2005). Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. *The Journal of Neuroscience*, *25*(35), 8010-8016.

Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography.*Experimental Brain Research*, *112*(1), 103-111.

Graves WW, Binder JR, Desai RH, Conant LL, Seidenberg MS. (2010). Neural correlates of implicit and explicit combinatorial semantic processing. *NeuroImage 53*, 638-46

Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, *338*(6213), 334-337.

Gray, J., Breazeal, C., Berlin, M., Brooks, A., & Lieberman, J. (2005). Action parsing and goal inference using self as simulator. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on* (pp. 202-209). IEEE.

Greenberg, D. L., & Verfaellie, M. (2010). Interdependence of episodic and semantic memory: Evidence from neuropsychology. *Journal of the International Neuropsychological Society*, *16*(5), 748.

Greenberg, D. L., Keane, M. M., Ryan, L., & Verfaellie, M. (2009). Impaired category fluency in medial temporal lobe amnesia: The role of episodic memory. *The Journal of Neuroscience*, *29*(35), 10900-10908.

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, *100*(1), 253-258.

Grezes, J. (1998). Top down effect of strategy on the perception of human biological motion: A PET investigation. *Cognitive Neuropsychology*, *15*(6-8), 553-582.

Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis.*Human brain mapping*, *12*(1), 1-19.

Grèzes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurones in the human brain: an fMRI study.*Neuroimage*, *18*(4), 928-937.

Grèzes, J., Tucker, M., Armony, J., Ellis, R., & Passingham, R. E. (2003). Objects automatically potentiate action: an fMRI study of implicit processing.*European Journal of Neuroscience*, *17*(12), 2735-2740.

Grill-Spector, K. (2003). The neural basis of object perception. *Current opinion in neurobiology*, *13*(2), 159-166.

Gronau, N., Neta, M., & Bar, M. (2008). Integrated contextual representation for objects' identities and their locations. *Journal of Cognitive Neuroscience*, *20*(3), 371-388.

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the Macaque. *Journal of neurophysiology*.

Grossman, M., Anderson, C., Khan, A., Avants, B., Elman, L., & McCluskey, L. (2008). Impaired action knowledge in amyotrophic lateral sclerosis.*Neurology*, *71*(18), 1396-1401.

Gruber, T., Müller, M. M., Keil, A., & Elbert, T. (1999). Selective visual-spatial attention alters induced gamma band responses in the human EEG. *Clinical Neurophysiology*, *110*(12), 2074-2085.

Gruber, T., Tsivilis, D., Giabbiconi, C. M., & Müller, M. M. (2008). Induced electroencephalogram oscillations during source memory: familiarity is reflected in the gamma band, recollection in the theta band. *Journal of cognitive neuroscience*, *20*(6), 1043-1053.

Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology*, *34*(6), 660-676.

Ha Duy Thuy, D., Matsuo, K., Nakamura, K., Toma, K., Oga, T., Nakai, T., ... & Fukuyama, H. (2004). Implicit and explicit processing of kanji and kana words and non-words studied with fMRI. *NeuroImage*, *23*(3), 878-889.

Hagoort, P. (2003). How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *Neuroimage*, *20*, S18-S29.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in cognitive sciences*, *9*(9), 416-423.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and cognitive processes*, *8*(4), 439-483.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and cognitive processes*, *8*(4), 439-483.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes.*Journal of Cognitive Neuroscience*, *11*(2), 194-205.

Hald, L. A., Bastiaansen, M., & Hagoort, P. (2006). EEG theta and gamma responses to semantic violations in online sentence processing. *Brain and language*, *96*(1), 90-105.

Halligan, P. W., Marshall, J. C., & Wade, D. T. (1995). Unilateral somatoparaphrenia after right hemisphere stroke: a case description. *Cortex*,*31*(1), 173-182.

Halsband, U., & Lange, R. K. (2006). Motor learning in man: a review of functional and clinical studies. *Journal of Physiology-Paris*, *99*(4), 414-424.

Hamilton, A. F. D. C., & Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *The Journal of Neuroscience*, *26*(4), 1133-1137.

Hamilton, A. F., Wolpert, D. M., Frith, U., & Grafton, S. T. (2006). Where does your own action influence your perception of another person's action in the brain?. *NeuroImage*, *29*(2), 524-535.

Hamm, J. P., Johnson, B. W., & Kirk, I. J. (2002). Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology*, *113*(8), 1339-1350.

Hamzei, F., Rijntjes, M., Dettmers, C., Glauche, V., Weiller, C., & Büchel, C. (2003). The human action recognition system and its relationship to Broca's area: an fMRI study. *Neuroimage*, *19*(3), 637-644.

Hare, B., & Tomasello, M. (2005). Human-like social skills in dogs?. *Trends in cognitive sciences*, *9*(9), 439-444.

Hari R, Forss N, Avikainen S, Kirveskari E, Salenius S, Rizzolatti G (1998) Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proc Natl Acad Sci* U S A 95:15061–15065.

Harrison, B. J., Pujol, J., López-Solà, M., Hernández-Ribas, R., Deus, J., Ortiz, H., ... & Cardoner, N. (2008). Consistency and functional specialization in the default mode brain network. *Proceedings of the National Academy of Sciences*,*105*(28), 9781-9786.

Harvey, D. Y., Wei, T., Ellmore, T. M., Cris Hamilton, A., & Schnur, T. T. (2013). Neuropsychological evidence for the functional role of the uncinate fasciculus in semantic control. *Neuropsychologia*.

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in cognitive sciences*, *11*(7), 299-306.

Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using imagination to understand the neural basis of episodic memory. *The Journal of Neuroscience*,*27*(52), 14365-14374.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *science*, *303*(5664), 1634-1640.

Hauk, O., & Pulvermüller, F. (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Human brain mapping*, *21*(3), 191-201.

Hauk, O., & Pulvermüller, F. (2011). The lateralization of motor cortex activation to action-words. *Frontiers in human neuroscience*, *5*.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301-307.

Havasi, C., Pustejovsky, J., Speer, R., & Lieberman, H. (2009). Digital intuition: Applying common sense using dimensionality reduction. *Intelligent Systems, IEEE*, *24*(4), 24-35.

Hein, G., & Knight, R. T. (2008). Superior temporal sulcus—it's my area: or is it?. *Journal of Cognitive Neuroscience*, *20*(12), 2125-2136.

Helbig, H. B., Graf, M., & Kiefer, M. (2006). The role of action representations in visual object recognition. *Experimental Brain Research*, *174*(2), 221-228.

Helbig, H. B., Steinwender, J., Graf, M., & Kiefer, M. (2010). Action observation can prime visual object recognition. *Experimental brain research*, *200*(3-4), 251-258.

Henry, R. G., Berman, J. I., Nagarajan, S. S., Mukherjee, P., & Berger, M. S. (2004). Subcortical pathways serving cortical language sites: initial experience with diffusion tensor imaging fiber tracking combined with intraoperative language mapping. *Neuroimage*, *21*(2), 616-622.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in cognitive sciences*, *4*(4), 131-138.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*(1), 67-99.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393-402.

Hillis, A. E., Rapp, B., & Caramazza, A. (1995). Constraining claims about theories of semantic memory: More on unitary versus multiple semantics.*Cognitive Neuropsychology*, *12*(2), 175-186.

Hillyard, S. A., & Kutas, M. (1983). Electrophysiology of cognitive processing.*Annual review of psychology*, *34*(1), 33-61.

Hinaut, X., & Dominey, P. F. (2013). Real-time parallel processing of grammatical structure in the fronto-striatal system: A recurrent network simulation study using reservoir computing. *PloS one*, *8*(2), e52946.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*(5), 791-804.

Hodges, J. R., & Patterson, K. (1996). Nonfluent progressive aphasia and semantic dementia: a comparative neuropsychological study. *JOURNAL-INTERNATIONAL NEUROPSYCHOLOGICAL SOCIETY*, *2*, 511-524.

Hodges, J. R., Patterson, K., Graham, N., & Dawson, K. (1996). Naming and knowing in dementia of Alzheimer's type. *Brain and language*, *54*(2), 302-325.

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia progressive fluent aphasia with temporal lobe atrophy. *Brain*, *115*(6), 1783-1806.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature neuroscience*, *3*(1), 80-84.

Hoffman, G. (2012). Embodied Cognition for Autonomous Interactive Robots.*Topics in cognitive science*, *4*(4), 759-772.

Holcomb, P. J., Anderson, J., & Grainger, J. (2005). An electrophysiological study of cross-modal repetition priming. *Psychophysiology*, *42*(5), 493-507.

Holz, E. M., Doppelmayr, M., Klimesch, W., & Sauseng, P. (2008). EEG correlates of action observation in humans. *Brain topography*, *21*(2), 93-99.

Hornak, J.P. (1996). The basic of MRI. Documentation en ligne : http://www.cis.rit.edu/htbooks/mri/

Horwitz, B., Rumsey, J. M., & Donohue, B. C. (1998). Functional connectivity of the angular gyrus in normal reading and dyslexia. *Proceedings of the National Academy of Sciences*, *95*(15), 8939-8944.

Houdé, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2002). Cerveau et psychologie. *Presses Universitaires de France*, 164.

Huijbers, W., Pennartz, C., Rubin, D. C., & Daselaar, S. M. (2011). Imagery and retrieval of auditory and visual information: neural correlates of successful and unsuccessful performance. *Neuropsychologia*, *49*(7), 1730-1740.

Humphreys, G. W., & Forde, E. M. (2001). Hierarchies, similarity, and interactivity in object recognition:" Category-specific" neuropsychological deficits. *Behavioral and Brain Sciences*, *24*(3), 453-475.

Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, *76*(6), 1210-1224.

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience*, *7*(12), 942-951.

Iacoboni, M., & Mazziotta, J. C. (2007). Mirror neuron system: basic findings and clinical applications. *Annals of neurology*, *62*(3), 213-218.

Iacoboni, M., Koski, L. M., Brass, M., Bekkering, H., Woods, R. P., Dubeau, M. C., ... & Rizzolatti, G. (2001). Reafferent copies of imitated actions in the right superior temporal cortex. *Proceedings of the National Academy of Sciences*, *98*(24), 13995-13999.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS biology*, *3*(3), e79.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*,*286*(5449), 2526-2528.

Ionta, S., Heydrich, L., Lenggenhager, B., Mouthon, M., Fornari, E., Chapuis, D., ... & Blanke, O. (2011). Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective. *Neuron*, *70*(2), 363-374.

Irish, M., & Piguet, O. (2013). The pivotal role of semantic memory in remembering the past and imagining the future. *Frontiers in behavioral neuroscience*, *7*.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*(5), 367-371.

Ivry, R. B. (1996). The representation of temporal information in perception and motor control. *Current opinion in neurobiology*, *6*(6), 851-857.

Jackson, P. L., Meltzoff, A. N., & Decety, J. (2006). Neural circuits involved in imitation and perspective-taking. *Neuroimage*, *31*(1), 429-439.

James, K. H., & Maouene, J. (2009). Auditory verb perception recruits motor systems in the developing brain: an fMRI investigation. *Developmental Science*,*12*(6), F26-F34.

James, W. (2011). *The principles of psychology*. Digireads.com.

Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain sciences*, *17*(2), 187-201.

Jeannerod, M. (1997). *The cognitive neuroscience of action*. Blackwell Publishing.

Jeannerod, M. (2004). Visual and action cues contribute to the self–other distinction. *Nature Neuroscience*, *7*(5), 422-423.

Jellema, T., Baker, C. I., Oram, M. W., & Perretl, D. I. (2002). 15 Cell populations in the banks of the superior temporal sulcus of the macaque and imitation. *The imitative mind: development, evolution, and brain bases*, *6*, 267.

Jellema, T., Baker, C. I., Wicker, B., & Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain and cognition*, *44*(2), 280-302.

Jirak, D., Menz, M. M., Buccino, G., Borghi, A. M., & Binkofski, F. (2010). Grasping language–A short story on embodiment. *Consciousness and cognition*, *19*(3), 711-720.

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. *Neuroimage*,*20*(2), 693-712.

Johnson, R. (1988). The amplitude of the P300 component of the event-related potential: Review and synthesis. *Advances in psychophysiology*, *3*, 69-137.

Johnson-Frey, S. H., Maloof, F. R., Newman-Norlund, R., Farrer, C., Inati, S., & Grafton, S. T. (2003). Actions or hand-object interactions? Human inferior frontal cortex and action observation. *Neuron*, *39*(6), 1053-1058.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness* (Vol. 6). Harvard University Press.

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proceedings of the National Academy of Sciences*, *107*(43), 18243-18250.

Josse, G., Mazoyer, B., Crivello, F., & Tzourio-Mazoyer, N. (2003). Left planum temporale: an anatomical marker of left hemispheric specialization for language comprehension. *Brain Res Cogn*, 18, 1-14.

Joubert, O. R., Rousselet, G. A., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision research*, *47*(26), 3286-3297.

Jouen, A. L., Verwey, W. B., Van Der Helden, J., Scheiber, C., Neveu, R., Dominey, P. F., & Ventre-Dominey, J. (2013). Discrete sequence production with and without a pause: the role of cortex, basal ganglia, and cerebellum. *Frontiers in human neuroscience*, *7*.

Jurgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and biobehavioral reviews*, *26*(2), 235-258.

Just, M, and Carpenter, P. (1980) "A theory of reading: From eye fixations to comprehension." *Psychological review* 87 329-354.

Kable, J. W., Kan, I. P., Wilson, A., Thompson-Schill, S. L., & Chatterjee, A. (2005). Conceptual representations of action in the lateral temporal cortex*Journal of Cognitive Neuroscience*, *17*(12), 1855-1870.

Kable, J. W., Lease-Spellmeyer, J., & Chatterjee, A. (2002). Neural substrates of action event knowledge. *Journal of Cognitive Neuroscience*, *14*(5), 795-805.

Kahlaoui, K., Baccino, T., Joanette, Y., & Magnié, M. N. (2007). Bases électro-physiologiques de la sémantique des objets. *médecine/sciences*, *23*, 193-7.

Kaiser, J., Ripper, B., Birbaumer, N., & Lutzenberger, W. (2003). Dynamics of gamma-band activity in human magnetoencephalogram during auditory pattern working memory. *Neuroimage*, *20*(2), 816-827.

Kamali, A., Kramer, L. A., & Hasan, K. M. (2010). Feasibility of prefronto-caudate pathway tractography using high resolution diffusion tensor tractography data at 3T. *Journal of neuroscience methods*, *191*(2), 249-254.

Kamide, Y., Altmann, G., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133-156.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302-4311.

Kaplan, F., & Oudeyer, P. Y. (2007). In search of the neural circuits of intrinsic motivation. *Frontiers in Neuroscience*, *1*(1), 225.

Kaplan, F., Oudeyer, P. Y., & Bergen, B. (2008). Computational models in the debate over language learnability. *infant and child development*, *17*(1), 55-80.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition*, *94*(3), B79-B89.

Kawato, M. (2008). From 'understanding the brain by creating the brain'towards manipulative neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1500), 2201-2214.

Keil, A., Müller, M. M., Ray, W. J., Gruber, T., & Elbert, T. (1999). Human gamma band activity and perception of a gestalt. *The Journal of neuroscience*,*19*(16), 7152-7161.

Kemmerer, D., & Gonzalez-Castillo, J. (2010). The two-level theory of verb meaning: An approach to integrating the semantics of action with the mirror neuron system. *Brain and language*, *112*(1), 54-76.

Kemmerer, D., & Tranel, D. (2003). A double dissociation between the meanings of action verbs and locative prepositions. *Neurocase*, *9*(5), 421-435.

Kemmerer, D., Castillo, J. G., Talavage, T., Patterson, S., & Wiley, C. (2008). Neuroanatomical distribution of five semantic components of verbs: evidence from fMRI. *Brain and language*, *107*(1), 16-43.

Kessler, K., & Kiefer, M. (2005). Disturbing visual working memory: electrophysiological evidence for a role of the prefrontal cortex in recovery from interference. *Cerebral Cortex*, *15*(7), 1075-1087.

Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental brain research*, *153*(4), 628-636.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. *Cortex*, *48*(7), 805–25.

Kiefer, M., & Spitzer, M. (2001). The limits of a distributed account of conceptual knowledge. *Trends in Cognitive Sciences*, *5*(11), 469-471.

Kiefer, M., Schuch, S., Schenck, W., & Fiedler, K. (2007). Mood states modulate activity in semantic brain areas during emotional word encoding.*Cerebral Cortex*, *17*(7), 1516-1530.

Kiefer, M., Sim, E. J., Helbig, H., & Graf, M. (2011). Tracking the time course of action priming on object recognition: evidence for fast and slow influences of action on perception. *Journal of cognitive neuroscience*, *23*(8), 1864-1874.

Kiefer, M., Sim, E. J., Liebich, S., Hauk, O., & Tanaka, J. (2007). Experience-dependent plasticity of conceptual representations in human sensory-motor areas. *Journal of Cognitive Neuroscience*, *19*(3), 525-542.

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *The Journal of Neuroscience*, *29*(32), 10153-10159.

Kim, H. (2012). A dual-subsystem model of the brain's default network: Self-referential processing, memory retrieval processes, and autobiographical memory retrieval. *Neuroimage*, *61*(4), 966-977.

Kircher, T., Sass, K., Sachs, O., & Krach, S. (2009). Priming words with pictures: Neural correlates of semantic associations in a cross-modal priming task using fMRI. *Human brain mapping*, *30*(12), 4116-4128.

Kliegl, Reinhold, Antje Nuthmann, and Ralf Engbert. "Tracking the mind during reading: the influence of past, present, and future words on fixation durations."*Journal of Experimental Psychology: General* 135.1 (2006): 12.

Knauff, M., Fangmeier, T., Ruff, C. C., & Johnson-Laird, P. N. (2003). Reasoning, models, and images: Behavioral measures and cortical activity.*Journal of cognitive neuroscience*, *15*(4), 559-573.

Knoll, A., Bekey, G., & Henderson, T. C. (2001). Cui bono robo sapiens?.*Robotics and autonomous systems*, *37*(2), 73-80.

Koch, G., Cercignani, M., Pecchioli, C., Versace, V., Oliveri, M., Caltagirone, C., ... & Bozzali, M. (2010). In vivo definition of parieto-motor connections involved in planning of grasping movements. *Neuroimage*, *51*(1), 300-312.

Kohler, E., Keysers, C., Umilta, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, *297*(5582), 846-848.

Kolers, P. A. (1976). Buswell's discoveries. *Eye movements and psychological processes*, 371-395.

Koski, L., Iacoboni, M., Dubeau, M. C., Woods, R. P., & Mazziotta, J. C. (2003). Modulation of cortical activity during different imitative behaviors. *Journal of Neurophysiology*, *89*(1), 460-471.

Koski, L., Wohlschläger, A., Bekkering, H., Woods, R. P., Dubeau, M. C., Mazziotta, J. C., & Iacoboni, M. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cerebral cortex*, *12*(8), 847-855.

Kosslyn, S. M. (1995). Mental imagery. *Visual cognition: An invitation to cognitive science*, *2*, 267-296.

Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of cognitive neuroscience*, *12*(1), 48-55.

Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science*, *293*(5534), 1506-1509.

Kovalenko, L. Y., Chaumon, M., & Busch, N. A. (2012). A Pool of Pairs of Related Objects (POPORO) for Investigating Visual Semantic Integration: Behavioral and Electrophysiological Validation. *Brain topography*, *25*(3), 272-284.

Krams, M., Rushworth, M. F. S., Deiber, M. P., Frackowiak, R. S. J., & Passingham, R. E. (1998). The preparation, execution and suppression of copied movements in the human brain. *Experimental Brain Research*, *120*(3), 386-398.

Kraskov, A., Dancause, N., Quallo, M. M., Shepherd, S., & Lemon, R. N. (2009). Corticospinal neurons in macaque ventral premotor cortex with mirror properties: a potential mechanism for action suppression?. *Neuron*, *64*(6), 922-930.

Krueger F, Barbey AK, Grafman J. (2009). The medial prefrontal cortex mediates social event knowledge. *Trends in cognitive sciences* 13: 103-9

Krueger F, Spampinato MV, Barbey AK, Huey ED, Morland T, Grafman J. (2009). The frontopolar cortex mediates event knowledge complexity: a parametric functional MRI study. *Neuroreport* 20: 1093-7

Kumashiro, M., Ishibashi, H., Uchiyama, Y., Itakura, S., Murata, A., & Iriki, A. (2003). Natural imitation induced by joint attention in Japanese monkeys.*International Journal of Psychophysiology*, *50*(1), 81-99.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23-49.

Kuperberg, G. R., Sitnikova, T., & Lakshmanan, B. M. (2008). Neuroanatomical distinctions within the semantic system during sentence comprehension: evidence from functional magnetic resonance imaging. *Neuroimage*, *40*(1), 367-388.

Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes*, *8*(4), 533-572.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*, *4*(12), 463-470.

Kutas, M., & Federmeier, K. D. (2009). N400. *Scholarpedia*, *4*(10), 7790.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP).*Annual review of psychology*, *62*, 621-647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, *11*(5), 539-550.

Kutas, M., & Kluender, R. (1994). What is who violating? A reconsideration of linguistic violations in light of event-related brain potentials. In *Cognitive electrophysiology* (pp. 183-210). Birkhäuser Boston.

Kutas, M., & Van Petten, C. (1988). Event-related brain potential studies of language. *Advances in psychophysiology*, *3*, 139-187.

Kutas, M., & Van Petten, C. (1994). Psycholinguistics electrified. *Handbook of psycholinguistics*, 83-143.

Kutas, M., Lindamood, T. E., & Hillyard, S. A. (1984). Word expectancy and event-related brain potentials during sentence processing. *Preparatory states and processes*, 217-237.

Kwong, K.K., Belliveau, J.W., Chesler, D.A., (1992). Dynamic MRI of human brain activity during primary sensory stimulation, *Proceedings of National Academy of Sciences of the U.S.A.*, Vol 89 5675-5679.

Lakoff, G., & Johnson, M. (1999). Philosophy in the flesh: The embodied mind and its challenge to western thought. *Basic books.*

Lallée, S., Hamann, K., Steinwender, J., Warneken, F., Martienz, U., Barron-Gonzales, H., ... & Dominey, P. F. (2013) Cooperative Human Robot Interaction Systems: IV. Communication of Shared Plans with Naïve Humans using Gaze and Speech.

Lallee, S., Madden, C., Hoen, M., & Dominey, P. F. (2010). Linking language with embodied and teleological representations of action for humanoid cognition. *Frontiers in neurorobotics*, *4*.

Lallée, S., Pattacini, U., Boucher, J. D., Lemaignan, S., Lenz, A., Melhuish, C., ... & Dominey, P. F. (2011, September). Towards a platform-independent cooperative human-robot interaction system: Ii. perception, execution and imitation of goal directed actions. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on* (pp. 2895-2902). IEEE.

Lallée, S., Pattacini, U., Lemaignan, S., Lenz, A., Melhuish, C., Natale, L., ... & Dominey, P. F. (2012). Towards a Platform-Independent Cooperative Human Robot Interaction System: III. An Architecture for Learning and Executing Actions and Shared Plans.

Lambon Ralph, M. A., Graham, K. S., Patterson, K., & Hodges, J. R. (1999). Is a picture worth a thousand words? Evidence from concept definitions by patients with semantic dementia. *Brain and language*, *70*(3), 309-335.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, *104*(2), 211.

Le Bihan, D. (2003). Looking into the functional architecture of the brain with diffusion MRI. *Nature Reviews Neuroscience*, *4*(6), 469-480.

Le Bihan, D., Mangin, J. F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., & Chabriat, H. (2001). Diffusion tensor imaging: concepts and applications.*Journal of magnetic resonance imaging*, *13*(4), 534-546.

Leather, C. V., & Henry, L. A. (1994). Working memory span and phonological awareness tasks as predictors of early reading ability. *Journal of Experimental Child Psychology*, *58*(1), 88-111.

Lee, J. K., & Breazeal, C. (2010). Human social response toward humanoid robot's head and facial features. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems* (pp. 4237-4242). ACM.

Lefebvre, C. D., Marchand, Y., Eskes, G. A., & Connolly, J. F. (2005). Assessment of working memory abilities using an event-related brain potential (ERP)-compatible digit span backward task. *Clinical neurophysiology*, *116*(7), 1665-1680.

Lehericy, S., Cohen, L., Bazin, B., Samson, S., Giacomini, E., Rougetet, R., ... & Baulac, M. (2000). Functional MR evaluation of temporal and frontal language dominance compared with the Wada test. *Neurology*, *54*(8), 1625-1633.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation* (Vol. 348). Chicago: University of Chicago press.

Levine, B. (2004). Autobiographical memory and the self in time: Brain lesion effects, functional neuroanatomy, and lifespan development. *Brain and cognition*, *55*(1), 54-68.

Liebal, K., Call, J., & Tomasello, M. (2004). Use of gesture sequences in chimpanzees. *American Journal of Primatology*, *64*(4), 377-396.

Liljeström, M., Tarkiainen, A., Parviainen, T., Kujala, J., Numminen, J., Hiltunen, J., ... & Salmelin, R. (2008). Perceiving and naming actions and objects. *Neuroimage*, *41*(3), 1132-1141.

Lindell, A. K. (2006). In your right mind: Right hemisphere contributions to language processing and production. *Neuropsychology review*, *16*(3), 131-148.

Lio, G., & Boulinguez, P. (2012). Greater robustness of second order statistics than higher order statistics algorithms to distortions of the mixing matrix in blind source separation of human EEG: Implications for single-subject and group analyses. *NeuroImage*.

Lissauer H. 1890/1988. A case of visual agnosia with a contribution to theory. Transl. M Jackson. *Cogn. Neuropsychol*. 5:157–92

Liu, H., & Singh, P. (2004). Commonsense reasoning in and over natural language. In *Knowledge-based intelligent information and engineering systems* (pp. 293-306). Springer Berlin Heidelberg.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(4), 565.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*(3), 291-308.

Lungarella, M., Metta, G., Pfeifer, R., & Sandini, G. (2003). Developmental robotics: a survey. *Connection Science*, *15*(4), 151-190.

Lutzenberger, W., Ripper, B., Busse, L., Birbaumer, N., & Kaiser, J. (2002). Dynamics of gamma-band activity during an audiospatial working memory task in humans. *The Journal of neuroscience*, *22*(13), 5630-5638.

MacDonald, P. A., & Paus, T. (2003). The role of parietal cortex in awareness of self-generated movements: a transcranial magnetic stimulation study.*Cerebral Cortex*, *13*(9), 962-967.

Machery, E. (2007). Concept empiricism: a methodological critique. *Cognition*,*104*(1), 19-46.

Madden, C. J., & Ferretti, T. R. (2009). Verb aspect and the mental representation of situations. *The expression of time*, *3*, 217.

Madden, C. J., & Therriault, D. J. (2009). Verb aspect and perceptual simulations. *The Quarterly Journal of Experimental Psychology*, *62*(7), 1294-1303.

Madden, C. J., & Zwaan, R. A. (2003). How does verb aspect constrain event representations?. *Memory & Cognition*, *31*(5), 663-672.

Madden, C. J., & Zwaan, R. A. (2006). Perceptual representation as a mechanism of lexical ambiguity resolution: An investigation of span and processing time. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(6), 1291.

Madden, C., Hoen, M., & Dominey, P. F. (2010). A cognitive neuroscience perspective on embodied language for human–robot cooperation. *Brain and language*, *112*(3), 180-188.

Maeda, F., Kleiner-Fisman, G., & Pascual-Leone, A. (2002). Motor facilitation while observing hand actions: specificity of the effect and role of observer's orientation. *Journal of Neurophysiology*, *87*(3), 1329-1335.

Magliano, J. P., & Schleich, M. C. (2000). Verb aspect and situation models.*Discourse processes*, *29*(2), 83-112.

Maguire, E. A. (2001). Neuroimaging studies of autobiographical event memory.*Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1413), 1441-1451.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences*, *97*(8), 4398-4403.

Maguire, E. A., Spiers, H. J., Good, C. D., Hartley, T., Frackowiak, R. S., & Burgess, N. (2003). Navigation expertise and the human hippocampus: a structural brain imaging analysis. *Hippocampus*, *13*(2), 250-259.

Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus*,*16*(12), 1091-1101.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content.*Journal of Physiology-Paris*, *102*(1), 59-70.

Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M., & Caramazza, A. (2009). Category-specific organization in the human brain does not require visual experience. *Neuron*, *63*(3), 397-405.

Makris, N., & Pandya, D. N. (2009). The extreme capsule in humans and rethinking of the language circuitry. *Brain Structure and Function*, *213*(3), 343-358.

Makris, N., Papadimitriou, G. M., Kaiser, J. R., Sorg, S., Kennedy, D. N., & Pandya, D. N. (2009). Delineation of the middle longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. *Cerebral Cortex*, *19*(4), 777-785.

Makris, N., Worth, A. J., Papadimitriou, G. M., Stakes, J. W., Caviness, V. S., Kennedy, D. N., ... & Davis, T. L. (1997). Morphometry of in vivo human white matter association pathways with diffusion-weighted magnetic resonance imaging. *Annals of neurology*, *42*(6), 951-962.

Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *Journal of Vision*, *10*(2).

Maldonado, I. L., Champfleur, N. M., Velut, S., Destrieux, C., Zemmoura, I., & Duffau, H. (2013). Evidence of a middle longitudinal fasciculus in the human brain from fiber dissection. *Journal of anatomy*.

Maldonado, I. L., Mandonnet, E., & Duffau, H. (2012). Dorsal Fronto-Parietal Connections of the Human Brain: A Fiber Dissection Study of Their Composition and Anatomical Relationships. *The Anatomical Record*, *295*(2), 187-195.

Mandonnet, E., Nouet, A., Gatignol, P., Capelle, L., & Duffau, H. (2007). Does the left inferior longitudinal fasciculus play a role in language? A brain stimulation study. *Brain*, *130*(3), 623-629.

Maouene, J., Hidaka, S., & Smith, L. B. (2008). Body Parts and Early-Learned Verbs. *Cognitive Science*, *32*(7), 1200-1216.

Marilyn, E (2009) "MRIs reveal possible source of woman's super-memory". USA Today

Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., & Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron*, *38*(3), 487-497.

Marocco, D., Cangelosi, A., Fischer, K., & Belpaeme, T. (2010). Grounding action words in the sensorimotor interaction with the world: experiments with a simulated iCub humanoid robot. *Frontiers in neurorobotics*, *4*.

Marques, J. F., Canessa, N., & Cappa, S. (2009). Neural differences in the processing of true and false sentences: Insights into the nature of 'truth'in language comprehension. *cortex*, *45*(6), 759-768.

Marrelec, G., Bellec, P., Krainik, A., Duffau, H., Pélégrini-Issac, M., Lehéricy, S., ... & Doyon, J. (2008). Regions, systems, and the brain: hierarchical measures of functional integration in fMRI. *Medical image analysis*, *12*(4), 484-496.

Marshall, J. C., & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of psycholinguistic research*, *2*(3), 175-199.

Martin, A. (2007). The representation of object concepts in the brain. *Annu. Rev. Psychol.*, *58*, 25-45.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Current opinion in neurobiology*, *11*(2), 194-201.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, *270*(5233), 102-105.

Martino, J., Brogna, C., Robles, S. G., Vergani, F., & Duffau, H. (2010). Anatomic dissection of the inferior fronto-occipital fasciculus revisited in the lights of brain stimulation data. *Cortex*, *46*(5), 691-699.

Martino, J., Vergani, F., Robles, S. G., & Duffau, H. (2010). New insights into the anatomic dissection of the temporal stem with special emphasis on the inferior fronto-occipital fasciculus: implications in surgical approach to left mesiotemporal and temporoinsular structures. *Neurosurgery*, *66*(3), ons4-ons12.

Masataka, N. (2001). Why early linguistic milestones are delayed in children with Williams syndrome: late onset of hand banging as a possible rate–limiting constraint on the emergence of canonical babbling. *Developmental Science*,*4*(2), 158-164.

Mashal, N., Faust, M., Hendler, T., & Jung-Beeman, M. (2007). An fMRI investigation of the neural correlates underlying the processing of novel metaphoric expressions. *Brain and language*, *100*(2), 115-126.

Mason, M. F., Banfield, J. F., & Macrae, C. N. (2004). Thinking about actions: the neural substrates of person knowledge. *Cerebral Cortex*, *14*(2), 209-214.

Matelli, M., Camarda, R., Glickstein, M., & Rizzolatti, G. (1986). Afferent and efferent projections of the inferior area 6 in the macaque monkey. *Journal of Comparative Neurology*, *251*(3), 281-298.

Maunsell, J. H., & van Essen, D. C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *The Journal of neuroscience*, *3*(12), 2563-2586.

Mavridis, N., & Roy, D. (2006). Grounded situation models for robots: Where words and percepts meet. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on* (pp. 4690-4697). IEEE.

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., ... & Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, *5*(4), 467-479.

McAdam, D.W., Whitaker, H.A. (1971). Language production: electroencephalographic localization in the normal human brain. *Science*, 172(982), 499-502.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in cognitive sciences*,*7*(7), 293-299.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*(4), 310-322.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological review*, *102*(3), 419.

McIntosh, A. R., & Gonzalez-Lima, F. (1991). Structural modeling of functional neural pathways mapped with 2-deoxyglucose: effects of acoustic startle habituation on the auditory system. *Brain Research*, *547*(2), 295-302.

McNamara, A., Buccino, G., Menz, M. M., Gläscher, J., Wolbers, T., Baumgärtner, A., & Binkofski, F. (2008). Neural Dynamics of Learning Sound—Action Associations. *PLoS One*, *3*(12), e3845.

McPherson, W. & Holcomb, P J (1999). An electrophysiological investigation of semantic priming with pictures of real objects.*Psychophysiology*, 1999, vol. 36, no 01, p. 53-65.

Meier, J. D., Aflalo, T. N., Kastner, S., & Graziano, M. S. (2008). Complex organization of human primary motor cortex: a high-resolution fMRI study.*Journal of neurophysiology*, *100*(4), 1800-1812.

Meister, I. G., & Iacoboni, M. (2007). No language-specific activation during linguistic processing of observed actions. *PLoS ONE*, 9, e891.

Mellet, E., Bricogne, S., Crivello, F., Mazoyer, B., Denis, M., & Tzourio-Mazoyer, N. (2002). Neural basis of mental scanning of a topographic representation built from a text. *Cerebral Cortex*, *12*(12), 1322-1330.

Meltzoff, A. N., & Decety, J. (2003). What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 491-500.

Meltzoff, A. N., & Gopnik, A. (1993). The role of imitation in understanding persons and developing a theory of mind. *Understanding other minds: Perspectives from autism*, 335-366.

Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child development*, 702-709.

Meltzoff, A. N., & Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Early development and parenting*, *6*(34), 179-192.

Menjot de Champfleur, N., Lima Maldonado, I., Moritz-Gasser, S., Machi, P., Le Bars, E., Bonafé, A., & Duffau, H. (2012). Middle longitudinal fasciculus delineation within language pathways: a diffusion tensor imaging study in human. *European journal of radiology*.

Menz, M. M., & Binkofski, F. (2008). Broca's region beyond language. The omnifariousness of Broca's region. *Rivista Medica*, *14*, 51-55.

Merleau-Ponty, M. (1942). La structure du comportement.

Merleau-Ponty, M. (1969). Phénoménologie de la perception (Vol. 194). *Paris: Gallimard.*

Metusalem, R., Kutas, M., Urbach, T. P., Hare, M., McRae, K., & Elman, J. L. (2012). Generalized event knowledge activation during online sentence comprehension. *Journal of memory and language*, *66*(4), 545-567.

Meyer, K., & Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in neurosciences*, *32*(7), 376-382.

.Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, *24*(1), 167-202.

Miltner, W. H., Braun, C., Arnold, M., Witte, H., & Taub, E. (1999). Coherence of gamma-band EEG activity as a basis for associative learning. *Nature*,*397*(6718), 434-436.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in neurosciences*, *6*, 414-417.

Mitz, A. R., Godschalk, M., & Wise, S. P. (1991). Learning-dependent neuronal activity in the premotor cortex: activity during the acquisition of conditional motor associations. *The Journal of neuroscience*, *11*(6), 1855-1872.

Molnar-Szakacs, I., Kaplan, J., Greenfield, P. M., & Iacoboni, M. (2006). Observing complex action sequences: the role of the fronto-parietal mirror neuron system. *Neuroimage*, *33*(3), 923-935.

Moore, C. J., & Price, C. J. (1999). Three distinct ventral occipitotemporal regions for reading and object naming. *Neuroimage*, *10*(2), 181-192.

Mori, S. (2007). *Introduction to diffusion tensor imaging*. Access Online via Elsevier.

Morrell, L.K., Huntington, D.A. (1971). Electrocortical localization of language production. *Science*. 24;174(16):1359-61.

Morrison, R (1984). "Manipulation of stimulus onset delay in reading: evidence for parallel programming of saccades." *Journal of Experimental Psychology: Human Perception and Performance* 10.5 : 667.

Muakkassa, K. F., & Strick, P. L. (1979). Frontal lobe inputs to primate motor cortex: evidence for four somatotopically organized 'premotor'areas. *Brain research*, *177*(1), 176-182.

Mudrik, L., Lamy, D., & Deouell, L. Y. (2010). ERP evidence for context congruity effects during simultaneous object–scene processing.*Neuropsychologia*, *48*(2), 507-517.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current biology*, *20*(8), 750-756.

Mummery, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S. J., & Hodges, J. R. (2000). A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory.*Annals of neurology*, *47*(1), 36-45.

Münte, T. F., Heinze, H. J., Matzke, M., Wieringa, B. M., & Johannes, S. (1998). Brain potentials and syntactic violations revisited: No evidence for specificity of the syntactic positive shift. *Neuropsychologia*, *36*(3), 217-226.

Muthukumaraswamy, S. D., & Johnson, B. W. (2004). Primary motor cortex activation during action observation revealed by wavelet analysis of the EEG.*Clinical Neurophysiology*, *115*(8), 1760-1766.

Myung, J. Y., Blumstein, S. E., & Sedivy, J. C. (2006). Playing on the typewriter, typing on the piano: manipulation knowledge of objects. *Cognition*,*98*(3), 223-243.

Nation, K., & Snowling, M. J. (1998). Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child development*, *69*(4), 996-1011.

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language*, *39*(1), 85-101.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*(4), 342-356.

Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., & Lee, R. G. (2000). The syntax of American sign language. Cambridge, MA: The MIT Press.

Neininger, B., & Pulvermüller, F. (2003). Word-category specific deficits after lesions in the right hemisphere. *Neuropsychologia*, *41*(1), 53-70.

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, *28*(4), 661.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE™//A lexical database for contemporary french: LEXIQUE™. *L'année Psychologique*, *101*(3), 447-462.

Newman, S. D., Ratliff, K., Muratore, T., & Burns Jr, T. (2009). The effect of lexical priming on sentence comprehension: An fMRI study. *Brain research*,*1285*, 99-108.

Nigam, A. Hoffman, j. E., and Simons, RF (1992).“N400 to semantically anomalous pictures and words”. Journal of Cognitive Neuroscience, 4(1), 15-22.

Nigam, A., Hoffman, J. E., & Simons, R. F. (1992). N400 to semantically anomalous pictures and words. *Journal of cognitive neuroscience*, *4*(1), 15-22.

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences*, *97*(2), 913-918.

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences*, *97*(2), 913-918.

Nishitani, N., Avikainen, S., & Hari, R. (2004). Abnormal imitation-related cortical activation sequences in Asperger's syndrome. *Annals of neurology*,*55*(4), 558-562.

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision research*, *45*(17), 2201-2217.

Oakhill, J. (1982). Constructive processes in skilled and less skilled comprehenders' memory for sentences. *British Journal of Psychology*, *73*(1), 13-20.

Oakhill, J. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology*, *54*(1), 31-39.

Oakhill, J., Cain, K., & Yuill, N. (1998). Individual differences in children's comprehension skill: Toward an integrated model. *Reading and spelling: Development and disorders*, 343-367.

Oakhill, J., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and writing*,*18*(7-9), 657-686.

Oberman, L. M., & Ramachandran, V. S. (2007). The simulating social mind: the role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychological bulletin*,*133*(2), 310.

Oberman, L. M., Hubbard, E. M., McCleery, J. P., Altschuler, E. L., Ramachandran, V. S., & Pineda, J. A. (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Cognitive Brain Research*, *24*(2), 190-198.

Oberman, L. M., Pineda, J. A., & Ramachandran, V. S. (2007). The human mirror neuron system: A link between action observation and social skills.*Social Cognitive and Affective Neuroscience*, *2*(1), 62-66.

Oberman, L. M., Ramachandran, V. S., & Pineda, J. A. (2008). Modulation of mu suppression in children with autism spectrum disorders in response to familiar or unfamiliar stimuli: the mirror neuron hypothesis. *Neuropsychologia*,*46*(5), 1558-1565.

Ogawa, K., & Inui, T. (2011). Neural representation of observed actions in the parietal and premotor cortex. *Neuroimage*, *56*(2), 728-735.

Ogawa, K., & Inui, T. (2012). Multiple neural representations of object-directed action in an imitative context. *Experimental brain research*, *216*(1), 61-69.

Ogawa, S., Lee, T.M. (1990). Magnetic Resonance Imaging of blood vessels at high fields: in vivo and in vitro measurements and image simulation, *Magn. Res. Med*., Vol 16 9-18.

Ogawa, S., Lee, T.M., Kay, A.R. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation, *Proceedings of National Academy of Sciences of the U.S.A*., Vol 87 9868-9872.

Ogawa, S., Lee, T.M., Nayak, A., Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high fields, *Magn. Res. Med*., Vol 14 68-78.

Ogawa, S., Menon, R.S., Kim, S.G., Ugurbil, K. (1998). On the characteristics of functional magnetic resonance imaging of the brain. *Annu Rev Biophys Biomol Struct*. 27:447-74.

Ogawa, S., Tank, D.W., Menon, R., Ellermann, J.M., Kim, S.G., Merkle, H., Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proceedings of National Academy of Sciences of the U.S.A*. 1;89(13):5951-5.

Okada, T., Tanaka, S., Nakai, T., Nishizawa, S., Inui, T., Sadato, N., ... & Konishi, J. (2000). Naming of animals and tools: a functional magnetic resonance imaging study of categorical differences in the human brain areas commonly used for naming visually presented objects. *Neuroscience letters*,*296*(1), 33-36.

Okuda, J., Fujii, T., Ohtake, H., Tsukiura, T., Tanji, K., Suzuki, K., ... & Yamadori, A. (2003). Thinking of the future and past: The roles of the frontal pole and the medial temporal lobes. *Neuroimage*, *19*(4), 1369-1380.

Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive psychology*, *41*(2), 176-210.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in brain research*, *155*, 23-36.

Olivares, E. I., Iglesias, J., & Antonieta Bobes, M. (1999). Searching for face-specific long latency ERPs: a topographic study of effects associated with mismatching features. *Cognitive Brain Research*, *7*(3), 343-356.

Oreja-Guevara, C., Kleiser, R., Paulus, W., Kruse, W., Seitz, R. J., & Hoffmann, K. P. (2004). The role of V5 (hMT+) in visually guided hand movements: an fMRI study. *European Journal of Neuroscience*, *19*(11), 3113-3120.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of memory and language*, *31*(6), 785-806.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of memory and language*, *31*(6), 785-806.

Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, *8*(4), 413-437.

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and language*, *34*(6), 739-773.

Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies.*Language and Cognitive Processes*, *14*(3), 283-317.

Oudeyer, P. Y., & Kaplan, F. (2006). Discovering communication. *Connection Science*, *18*(2), 189-206.

Oudeyer, P. Y., & Kaplan, F. (2007). What is intrinsic motivation? a typology of computational approaches. *Frontiers in neurorobotics*, *1*.

Oudeyer, P. Y., & Kaplan, F. (2008). How can we define intrinsic motivation?. In*proceedings of the 8th international conference on epigenetic robotics: modeling cognitive development in robotic systems*.

Paivio, A. (1971). Imagery and verbal processes.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *45*(3), 255.

Parasuraman, R. (1980). Effects of information processing demands on slow negative shift latencies and N100 amplitude in selective and divided attention.*Biological psychology*, *11*(3), 217-233.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of cognitive neuroscience*, *10*(6), 717-733.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), 976-987.

Patuzzo, S., Fiaschi, A., & Manganotti, P. (2003). Modulation of motor cortex excitability in the left hemisphere during action observation: a single-and paired-pulse transcranial magnetic stimulation study of self-and non-self-action observation. *Neuropsychologia*, *41*(9), 1272-1278.

Paukner, A., Anderson, J. R., Borelli, E., Visalberghi, E., & Ferrari, P. F. (2005). Macaques (Macaca nemestrina) recognize when they are being imitated. *Biology Letters*, *1*(2), 219-222.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory.

Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*, *93*(1), 603-608.

Peelen, M. V., Wiggett, A. J., & Downing, P. E. (2006). Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron*, *49*(6), 815-822.

Peigneux, P., Salmon, E., Van Der Linden, M., Garraux, G., Aerts, J., Delfiore, G., ... & Franck, G. (2000). The role of lateral occipitotemporal junction and area MT/V5 in the visual analysis of upper-limb postures. *Neuroimage*, *11*(6), 644-655.

Pelphrey, K. A., Mitchell, T. V., McKeown, M. J., Goldstein, J., Allison, T., & McCarthy, G. (2003). Brain activity evoked by the perception of human walking: controlling for meaningful coherent motion. *The Journal of Neuroscience*,*23*(17), 6819-6825.

Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: an fMRI study of eye, mouth and hand movements. *Cerebral Cortex*,*15*(12), 1866-1876.

Penfield, W., & Boldrey, E. (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain: A journal of neurology*.

Penny, W. D., Stephan, K. E., Mechelli, A., & Friston, K. J. (2004). Comparing dynamic causal models. *NeuroImage*, *22*(3), 1157-1172.

Péran, P., Démonet, J. F., Cherubini, A., Carbebat, D., Caltagirone, C., & Sabatini, U. (2010). Mental representations of action: the neural correlates of the verbal and motor components. *Brain research*, *1328*, 89-103.

Perani, D., Schnur, T., Tettamanti, M., Cappa, S. F., & Fazio, F. (1999). Word and picture matching: a PET study of semantic category effects.*Neuropsychologia*, *37*(3), 293-306.

Perlbarg, V., & Marrelec, G. (2008). Contribution of exploratory methods to the investigation of extended large-scale brain networks in functional MRI: methodologies, results, and challenges. *Journal of Biomedical Imaging*, *2008*, 4.

Perlbarg, V., & Marrelec, G. (2008). Contribution of exploratory methods to the investigation of extended large-scale brain networks in functional MRI: methodologies, results, and challenges. *Journal of Biomedical Imaging*, *2008*, 4.

Perlbarg, V., Bellec, P., Anton, J. L., Pélégrini-Issac, M., Doyon, J., & Benali, H. (2007). CORSICA: correction of structured noise in fMRI by automatic identification of ICA components. *Magnetic resonance imaging*, *25*(1), 35-46.

Perlbarg, V., Marrelec, G., Bellec, P., Coynel, D., Pélégrini-Issac, M., & Benali, H. (2009). NetBrainWork: a toolbox for studying functional interactions in large-scale brain networks in fMRI. *Neuroimage*, *47*, S80.

Perlbarg, V., Marrelec, G., Doyon, J., Pélégrini-Issac, M., Lehericy, S., & Benali, H. (2008, May). NEDICA: Detection of group functional networks in fMRI using spatial independent component analysis. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on* (pp. 1247-1250). IEEE.

Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., ... & Ortega, J. E. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology*,*146*(1), 87-113.

Perrett, D. I., Mistlin, A. J., Harries, M. H., & Chitty, A. J. (1990). Understanding the visual appearance and consequence of hand actions. *Vision and action: the control of grasping*, 163-180.

Petit, M., Lallée, S., Boucher, J., Pointeau, G., Cheminade, P., Ognibene, D., ... & Dominey, P. (2013). The Coordinating Role of Language in Real-Time Multi-Modal Learning of Cooperative Tasks.

Petrides, M., & Pandya, D. N. (1984). Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. *Journal of Comparative Neurology*, *228*(1), 105-116.

Petrides, M., & Pandya, D. N. (1999). Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *European Journal of Neuroscience*,*11*(3), 1011-1036.

Petrides, M., & Pandya, D. N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience*, *16*(2), 291-310.

Pfeifer, R., & Gomez, G. (2005). Interacting with the real world: design principles for intelligent systems. *Artificial life and Robotics*, *9*(1), 1-6.

Pfeifer, R., & Scheier, C. (1997). Sensory—motor coordination: The metaphor and beyond. *Robotics and autonomous systems*, *20*(2), 157-178.

Phillips, S., & Takeda, Y. (2009). An EEG/ERP study of efficient versus inefficient visual search. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society, Amsterdam, The Netherlands*.

Picard, L., Eustache, F., & Piolino, P. (2009). De la mémoire épisodique à la mémoire autobiographique: approche développementale. *L'Année psychologique*, *109*(02), 197-236.

Pika, S., Liebal, K., & Tomasello, M. (2003). Gestural communication in young gorillas (Gorilla gorilla): gestural repertoire, learning, and use. *American Journal of Primatology*, *60*(3), 95-111.

Pineda, J. A. (2005). The functional significance of mu rhythms: translating “seeing” and “hearing” into “doing”. *Brain Research Reviews*, *50*(1), 57-68.

Pineda, J. A., & Hecht, E. (2009). Mirroring and mu rhythm involvement in social cognition: are there dissociable subcomponents of theory of mind?.*Biological psychology*, *80*(3), 306-314.

Piolino, P., Desgranges, B., & Eustache, F. (2009). Episodic autobiographical memories over the course of time: cognitive, neuropsychological and neuroimaging findings. *Neuropsychologia*, *47*(11), 2314-2329.

Plaut, D. C. (2002). Graded modality-specific specialisation in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology*, *19*(7), 603-639.

Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2010). Category-specific versus category-general semantic impairment induced by transcranial magnetic stimulation. *Current Biology*, *20*(10), 964-968.

Polich, J. (1985). N400s from sentences, semantic categories, number and letter strings?. *Bulletin of the Psychonomic Society*.

Polich, J., & Comerchero, M. D. (2003). P3a from visual stimuli: typicality, task, and topography. *Brain topography*, *15*(3), 141-152.

Polich, J. (2003). Theoretical overview of P3a and P3b. In *Detection of Change*(pp. 83-98). Springer US.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology*, *118*(10), 2128-2148.

Pollick, A. S., & de Waal, F. B. (2007). Ape gestures and language evolution.*Proceedings of the National Academy of Sciences*, *104*(19), 8184-8189.

Pravdich-Neminsky, V. V. (1913). Ein Versuch der Registrierung der elektrischen Gehirnerscheinungen. *Zbl Physiol*, *27*, 951-960.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and brain sciences*, *1*(04), 515-526.

Press, C. (2011). Action observation and robotic agents: learning and anthropomorphism. *Neuroscience & Biobehavioral Reviews*, *35*(6), 1410-1418.

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, *1191*(1), 62-88.

Price, C. J., Wise, R. J. S., Watson, J. D., Patterson, K., Howard, D., & Frackowiak, R. S. (1994). Brain activity during reading The effects of exposure duration and task. *Brain*, *117*(6), 1255-1269.

Pritchard, W. S., Shappell, S. A., & Brandt, M. E. (1991). Psychophysiology of N200/N400: A review and classification scheme. *Advances in psychophysiology*, *4*, 43-106.

Proverbio, A. M., Riva, F., & Zani, A. (2009). Observation of static pictures of dynamic actions enhances the activity of movement-related brain areas. *PLoS One*, *4*(5), e5389.

Puce, A., & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 435-445.

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *The Journal of Neuroscience*, *16*(16), 5205-5215.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *The Journal of Neuroscience*, *18*(6), 2188-2199.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., ... & Shaywitz, B. A. (2001). Neurobiological studies of reading and reading disability. *Journal of communication disorders*, *34*(6), 479-492.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*(7), 576-582.

Pulvermüller, F. (2008). Brain embodiment of category-specific semantic memory circuits.

Pulvermüller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, *11*(5), 351-360.

Pulvermüller, F., Härle, M., & Hummel, F. (2001). Walking or talking?: Behavioral and neurophysiological correlates of action verb processing. *Brain and language*, *78*(2), 143-168.

Pulvermüller, F., Lutzenberger, W., & Preissl, H. (1999). Nouns and verbs in the intact brain: Evidence from event-related potentials and high-frequency cortical responses. *Cerebral Cortex*, *9*(5), 497-506.

Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience*,*17*(6), 884-892.

Pylyshyn, Z. W. (1984). Computation and cognition: Towards a foundation for cognitive science. *Cambridge, MA*.

Quillian, M. R. (1969). The teachable language comprehender: a simulation program and theory of language. *Communications of the ACM*, *12*(8), 459-476.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, *435*(7045), 1102-1107.

Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly*, 29-53.

Raichle, M. E., & Snyder, A. Z. (2007). A default mode of brain function: a brief history of an evolving idea. *Neuroimage*, *37*(4), 1083-1090.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(2), 676-682.

Ralph, M. A. L., Patterson, K., & Hodges, J. R. (1997). The relationship between naming and semantic knowledge for different categories in dementia of Alzheimer's type. Neuropsychologia, 35(9), 1251-1260.

Ramachandran, V. S. (2000). Mirror neurons and imitation learning as the driving force behind “the great leap forward” in human evolution. *Edge Website article http://www. edge. org/3rd_culture/ramachandran/ramachandran_p1. html*.

Raposo, A., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences.*Neuropsychologia*, *47*(2), 388-396.

Rapp, B. C., Hillis, A. E., & Caramazza, A. (1993). The role of representations in cognitive theory: More on multiple semantics and the agnosias. *Cognitive Neuropsychology*, *10*(3), 235-249.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin*, 85, 618-660.

Rayner, K. (1979) "Eye guidance in reading: Fixation locations within words." *Perception* 8.1 21-30.

Rebai, M., Bernard, C., & Lannou, J. (1997). The Stroop's test evokes a negative brain potential, the N400. *International Journal of Neuroscience*, *91*(1-2), 85-94.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The EZ Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and brain sciences*, *26*(4), 445-476.

Reinholz, J., & Pollmann, S. (2005). Differential activation of object-selective visual areas by passive viewing of pictures and words. *Cognitive brain research*,*24*(3), 702-714.

Reinitz, M. T., Wright, E., & Loftus, G. R. (1989). Effects of semantic priming on visual encoding of pictures. *Journal of Experimental Psychology: General*,*118*(3), 280.

Riby, D. M., & Hancock, P. J. (2008). Viewing it differently: Social scene perception in Williams syndrome and autism. *Neuropsychologia*, *46*(11), 2855-2860.

Riby, D. M., & Hancock, P. J. (2009). Do faces capture the attention of individuals with Williams syndrome or autism? Evidence from tracking eye movements. *Journal of Autism and Developmental Disorders*, *39*(3), 421-431.

Rieger, J. W., Braun, C., Bülthoff, H. H., & Gegenfurtner, K. R. (2005). The dynamics of visual pattern masking in natural scene processing: A magnetoencephalography study. *Journal of Vision*, *5*(3).

Riek, L. D., Paul, P. C., & Robinson, P. (2010). When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry. *Journal on Multimodal User Interfaces*, *3*(1-2), 99-108.

Rizzolatti G, Luppino G. (2001).The cortical motor system. *Neuron* 31:889–901

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in neurosciences*, *21*(5), 188-194.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*, 169-192.

Rizzolatti, G., & Craighero, L. (2007). Language and mirror neurons. *Oxford Handbook of Psycholinguistics, Oxford University Press, Oxford*.

Rizzolatti, G., & Sinigaglia, C. (2008). *Mirrors in the brain: How our minds share actions and emotions*. Oxford University Press.

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*(4), 264-274.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research*, *3*(2), 131-141.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, *2*(9), 661-670.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2006). Mirrors in the mind. *Scientific American*, *295*(5), 54-61.

Rizzolatti, G., Luppino, G., & Matelli, M. (1998). The organization of the cortical motor system: new concepts. *Electroencephalography and clinical neurophysiology*, *106*(4), 283-296.

Robertson, E. M., Tormos, J. M., Maeda, F., & Pascual-Leone, A. (2001). The role of the dorsolateral prefrontal cortex during sequence learning is specific for spatial information. *Cerebral Cortex*, *11*(7), 628-635.

Robins, B., Dautenhahn, K., & Dickerson, P. (2012). Embodiment and Cognitive Learning–Can a Humanoid Robot Help Children with Autism to Learn about Tactile Social Behaviour?. In *Social Robotics* (pp. 66-75). Springer Berlin Heidelberg.

Robinson, J. L., Laird, A. R., Glahn, D. C., Blangero, J., Sanghera, M. K., Pessoa, L., ... & Fox, P. T. (2012). The functional connectivity of the human caudate: An application of meta-analytic connectivity modeling with behavioral filtering. *NeuroImage*, *60*(1), 117-129.

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*,*15*(8), 1261-1269.

Rodríguez-Ferreiro, J., Menéndez, M., Ribacoba, R., & Cuetos, F. (2009). Action naming is impaired in Parkinson disease patients. *Neuropsychologia*,*47*(14), 3271-3274.

Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychological review*, *111*(1), 205.

Rosenbaum, D. A. (2005). The Cinderella of psychology: the neglect of motor control in the science of mental life and behavior. *American Psychologist*, *60*(4), 308.

Rousselet, G. A., Macé, M. J. M., & Fabre-Thorpe, M. (2003). Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes.*Journal of Vision*, *3*(6).

Rousselet, G., Joubert, O., & Fabre-Thorpe, M. (2005). How long to get to the “gist” of real-world natural scenes?. *Visual Cognition*, *12*(6), 852-877.

Roy, D., Hsiao, K. Y., & Mavridis, N. (2003). Conversational robots: building blocks for grounding word meaning. In *Proceedings of the HLT-NAACL 2003 workshop on Learning word meaning from non-linguistic data-Volume 6*(pp. 70-77). Association for Computational Linguistics.

Roy, D., Hsiao, K. Y., & Mavridis, N. (2004). Mental imagery for a conversational robot. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, *34*(3), 1374-1383.

Rueschemeyer, S. A., Brass, M., & Friederici, A. D. (2007). Comprehending prehending: neural correlates of processing verbs with motor stems. *Journal of Cognitive Neuroscience*, *19*(5), 855-865.

Rueschemeyer, S. A., van Rooij, D., Lindemann, O., Willems, R. M., & Bekkering, H. (2010). The function of words: Distinct neural correlates for words denoting differently manipulable objects. *Journal of Cognitive Neuroscience*,*22*(8), 1844-1851.

Rugg, M. D., & Coles, M. G. (1995). *Electrophysiology of mind: Event-related brain potentials and cognition*. Oxford University Press.

Rumelhart, D. E., & McClelland, J. L. (1986). Parallel distributed processing: explorations in the microstructure of cognition. Volume 1. Foundations.

Rushworth, M. F. S., Behrens, T. E. J., & Johansen-Berg, H. (2006). Connection patterns distinguish 3 regions of human parietal cortex. *Cerebral Cortex*, *16*(10), 1418-1430.

Saccuman, M. C., Cappa, S. F., Bates, E. A., Arevalo, A., Della Rosa, P., Danna, M., & Perani, D. (2006). The impact of semantic reference on word class: An fMRI study of action and object naming. *Neuroimage*, *32*(4), 1865-1878.

Saffran, E. M., Coslett, H., Martin, N., & Boronat, C. B. (2003). Access to knowledge from pictures but not words in a patient with progressive fluent aphasia. *Language and cognitive processes*, *18*(5-6), 725-757.

Santos, L. R., & Nissen, A. G. Ferrugia, 1. A.(2006). Rhesus monkeys, Macaca mulatta, know what others can and cannot hear. *Animal Behavior*, *71*, 1175-81.

Sarubbo, S., De Benedictis, A., Maldonado, I. L., Basso, G., & Duffau, H. (2013). Frontal terminations for the inferior fronto-occipital fascicle: anatomical dissection, DTI study and functional considerations on a multi-component bundle. *Brain Structure and Function*, *218*(1), 21-37.

Sass, K., Krach, S., Sachs, O., & Kircher, T. (2009). Lion–tiger–stripes: neural correlates of indirect semantic priming across processing modalities.*Neuroimage*, *45*(1), 224-236.

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M. S., ... & Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, *105*(46), 18035-18040.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: the role of the temporo-parietal junction in “theory of mind”. *Neuroimage*, *19*(4), 1835-1842.

Saxe, R., Jamal, N., & Powell, L. (2006). My body or yours? The effect of visual perspective on cortical body representations. *Cerebral Cortex*, *16*(2), 178-182.

Saygın, A. P., Wilson, S. M., Dronkers, N. F., & Bates, E. (2004). Action comprehension in aphasia: linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychologia*, *42*(13), 1788-1804.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews Neuroscience*, *8*(9), 657-661.

Schmahmann, J. D., Pandya, D. N., Wang, R., Dai, G., D'Arceuil, H. E., de Crespigny, A. J., & Wedeen, V. J. (2007). Association fibre pathways of the brain: parallel observations from diffusion spectrum imaging and autoradiography. *Brain*, *130*(3), 630-653.

Schmidt, G. L., DeBuse, C. J., & Seger, C. A. (2007). Right hemisphere metaphor processing? Characterizing the lateralization of semantic processes.*Brain and language*, *100*(2), 127-141.

Schmitt, B., Münte, T, and Kutas, M. (2000) "Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming." *Psychophysiology* 37.4 473-484.

Schwanenflugel, P. J., & LaCount, K. L. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(2), 344.

Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *The Journal of Neuroscience*, *25*(47), 11055-11059.

Scorolli, C., & Borghi, A. M. (2007). Sentence comprehension and action: Effector specific modulation of the motor system. *Brain research*, *1130*, 119-124.

Scorolli, C., Borghi, A. M., & Glenberg, A. (2009). Language-induced motor activity in bi-manual object lifting. *Experimental Brain Research*, *193*(1), 43-53.

Seghier, M. L., Lazeyras, F., Pegna, A. J., Annoni, J. M., Zimine, I., Mayer, E., ... & Khateb, A. (2004). Variability of fMRI activation during a phonological and semantic language task in healthy subjects. *Human brain mapping*, *23*(3), 140-155.

Seltzer, B., & Pandya, D. N. (1989). Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. *Journal of Comparative Neurology*,*281*(1), 97-113.

Seltzer, B., & Pandya, D. N. (1994). Parietal, temporal, and occipita projections to cortex of the superior temporal sulcus in the rhesus monkey: A retrograde tracer study. *Journal of Comparative Neurology*, *343*(3), 445-463.

Senior, C., Barnes, J., Giampietroc, V., Simmons, A., Bullmore, E. T., Brammer, M., & David, A. S. (2000). The functional neuroanatomy of implicit-motion perception or 'representational momentum'. *Current Biology*, *10*(1), 16-22.

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in cognitive sciences*, *7*(11), 489-493.

Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a time-line of word recognition: evidence from eye movements and event-related potentials.*NeuroReport*, *9*(10), 2195-2200.

Sergent, J., Zuck, E., Lévesque, M., & MacDonald, B. (1992). Positron emission tomography study of letter and object processing: Empirical findings and methodological considerations. *Cerebral Cortex*, *2*(1), 68-80.

Shafy, S (2011) "An Infinite Loop in the Brain". The Science of Memory. Spiegel Online.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge University Press.

Shallice, T., Warrington, E. K., & McCarthy, R. (1983). Reading without semantics. The Quarterly Journal of Experimental Psychology, 35(1), 111-138.

Shapiro, K. A., Moo, L. R., & Caramazza, A. (2006). Cortical signatures of noun and verb production. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(5), 1644-1649.

Shapiro, K. A., Mottaghy, F. M., Schiller, N. O., Poeppel, T. D., Flüß, M. O., Müller, H. W., ... & Krause, B. J. (2005). Dissociating neural correlates for nouns and verbs. *Neuroimage*, *24*(4), 1058-1067.

Shapiro, L. P., Brookins, B., Gordon, B., & Nagel, N. (1991). Verb effects during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(5), 983.

Shimazu, H., et al. "Macaque ventral premotor cortex exerts powerful facilitation of motor cortex outputs to upper limb motoneurons." *The Journal of neuroscience* 24.5 (2004): 1200-1211.

Silveri, M. C., & Ciccarelli, N. (2007). The deficit for the word-class "verb" in corticobasal degeneration: Linguistic expression of the movement disorder?.*Neuropsychologia*, *45*(11), 2570-2579.

Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: Reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, *20*(3-6), 451-486.

Simmons, W. K., & Martin, A. (2009). The anterior temporal lobes and the functional architecture of semantic memory. *Journal of the International Neuropsychological Society*, *15*(5), 645.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, *303*(5661), 1157-1162.

Siri, S., Tettamanti, M., Cappa, S. F., Della Rosa, P., Saccuman, C., Scifo, P., & Vigliocco, G. (2008). The neural substrate of naming events: Effects of processing demands but not of grammatical class. *Cerebral Cortex*, *18*(1), 171-177.

Sitnikova, T., Holcomb, P. J., Kiyonaga, K. A., & Kuperberg, G. R. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience*, *20*(11), 2037-2057.

Sitnikova, T., Kuperberg, G., & Holcomb, P. J. (2003). Semantic integration in videos of real–world events: An electrophysiological investigation.*Psychophysiology*, *40*(1), 160-164.

Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex*, *17*(10), 2387-2399.

Small, S. L., Noll, D. C., Perfetti, C. A., Hlustik, P., Wellington, R., & Schneider, W. (1996). Localizing the lexicon for reading aloud: Replication of a PET study using fMRI. *NeuroReport*, *7*(4), 961-965.

Smilek, D., Birmingham, E., Cameron, D., Bischof, W., & Kingstone, A. (2006). Cognitive ethology and exploring attention in real-world scenes. *Brain Research*,*1080*(1), 101-119.

Smith, E. E., & Medin, D. L. (1981). Categories and concepts (pp. 51-83). Cambridge, MA: Harvard University Press.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological review*, *81*(3), 214.

Smith, F. (1985). *Reading without nonsense* . Teachers College Press.

Sommerville, J. A., & Decety, J. (2006). Weaving the fabric of social interaction: Articulating developmental psychology and cognitive neuroscience in the domain of motor cognition. *Psychonomic Bulletin & Review*, *13*(2), 179-200.

Spencer, K. M., & Polich, J. (1999). Poststimulus EEG spectral analysis and P300: attention, task, and probability. *Psychophysiology*, *36*(2), 220-232.

Spiridon, M., Fischl, B., & Kanwisher, N. (2006). Location and spatial profile of category-specific regions in human extrastriate cortex. *Human brain mapping*,*27*(1), 77-89.

Spreng, R. N., Mar, R. A., & Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *Journal of cognitive neuroscience*,*21*(3), 489-510.

Sprenger-Charolles, L. (1989). L'apprentissage de la lecture et ses difficultés: approche psycho-linguistique. *Revue française de pédagogie*, 77-106.

Springer, J. A., Binder, J. R., Hammeke, T. A., Swanson, S. J., Frost, J. A., Bellgowan, P. S., ... & Mueller, W. M. (1999). Language dominance in neurologically normal and epilepsy subjects A functional MRI study. *Brain*,*122*(11), 2033-2046.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man.*Electroencephalography and clinical neurophysiology*, *38*(4), 387-401.

Sreenivasa, M., Souères, P., & Laumond, J. (2012). Walking to grasp: Modeling of human movements as invariants and an application to humanoid robotics. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, *42*(4), 880-893.

Stasse, O., Saïdi, F., Yokoi, K., Verrelst, B., Vanderborght, B., Davison, A., ... & Esteves, C. (2008). Integrating walking and vision to increase humanoid autonomy. *International Journal of Humanoid Robotics*, *5*(02), 287-310.

Steels, L. (1998). The origins of syntax in visually grounded robotic agents.*Artificial Intelligence*, *103*(1), 133-156.

Steels, L. (2001). Language games for autonomous robots. *Intelligent Systems, IEEE*, *16*(5), 16-22.

Steels, L. (2003). Evolving grounded communication for robots. *Trends in cognitive sciences*, *7*(7), 308-312.

Steels, L., & Kaplan, F. (2002). Aibos first words: The social learning of language and meaning. *Evolution of communication*, *4*(1), 3-32.

Strafella, A. P., & Paus, T. (2000). Modulation of cortical excitability during action observation: a transcranial magnetic stimulation study. *Neuroreport*,*11*(10), 2289-2292.

Stuss, D. T., & Knight, R. T. (2002). Principles of frontal lobe functioning.

Subiaul, F., Cantlon, J. F., Holloway, R. L., & Terrace, H. S. (2004). Cognitive imitation in rhesus macaques. *Science*, *305*(5682), 407-410.

Sugita, Y., & Tani, J. (2005). Learning semantic combinatoriality from the interaction between linguistic and behavioral processes. *Adaptive Behavior*,*13*(1), 33-52.

Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2009). Cortical midline involvement in autobiographical memory. *Neuroimage*, *44*(3), 1188-1200.

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: a meta-analysis. *Neuropsychologia*,*44*(12), 2189-2208.

Swanson, H. L., & Berninger, V. (1995). The role of working memory in skilled and less skilled readers' comprehension. *Intelligence*, *21*(1), 83-108.

Swinney, D. A. (1979). Lexical access during sentence comprehension:(Re) consideration of context effects. *Journal of verbal learning and verbal behavior*,*18*(6), 645-659.

Szpunar, K. K., Watson, J. M., & McDermott, K. B. (2007). Neural substrates of envisioning the future. *Proceedings of the National Academy of Sciences*,*104*(2), 642-647.

Taïx, M., Flavigné, D., & Ferré, E. (2012). Human Interaction with Motion Planning Algorithm. *Journal of Intelligent & Robotic Systems*, *67*(3-4), 285-306.

Talairach, J.&Tournoux, P. (1988) Co-Planar Stereotaxic Atlas of the Human Brain (Thieme, New York).

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in cognitive sciences*, *3*(4), 151-162.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *The Journal of Neuroscience*, *16*(13), 4240-4249.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1997). Oscillatory γ-band (30–70 Hz) activity induced by a visual search task in humans. *The Journal of Neuroscience*, *17*(2), 722-734.

Tallon-Baudry, C., Bertrand, O., Peronnet, F., & Pernier, J. (1998). Induced γ-band activity during the delay of a visual short-term memory task in humans.*The Journal of neuroscience*, *18*(11), 4244-4254.

Tanaka, S., & Inui, T. (2002). Cortical involvement for action imitation of hand/arm postures versus finger configurations: an fMRI study. *Neuroreport*,*13*(13), 1599-1602.

Tattersall, I. (2010). Human evolution and cognition. *Theory in Biosciences*,*129*(2-3), 193-201.

Taylor, K., Mandon, S., Freiwald, W. A., & Kreiter, A. K. (2005). Coherent oscillatory activity in monkey area V4 predicts successful allocation of attention. *Cerebral Cortex*, *15*(9), 1424-1437.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., ... & Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, *17*(2), 273-281.

Theoret, H., Halligan, E., Kobayashi, M., Fregni, F., Tager-Flusberg, H., & Pascual-Leone, A. (2005). Impaired motor facilitation during action observation in individuals with autism spectrum disorder. *Current Biology*, *15*(3), R84-R85.

Thiebaut de Schotten, M., Dell'Acqua, F., Valabregue, R., & Catani, M. (2012). Monkey to human comparative anatomy of the frontal lobe association tracts.*Cortex*, *48*(1), 82-96.

Thiebaut De Schotten, M., Cohen, L., Amemiya, E., Braga, L. W., & Dehaene, S. (2012). Learning to Read Improves the Structure of the Arcuate Fasciculus.*Cerebral Cortex.*

Thirioux, B., Mercier, M. R., Jorland, G., Berthoz, A., & Blanke, O. (2010). Mental Imagery of Self-Location during Spontaneous and Active Self–Other Interactions: An Electrical Neuroimaging Study. *The Journal of Neuroscience*,*30*(21), 7202-7214.

Thompson-Schill, S. L., D'Esposito, M., & Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation.*Neuron*, *23*(3), 513-522.

Todorov, A., Harris, L. T., & Fiske, S. T. (2006). Toward socially inspired social neuroscience. *Brain research*, *1079*(1), 76-85.

Tokuno, H., & Nambu, A. (2000). Organization of nonprimary motor cortical inputs on pyramidal and nonpyramidal tract neurons of primary motor cortex: an electrophysiological study in the macaque monkey. *Cerebral Cortex*, *10*(1), 58-68.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition.*Behavioral and brain sciences*, *28*(5), 675-690.

Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., ... & Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, *15*(4), 3215.

Tranel, D., Adolphs, R., Damasio, H., & Damasio, A. R. (2001). A neural basis for the retrieval of words for actions. *Cognitive Neuropsychology*, *18*(7), 655-674.

Tremblay, P., & Small, S. L. (2011). From language comprehension to action understanding and back again. *Cerebral Cortex*, *21*(5), 1166-1177.

Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta psychologica*, *116*(2), 185-203.

Tulving, E. (1972). Episodic and Semantic Memory1. *Organization of memory*, 381-402.

Tulving, E. (2005). *Episodic memory and autonoesis: Uniquely human* (pp. 3-56). New York: Oxford University Press.

Tulving, E., & Szpunar, K. K. (2009). Episodic memory. *Scholarpedia*, *4*(8), 3332.

Tunmer, W. E., & Hoover, W. A. (1992). Cognitive and linguistic factors in learning to read.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433-460.

Turken, U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. *Frontiers in systems neuroscience*, *5*.

Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in cognitive sciences*, *5*(6), 244-252.

Uddin, L. Q., Supekar, K., Amin, H., Rykhlevskaia, E., Nguyen, D. A., Greicius, M. D., & Menon, V. (2010). Dissociable connectivity within human angular gyrus and intraparietal sulcus: evidence from functional and structural connectivity. *Cerebral Cortex*, *20*(11), 2636-2646.

Umilta, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study.*Neuron*, *31*(1), 155-165.

Urgesi, C., Moro, V., Candidi, M., & Aglioti, S. M. (2006). Mapping implied body actions in the human motor system. *The Journal of Neuroscience*, *26*(30), 7942-7949.

Van Berkum, J. J. (2008). Understanding Sentences in Context What Brain Waves Can Tell Us. *Current Directions in Psychological Science*, *17*(6), 376-380.

Van Berkum, J. J., Van Den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of cognitive neuroscience*, *20*(4), 580-591.

Van Berkum, J. J., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research*, *17*(3), 701-718.

Van Dam, W. O., van Dijk, M., Bekkering, H., & Rueschemeyer, S. A. (2012). Flexibility in embodied lexical-semantic representations. *Human brain mapping*,*33*(10), 2322-2333.

Van Doren, L., Dupont, P., De Grauwe, S., Peeters, R., & Vandenberghe, R. (2010). The amodal system for conscious word and picture identification in the absence of a semantic task. *Neuroimage*, *49*(4), 3295-3307.

Van Doren, L., Schrooten, M., Adamczuk, K., Dupont, P., & Vandenberghe, R. (2012). Chronometry of word and picture identification: Common and modality-specific effects. *NeuroImage*, *59*(4), 3701-3712.

Van Elk, M., van Schie, H. T., Zwaan, R. A., & Bekkering, H. (2010). The functional role of motor activation in language processing: motor cortical oscillations support lexical-semantic retrieval. *Neuroimage*, *50*(2), 665-677.

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis.*Human brain mapping*, *30*(3), 829-858.

Van Petten, C. (1993). A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes*, *8*(4), 485-531.

Van Petten, C. (1995). Words and sentences: Event-related brain potential measures. *Psychophysiology*, *32*(6), 511-525.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: benefits, costs, and ERP components. *International Journal of Psychophysiology*, *83*(2), 176-190.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language.*Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 394.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, *383*(6597), 254-6.

Vander Wyk, B. C., Hudac, C. M., Carter, E. J., Sobel, D. M., & Pelphrey, K. A. (2009). Action understanding in the superior temporal sulcus region.*Psychological Science*, *20*(6), 771-777.

Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do?. *Nature Reviews Neuroscience*, *10*(11), 792-802.

Varela, F., Thompson, E., & Rosch, E. (1993). L'inscription corporelle de l'esprit. *Seuil, Paris*.

Vermeulen, N., Corneille, O., & Niedenthal, P. M. (2008). Sensory load incurs conceptual processing costs. *Cognition*, *109*(2), 287-294.

Vernon, D., Metta, G., & Sandini, G. (2007). A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents. *Evolutionary Computation, IEEE Transactions on*, *11*(2), 151-180.

Vidal, J. R., Freyermuth, S., Jerbi, K., Hamamé, C. M., Ossandon, T., Bertrand, O., ... & Lachaux, J. P. (2012). Long-distance amplitude correlations in the high gamma band reveal segregation and integration within the reading network. *The Journal of Neuroscience*, *32*(19), 6421-6434.

Vigliocco, G., & Kita, S. (2006). Language-specific properties of the lexicon: Implications for learning and processing. *Language and Cognitive Processes*,*21*(7-8), 790-816.

Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: a review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience & Biobehavioral Reviews*, *35*(3), 407-426.

Vigneau, M., Beaucousin, V., Herve, P. Y., DuVau, H., Crivello, F., Houde, O., Mazoyer, B. & Tzourio-Mazoyer, N. (2006).Meta- analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage*, 30, 1414 – 1432

Vigneau, M., Beaucousin, V., Hervé, P. Y., Jobard, G., Petit, L., Crivello, F., ... & Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing?: Insights from a meta-analysis. *NeuroImage*, *54*(1), 577-593.

Vilberg, K. L., & Rugg, M. D. (2008). Memory retrieval and the parietal cortex: a review of evidence from a dual-process perspective. *Neuropsychologia*, *46*(7), 1787-1799.

Villalobos, M. E., Mizuno, A., Dahl, B. C., Kemmotsu, N., & Müller, R. A. (2005). Reduced functional connectivity between V1 and inferior frontal cortex associated with visuomotor performance in autism. *Neuroimage*, *25*(3), 916-925.

Visser, M., Jefferies, E., & Ralph, M. L. (2010). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *Journal of Cognitive Neuroscience*, *22*(6), 1083-1094.

Vogeley, K., & Bente, G. (2010). "Artificial humans": Psychology and neuroscience perspectives on embodiment and nonverbal communication.*Neural Networks*, *23*(8), 1077-1090.

Wagner, A. D., Desmond, J. E., Demb, J. B., Glover, G. H., & Gabrieli, J. D. (1997). Semantic repetition priming for verbal and pictorial knowledge: A functional MRI study of left inferior prefrontal cortex. *Journal of Cognitive Neuroscience*, *9*(6), 714-726.

Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in cognitive sciences*, *9*(9), 445-453.

Wagner, D. D., Dal Cin, S., Sargent, J. D., Kelley, W. M., & Heatherton, T. F. (2011). Spontaneous action representation in smokers when watching movie characters smoke. *The Journal of Neuroscience*, *31*(3), 894-898.

Wallentin, M., Lund, T. E., Östergaard, S., Östergaard, L., & Roepstorff, A. (2005). Motion verb sentences activate left posterior middle temporal cortex despite static context. *NeuroReport*, *16*(6), 649-652.

Wang, L., Zhu, Z., & Bastiaansen, M. (2012). Integration or predictability? A further specification of the functional role of gamma oscillations in language comprehension. *Frontiers in psychology*, *3*.

Wang, W. S. Y. (1973). The Chinese language. *Scientific American*, *228*, 50-60.

Wang, Y. W., et al. "Person perception precedes theory of mind: an event related potential analysis." *Neuroscience* 170.1 (2010): 238-246.

Warburton, E., Wise, R. J., Price, C. J., Weiller, C., Hadar, U., Ramsay, S., & Frackowiak, R. S. (1996). Noun and verb retrieval by normal subjects studies with PET. *Brain*, *119*(1), 159-179.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301-1303.

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child development*, *77*(3), 640-663.

Warrington, E. K. (1975). The selective impairment of semantic memory. *The Quarterly journal of experimental psychology*, *27*(4), 635-657.

Warrington, E. K., & McCarthy, R. A. (1994). Multiple meaning systems in the brain: A case for visual semantics. *Neuropsychologia*, *32*(12), 1465-1473.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, *107*(3), 829-853.

Watkins, K. E., Dronkers, N. F., & Vargha-Khadem, F. (2002). Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain*, *125*(3), 452-464.

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*,*41*(8), 989-994.

Weiskopf, D. A. (2010). Embodied cognition and linguistic comprehension. *Studies In History and Philosophy of Science Part A*, *41*(3), 294-304.

Weiss, S., & Mueller, H. M. (2003). The contribution of EEG coherence to the investigation of language. *Brain and language*, *85*(2), 325-343.

West, W. C., & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research*, *13*(3), 363-375.

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of Us Disgusted in My Insula: The Common Neural Basis of Seeing and Feeling Disgust. *Neuron*, *40*(3), 655-664.

Wiggs, C. L., Weisberg, J., & Martin, A. (1998). Neural correlates of semantic and episodic memory retrieval. *Neuropsychologia*, *37*(1), 103-118.

Willems, R. M., & Casasanto, D. (2011). Flexibility in embodied language understanding. *Frontiers in psychology*, *2*.

Willems, R. M., & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: A review. *Brain and language*, *101*(3), 278-289.

Willems, R. M., Hagoort, P., & Casasanto, D. (2010). Body-specific representations of action verbs neural evidence from right-and left-handers.*Psychological Science*, *21*(1), 67-74.

Willems, R. M., Özyürek, A., & Hagoort, P. (2008). Seeing and hearing meaning: ERP and fMRI evidence of word versus picture integration into a sentence context. *Journal of Cognitive Neuroscience*, *20*(7), 1235-1249.

Williams, A. L., & Wright, M. J. (2009). Static representations of speed and their neural correlates in human area MT/V5. *NeuroReport*, *20*(16), 1466-1470.

Witt, J. K., Kemmerer, D., Linkenauger, S. A., & Culham, J. (2010). A functional role for motor simulation in identifying tools. *Psychological Science*,*21*(9), 1215-1219.

Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *Neuroimage*, *62*(1), 356-366.

Woods, D. L. (1990). Selective auditory attention: Complex processes and complex ERP generators. *Behavioral and Brain Sciences*, *13*(02), 260-261.

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: emergent features of word, sentence, and narrative comprehension.*Neuroimage*, *25*(3), 1002-1015.

Yokoyama, S., Miyamoto, T., Riera, J., Kim, J., Akitsuki, Y., Iwata, K., ... & Kawashima, R. (2006). Cortical mechanisms involved in the processing of verbs: an fMRI study. *Journal of Cognitive Neuroscience*, *18*(8), 1304-1313.

Yu, C. (2005). The emergence of links between lexical acquisition and object categorization: A computational study. *Connection Science*, *17*(3-4), 381-397.

Yuill, N., & Oakhill, J. (1991). *Children's problems in text comprehension: an experimental inverstigation*. Cambridge University Press.

Yuill, N., Oakhill, J., & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British journal of psychology*, *80*(3), 351-361.

Yum, Y. N., Holcomb, P. J., & Grainger, J. (2011). Words and pictures: An electrophysiological investigation of domain specific processing in native Chinese and English speakers. *Neuropsychologia*, *49*(7), 1910-1922.

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *The Journal of neuroscience*, *11*(3), 641-649.

Zhu, Z., Hagoort, P., Zhang, J. X., Feng, G., Chen, H. C., Bastiaansen, M., & Wang, S. (2012). The anterior left inferior frontal gyrus contributes to semantic unification. *NeuroImage*, *60*(4), 2230-2237.

Zwaan, R. A. (2003). The immersed experiencer: Toward an embodied theory of language comprehension. *Psychology of learning and motivation*, *44*, 35-62.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*(1), 1.

Annexes

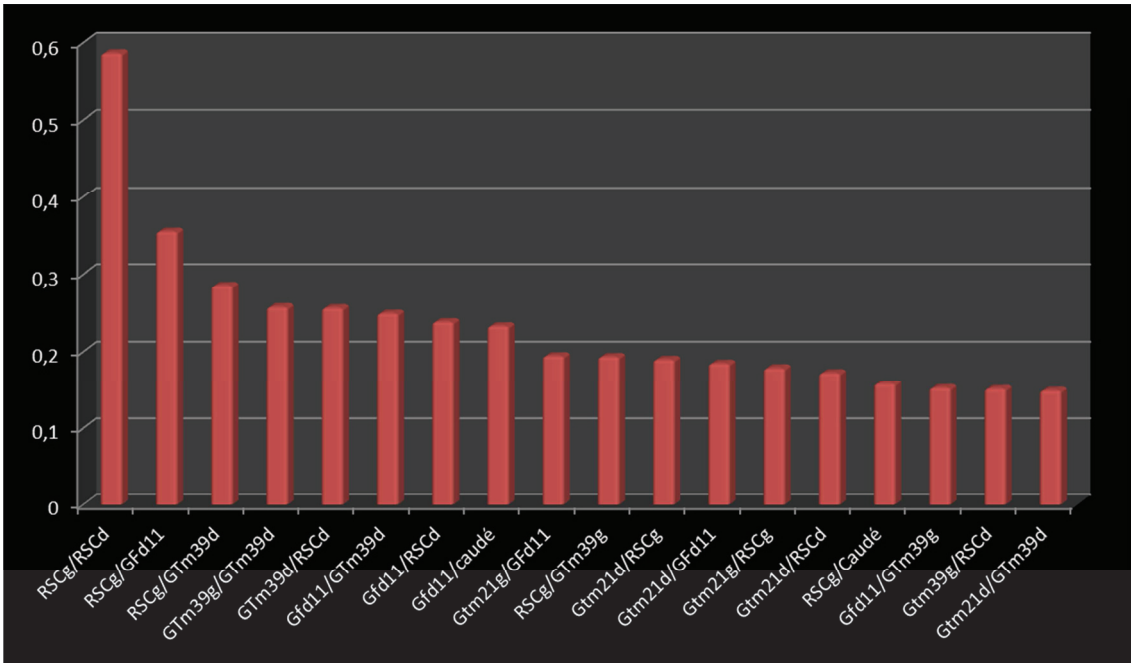
Annexe 1

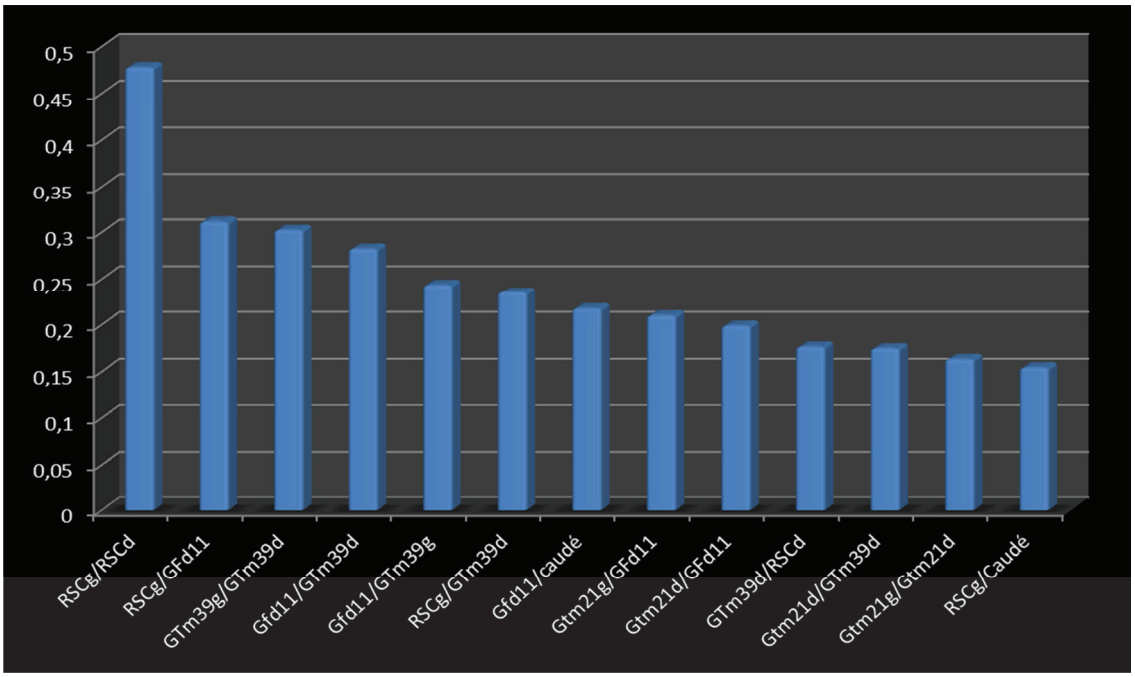
Répartition équilibrée des stimuli Phrases pour le protocole IRMf-DTI

	SET A session 1	SET A Session 2	SET B session 1	SET B Session 2
Résultats pré-tests				
Facilité de représentation	4,64	4,71	4,45	4,72
Degré de dynamisme (phrases dynamiques)	4,45	4,43	4,68	4,44
Degré de dynamisme (phrases statiques)	1,40	1,33	1,33	1,30
Fréquences Lexique				
Fréquence Films	227,47	267,09	253,90	263,91
Fréquences Livres	250,03	266,02	261,21	258,34
Caractéristiques des phrases				
Nombre moyen de mots	7,99	7,50	7,50	7,48
Longueur (nb caractères)	39,78	38,88	38,06	39,23
Verbes intransitifs	16	18	18	17
Verbes transitifs	27	25	25	26
Verbes réflexifs	7	7	7	6
Adverbes	4	1	1	2
Prépositions	45	44	43	44
Adjectifs subjectifs	29	25	26	27
Adjectifs objectifs	7	4	4	5
Caractéristiques des personnages				
Petite fille	12	12	13	11
Petit garçon	12	5	7	13
Femme	11	19	18	10
Homme	15	14	12	16

Annexe 2

Corrélations globales significatives (p=0.01) issues de l'analyses ICA pour les réseaux Phrases (figure du haut) et Images (figure du bas)

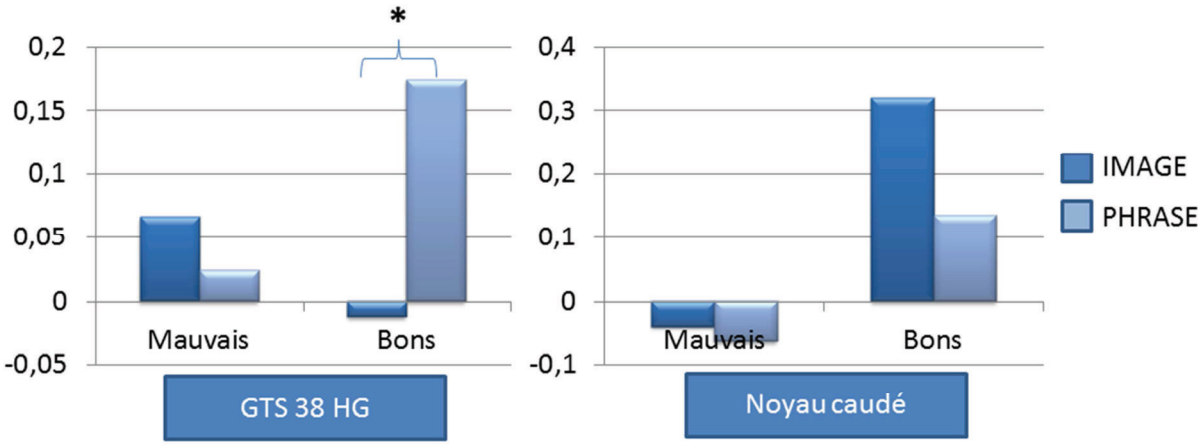




Annexe 3

Résultats complémentaires IRMf : analyse statistiques sur les β (pourcentage de changement du signal IRMf)

Nous avons réalisé une analyse statistique supplémentaire sur les β (pourcentage de changement du signal IRMf) des ROIs issues de l'analyse de conjonction. L'analyse était de type Anova , elle comprenait deux facteurs intra-sujets (le facteur Modalité (Images/Phrases) et le facteur Stimuli (Stimuli Dynamiques/Stimuli statiques/Simuli Contrôles) et un facteur inter-sujets déterminés par le COR (bons vs mauvais représentateurs). Les résultats ont confirmé qu'il n'y avait pas de différence entre Stimuli Dynamiques et Stimuli Statiques et peu de différence entre entre Images et Phrases. En revanche, dans le cadre de la comparaison « bons » vs « mauvais représentateurs », ces analyses ont apporté quelques éléments intéressants. Parmi toutes les ROIs étudiées, seules deux ROIs présentent des différences entre les deux groupes : le noyau caudé et le gyrus temporal supérieur de l'HG. Pour le caudé, on trouve un effet global du facteur Groupe qui est significatif : cette région est plus activée chez les « bons » que chez les « mauvais représentateurs », quelle que soit la condition (image ou phrase), même si on note une tendance plus forte pour les images. Tandis que pour le gyrus temporal supérieur, il y a une interaction significative entre les facteurs Groupe et Condition : cette région est particulièrement activée chez les « bons représentateurs » lorsqu'on leur présente des phrases**.**



Annexe 4

Répartition équilibrée des stimuli pour le protocole comportemental

Les stimuli ont été répartis entre les sets grâce aux critères de la base Lexique (FreqFilms et FreqLivres), il a aussi été vérifié que chaque session et condition contenait une répartition proche (homme/femme, nombre de lettres…).

D'autres critères (qui ne sont pas détaillés ici) ont été pris en compte pour la répartition, notamment le nombre de personnages faisant l'action (1 ou 2 personnages), répétitions d'un même verbe dans une même condition, etc…

	SET A visu ou SET B audio	SET A visu ou SET B audio	SET B visu ou SET A audio	SET B visu ou SET A audio
Condition	**SUITE**	**NON SUITE**	**SUITE**	**NON SUITE**
Fréquences Lexique				
Fréquence Films	146,39	153,19	149,72	156,22
Fréquences Livres	146,43	154,32	144,13	165,68
Nb caractères	70,88	71,28	71,36	71,12
Personnages				
TOTAL Hommes	13	15	13	12
TOTAL Femmes	14	13	15	16
TOTAL Adultes	17	18	18	20
TOTAL Enfants	11	10	10	8

Annexe 5

Questionnaire post-test du protocole comportemental

Date

Sujet N° :

NOM :

Prénom :

Date de naissance :

Sexe :

Latéralisation :

Niveau d'études :

Liste :

Questions

- Aimez-vous lire ?
- Combien de romans lisez-vous par an environ ?

Zéro <5 5-10 10-20 20-30 >30

- Sport

Fréquence (par semaine) : 1 2 3 4 5 6 7

- Est-ce que l'expérience vous a paru difficile ? OUI NON

- Est-ce qu'une modalité vous a paru plus difficile ? AUDIO VISION

- Autre remarques sur l'expérience :

……
……
……
……
……
……

- Comment vous représentiez-vous les phrases de cette expérience ? Aviez-vous une image ? Autre ?

……
……
……
……
……
……

Annexe 6

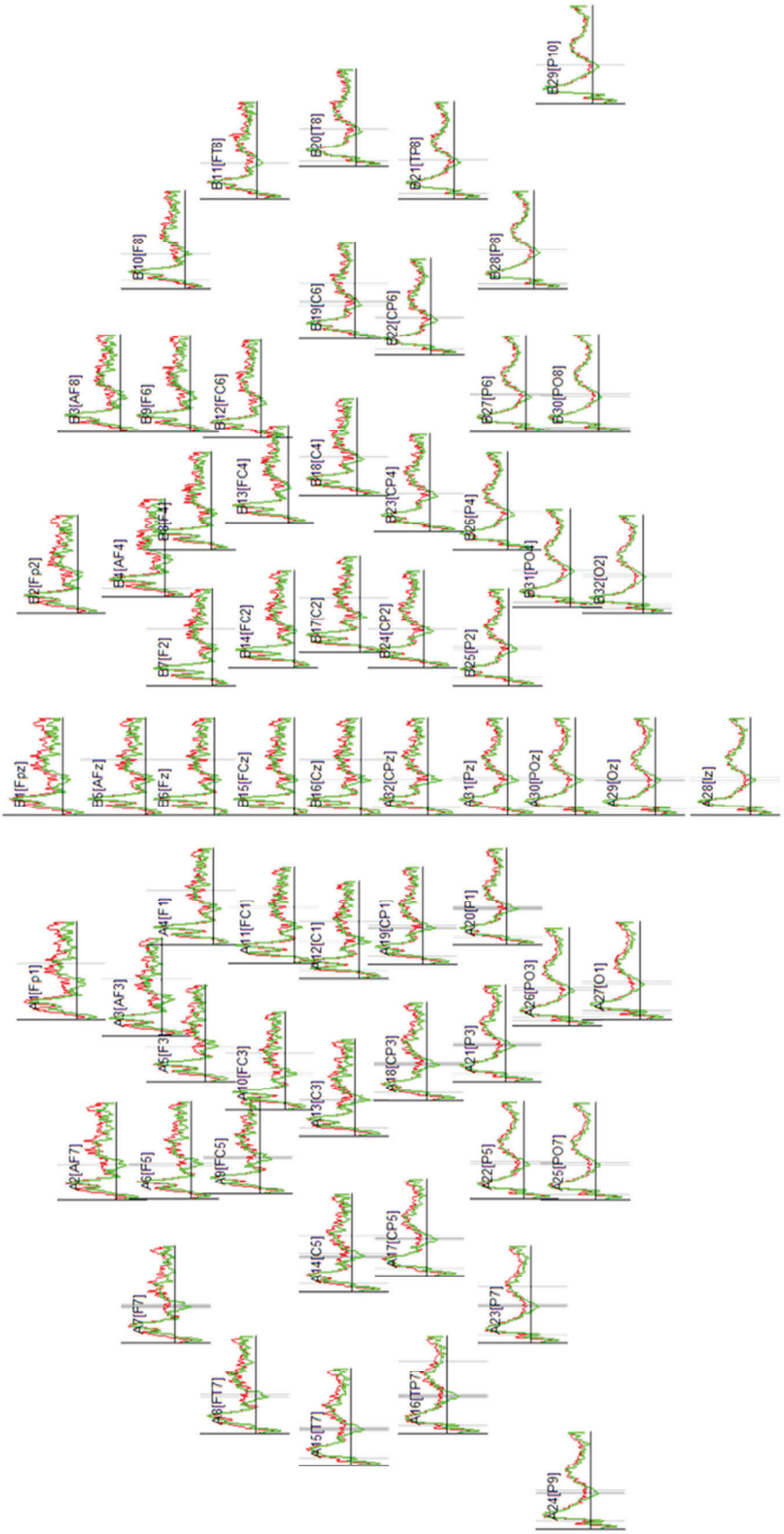
Répartition équilibrée des stimuli pour le protocole EEG

	Phrases Session 1	Phrases Session 2	Images Session 1	Images Session 2
Fréquences Lexique				
Fréquence Films	346,72	346,66		
Fréquences Livres	352,90	338,26		
Nb caractères	73,80	73,87		
Orientation Images				
Horizontale			25	25
Verticale			33	33
Carrée			2	2
Personnages				
TOTAL Hommes	39	40	41	40
TOTAL Femmes	41	40	52	52
TOTAL Adultes	49	49	61	64
TOTAL Enfants	31	32	31	28

Annexe 7

Visualisation des ERPs pour les 64 électrodes (scalp map) pour la condition IMAGE 1.

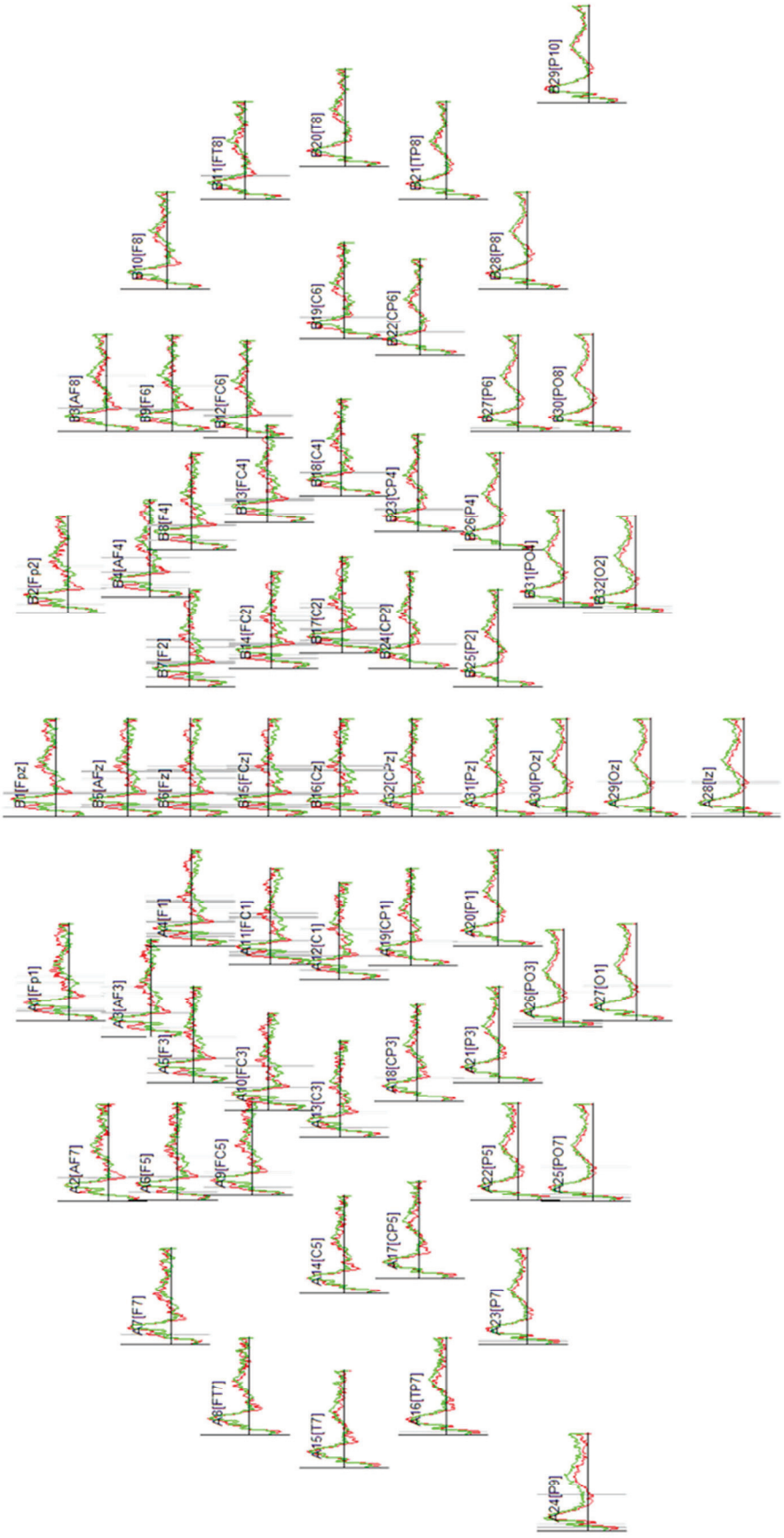
La condition Image 1 Suite est en vert, la condition Image 1 Non Suite est en rouge. Les périodes de temps où il y a une différence significative entre les deux conditions apparaissent en grisé.



Annexe 8

Visualisation des ERPs pour les 64 électrodes (scalp map) pour la condition IMAGE 2.

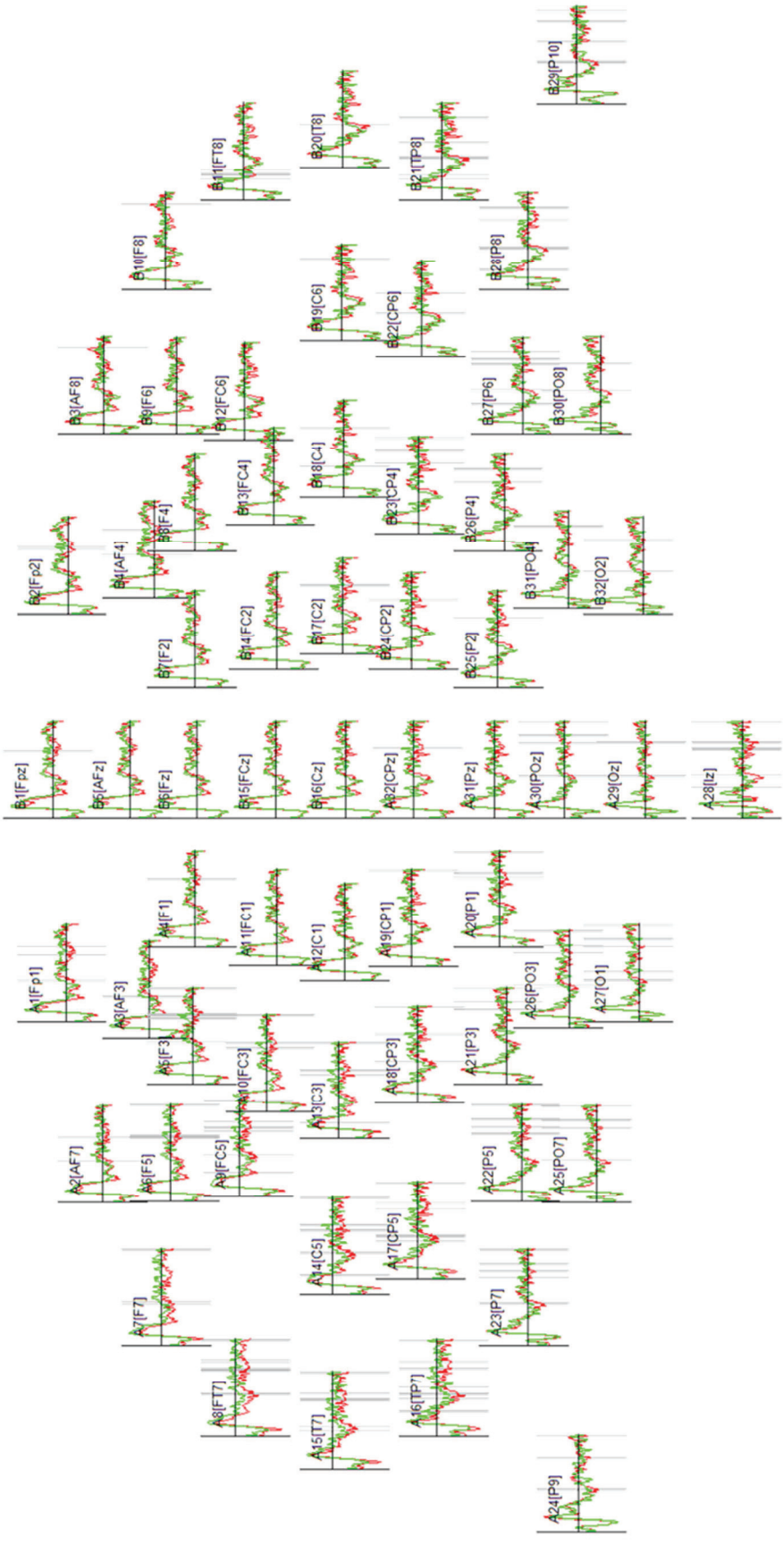
La condition Image 2 Suite est en vert, la condition Image 2 Non Suite est en rouge. Les périodes de temps où il y a une différence significative entre les deux conditions apparaissent en grisé.



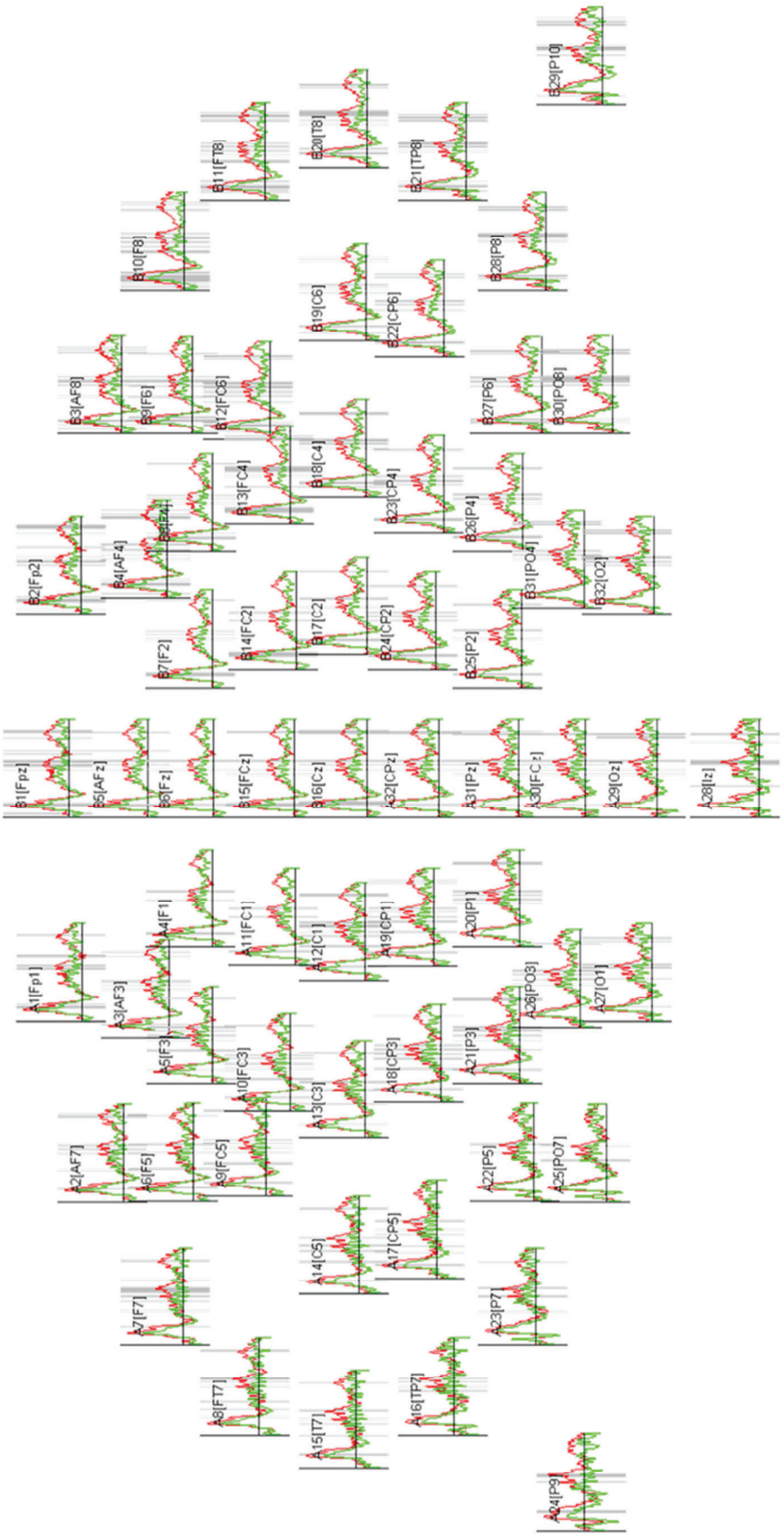
Annexe 9

Visualisation des ERPs pour les 64 électrodes (scalp map) pour la condition PHRASE 1.

La condition Phrase 1 Suite est en vert, la condition Phrase 1 Non Suite est en rouge. Les périodes de temps où il y a une différence significative entre les deux conditions apparaissent en grisé.



Annexe 10

Visualisation des ERPs pour les 64 électrodes (scalp map) pour la condition PHRASE 2.

La condition Phrase 2 Suite est en vert, la condition Phrase 2 Non Suite est en rouge. Les périodes de temps où il y a une différence significative entre les deux conditions apparaissent en grisé.