



**ISBN 978-2-87996-720-2**

Imprimé à Munsbach, mai 2007

© 2007 Carlos Paulos , Luxembourg (L) / Strasbourg (F)  
cpaulos@cept.lu

Le code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3°a, d'une part, que les «copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4). Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

Avec le soutien financier du  
fonds national de la recherche  
du Luxembourg (FNR)



UNIVERSITÉ LOUIS PASTEUR - STRASBOURG I  
Ecole Doctorale des Sciences de la Vie et de la Santé

**THESE**

présentée pour obtenir le grade de  
Docteur de l'Université Louis Pasteur Strasbourg I  
Discipline: Sciences de la Vie  
Spécialité: Neurosciences Cognitives

par

Carlos PAULOS

Les représentations visuelles d'objets et d'action :  
la spécificité du corps humain.

Soutenue publiquement le 16 mars 2007 devant la commission d'examen  
constituée par :

Directeur de Thèse : M. Claude BONNET  
Rapporteur Interne : Mme Marie-Noëlle METZ-LUTZ  
Rapporteur Externe : M. Michel-Ange AMORIM  
Rapporteur Externe : M. Bruno ROSSION  
Examineur : M. Christian SCHEIBER  
Membre Invité : M. Daniel GRUCKER

## Remerciements

Je me trouve enfin devant ma dernière page de cette thèse, ce moment « inoubliable » que j'ai tant attendu et que je n'aurais jamais atteint sans le soutien et l'accompagnement de toutes ces personnes ...

Mes premiers remerciements vont à mon directeur de thèse, le Pr. Claude Bonnet, pour son initiation au métier de chercheur, son encadrement de qualité, sa rigueur et sa passion envers la recherche. Claude, MERCI, pour votre soutien, vos encouragements et votre présence tout au long de ces dernières années.

Je tiens également à remercier Christian Scheiber, pour ton introduction à l'imagerie par résonance magnétique et tes avertissements aux traitements d'images, pour « ma première fois » dans le scanner et de m'avoir transmis la « magie » de la technique ...

Je remercie Lilianne Manning pour ses encouragements, son aide et ses conseils tout au long de ce travail.

Un grand merci aux membres du jury, Marie-Noëlle Metz-Lutz, Michel-Ange Amorim et Bruno Rossion d'avoir accepté d'examiner ce travail.

Merci au Pr. Daniel Grucker pour son accueil chaleureux au sein de son laboratoire. Tout particulièrement, je tiens à remercier Nathalie Heider (entre autre pour ses conseils d'anglais), Corinne Marrer pour le pilotage du T2, Daniel Gounot pour ses conseils et son aide autour de l'IRM, Alain, Anne-Claire, Jack et Mario pour toutes les discussions autour d'un café, et tous les membres de l'UMR 7004 pour leur intérêt les vendredis matin.

Je tiens également à remercier tout particulièrement Thierry Lustrement pour toutes ses interventions de « sauvetages » et pour son immense sympathie.

Merci à Elisabeth Demont pour m'avoir soutenu et encouragé lors de différents projets, et de m'avoir donné le goût à l'enseignement.

Merci à tous les sujets d'avoir participé et contribué à toutes ces expériences. Merci aussi aux étudiants de maîtrise Emilie, Nathalie, Raaj, Christelle, Mitsouko et Christophe pour leurs contributions et discussions autour de ce travail. Special thanks to Cindy Di Nicolo pour son travail artistique sur les stimuli et autres bricolages.

Un gigantesque merci à tous mes amis et amies de m'avoir accompagné et soutenu durant toutes ces années, et de m'avoir prêté leurs canapés lors d'innombrables soirées : Caroline, Bertrand, Damien aka « my ball », RV et Patricia (et spécialement MERCI pour toutes les relectures et corrections du manuscrit), Annabelle et Jean, Pauline, GeGe, Steve, Dali et Dom, Stéphane « wolf », Caroline, Jérôme, Karine, Steve Riot, Pol, Sander et Yves. Je remercie également Astrid et Nicole de m'avoir motivé « tous les jours » pendant ces 2 dernières années à achever ce travail. Un immense merci à Aliette Lauff pour la relecture et les corrections du manuscrit.

Merci à Romain Thill pour tous les « artmails » échangés ces dernières années. Merci à Anne-Marie Schuller pour les passionnants « brain-walking » et les échanges inspirants.

Merci à Catherine Sinner d'avoir « élargi mes horizons », pour ta patience tout au long de ce travail et pour toutes les escapades partagées ensemble. Merci à Lisza pour toutes les longues promenades.

En dernier lieu, je remercie toute ma famille, merci à Nicole, Rhett, Sophie, Véronique et Claude, et particulièrement mes parents, Maria de Fatima et Damião, de m'avoir toujours motivé à chercher mon chemin, pour leur confiance, leurs encouragements et leur accompagnement.

Dedicado os meus pais.

## Résumé

Ce travail de recherche étudie la spécificité du corps humain à générer des représentations visuelles dynamiques à partir d'images statiques (cf. mouvement implicite).

Lors d'une expérience en IRMf il apparaît que deux régions cérébrales jouent un rôle majeur dans les processus de traitement du corps humain : le gyrus fusiforme latéral (GFL) et l'aire du mouvement MT / V5. Nos résultats suggèrent que le GFL serait impliqué dans les traitements de représentations du corps humain, tandis que l'aire MT / V5 dans les traitements de représentations de mouvement.

Dans une seconde expérience en IRMf événementiel, l'introduction d'une tâche de catégorisation montre une activation cérébrale de l'aire MT / V5 lorsque différentes postures du corps humain sont jugées étant en mouvement, et différents objets inanimés étant des objets de mouvement – action.

Une série d'expériences comportementales montre dans certaines conditions, l'utilisation spontanée de représentations visuelles spécifiques au corps humain en respectant les contraintes biomécaniques de celui-ci. Ces connaissances sont transposées à d'autres objets inanimés anthropomorphes.

Une dernière série d'expériences a utilisé le phénomène de mouvement apparent présentant des séquences d'action à partir de différentes postures du corps humain. Un effet modulateur apparaît lorsque les différentes postures du corps humain constituent une séquence d'action cohérente.

Ces recherches mettent en évidence l'implication de traitements visuels spécifiques lorsque les stimuli représentent un corps humain. Particulièrement, elles montrent la mise en place de toute une panoplie de représentations de mouvement – action spécifiques au corps humain faisant le lien avec les actions possibles que celui-ci peut exécuter.

## SOMMAIRE

|                                                                                                                                             |    |
|---------------------------------------------------------------------------------------------------------------------------------------------|----|
| A. Introduction .....                                                                                                                       | 6  |
| I – La duplicité du système visuel .....                                                                                                    | 8  |
| II - Le traitement du mouvement visuel .....                                                                                                | 9  |
| III - Le mouvement biologique, un cas particulier du mouvement visuel .....                                                                 | 13 |
| IV - Les représentations visuelles du mouvement : le mouvement implicite à partir de la<br>voie ventrale .....                              | 15 |
| V - Les visages et le corps humain : des stimuli spécifiques ? .....                                                                        | 17 |
| B. Synthèse des travaux expérimentaux .....                                                                                                 | 23 |
| Exploration en imagerie par résonance magnétique fonctionnelle (IRMf) : .....                                                               | 23 |
| I - Représentations visuelles du corps humain et d'objets manufacturés : animés vs<br>inanimés et mouvement implicite vs statique .....     | 23 |
| II - De la perception du mouvement aux représentations visuelles liées au mouvement et<br>à l'action : activation de l'aire MT / V5 ? ..... | 37 |
| Expériences comportementales sur la spécificité du corps humain: .....                                                                      | 50 |
| I - Représentations visuelles d'actions à partir d'images statiques .....                                                                   | 50 |
| II - La spécificité du corps humain dans un paradigme de mouvement apparent .....                                                           | 63 |
| C. Discussion générale et perspectives .....                                                                                                | 77 |
| D. Bibliographie .....                                                                                                                      | 81 |
| E. Annexes .....                                                                                                                            | 91 |

## A. Introduction

L'objectif général de ce travail de recherche est d'étudier a) les conditions d'évocation de représentations de mouvement par des stimuli statiques (*cf. mouvement implicite*), b) l'existence de représentations biomécaniques dans l'évaluation des conséquences d'une posture corporelle c) la possibilité que des représentations sémantiques de ces propriétés puissent modifier les conséquences de traitements perceptifs précoces. Dans ce cadre d'investigation, la spécificité du corps humain à générer de telles représentations visuelles dynamiques a été comparée à différents objets inanimés.

La base théorique de ce travail est fondée sur les travaux de Freyd (1983). En effet, Freyd parvint à démontrer que la présentation de deux photographies d'une même situation, prises à deux moments différents, induit des *représentations visuelles du mouvement* reliant mentalement ces deux événements.

Avec l'émergence des nouvelles techniques d'imagerie cérébrale, en particulier l'imagerie par résonance magnétique fonctionnelle (IRMf), de nouvelles perspectives d'investigation des représentations visuelles du mouvement se sont ouvertes. Ainsi, Kourtzi & Kanwisher (2000) et Senior et al. (2000), ont démontré, grâce à la technique d'IRMf, que la présentation visuelle de photographies figeant un contexte particulier dynamique (i.e. un athlète en train de réaliser un exercice physique, une tasse en train de tomber d'une table), induit une activation bilatérale des régions médio-temporales (MT / V5) du cortex, régions spécifiques de la présentation visuelle de telles photographies. Cette même région MT / V5, qui a suscité un grand intérêt dans les études du traitement du mouvement visuel, est principalement impliquée dans l'intégration globale des informations du mouvement visuel (e.g. pour une revue Zeki (1993) ; Farah (2000)). Ainsi, en absence de mouvement visuel perceptif, Kourtzi & Kanwisher (2000), et Senior et al. (2000), proposent qu'à partir des stimuli présentés dans ces deux études, il y aurait évocation de représentations visuelles de mouvement (*c.f. implied motion* ou *mouvement implicite*), et que ces processus engendrés partageraient (au moins partiellement) les mêmes régions du cortex visuel spécialisées dans l'analyse du mouvement visuel.



### **Représentations visuelles de mouvement**

Connaissances (= traces mnésiques), acquises à travers l'expérience visuelle, sur la manière dont les différents objets de notre environnement se déplacent dans celui-ci, en prenant en compte leurs caractéristiques physiques. Ces traces mnésiques sont localisables dans le cerveau et, s'agissant du mouvement, dans les structures qui sont 'normalement' activées par un mouvement visuel (c'est-à-dire rétinien).

### **Mouvement implicite (implied motion)**

Informations relatives au mouvement (d'un objet) contenues dans une image figée. Le mouvement n'est pas perceptible directement, mais évoqué à partir des connaissances disponibles en mémoire : représentations visuelles de mouvement.

Dans la continuité des travaux princeps de Freyd (1983) sur les représentations visuelles de mouvement, Shiffrar et Freyd (1993) ont proposé d'étudier ces représentations dans le cadre d'un paradigme de mouvement apparent. Le mouvement apparent se caractérise par la présentation répétée de deux images représentant un même objet à deux positions spatiales distinctes, p. ex. A et B de chaque image. En présentant assez rapidement les deux images, le mouvement de la trajectoire la plus courte de la position A vers la position B, et vice-versa, est perçu. Dans le cadre de ce paradigme, différents objets ont été présentés (i.e. soit différentes parties du corps humain ou d'un mannequin en bois). De manière générale, la « règle » du chemin le plus court est valide, hormis lorsque l'objet se déplaçant est une partie du corps humain (i.e. un bras, une jambe). Un exemple illustrant ce phénomène spécifique est la présentation de deux images d'un même personnage ayant sa main de part et d'autre de sa tête. Lorsque ces deux images sont répétées alternativement avec des durées d'expositions relativement longues (600-800ms), un mouvement circulaire du bras en train de contourner la tête est perçu. Dans cet exemple spécifique, le chemin le plus court, qui correspond à la main traversant la tête, n'est pas perçu. Selon les auteurs, ces résultats laissent suggérer que lorsque l'objet présenté est un personnage ou une partie du corps humain (contrairement à un objet manufacturé), les contraintes biologiques du fonctionnement du corps humain sont prises en compte. Cette différence reposerait sur les effets induits par les représentations visuelles dynamiques du corps humain sur les traitements visuels effectués. Depuis ces expériences sur les représentations visuelles du mouvement peu d'intérêt a été porté à ces processus.

## I – La duplicité du système visuel

Le traitement de l'information visuelle est une tâche hautement complexe qui se réalise automatiquement et rapidement chez les sujets normaux. Dans un premier temps, ce travail de recherche s'inscrit dans le cadre général théorique de la séparation de deux systèmes de traitement de l'information visuelle. Il s'intéressera ensuite aux interconnexions possibles entre ces deux systèmes (à un niveau tardif du traitement de l'information, haut niveau ou *cognitif*; c'est-à-dire après les traitements visuels de bas niveau ou *sensoriels*).

Lorsqu'un objet est perçu, l'information visuelle est décomposée en différentes unités à partir de la rétine puis projetée, à travers le corps genouillé latéral (CGL), vers les différentes régions du cortex visuel. La mise en commun des différentes informations disponibles forme un percept : la **représentation visuelle** de l'objet.

Les processus de reconnaissance visuelle peuvent ainsi être décrits en plusieurs étapes de traitement de l'information jusqu'à l'identification de l'objet. Cette activité sensorielle et mentale (cognitive) se réalise en quelques centaines de millisecondes, donnant l'impression d'une perception directe.

D'un point de vue du traitement de l'information, la reconnaissance visuelle d'objets et de leurs localisations spatiales implique des mécanismes de traitement complexes qui sont fonctionnellement séparables : la voie ventrale répond à l'identification de l'objet (parvocellulaire, **PC**) et la voie dorsale à sa localisation spatiale (magnocellulaire, **MC**). Classiquement, après stimulation visuelle, on attribue le traitement visuel de la forme à la voie ventrale et celui du mouvement à la voie dorsale (e.g. Milner & Goodale, 1995). Néanmoins, ces traitements, fonctionnellement séparables, se réalisent en permanence de manière parallèle et avec une multitude d'interconnexions utilisant des processus de *feedback* et de *feedforward* (e.g. Bonnet & Lestienne, 2003).

Dans une première partie, nous nous intéresserons au traitement du mouvement visuel classiquement attribué à la voie dorsale, et plus particulièrement au traitement visuel du mouvement biologique (cf. Johansson, 1973), suggérant à partir des traitements effectués par la voie dorsale des projections vers la voie ventrale n'intervenant pas a priori dans le traitement du mouvement visuel.

Dans une deuxième partie, nous nous intéresserons aux traitements visuels de la voie ventrale (à partir de la présentation de formes) ainsi qu'aux activations possibles de celle-ci dans des régions cérébrales classiquement attribuées à la voie dorsale.

## II - Le traitement du mouvement visuel

La figure 2.1 illustre schématiquement l'exemple du traitement du mouvement visuel. Lorsqu'un objet se déplace dans le champ visuel (mouvement rétinien), l'information relative à son déplacement spatial ou à son mouvement est codée séparément par la voie dorsale, tandis que l'information relative à sa forme et à son identité est codée par la voie ventrale (e.g. pour une revue Milner & Goodale, 1995).

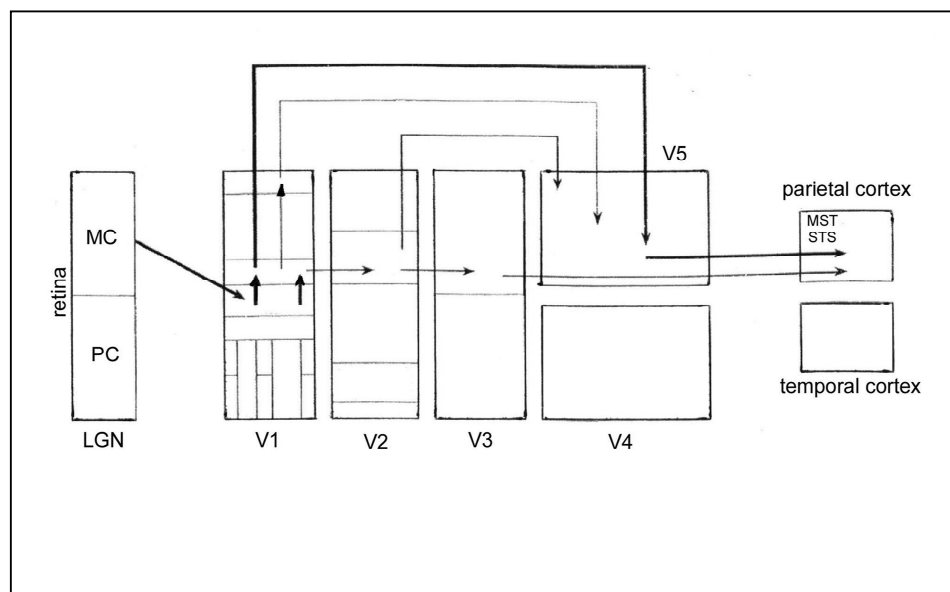


Figure 2.1 : Schéma fonctionnel simplifié des différentes étapes du traitement de l'information visuelle du mouvement partant de la rétine, à travers le corps genouillé latéral (LGN) vers le cortex occipital (V1, V2) pour se projeter vers le cortex médio temporal (MT / V5), médio temporal supérieur (MST) et pariétal (STS, sillon temporal supérieur). Issu de Milner & Goodale (1995).

A partir de la rétine, les informations relatives au mouvement visuel sont transmises à travers le corps genouillé latéral (cellules M, voie magnocellulaire, plus large et rapide) vers les régions occipitales du cerveau : le *cortex visuel primaire* (V1). Les études neurophysiologiques (enregistrements unitaires) chez le macaque ont mis en évidence, à partir de V1, des populations de neurones qui répondent préférentiellement à l'information de l'orientation et d'un mouvement (uni-) directionnel (cf. mouvement simple) d'un stimulus présenté visuellement (e.g. Newsome & Pare, 1988 ; Snowden et al., 1991 ; Zeki et al., 1993).

Récemment, des études en neuroimagerie fonctionnelle réalisées chez l'homme sain ont montré l'implication de tout un réseau neuronal du traitement de l'information du mouvement (e.g. Suneart et al., 1999 (figure 2.2) ; Dukelow et al., 2001). Ce réseau partirait du cortex occipital (V1), et passant par les régions inféro-pariétales (V3), il se poursuivrait vers les régions du cortex médio-temporal (MT / V5 ; MST ; STS) et du cortex pariétal (Ant et Occ SIP). Néanmoins, ce n'est qu'à partir de l'aire MT / V5 que les différentes informations, p. ex. (multi-) directionnelles (cf. mouvement complexe) sont intégrées à un niveau global de la perception visuelle du mouvement (e.g. Huk et al., 2001, 2002).

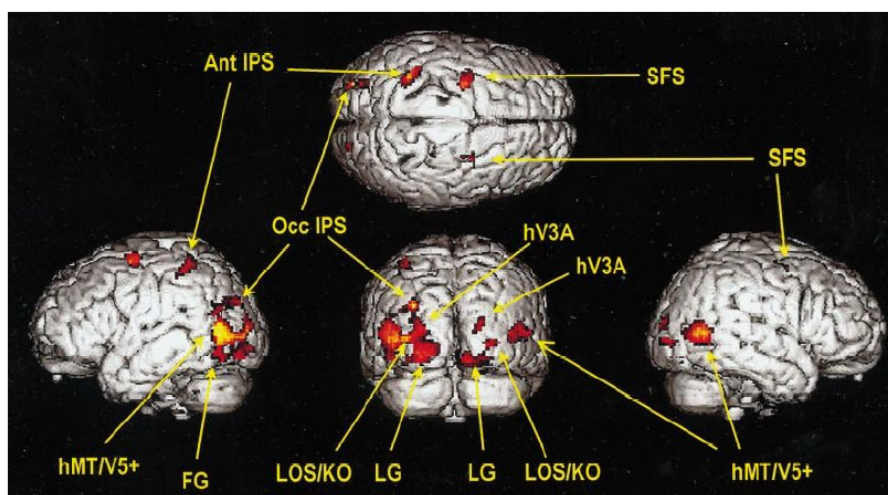


Figure 2.2 : Activation cérébrale du groupe (n=6) des régions activées lors d'une stimulation de points en mouvement complexe comparées aux mêmes points stationnaires (Suneart et al., 1999). Le changement de signal de la réponse BOLD est indiqué par les différences de couleurs (rouge = faible et jaune = importante). La région MT / V5 présente une activation maximale localisée bilatéralement à HG :  $x = -42$   $y = -66$   $z = 2$  et HD :  $x = 42$   $y = -62$   $z = 6$  ( $p < 0.0001$ , corrigé).

Plus particulièrement, Senior et ses collaborateurs (2002) ont montré, en utilisant la stimulation magnétique transcrânienne répétée (SMTr), le rôle crucial que joue l'aire MT / V5 dans l'intégration des informations d'un mouvement complexe chez l'homme sain. En effet, lorsqu'une lésion fonctionnelle réversible est induite par SMTr les sujets deviennent incapables d'intégrer correctement l'information visuelle d'un mouvement complexe.

Dans l'état actuel des connaissances, l'aire MT / V5 est considérée comme « le centre du mouvement visuel » (e.g. pour une revue Zeki, 1993 ; Farah, 2000, 2004). Les fondements de

l'implication majeure de l'aire MT / V5 dans le traitement du mouvement visuel proviennent tout d'abord des nombreuses études menées chez le singe (e.g. Newsome & Pare, 1988). Dès lors, le traitement du mouvement visuel a été également largement étudié chez l'homme à l'aide de différentes techniques de neuroimagerie (e.g. Sunaert et al., 1999 ; Dukelov et al. ; 2001 ; Paradis et al., 2000). Récemment, en IRMf, Orban et ses collaborateurs (2003) ont mené une étude comparative chez le macaque et l'homme qui leur a permis de mettre en évidence les activations d'un système spécialisé dans le traitement du mouvement visuel et la spécificité de l'aire MT / V5 chez le macaque, et de son homologue chez l'homme (cf. figure 2.3).

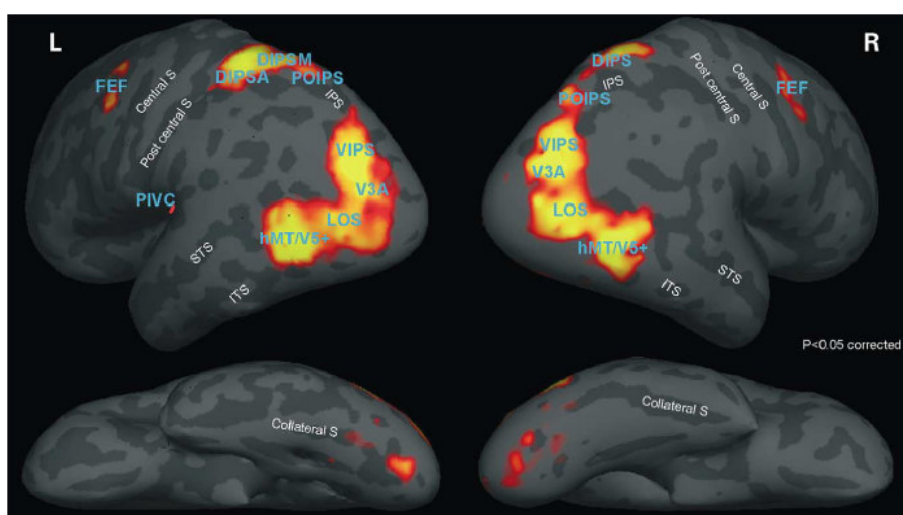


Figure 2.3 : Activation de l'aire MT / V5 chez le macaque (en bas) et de son homologue chez l'homme (en haut) faisant partie d'un réseau distribué du cortex pariétal ( $p < 0.05$ , corrigé). Issu de Orban et ses collaborateurs (2003).

Conjointement aux études expérimentales, les données de la neuropsychologie clinique soutiennent également l'hypothèse de l'aire MT / V5 comme « le centre du mouvement ». Le cas le mieux documenté de la littérature est celui de la patiente L.M. diagnostiquée comme

étant atteinte d'une agnosie visuelle du mouvement et suivie par Zihl et ses collègues depuis 1983.

**Étude de cas :** agnosie visuelle du mouvement (cf. visual akinesia)

Une dimension de l'information visuelle qui peut être sélectivement atteinte est celle du mouvement visuel. Le cas le mieux documenté dans la littérature est celui de la patiente L.M. qui présente une lésion dont la conséquence est un déficit de la perception visuelle du mouvement, décrit pour la première fois au début des années 80 par Zihl (e.g. Zihl et al., 1983, Zihl et al., 1991). Cette patiente, suite à une lésion bilatérale des régions médio temporales (cf. région MT), présente un diagnostic fonctionnel satisfaisant dans l'intégralité des fonctions cognitives avec une incapacité sélective à percevoir du mouvement visuel.

Lors de sa première consultation en 1980, L.M. se plaint de ne plus pouvoir réaliser certaines tâches de la vie quotidienne. Notamment, elle décrit : « lorsque je veux me verser une tasse de thé, je n'arrive pas à verser correctement le thé dans la tasse. Je ne vois plus le thé *fluide* couler de la théière, mais c'est comme si des morceaux de glaçons tombaient un par un dans la tasse ... » ... « ... aussi lorsque je m'apprêtais à traverser une rue, en regardant à gauche et à droite avant de traverser, j'aperçois une voiture au loin et m'apprêtant à traverser, soudainement la voiture était très proche et a failli me renverser ... ». Depuis sa première consultation, toutes les fonctions visuelles et cognitives ont été largement testées chez la patiente L.M.. Hormis pour la perception du mouvement visuel complexe, elle a obtenu des scores tout à fait satisfaisant à la moyenne. L. M. présente, après lésion bilatérale des régions médio temporales, un déficit sélectif de la perception du mouvement : une agnosie visuelle du mouvement. Cette patiente se trouve dans l'incapacité d'intégrer des informations visuelles relatives au mouvement. On pourrait ainsi décrire sa perception visuelle comme une succession d'images figées. La patiente décrit elle-même sa perception comme étant sous une lumière stroboscopique, c'est-à-dire comme la perception d'instantanés discrets sans aucune continuité. Par ailleurs, cette patiente a été suivie pendant une vingtaine d'années (e. g. Zihl, 1991 ; Schenk et al., 2000) sans que le diagnostic n'évolue au cours de cette période.

L'étude de cas de la patiente L.M. montre que le traitement du mouvement visuel peut être sélectivement déficitaire suite à une lésion bilatérale de l'aire MT / V5. Ces observations soulignent l'importance cruciale de l'aire MT / V5 à l'intérieur d'un réseau neuronal distribué

traitant l'information visuelle du mouvement à différents niveaux (e.g. thèse de doctorat de Sunaert, 2001).

### **III - Le mouvement biologique, un cas particulier du traitement du mouvement visuel**

En 1973, le suédois Johansson fut le premier à décrire le phénomène du *mouvement biologique* et sa particularité. Dans son expérience princeps, douze petites lampes sont fixées aux principales articulations du corps (épaules, bras, mains, hanches, jambes et pieds) d'un acteur vêtu d'un costume noir. L'acteur, en train de réaliser une action (p. ex. : marcher, courir, sauter), est filmé sur un fond noir. De cette manière, seuls les points lumineux étaient visibles. Lorsqu'on présente une seule image du film à des sujets, une configuration aléatoire de points lumineux est perçue. Dès que le film est mis en marche, le mouvement des différents points lumineux permet d'identifier visuellement l'action réalisée par le corps humain: *mouvement biologique*. De nombreux travaux ont été réalisés dans le but d'étudier différents aspects du mouvement biologique. Ainsi, Kozlowski & Cutting (1977) ont montré la capacité à différencier les sexes (i.e. homme ou femme). Les sujets sont également capables de reconnaître des proches et de se reconnaître eux-mêmes à partir de stimuli de mouvement biologique (pour une revue Decety & Grèze, 1999). La reconnaissance et l'identification visuelles, à partir de stimuli de type mouvement biologique, supposent l'accès aux régions cérébrales de la voie ventrale. Ce phénomène de reconnaissance visuelle à partir du seul mouvement d'une configuration de points est décrit dans la littérature qui propose le concept de *form from motion* (FFM). Ainsi, la perception de configurations particulières de points en mouvement (i.e. un mouvement biologique), dépourvues de forme visuelle explicite, permettrait l'accès à son identification. Celle-ci serait attribuée aux processus de traitements de l'information de la voie ventrale.

Récemment, Pavlova et al. (2004) ont montré, en magnétoencéphalographie (MEG), le décours temporel pour les stimuli de type mouvement biologique en soulignant l'implication des régions du sillon temporal supérieur (STS). En se basant sur des données d'IRMf, Vaina et al. (2001) proposent un réseau neuronal distribué qui s'étend du cortex occipital vers les

régions pariéto-temporales avec un maximum d'activation locale dans STS lors de la présentation de stimuli de type mouvement biologique.

Ces études réalisées par imagerie attribuent au sillon temporal supérieur (STS) un rôle crucial dans les processus de traitement de l'information de stimuli de type mouvement biologique, pour des images de corps humain entier, mais également pour des parties du corps (cf. la bouche en lecture labiale) et des animaux (e.g. un chien ou un cheval en train de courir) (e.g. Pinto et al., 2001, 2003).

Le point commun à ces différents processus cognitifs est l'accès à la voie ventrale (*à une représentation mentale sémantique*) à partir d'un stimulus dynamique visuellement appauvri, c'est-à-dire un mouvement visuel complexe d'une configuration de points lumineux dépourvue de forme explicite. Toutes ces expériences montrent bien, à partir des traitements visuels de mouvement (cf. voie dorsale), un accès à des informations (non perçues) traitées a priori par des processus de la voie ventrale (i.e. l'identification de l'objet). L'hypothèse des processus cognitifs sous-jacents est qu'à partir de la région STS, il y aurait des projections vers différentes régions du cortex médio temporal donnant accès à la représentation de la forme absente dans ce type de stimuli.

Du point de vue neuropsychologique, Batelli et al. (2003) présentent une étude incluant 3 patients présentant des atteintes pariétales. Pour ces 3 patients, il s'agit d'une lésion unilatérale localisée dans l'hémisphère droit pour 2 patients et dans l'hémisphère gauche pour le troisième. Les performances des patients sont évaluées à l'aide d'un éventail de tests de perception de mouvements visuels incluant des stimuli de mouvement biologique. Ces patients sont capables de discriminer l'orientation de rectangles présentés à partir d'un mouvement cohérent de points lumineux : accès à la forme (FFM). Néanmoins, ces mêmes patients présentent un déficit sélectif quand les stimuli présentés sont de type mouvement biologique. Les auteurs suggèrent que ces patients présentent un déficit dans le traitement du mouvement visuel de haut niveau, ne permettant pas soit l'activation de STS, soit des projections de la région STS vers les régions médio temporales. Ces résultats apportent des arguments supplémentaires quant au rôle prédominant de STS dans l'interprétation d'une stimulation visuelle de mouvement à caractère communicatif.



#### **IV - Les représentations visuelles du mouvement : le mouvement implicite à partir de la voie ventrale**

En 1983, Jennifer Freyd émet l'hypothèse de représentations relatives au mouvement à partir d'images figées. Dans son étude princeps, des photographies représentant une séquence d'action à deux moments différents, *avant et après*, (e.g. un personnage humain sautant d'un mur) sont présentées, soit dans l'ordre chronologique de l'événement, soit dans l'ordre inverse. Les sujets doivent juger, en temps de réaction (TR), si les deux photographies présentées font partie ou non de la même séquence. Les performances des sujets sont meilleures (TR les plus courts et % BR les plus élevés) pour la condition où les photographies sont présentées dans l'ordre chronologique des événements. Ces résultats sont interprétés par la proposition d'effets de représentations visuelles de mouvement reliant deux événements distincts représentés sur des photographies. Les fondements de représentations de mouvements étant ainsi posés, ils ont permis d'envisager de nouvelles perspectives de recherche portant sur la possibilité d'évoquer des représentations relatives au mouvement à partir d'images figées (cf. *mouvement implicite*), les propriétés de ces représentations dynamiques et leurs organisations, enfin les mécanismes neuronaux sous-tendant ce type de représentations (e.g. Verfaillie & d'Ydewalle, 1991; Freyd & Pantzer, 1995).

Dans la continuité de ces recherches, Maggie Shiffrar et ses collaborateurs (e.g. Shiffrar & Freyd, 1993 ; Chatterjee et al., 1996 ; Kourtzi & Shiffrar, 1999) reprennent l'idée de représentations dynamiques en utilisant dans un premier temps un paradigme de mouvement apparent *biologique*.

##### **Le mouvement apparent**

Le mouvement apparent se caractérise par la présentation successive et répétée de deux images représentant par exemple un même objet dans deux positions spatiales différentes, donnant ainsi lieu à la perception d'un mouvement. Le mouvement apparent de l'objet suit la trajectoire la plus courte (i.e. une ligne droite).

Plus précisément, Shiffrar & Freyd (1993) présentent deux photographies de la partie supérieure du corps humain avec la main positionnée de part et d'autre de la tête. Dans cette expérience l'intervalle inter stimulus (ISI) et la durée de présentation du stimulus (150ms) sont alternées par tranches de 50ms donnant lieu à 6 niveaux de stimulus onset asynchrony (SOA) allant de 150ms à 650ms. La tâche des sujets était de décrire la trajectoire du mouvement perçu (i.e. ligne droite ou mouvement circulaire). De manière générale, les résultats de cette expérience montrent que les sujets ne sont pas sensibles aux variations temporelles de présentations des stimuli dans un paradigme de mouvement apparent biologique. Ces résultats suggèrent que les sujets prennent en compte les contraintes biomécaniques du corps humain.

Dans le même groupe de recherche, l'expérience de Chatterjee et al. (1996) comparent le mouvement apparent biologique (i.e. des parties du corps humain) avec un mouvement apparent d'objets manufacturés (i.e. une poutre en bois, un mannequin en bois) (cf. figure 4.1 & figure 4.2).

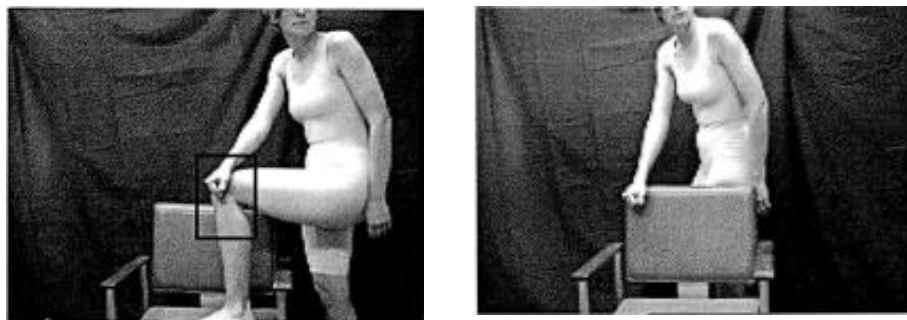


Figure 4.1 : Une paire de photographies donnant lieu à un mouvement apparent autour de la jambe (à gauche) et un mouvement apparent autour d'une chaise (à droite). Issu de Chatterjee et al. (1996).

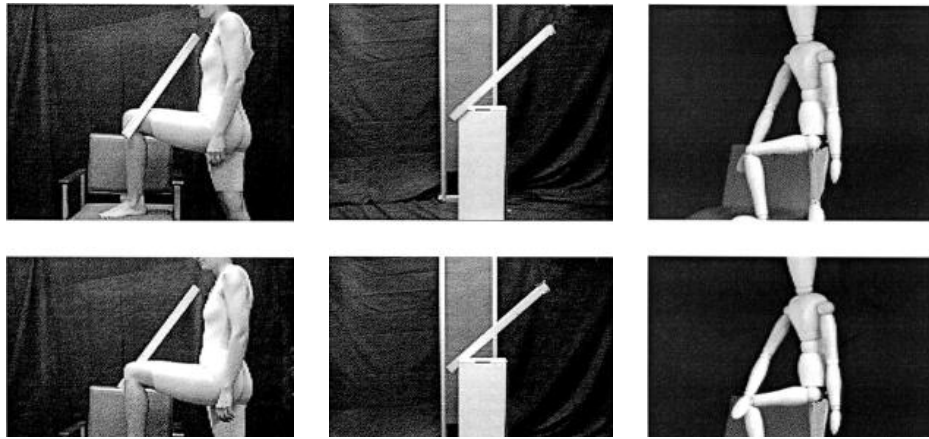


Figure 4.2 : Trois paires de photographies donnant lieu à (de gauche à droite) un mouvement apparent d'une poutre *autour* de la jambe, d'une poutre *traversant* un socle, et le mouvement apparent d'un bras d'un mannequin en bois *autour* de sa jambe. Issu de Chatterjee et al. (1996).

De manière générale, ces expériences montrent que, lorsque l'image représente une partie du corps humain, les contraintes biomécaniques et physiquement possibles sont prises automatiquement en compte. Ces résultats mettent en évidence des effets de représentations dynamiques spécifiques au corps humain (ou des parties du corps humain). Effectivement, lorsque les objets présentés sont des objets manufacturés (inanimés), le chemin le plus court de la trajectoire est perçu. A partir de ces expériences, l'hypothèse de représentations dynamiques spécifiques de parties du corps humain (i.e. un bras, une main et une jambe), par rapport à des objets manufacturés, est renforcée.

En revanche, la spécificité de représentations dynamiques relatives au corps humain entier a été peu étudiée. Mais une autre partie du corps humain a connu un intérêt particulier : le visage humain.

## V - Les visages et le corps humain : des stimuli spécifiques ?

Depuis une trentaine d'année, de nombreuses études se sont intéressées à la particularité ou à la spécificité du traitement perceptif et cognitif des visages humains.

Les premiers travaux en psychologie du développement ont montré, à l'aide d'un paradigme de préférence visuel, que les nouveaux nés, âgés de quelques semaines seulement, étaient attirés plus longtemps par des stimuli visuels de visages humains ou même par une simple configuration schématique représentant un visage (2 points, 1 trait) que par d'autres objets. Récemment, une étude portant sur la configuration schématique du corps humain entier

(Slaughter et al., 2002) a également révélé une préférence visuelle lorsque les stimuli (i.e. des dessins du corps humain) étaient présentés dans leur configuration structurelle naturelle par rapport à une configuration mélangée (scrambled).

En psychologie comparative, des études récentes ont également mis en évidence des différences entre les processus cognitifs en jeu pour le traitement des visages de la même espèce et entre les espèces (e.g. Pascalis et Bachevalier, 1998). Ce résultat d'une spécificité dans le traitement des visages intra-espèce a été trouvé chez des primates humains et non humains. Ces travaux ont abouti à un résultat particulièrement intéressant qui met en évidence le caractère général de l'expertise de reconnaissance visuelle de visages à l'intérieur d'une espèce, mais également la capacité, chez les primates humains, à devenir experts de visages de primates non humains ou même d'autres animaux (e.g. Pascalis et al., 2001, 2002). Ces auteurs montrent alors chez des experts l'apparition d'un effet d'inversion similaire à celui des visages humains en présentant des têtes de moutons. L'effet d'inversion montre une baisse considérable de reconnaissance lorsque le stimulus est présenté en position de haut en bas.

En *neuropsychologie*, le cas de la prosopagnosie définie par l'incapacité à reconnaître des visages familiers (pour une revue Farah, 2004 ; Manning, 2006), mais par la préservation des autres fonctions cognitives, soutient l'hypothèse de processus spécifiquement dédiés au traitement visuel de visages humain. Ces observations soutiennent l'idée que le traitement visuel des visages se fait par un système indépendant ou partiellement indépendant du traitement visuel d'autres objets. Cette idée est renforcée par la dissociation inverse, c'est-à-dire, le cas du patient CK, qui présente une agnosie visuelle pour tous les objets accompagnée d'une reconnaissance visuelle préservée pour les visages humains (Moscovitch et al., 1997 ; Moscovitch & Moscovitch, 2000).

La spécificité des visages humains a été largement argumentée à travers un grand nombre de travaux expérimentaux. Des études comportementales ont mis en évidence cette spécificité en démontrant un effet d'inversion particulier aux visages humains (e.g. Yin, 1969 ; pour une revue Valentine, 1988 ; Rossion et al., 2000 ; Maurer et al., 2002) par rapport à une multitude d'autres objets manufacturés. Cet effet se caractérise par un abaissement significatif des performances en reconnaissance visuelle lorsque les visages sont présentés dans une orientation non canonique. En ce qui concerne le corps humain entier, ce n'est que récemment

(Reed et al., 2003) qu'un effet d'inversion a été montré, avec des résultats similaires à ceux obtenus pour les visages humains.

### **L'effet d'inversion**

Dans l'expérience princeps, Yin (1969) énonce l'effet d'inversion spécifique au visage humain. Cet effet se caractérise par l'incapacité à reconnaître un visage familier lorsqu'il est présenté en position de haut en bas (rotation de 180°). En particulier, lorsque l'objet est, par exemple, un visage humain ou un corps humain, le temps de décision des sujets, ainsi que le nombre de mauvaises réponses, augmentent considérablement (e.g. Reed et al., 2003).

Le traitement visuel des visages a également reçu un intérêt particulier parmi les travaux d'imagerie cérébrale. Les études réalisées en neuroimagerie montrent qu'une région bilatérale du cerveau est particulièrement sensible à la présentation de visages humains. Kanwisher et al. (1997) localisent, grâce à l'IRMf, une région du gyrus fusiforme latéral (avec une prédominance dans l'hémisphère droit) activée par la présentation de visages humains. Ils lui attribuent l'implication exclusive dans les traitements de ces visages en l'étiquetant : fusiform face area (FFA). Un argument à cette interprétation des résultats est soutenu par les données de la neuropsychologie clinique présentant des patients prosopagnosiques. Néanmoins, la même année, d'autres auteurs, Gauthier et al. (1997), interprètent l'implication de cette FFA en terme d'expertise fonctionnelle. En effet, ces auteurs montrent, en présentant différents stimuli (p.ex. : des greebles, des oiseaux ou des voitures, (Gauthier et al., 2000)) à des experts respectifs, des activations dans le gyrus fusiforme préalablement défini comme FFA et rebaptisent cette zone en « flexible fusiform area ». Un grand débat (toujours d'actualité) s'installe entre ces deux interprétations fonctionnelles concernant cette région et l'organisation cérébrale des représentations mentales de différents objets dans le cortex postérieur. Compte tenu des différents travaux réalisés jusqu'à ce jour, il est difficile de défendre une approche strictement modulaire, comme l'hypothèse d'une FFA pour les visages humains. Pour aboutir à un compromis entre ces deux interprétations, des efforts ont été réalisés par le groupe de Haxby et ses collaborateurs. En effet, ils émettent l'hypothèse des représentations visuelles en patterns de réseaux distribués (e.g. Haxby et al., 2001 ; Ishai et al. 1999). Selon ces auteurs, il existerait, dans un réseau distribué de différentes régions cérébrales, une région donnant une activité maximale pour une catégorie de stimuli (telle que

les visages), mais ils n'excluent pas d'autres régions cérébrales participant également au traitement de l'information du stimulus présenté. Ce point de vue n'exclue pas la proposition modulaire spécifique d'une catégorie sémantique. Il propose plutôt de nuancer la définition de module. En effet, l'hypothèse d'une région répondant préférentiellement à une catégorie sémantique de stimuli est maintenue tout en incluant une multitude d'autres régions distribuées qui participeraient aux traitements cognitifs mis en œuvre. Ainsi, des configurations différentes de réseaux de neurones ou des patterns d'activation pourraient refléter l'activation spécifique d'une stimulation.

Dans son modèle, Haxby et al. (2000) proposent l'exemple du traitement visuel des visages. Le traitement visuel des visages se réaliserait dans, au moins, deux systèmes distincts : un système principal (cf. core system) et un système d'extension (cf. extended system). Dans un premier temps, après les traitements visuels sensoriels dans le cortex occipital inférieur, les aspects variants (i.e. les mouvements faciaux, reconnaissance d'une émotion) seraient traités dans les régions du sillon temporel supérieur (STS) et les aspects invariants (i.e. configuration faciale, reconnaissance individuelle) dans le gyrus fusiforme latéral (GFL). D'autres régions du cerveau seraient également impliquées dans un système d'extension contribuant et complétant les traitements sensoriels et cognitifs du système principal.

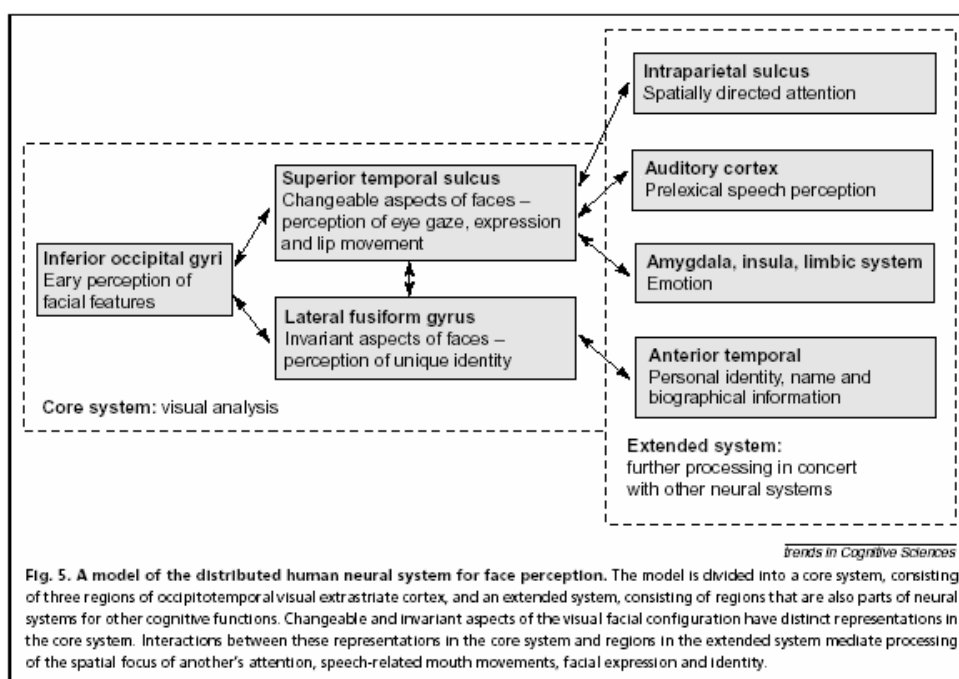


Figure 5.1 : Modèle du système neuronal distribué de la perception des visages chez l'homme. Issu de Haxby et al. (2000).

Malheureusement, ce modèle se focalise davantage sur un ensemble de différentes structures cérébrales activées et laisse de côté les nombreux apports de modèles fonctionnels de traitements des visages (e.g. Bruce et al., 1992 ; Hancock et al., 1996). De même, lors des traitements visuels du corps humain, se pose la question de la localisation de ces traitements : soit dans un module spécifique du cortex, soit dans plusieurs régions distribuées différenciant certaines caractéristiques types du corps humain (i.e. la forme, le mouvement).

Le point de départ de ce travail de recherche est formé par les études qui portent sur le traitement du mouvement visuel perceptif. Différentes données expérimentales et cliniques suggèrent que l'aire MT / V5 est la région critique (« le centre du mouvement ») à partir de laquelle les différentes informations relatives à un mouvement visuel complexe sont intégrées. Les expériences concernant les représentations du mouvement (e.g. Freyd, 1983) et le mouvement apparent (e.g. Shiffrar et al., 1993) suggèrent d'une part l'existence de représentations du mouvement et d'autre part des représentations dynamiques spécifiques à des parties du corps humain tenant compte des contraintes biomécaniques.

Les expériences réalisées en IRMf par Kourtzi & Kanwisher (2000) et Senior et ses collaborateurs (2000) montrent une activation bilatérale de l'aire MT / V5 lorsqu'un mouvement est évoqué (cf. mouvement implicite) à partir de la présentation de photographies de différents objets. En particulier, Kourtzi & Kanwisher (2000) montrent une activation maximale dans l'aire MT / V5 lorsque la photographie représente un athlète en train d'effectuer un exercice physique comparativement à une photographie du même athlète au repos. Néanmoins, les stimuli utilisés dans ces deux études contiennent un grand nombre d'informations visuelles de contexte difficilement contrôlables. Les expériences que nous avons réalisées visent en premier lieu à définir la stimulation minimale nécessaire et suffisante à évoquer un mouvement. Dans les expériences de Kourtzi & Kanwisher (2000) et Senior et al. (2000) les stimuli présentés étaient des photographies, images complexes véhiculant entre autres des informations de contexte. Le choix des stimuli étant critique, nous avons choisi de présenter des dessins au trait minimisant la complexité des stimuli. En considérant les quelques résultats présents dans la littérature, des dessins au trait de différentes postures d'un corps humain ont été considérées comme particulièrement « puissants » à évoquer un mouvement ou une action. Le corps humain présente une caractéristique intrinsèque à pouvoir se déplacer. Les dessins au trait du corps humain ont été comparés à des dessins au trait de différents objets manufacturés redessinés à partir de la

collection de stimuli de Snodgrass et Vandewart (1986). Se focalisant sur l'hypothèse de la spécificité de représentation du corps humain, différents paradigmes viseront à apporter quelques éléments supplémentaires suggérant la particularité et la nécessité de représentations visuelles spécifiques au corps humain.



## **B. Synthèse des travaux expérimentaux**

### **Exploration en imagerie par résonance magnétique fonctionnelle (IRMf) :**

Deux expériences ont été menées afin de tester la capacité de différents objets à évoquer des représentations de mouvement et d'action, et afin d'explorer leurs corrélats cérébraux avec les régions cérébrales activées lors du traitement de l'information d'un mouvement visuel.

Les expériences présentées ici ont été soumises à évaluation et acceptées par le comité consultatif de la protection des personnes dans la recherche biomédicale (CCPPRB Alsace – Numéro d'enregistrement CNRS : 01027).

### **I - Représentations visuelles du corps humain et d'objets manufacturés : animés vs inanimés et mouvement implicite vs statique**

Dans une première expérience réalisée en IRMf, des dessins aux traits de deux catégories sémantiques ont été présentés aux sujets: différentes postures d'un corps humain (Paulos, 2001) et des objets manufacturés (redessinés à partir de Snodgrass & Vanderwert, 1986). Cette expérience comprend deux objectifs. D'une part, elle vise à différencier les aires corticales impliquées dans le traitement visuel de chaque catégorie sémantique d'objets animés et inanimés et d'autre part à différencier des représentations visuelles de mouvement et d'action susceptibles d'être évoquées (en l'absence de mouvement perceptif) par les deux catégories sémantiques présentées. En effet, deux études de neuroimagerie révèlent que la présentation de photographies contenant ce que leurs auteurs appellent du *mouvement implicite* (Kourtzi & Kanwisher, 2000 ; Senior et al., 2000), permet d'induire des activations dans l'aire du mouvement MT / V5. Ces photographies représentent un contexte sémantique (une scène) décrivant soit une action d'un objet animé (i.e. un athlète effectuant un exercice physique), soit un mouvement, c'est-à-dire l'action d'un objet inanimé (i.e. une tasse en train de tomber d'une table).

Cette première expérience vise à vérifier ces résultats à partir de stimulations visuelles moins complexes que des photographies, c'est-à-dire en l'absence de contexte sémantique (i.e. des dessins aux traits) créé soit par la présence d'une scène visuelle, soit par l'introduction d'une tâche expérimentale.

## Matériel et Méthode

Dans une étude préliminaire, des dessins aux traits de différentes postures du corps humain ont été réalisés et présentés lors de quelques expériences psychophysiques (Paulos, 2001). Lors d'une tâche de catégorisation, les sujets devaient juger en temps de réaction si la posture du corps humain leur évoquait ou non une action. Cette tâche ne présentait aucune difficulté pour les sujets. A partir de ces résultats, deux sous catégories d'images du corps humain ont été créées : des postures du corps humain évoquant une action et des postures du corps humain au repos (10 exemplaires pour chaque sous catégorie). De même, des objets manufacturés choisis parmi la collection de Snodgrass et Vandewart (1986) ont été présentés lors de quelques expériences psychophysiques (Paulos, 2001). Les sujets devaient juger en temps de réaction si les objets étaient manipulables lors d'une action (i.e. un ballon, un balaie, ...etc.,) ou bien s'ils ne l'étaient pas (i.e. un ancre, un frigidaire). Comme précédemment, à partir des résultats obtenus, deux sous catégories d'objets ont été créées : des objets manipulables évoquant une action et des objets statiques (10 exemplaires pour chaque sous catégorie).

L'expérience a été réalisée sur une population de 10 sujets (5 femmes et 5 hommes) donnant leur consentement écrit et n'ayant aucun antécédent neurologique ou psychologique (examen médical d'inclusion au protocole). Tous les sujets étaient droitiers et avaient une vue normale ou corrigée à la normale (lentilles). Les sujets étaient allongés dans un IRM corps entier S200 à 2 Tesla (Bruker, Karlsruhe, Allemagne) et un système de lunettes à fibres optiques (Avotec Inc., Florida, USA), relié à un ordinateur, présentait les stimuli. La présentation des stimuli était synchronisée par le scanner au début de chaque bloc de stimuli. La variation locale du taux d'oxygénation dépendante du flux sanguin, réponse BOLD, a été mesurée grâce à une séquence d'imagerie ultra rapide (EPI) balayant le cerveau entier. Trente-deux coupes de 4 mm ayant une résolution de 64x64 (taille d'un voxel 4x4x4 mm<sup>3</sup>), étaient obtenues avec un temps de répétition (TR) de 5000 ms (TE = 43 ms, FA = 90°). Chaque session expérimentale était suivie d'une acquisition morphologique réalisée à l'aide d'une séquence turbo-RARE pondérée en T2 et une résolution de 128x128x64 (taille des voxels 4x4x4 mm<sup>3</sup> ; TR = 15000 ms ; TE = 73.8ms ; Turbofactor = 10). Une séance (session expérimentale et morphologique) durait approximativement 30 minutes.

Les dessins sont présentés en traits blancs sur fond noir ayant une luminance de  $10 \text{ cd/m}^2$  et une taille angulaire approximative de  $10^\circ \times 6^\circ$  d'arc. Les deux catégories sémantiques comprenant deux sous catégories définies a priori (Paulos, 2001) sont reprises. Un bloc est constitué de 20 présentations faites au hasard et composées de 10 exemplaires par (sous) catégorie. Chaque exemplaire de stimulus est répété une fois à l'intérieur d'un bloc (figure 1.1).

A chaque essai, une image est présentée pendant 166 ms avec un intervalle inter stimulus (ISI) de 1s. Tous les blocs ont été égalisés à 25s et répétés 5 fois au cours d'une session. Tous les blocs sont contrebalancés et alternés avec une période de repos de même durée : présentation d'une croix de fixation au milieu de l'écran.

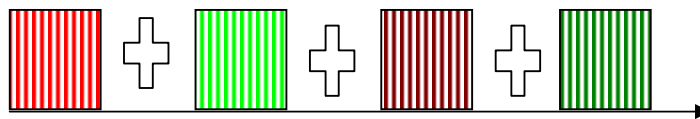


Figure 1.1 : Paradigme en bloc, présentation aléatoire des stimuli des deux catégories sémantiques à deux conditions définies a priori (Paulos, 2001) : corps humain en mouvement implicite (en rouge), corps humain statique (en rouge foncé), objets d'action implicite (vert) et objets statiques (vert foncé), périodes ON ; alternées avec la fixation d'une croix, périodes OFF. Toutes les périodes durent 25s.

Les sujets étaient avertis de la nature sémantique des stimuli (dessins aux traits du corps humain ou d'objets manufacturés ; cf. figure 1.2) et avaient pour seule consigne de bien maintenir leur regard vers le milieu du champ visuel (i.e. passive viewing).

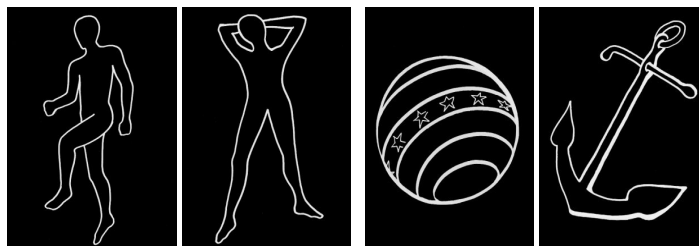


Figure 1.2 : Exemplaires des dessins aux traits présentés, de gauche à droite, selon la catégorie sémantique : corps humain et objets, et les sous catégories respectives : mouvement - au repos et mouvement - statique.

## Résultats

Les données obtenues en IRMf ont été analysées avec le logiciel SPM 99b (Wellcome Department of Cognitive Neurology, London ; <http://www.fil.ion.ucl.ac.uk/spm> ) suivant la méthode décrite dans Friston et al. (1995; 1997). Les scans de l'ensemble des sujets ont été prétraités en trois étapes : réalignement (superposition de toutes les images sur une image de référence), normalisation (transformation non linéaire des images dans l'espace anatomique de l'atlas de Talairach et Tournoux, 1984) et lissage spatial (filtre Gaussien FWHM de 4 mm).

Une première analyse de groupe (N=10) par comparaison multiple a été effectuée à l'aide d'un modèle statistique différenciant les 2 catégories sémantiques principales : corps humain et objets manufacturés.

Les données ont été ajustées et comparées au modèle statistique (figure 1.3) présentant une corrélation satisfaisante. L'évolution de la réponse BOLD mesurée suit celle prédite par le modèle statistique (réponse BOLD en box car). Des résultats similaires ont été obtenus auprès de tous les sujets.

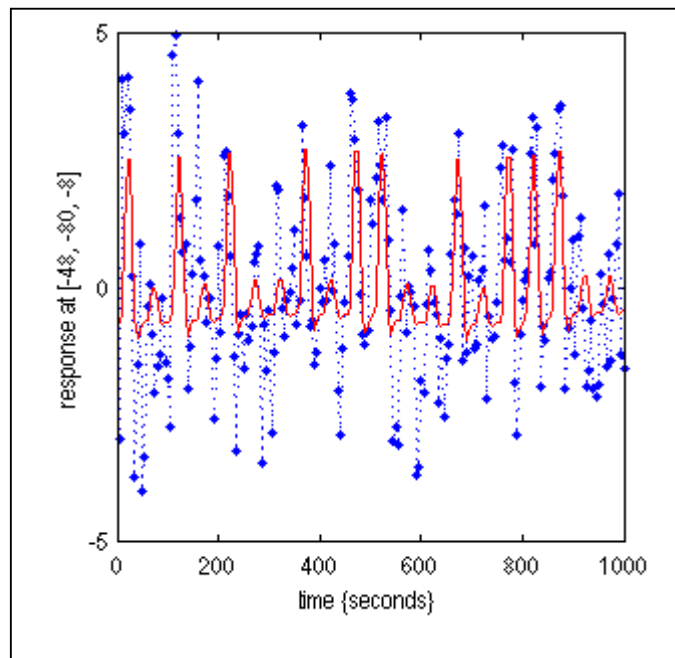


Figure 1.3 : Ajustement de l'évolution de la réponse BOLD (en bleu) au modèle statistique (en rouge) en fonction du dérouls temporel. Exemple des réponses d'un sujet pour la comparaison des deux catégories sémantiques corps humain vs objets manufacturés.

Après cette première étape de vérification des données, des cartes statistiques (individuelles et sur le groupe) ont été obtenues par comparaisons multiples ( $p < 0.001$ , non corrigé et un seuil cluster de  $k = 30$  voxels) en utilisant des contrastes par masquage exclusif ( $p < 0.05$  ; non corrigé) entre les deux catégories sémantiques principales. Afin de déterminer l'activation cérébrale spécifique à chacune des catégories sémantiques, deux contrastes ont été utilisés : contraste (H) *corps humain* > *objets manufacturés*, et le contraste inverse (O) *objets manufacturés* > *corps humain*. Des cartes statistiques d'activation ont été obtenues pour chaque sujet individuellement et une analyse de groupe ( $N=10$ ) par comparaison multiple a été réalisée.

Le tableau 1 présente les différentes régions d'intérêt (ROIs) révélées par l'analyse de groupe ( $N=10$ ) pour les deux contrastes spécifiques (H et O) au seuil  $p < 0.001$  (non corrigé) et  $k = 30$  voxels. Les régions cérébrales ont été identifiées avec le logiciel Talairach Daemon (<http://biad02.uthscsa.edu/projects> ; Research Imaging Center, University of Texas, San Antonio).

|         | Hémisphère Gauche |     |     |            |             | Hémisphère Droit |     |     |            |             |
|---------|-------------------|-----|-----|------------|-------------|------------------|-----|-----|------------|-------------|
| Région  | X                 | Y   | Z   | $K_E$      | $E$         | X                | Y   | Z   | $K_E$      | $E$         |
| GFL (H) | -42               | -48 | -22 | <b>49</b>  | <b>2.9</b>  | 42               | -54 | -18 | <b>92</b>  | <b>4.3</b>  |
| GFM (O) | -24               | -48 | -24 | <b>318</b> | <b>-4.9</b> | 26               | -36 | -24 | <b>216</b> | <b>-4.3</b> |
| MT (H)  | -52               | -72 | 2   | <b>215</b> | <b>4.0</b>  | 52               | -70 | 6   | <b>628</b> | <b>4.2</b>  |
| LPS (O) | -22               | -78 | 34  | <b>92</b>  | <b>-2.3</b> | 34               | -72 | 46  | <b>59</b>  | <b>-3.0</b> |
| GFS (O) | -48               | -60 | -14 | <b>53</b>  | <b>-2.7</b> |                  |     |     |            |             |
| LPI (O) |                   |     |     |            |             | 48               | -52 | 52  | <b>34</b>  | <b>-2.7</b> |

Tableau 1 : Coordonnées de Talairach du groupe des différentes ROIs identifiées pour les contrastes corps humain < objets manufacturés (H) et objets manufacturés < corps humain (O). GFL = gyrus fusiforme latéral, GFM = gyrus fusiforme médian, MT = aire du mouvement, LPS (Ba19) = lobule pariétal supérieur, GFS = gyrus fusiforme supérieur, LPI = lobe pariétal inférieur ;  $K_E$  = étendue du cluster,  $E$  = amplitude de l'activation

Sur la figure 1.4, un premier cluster d'activation bilatérale spécifique du corps humain a été identifié dans le gyrus fusiforme latéral (GFL). De même, une activation bilatérale spécifique du corps humain a été localisée dans les régions médio temporales, incluant l'aire du mouvement MT / V5.

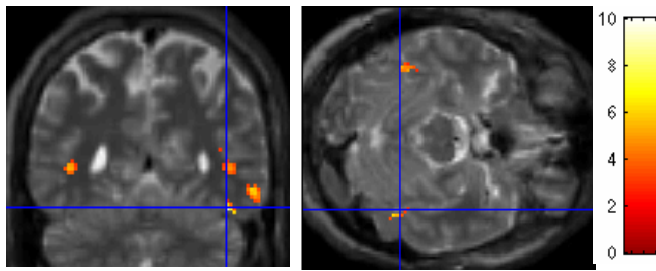


Figure 1.4 : Activation bilatérale dans le gyrus fusiforme latéral (GFL). L'intersection des lignes bleues se situe dans le GFL droit ( $z = -20$ ).

La figure 1.5 présente les activations bilatérales de l'aire MT sur une coupe coronale (A) et horizontale (B). L'étendue des clusters pour ces deux régions est plus large dans l'HD (GFL = 92 voxels ; MT = 628 voxels) que dans l'HG (GFL = 49 ; MT = 215 voxels).

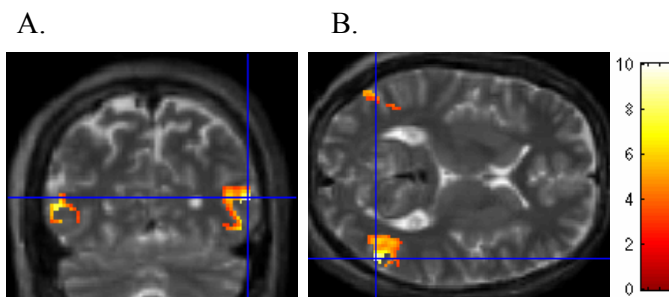


Figure 1.5 : Activation bilatérale dans l'aire MT / V5. L'intersection des lignes bleues montre l'activation dans l'aire MT / V5 droit ( $z = 6$ ).

Les activations spécifiques des objets ont été obtenues avec le contraste objets manufacturés < corps humain ( $p < 0.001$  ;  $k = 30$  voxels). La figure 1.6 présente une activation bilatérale spécifique des objets dans le gyrus fusiforme médian (GFM). Un deuxième site d'activation bilatérale (figure 1.7) est localisé dans le lobule pariétal supérieur (LPS). L'étendue de l'activation spécifique aux objets manufacturés est plus large dans l'hémisphère gauche (GFM = 318 voxels ; LPS = 92 voxels) que dans l'hémisphère droit (GFM = 216 voxels ; SPL = 59 voxels).

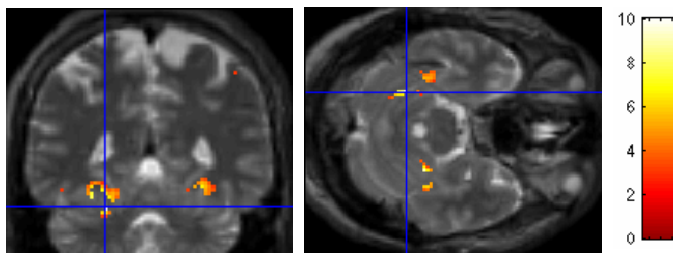


Figure 1.6 : Activation bilatérale spécifique aux objets manufacturés, identifiée dans le gyrus fusiforme médian (GFM). L'intersection des lignes bleues localise le GFM gauche ( $z = -24$ ).

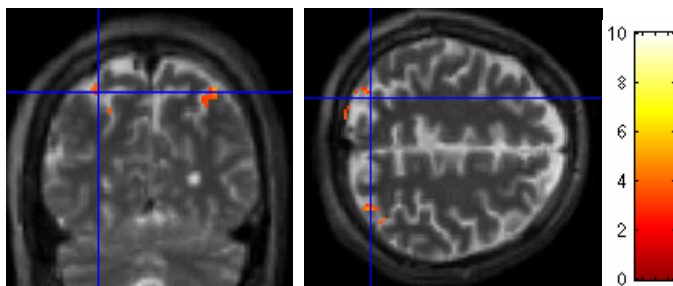


Figure 1.7 : Activation bilatérale spécifique aux objets manufacturés, identifiée dans le lobule pariétal supérieur (LPS). L'intersection des lignes bleues localise le LPS gauche ( $z = 44$ ).

Deux régions supplémentaires spécifiques des objets ont été identifiées, l'une dans le gyrus fusiforme supérieur gauche (GFS) et l'autre dans le lobe pariétal inférieur droit (LPI), l'étendue de l'activation étant respectivement de  $k = 53$  voxels et  $k = 34$  voxels ( $p < 0.001$ , non corrigé).

La figure 1.8 résume les résultats du groupe ( $N=10$  ;  $p < 0.001$  ;  $k = 30$ ) avec les activations spécifiques du corps humain (en rouge) et les activations spécifiques des objets manufacturés (en vert).

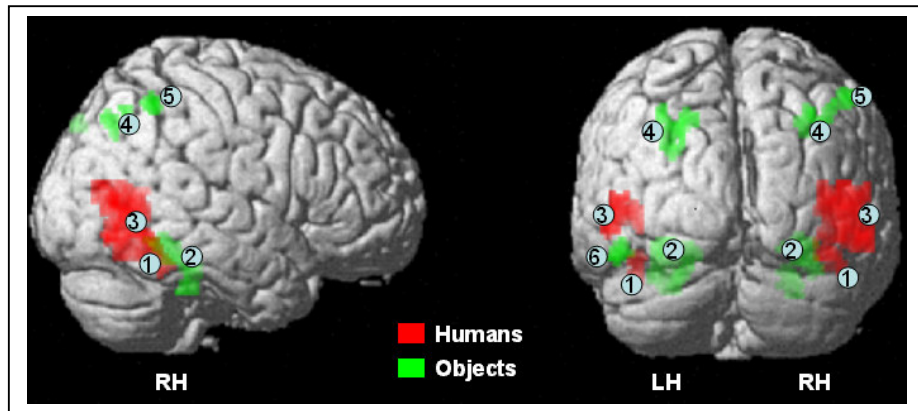


Figure 1.8 : Vue latérale (HD) et postérieure d'un cerveau modèle (SPM 99b) avec les activations spécifiques aux deux catégories sémantiques (en rouge pour le corps humain et en vert pour les objets) dans les différentes régions : 1-GFL ; 2-GFM ; 3-MT ; 4-LPS ; 5-LPI et 6-GFS.

Dans un deuxième temps, les résultats ont été analysés individuellement, sujet par sujet, afin d'estimer empiriquement la variabilité d'activation cérébrale inter sujet. Les cartes statistiques individuelles ont été obtenues avec SPM 99b en utilisant la même procédure et les mêmes seuils statistiques ( $p < 0.001$ , non corrigé;  $k = 30$  voxels) que ceux décrits auparavant.

Les différentes coordonnées des ROIs obtenues par l'analyse de groupe ont été prises comme points de référence afin de les comparer à ceux des sujets pris individuellement. L'activation maximale la plus proche spatialement de chaque ROI a été retenue pour chaque sujet afin de reconstruire des volumes cubiques individuels (VOIs) mesurant ainsi la dispersion spatiale de l'activation cérébrale. Suivant cette procédure pour les quatre ROIs principales et les deux contrastes (H et O), le tableau 2 présente les différentes VOIs avec leurs coordonnées moyennes et le nombre de sujets retenus pour la construction de chaque volume présentant une activation significative ( $p < 0.0001$ , non corrigé ;  $k = 30$ ).



| Hémisphère gauche |        |         |         |      | Hémisphère droit |        |         |        |      |      |
|-------------------|--------|---------|---------|------|------------------|--------|---------|--------|------|------|
| Région            | X      | Y       | Z       | N(H) | N(O)             | X      | Y       | Z      | N(H) | N(O) |
| GFL (H)           | -45(5) | -59(23) | -21(13) | 7    | 3                | 47(7)  | -61(15) | -17(5) | 7    | 0    |
| GFM (O)           | -30(8) | -49(19) | -18(6)  | 0    | 8                | 30(10) | -61(29) | -21(9) | 0    | 7    |
| MT (H)            | -47(5) | -74(10) | 4(12)   | 5    | 0                | 42(12) | -68(14) | 10(12) | 8    | 0    |
| LPS (O)           | -26(4) | -74(6)  | 42(12)  | 0    | 3                | 33(13) | -78(6)  | 41(9)  | 0    | 5    |

Tableau 2 : Les VOIs reconstruites à partir des résultats individuels avec leurs coordonnées moyennes, l'écart type, et le nombre de sujets (N) présentant une activation significative ( $p < 0.001$  ;  $k = 30$ ) à l'intérieur du volume.

Avec cette méthode, toutes les ROIs sont localisées à l'intérieur des VOIs. Ainsi, l'étendue des clusters des VOIs est à chaque fois plus large que celle des ROIs. L'hémisphère droit présente une plus grande sélectivité que l'hémisphère gauche.

En dernier lieu, afin de tester l'hypothèse d'une activation cérébrale différenciée en fonction des sous catégories, un nouveau modèle statistique, qui permet de différencier les postures du corps humain en mouvement implicite, les postures du corps humain au repos, les objets d'action implicite et les objets statiques, a été créé. La même procédure décrite ci-dessus utilisant les mêmes seuils ( $p < 0.001$ , non corrigé et  $k = 30$  voxels) a été utilisée pour une analyse de groupe ( $N=10$ ) et une analyse sujet par sujet. Les différentes cartes statistiques ont été obtenues en réalisant les contrastes par masquage exclusif ( $p < 0.05$ ) pour corps humain en mouvement implicite > corps humain en repos, et pour les objets d'action implicite > objets statiques aux seuils ( $p < 0.001$ , non corrigé et  $k = 30$  voxels).

Aucune activation significative n'a été observée ni pour l'analyse de groupe, ni pour les analyses individuelles. Afin de comprendre l'absence d'activation différenciée suivant le critère de mouvement – action implicite par rapport au repos – statique, l'évolution BOLD des différentes catégories et sous catégories a été analysée. La comparaison de l'évolution des réponses BOLD est présentée dans la figure 1.8 et des fonctions sinusoïdales ont été ajustées aux résultats.

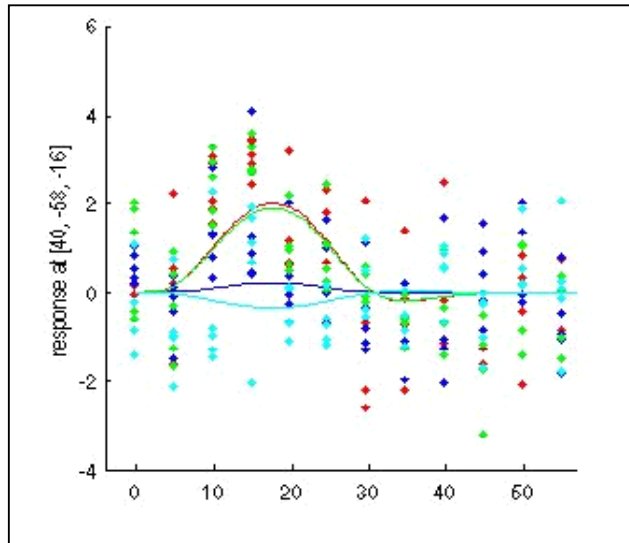


Figure 1.8 : Evolution des réponses BOLD dans le gyrus fusiforme latéral (GFL) (40, -58, -16) pour les sous catégories corps humain en mouvement implicite (rouge), corps humain au repos (vert), objets d'actions implicite (bleu) et objets statiques (bleu clair). Exemple de réponse obtenue pour 1 sujet.

La comparaison inverse est présentée sur la figure 1.9 pour les réponses hémodynamiques des différentes sous catégories ajustées à la réponse BOLD. L'activation maximale localisée dans le gyrus fusiforme médian (34, -50, -14) obtenue chez un sujet.

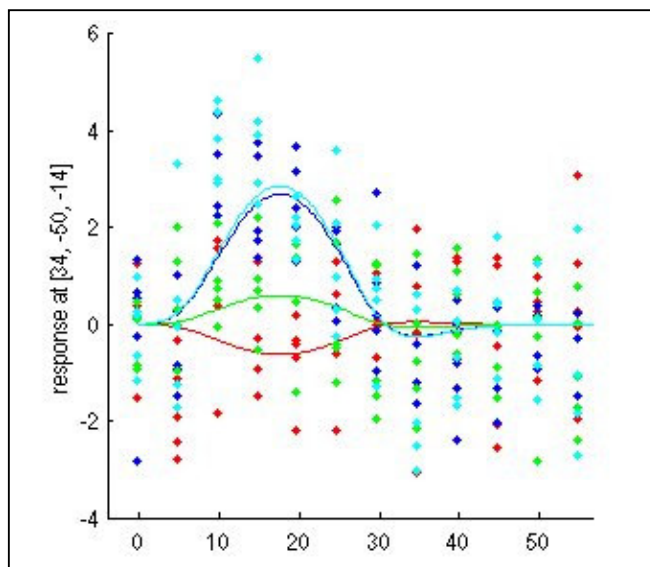


Figure 1.9 : Evolution des réponses BOLD dans le gyrus fusiforme médian (GFM) pour le contraste objets (bleu et bleu clair) < corps humain (rouge et vert) en fonction du décours temporel.

Les résultats présentés sur les figures 1.8 et 1.9 montrent une évolution identique de la réponse BOLD pour les sous catégories de corps humain, de même que pour les sous catégories d'objets manufacturés dans les deux sites d'activation, respectivement les gyri fusiformes latéral et médian. Cette évolution identique explique l'absence d'activation significative pour les contrastes à l'intérieur de chaque catégorie sémantique : corps humain en mouvement implicite < corps humain au repos et objets d'actions implicite < objets statiques. Des résultats très similaires ont été obtenus pour l'ensemble de la population sur les différentes régions d'intérêt.

## Discussion

Les gyri fusiforme latéral et médian (figure 1.8) présentent une activation cérébrale bilatérale sélective aux deux catégories sémantiques présentées : respectivement le corps humain (en rouge) et les objets manufacturés (en vert). Ces deux régions situées au niveau médio temporal inférieur sont classiquement attribuées à la voie ventrale (i.e. participant à la reconnaissance visuelle) et sont en accord avec un grand nombre de travaux expérimentaux et cliniques différenciant la reconnaissance visuelle de deux classes de stimuli : animés et inanimés (e. g. Beauchamps et al. (2003) ; Haxby et al. (2001) ; Ishai et al. (1999)). Ce premier résultat suggère que les stimuli ont bien été identifiés et différenciés par les sujets. Néanmoins, l'organisation cérébrale précise de ces différentes catégories sémantiques entretient toujours actuellement un débat très animé (e.g. Gauthier et al.; 1997 ; Kanwisher et al. 1997 ; Haxby et al. 2001).

Il est intéressant de noter que les coordonnées de l'activation bilatérale dans le gyrus fusiforme latéral (GFL) correspondent à une aire classiquement attribuée exclusivement aux processus participant à la reconnaissance visuelle des visages humains (i. e. fusiform face area : FFA ; Kanwisher et al. (1997)). Cette conception strictement modulaire est soutenue par des études de cas de patients présentant des lésions fonctionnelles sélectives de la reconnaissance des visages familiers (i.e. prosopagnosie ; e.g. Uthner, 2003). D'autres auteurs nuancent cette organisation cérébrale spécifique d'une catégorie sémantique en attribuant une fonction d'expertise visuelle à cette même aire (Tarr & Gauthier, 1997 ; Gauthier et al., 2000). En effet, ils démontrent, à partir de stimuli artificiels (i.e. des greebles) et en fonction du degré d'expertise, que le gyrus fusiforme latéral (GFL) ou FFA est impliqué dans les traitements visuels. Récemment, Gauthier et al. (2000), en présentant des photographies de voitures et d'oiseaux respectivement à des experts en voitures et ornithologues, retrouvent les

mêmes activations dans le gyrus fusiforme latéral (FFA ou « flexible fusiform area » ; cf. Gauthier & Tarr, 2001). Ces derniers résultats suggèrent l'implication du gyrus fusiforme latéral dans une multitude de différents stimuli visuels (animé et inanimé) en fonction du degré d'expertise. Les résultats obtenus dans cette expérience sont en accord avec cette approche en considérant que les dessins au trait représentant un corps humain en absence de visage (cf. figure 1.2) sont des stimuli particuliers faisant appel à des représentations expertes du corps humain nécessaires à toute interaction avec l'environnement physique et social (e.g. Grèze & Decety, 2001 ; Decety & Sommerville, 2003 ; Jeannerod, 2004). Jusqu'à présent, peu de travaux en neuroimagerie se sont intéressés à cet aspect à partir d'images du corps humain entier. Une expérience unique du groupe de recherche de Kanwisher propose une aire strictement dédiée au traitement du corps humain et la localise bilatéralement au niveau médio temporal du cortex extrastrié : « the extrastriate body area » (EBA ; Downing et al., 2001). Dans cette étude, réalisée par IRMf, les auteurs présentent des stimuli de partie du corps humain et du corps humain entier dans différentes postures et à partir de photographies, dessins au trait, fil de fer, silhouette et les comparent à des d'objets manufacturés présentés également en entier ou en partie. A partir de cette comparaison (corps humain < objet manufacturé), l'EBA a été identifiée et une série de différentes expériences ont comparé l'activation cérébrale uniquement dans cette région en utilisant un grand nombre de différents stimuli (i.e. corps humain, visages, animaux, objets manufacturés et pour dissocier l'aire MT / V5 un réseau de fréquences spatiales oscillatoires en mouvement et stationnaire). Dans les résultats présentés par Downing et al. (2001) il n'apparaît pas clairement de distinction entre l'activation spécifique du corps humain dans une aire de type EBA et une activation à l'intérieur de l'aire du mouvement MT / V5, résultant par exemple de représentations de mouvement – action évoquées par la présentation du corps humain entier ou d'une de ses partie (i.e. un bras levé, un pied).

Dans l'expérience présente, une activation bilatérale spécifique du corps humain a également été trouvée dans les régions médio temporales du cortex postérieur. Les coordonnées de cette activation coïncident bien avec les coordonnées reportées dans la littérature pour l'aire du mouvement MT / V5 et celles présentées par Downing et al. (2001) concernant l'hypothèse de l'EBA. Néanmoins, à partir de l'activation spécifique du corps humain observée dans le GFL qui joue un rôle crucial dans les traitements participant à la reconnaissance visuelle du stimulus (voie ventrale), nos résultats suggèrent que l'activation observée dans l'aire MT / V5 reflète plutôt l'évocation de représentation de mouvements liés au corps humain, mettant en

évidence la caractéristique de mouvement implicite (Kourtzi & Kanwisher (2000) et Senior et al. (2000) ; Paulos et al. ; 2004).

Ainsi, cette expérience suggère l'existence d'au moins deux régions bilatérales participant aux processus visuels du corps humain : le gyrus fusiforme latéral (GFL) qui participe aux traitements visuelles à l'intérieur de la voie ventrale, et l'aire MT / V5 qui participe aux traitements visuelles de la voie dorsale. Plus particulièrement, en l'absence de stimulation d'un mouvement visuel rétinien (i.e. dessins au trait), la présentation de nos stimuli du corps humain a engendré l'activation d'une partie d'une aire qui n'est classiquement pas activée par la nature physique de tel stimuli, et cela indifféremment de leur caractère sémantique implicite de postures de mouvement ou de repos. Dans l'expérience IRMf de Kourtzi & Kanwisher (2000), des photographies d'athlètes réalisant un exercice physique par rapport aux mêmes athlètes au repos induisent également des activations dans les région MT / MST. Ces auteurs attribuent cette activation au caractère de mouvement implicite contenu dans les photographies montrant par exemple un athlète en train de réaliser un exercice physique. Dans la présente expérience les stimuli sont des dessins au trait du corps humain qui possèdent des caractéristiques physiques beaucoup plus simples que des photographies et probablement plus difficiles à mettre dans un contexte. En effet, l'utilisation de la méthode de vision passive n'a pas permis de mettre en évidence, dans ces conditions expérimentales, une différenciation du caractère des postures du corps humain évoquant un mouvement implicite ou le repos. Nos résultats suggèrent que toutes les postures du corps humain à partir de dessins aux traits ont indifféremment évoqué des représentations de mouvement (cf. figure 1.10).

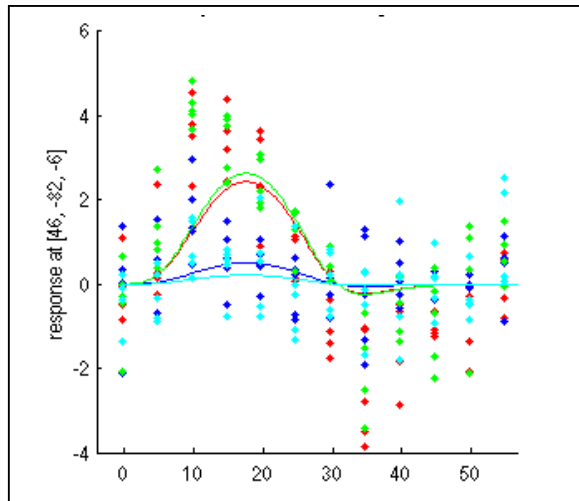


Figure 1.10 : Activation dans l'aire MT / V5 pour les conditions corps humain en mouvement implicite (rouge), corps humain au repos (vert), objets d'action implicite (bleu) et objets statiques (bleu clair).

La question de la tâche expérimentale sera de nouveau évoquée lors d'une deuxième expérience qui met en œuvre la technique d'imagerie par résonance fonctionnelle événementielle (inter stimulus) et qui inclue une tâche de catégorisation (Paulos, 2001) des mêmes stimuli : différentes postures du corps humain et des objets manufacturés.

## **II - De la perception du mouvement aux représentations visuelles liées au mouvement et à l'action : activation de l'aire MT / V5 ?**

Dans la première expérience, une activation bilatérale a été observée dans les régions de l'aire du mouvement MT / V5. Lors de cette première étude, des dessins au trait de différentes postures du corps humain et d'objets manufacturés étaient présentés à des sujets préalablement informés de la nature des stimuli et qui recevaient pour seule consigne de bien maintenir leur regard au milieu du champ visuel.

Une première différence d'activation apparaît lorsque l'on compare les stimuli des deux catégories sémantiques animées – inanimées, plus particulièrement des postures du corps humain et des objets manufacturés. Néanmoins, ces deux catégories sémantiques présentées comportaient chacune deux conditions qui avaient été définies a priori lors d'une pré étude (Paulos, 2001) : des postures du corps humain suggérant un mouvement versus des postures du corps humain au repos et des objets d'action versus des objets statiques. Cette différenciation n'a été retrouvée pour aucune des comparaisons de l'activité cérébrale des deux conditions respectives.

A partir de ces premiers résultats, une deuxième expérience a été menée en imagerie par résonance magnétique fonctionnelle événementielle (inter stimulus). Cette méthode permettra de préciser le rôle de la tâche expérimentale en introduisant une tâche de catégorisation, et ainsi de tester une différenciation d'activation cérébrale suivant la catégorisation des sujets d'images du corps humain et celles des objets manufacturés a posteriori.

Dans la première expérience en IRMf présentée ci-dessus, l'activation cérébrale (cf. aire MT / V5) engendrée par les différentes postures du corps humain présentées en bloc, en absence de tâche cognitive explicite (cf. vision passive), n'était pas significativement différenciable a priori pour les sous catégories mouvement implicite – au repos. De plus, pour les objets manufacturés, aucune des sous catégories (mouvement – statique) n'a montré d'activation significative dans l'aire MT / V5. Par contre, Kourtzi et Kanwisher (2000) montrent une activation de la région MT / MST lorsque le sujet visualise l'image d'un athlète le représentant en train de réaliser un exercice physique, alors que cette région reste silencieuse lorsque l'image le représente en posture de repos. Une différence majeure par rapport à nos expériences est que ces auteurs utilisent une tâche de détection de répétition d'image afin de maintenir l'attention des sujets constante. Dans l'étude de Senior et ses collaborateurs (2000),

une activation dans l'aire MT / V5 est induite lors de la présentation de photographies contenant un mouvement implicite représentant des objets manufacturés. Néanmoins, les images utilisées dans cette étude sont des photographies standardisées représentant, pour la condition de mouvement implicite, une tasse en train de tomber d'une table et une main essayant de la rattraper ; et pour la condition statique la même tasse sur la table avec la main reposant à côté. Une question suscitée par ce type de matériel (figure 2.1) est l'origine du caractère de mouvement implicite évoquée par la photographie : par exemple est-ce bien la tasse (objet inanimé) en train de tomber de la table ? ou bien l'action de la main essayant de rattraper la tasse ? ou encore l'ensemble de la scène visuelle décrivant le scénario qu'une tasse est en train de tomber d'une table et qu'une main est en train de la rattraper ? qui évoque des représentations de mouvement. Dans cette étude les sujets avaient pour consigne de bien regarder les stimuli (ou croix de fixation) dans le but d'être interrogé après la session d'acquisition (cf. maintien de l'attention constante).

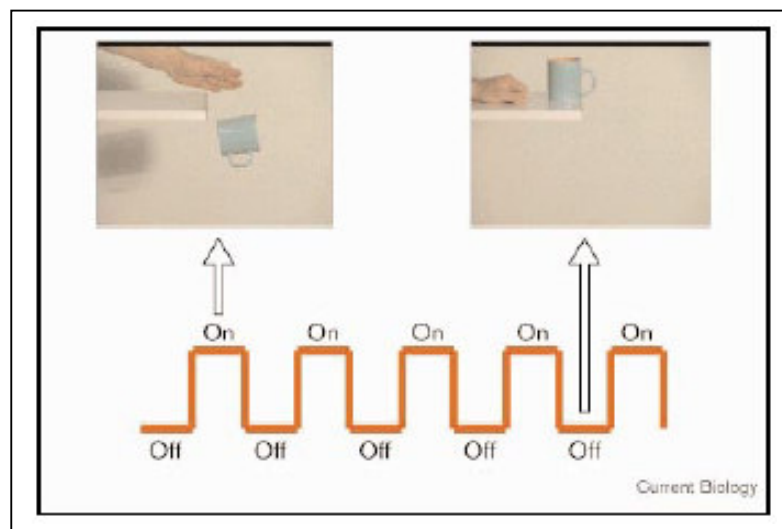


Figure 2.1 : Présentation de la condition ON : photographie de gauche contenant du mouvement implicite et de la condition OFF : photographie de droite. Issu de Senior et al. (2000).

Dans cette expérience, les mêmes stimuli sont présentés, c'est-à-dire des dessins au trait de différentes postures du corps humain et d'objets manufacturés. L'utilisation d'un paradigme événementiel permet d'introduire une tâche de catégorisation différenciant les stimuli en mouvement implicite versus statique. Dans cette étude, nous chercherons les régions corticales impliquées dans une tâche de catégorisation de dessins aux traits de différentes postures du corps humain et d'objets manufacturés.



Ainsi, selon notre hypothèse, les postures du corps humain ainsi que les objets catégorisés en tant que en mouvement implicite, évoqueraient des représentations de mouvement spécifiques de celles-ci et celles-ci seraient traitées, au moins partiellement, dans les mêmes régions cérébrales que celles impliquées dans la perception visuelle du mouvement, et plus particulièrement, l'aire du mouvement MT / V5.

### Matériel et Méthode

L'expérience a été réalisée sur une population de 10 sujets (5 femmes et 5 hommes ; âge moyen 28 ans), n'ayant pas participé à l'étude précédente, donnant leur consentement écrit et sans antécédent neurologique ou psychologique (examen médical d'inclusion au protocole). Tous les sujets étaient droitiers et avaient une vue normale ou corrigée à la normale (lentilles de contact). Les sujets étaient allongés dans un IRM corps entier S200 à 2 Tesla (Bruker, Karlsruhe, Allemagne) et un système de rétro projection sur un écran, placé derrière l'antenne et relié à un ordinateur, présentait les stimuli. La présentation de la première stimulation était synchronisée par le scanner au début de la session. La variation locale du taux d'oxygénation dépendante du flux sanguin, réponse BOLD, a été mesurée en utilisant une séquence d'imagerie ultra rapide (EPI) balayant le cerveau entier. Trente-deux coupes de 4 mm ayant une résolution de 64x64 (taille d'un voxel 4x4x4 mm<sup>3</sup>), étaient obtenues avec un temps de répétition (TR) de 2500 ms (TE = 43 ms, FA = 90°). Chaque session expérimentale était suivie d'une acquisition morphologique utilisant une séquence turbo-RARE pondérée en T2 et une résolution de 128x128x64 (taille des voxels 4x4x4 mm<sup>3</sup> ; TR = 15000 ms ; TE = 73.8 ms ; Turbofactor = 10). Une séance (une session de localisation, deux sessions expérimentales et une session morphologique) durait approximativement une quarantaine de minutes.

Les dessins sont présentés en traits blancs sur fond noir ayant une luminance d'environ 10 cd/m<sup>2</sup> et une taille angulaire approximative de 10° x 6° d'arc. La durée de présentation d'un stimulus est de 166 ms suivies d'une fenêtre de réponse variable de 500 ms (minimum) à 3000 ms (maximum). Après la réponse, donnée par le sujet, suit un intervalle inter stimulus (ISI) de 1 seconde avant la stimulation suivante. L'ensemble des stimuli (8 exemplaires par

catégorie ; c.f. figure 2.2) est présenté suivant un tirage au hasard. Celui-ci est répété 5 fois au cours d'une session expérimentale.

Toutes les séances débutaient avec la session de localisation et se terminaient par l'acquisition morphologique du sujet. L'ordre de présentation des deux sessions expérimentales est inversé pour la moitié du groupe des sujets. Lors de la session présentant le corps humain, les sujets devaient catégoriser en temps de réaction si la posture du corps humain leur évoquait soit une action – mouvement (i.e. un mouvement implicite) soit une posture de repos. Les objets manufacturés étaient catégorisés en objets d'action (i.e. mouvement implicite : un ballon, un rouleau à pâtisserie) versus objets statiques (i.e. un ancre, un lit).

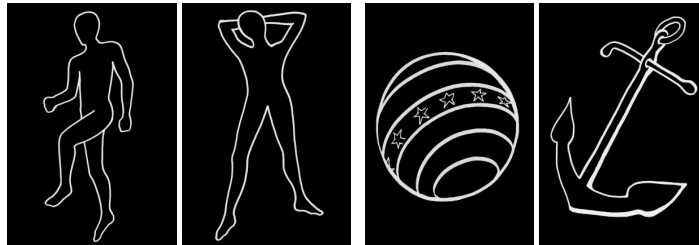


Figure 2.2 : Exemplaies des dessins aux traits catégorisés par les sujets, de gauche à droite, en posture du corps humain en mouvement, en repos, en objets d'action et en objets statiques.

Les stimuli présentés dans cette expérience sont les mêmes que ceux de l'expérience précédente et ceux qui ont obtenu préalablement une probabilité de catégorisation supérieure à  $p=0.89$  correspondant au critère d'inclusion aux différentes sous catégories respectives (Paulos, 2001). Ce choix a été fait afin de maximiser la balance entre les deux événements possibles pour chaque sujet. Le déroulement de l'expérience s'est fait en deux étapes consécutives d'acquisition. Dans un premier temps, la région MT / V5 a été localisée pour chaque sujet en présentant dans un paradigme en bloc alternativement un réseau de fréquences spatiales d'un mouvement unidirectionnel simple et un réseau de fréquences spatiales composé : sommation d'un mouvement vertical et horizontal (e.g. Huk et al. ; 2001, 2002). La combinaison de ces deux mouvements résulte en un mouvement unidirectionnel oblique (cf. figure 2.3). La comparaison de ces deux mouvements permet de localiser l'aire MT / V5, la région participant à l'intégration de mouvements visuels complexes (i.e. réseaux de

fréquences spatiales multidirectionnelles), ce que ne font pas les étapes antérieures du traitement (pour une revue Derrington, 2005).

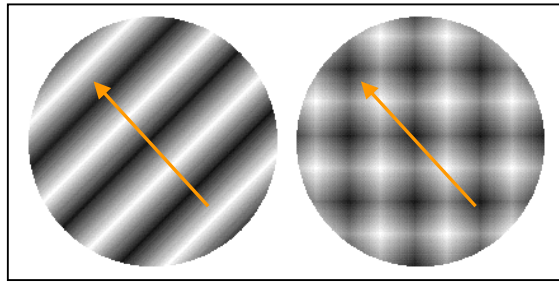


Figure 2.3 : A gauche un réseau de fréquences spatiales unidirectionnelles (cf. mouvement oblique). A droite un réseau de fréquences spatiales composées (cf. deux réseaux de fréquences spatiales superposés). La flèche indique à chaque fois la direction du mouvement perçu.

Après l'étape de localisation, les sujets passaient les deux sessions expérimentales successivement, l'une présentant les différentes postures du corps humain, l'autre des objets manufacturés. L'ordre des sessions de l'une ou de l'autre catégorie sémantique a été inversé entre les sujets.

Lors de la session corps humain, les sujets avaient pour consigne de catégoriser à chaque essai le corps humain en posture suggérant, soit une posture d'action – mouvement, soit une posture de repos. De même, pour la session d'objets manufacturés, pour chaque essai, la tâche des sujets consistait à catégoriser, en temps de réaction, les stimuli, soit en objets d'action (cf. mouvement implicite), soit en objet *statique*.

A partir de l'analyse de groupe de la session de localisation, l'aire du mouvement MT / V5 a pu être localisée expérimentalement pour le groupe de sujets étudiés. Pour les deux sessions expérimentales, l'activation cérébrale a été comparée en fonction des événements catégorisés en mouvement ou action par rapport à ceux catégorisés en repos ou statique, respectivement.

## Résultats

Les résultats de cette expérience ont été analysés suivant la même procédure (e.g. Friston et al., 1995 ; 1999) que celle utilisée lors de l'expérience en IRMf et décrite précédemment avec SPM99b (Wellcome Department, London). L'étape de prétraitement des images comportait un réalignement des coupes, une normalisation des données dans l'espace de Talairach ainsi qu'un lissage spatial (filtre Gaussien FWHM de 4 mm).

L'aire MT / V5 a été localisée, grâce à des comparaisons multiples sur le groupe expérimental ( $n=10$ ), en utilisant les contrastes mouvement complexe < mouvement simple ( $p < 0.001$ ,  $k = 16$  voxels). Plus précisément, une activation bilatérale (figure 1) dans les régions médio temporales a été identifiée avec les coordonnées dans l'hémisphère gauche (HG) :  $x = -48$  ;  $y = -76$  ;  $z = -4$  et dans l'hémisphère droit (HD) :  $x = 50$  ;  $y = -76$  ;  $z = -4$ . Ces coordonnées sont en accord avec celles rapportées dans un grand nombre d'études portant sur la perception du mouvement (e.g. Sinaert et al., 1999 ; Huk et al., 2001, 2002, Zeki et al., 2004).

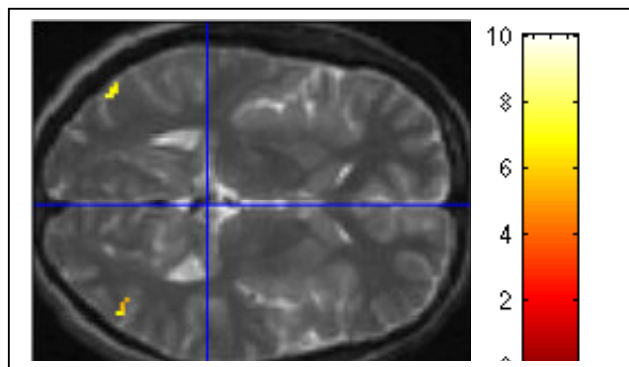


Figure 2.4 : Localisation de l'aire du mouvement MT / V5 du groupe ( $n=10$ ) sur une coupe horizontale. Les coordonnées de l'aire du mouvement MT / V5 dans l'HG se situe à  $x = -48$  ;  $y = -76$  ;  $z = -4$  ( $k = 43$  ;  $t = 5.31$  ;  $p < 0.001$ , non corrigé) et dans l'HD à  $x = 50$  ;  $y = -76$  ;  $z = -4$  ( $k = 16$  ;  $t=5.28$  ;  $p < 0.001$ , non corrigé).

Les deux sessions expérimentales ont été analysées en utilisant un paradigme événementiel. Afin de vérifier la cohérence d'appartenance des stimuli aux différentes catégories, définies a priori dans l'expérience précédente, les fréquences de catégorisation des différentes images présentées ont été calculées. La figure 2.5 présente les fréquences des stimuli catégorisés dans les deux sessions en posture du corps humain de mouvement – action versus au repos et les

objets d'action versus statiques. La fréquence moyenne obtenue pour la catégorisation du corps humain en mouvement implicite est de 0.92 (0.01) et au repos de 0.91 (0.02). Les résultats sont comparables pour la session objets manipulables : 0.96 (0.01) et 0.91 (0.03). Ce résultat suggère que tous les sujets catégorisent, avec peu de différence, tous les stimuli présentés à l'intérieur des sous catégories ; ces sous catégories ayant été définies a priori lors de l'étude préliminaire (Paulos, 2001).

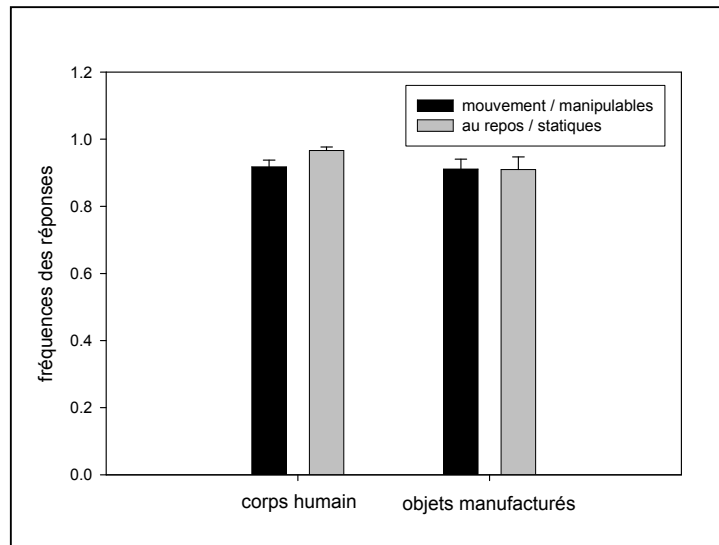


Figure 2.5 : Fréquences moyennes des réponses de catégorisation des sujets obtenues lors des deux sessions expérimentales : corps humain et objets manufacturés.

Un modèle statistique événementiel a été construit, pour chaque sujet, à partir des réponses de catégorisation qu'ils ont données. Ainsi, pour chaque sujet et chaque session, un modèle à deux facteurs a été obtenu : action – mouvement implicite < au repos et objets d'action < statiques. Les résultats étaient comparables pour les différents sujets.

Après cette analyse individuelle, une analyse de groupe a été réalisée pour chaque session. Pour ce faire, des contrastes ont été réalisés par masquage exclusif ( $p < 0.05$ , non corrigé) pour le corps humain : en mouvement < en repos ( $p < 0.001$ , non corrigé ;  $k = 30$ ) ainsi que pour les objets manufacturés : manipulables < statiques ( $p < 0.001$ , non corrigé ;  $k = 30$ ).

La figure 2.6 présente l'activation cérébrale du groupe obtenue lors de la session corps humain pour le contraste mouvement implicite < au repos ( $p < 0.001$ , non corrigé ;  $k = 30$ ).

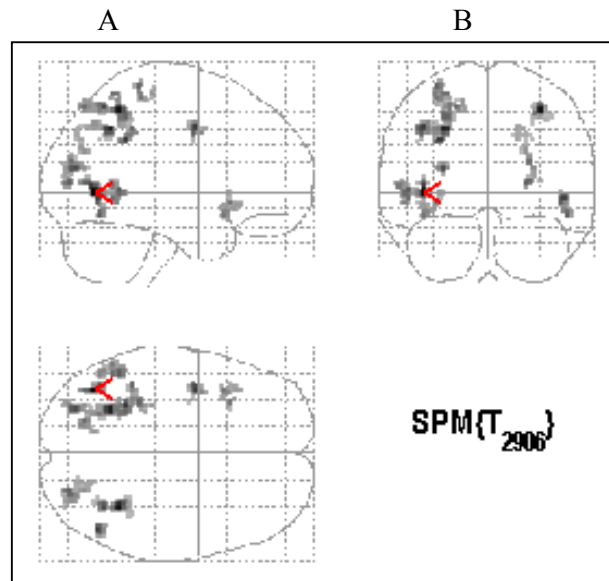


Figure 2.6 : Cartes statistiques du groupe (n=10) en présentation *glass brain* : (A) vue latérale, (B) coronale et (C) horizontale du contraste corps humain en mouvement < au repos ( $p < 0.001$ , non corrigé ;  $k = 30$ ). Le global maximum est indiqué par la flèche dans l'aire MT / V5 (-40, -66, 0).

Dans le tableau 2.1 figurent les différentes coordonnées de l'analyse de groupe (n=10) ainsi que leurs régions correspondantes identifiées avec le logiciel Talairach Daemon.

|         | Hémisphère Gauche |     |     |       |      | Hémisphère Droit |     |    |       |      |
|---------|-------------------|-----|-----|-------|------|------------------|-----|----|-------|------|
| Régions | X                 | Y   | Z   | $K_E$ | $t$  | X                | Y   | Z  | $K_E$ | $t$  |
| post    |                   |     |     |       |      |                  |     |    |       |      |
| MT / V5 | -40               | -66 | 0   | 43    | 6.84 | 48               | -64 | -4 | 31    | 5.48 |
| SPL /   | -28               | -56 | -40 | 43    | 6.12 |                  |     |    |       |      |
| BA7     |                   |     |     |       |      |                  |     |    |       |      |
| IPL /   | -30               | -46 | 34  | 35    | 5.62 | 34               | -50 | 52 | 59    | 6.45 |
| BA40    |                   |     |     |       |      |                  |     |    |       |      |
| ant     |                   |     |     |       |      |                  |     |    |       |      |
| BA 47   | -32               | 20  | -6  | 43    | 4.49 |                  |     |    |       |      |
| BA 6    | -40               | -2  | 42  | 33    | 5.57 |                  |     |    |       |      |

Tableau 2.1 : Coordonnées de Talairach obtenues lors de la session corps humain pour le contraste postures du corps humain en mouvement moins postures du corps humain au repos ( $p < 0.001$ , non corrigé) avec  $K_E$ , l'étendue du cluster et  $t$ , sa valeur associée.

La figure 2.7 représente les cartes statistiques des activations obtenues pour le contraste objets d'action moins objets statiques ( $p < 0.001$ , non corrigé ;  $k = 30$ ). En raison d'un dysfonctionnement technique du bouton réponse, survenu pendant l'acquisition, les données de deux sujets ont dû être écartées des analyses statistiques suivantes ( $n=8$ ).

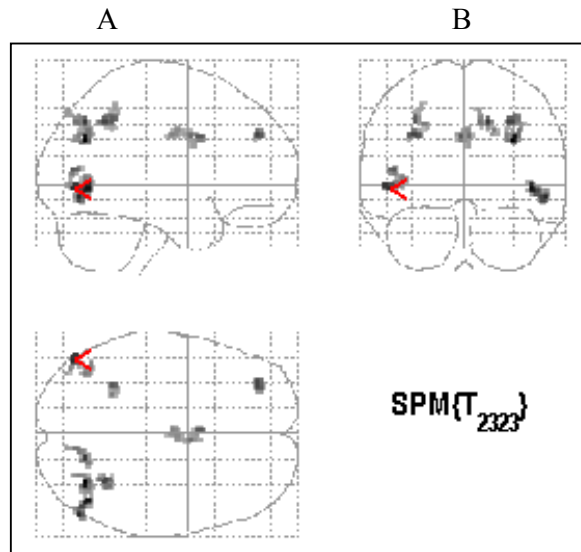


Figure 2.7 : Cartes statistiques du groupe ( $n=8$ ) en présentation *glass brain* : (A) vue latérale, (B) coronale et (C) horizontale du contraste objets d'action moins objets statiques ( $p < 0.001$ , non corrigé ;  $k = 30$ ). Le global maximum est indiqué par la flèche dans l'aire MT / V5 (-48, -76, -2).

Le tableau 2.2 résume les différentes coordonnées des activations pour le contraste objets d'action moins objets statiques et l'identification des régions avec le logiciel Talairach Daemon.

|         | Hémisphère Gauche |     |    |       |      | Hémisphère Droit |     |    |       |      |
|---------|-------------------|-----|----|-------|------|------------------|-----|----|-------|------|
| Regions | X                 | Y   | Z  | $K_E$ | $t$  | X                | Y   | Z  | $K_E$ | $t$  |
| post    |                   |     |    |       |      |                  |     |    |       |      |
| MT / V5 | -48               | -76 | -2 | 35    | 5.82 | 44               | -66 | -2 | 59    | 6.30 |
| SPL /   |                   |     |    |       |      | 30               | -54 | 42 | 44    | 5.33 |
| BA7     |                   |     |    |       |      |                  |     |    |       |      |
| IPL /   | -32               | -50 | 42 | 31    | 4.97 |                  |     |    |       |      |
| BA40    |                   |     |    |       |      |                  |     |    |       |      |
| ant     |                   |     |    |       |      |                  |     |    |       |      |
| BA9     | -34               | 46  | 34 | 31    | 5.35 |                  |     |    |       |      |

Tableau 2.2 : Coordonnées de Talairach obtenues lors de la session objets pour le contraste objets d'action moins objets statiques ( $p < 0.001$ , non corrigé) avec  $K_E$ , l'étendue du cluster et  $t$ , sa valeur associée.

Afin de comparer les différentes activations obtenues dans l'aire MT / V5, les résultats de la session de localisation et des sessions expérimentales corps humain et objets manufacturés ont été superposées sur un cerveau modèle. La figure 2.8 présente les différentes activations dans l'aire MT / V5 spécifique de la perception d'un mouvement complexe (en bleu), au corps humain catégorisé en tant que en mouvement (en rouge) et aux objets catégorisés en tant qu'objets d'action (en vert).

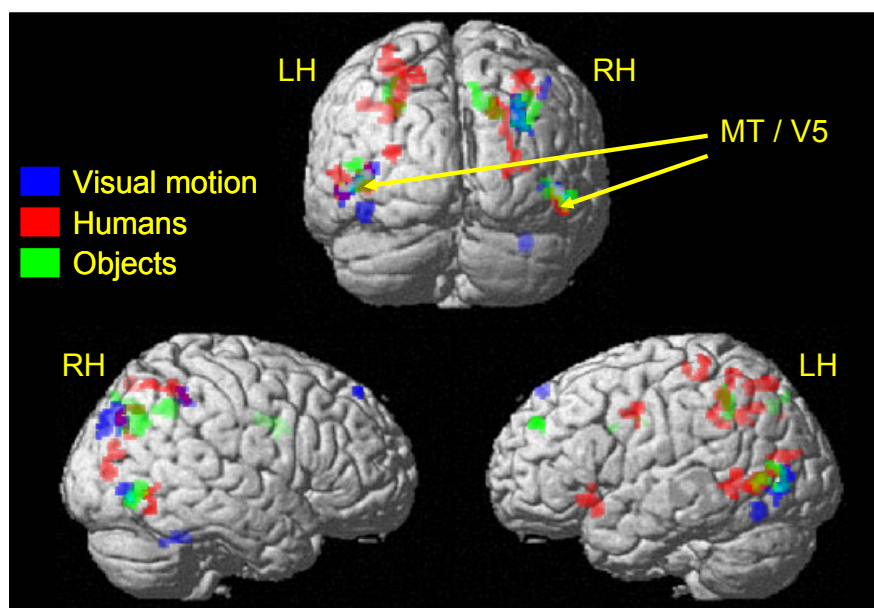


Figure 2.8 : Activation cérébrale dans l'aire MT / V5 spécifique de la perception du mouvement complexe (en bleue), à la catégorie corps humain en mouvement (en rouge) et à la catégorie objets manipulables (en vert). En haut, vue postérieure du cerveau modèle et deux vue latérales de chaque hémisphère en bas.

Le tableau 2.3 reprend les coordonnées moyennes de l'aire MT / V5 pour chaque session. La somme de l'étendue des clusters dans l'HG ( $k = 121$ ) est globalement plus large que celle dans l'hémisphère droit ( $k = 106$ ). Cependant, les clusters dans l'HG se recouvrent partiellement entre les trois sessions, alors qu'ils sont bien délimités dans l'HD (cf. figure 2.8).



|              | Hémisphère Gauche |     |    |           |             | Hémisphère Droit |     |    |           |             |
|--------------|-------------------|-----|----|-----------|-------------|------------------|-----|----|-----------|-------------|
| session      | X                 | Y   | Z  | $K_E$     | $t$         | X                | Y   | Z  | $K_E$     | $t$         |
| localisation | -48               | -76 | -4 | <b>43</b> | <b>5.82</b> | 50               | -76 | -4 | <b>16</b> | <b>5.28</b> |
| corps humain | -40               | -66 | 0  | <b>43</b> | <b>6.84</b> | 48               | -64 | -4 | <b>31</b> | <b>5.48</b> |
| objets       | -48               | -76 | -2 | <b>35</b> | <b>5.82</b> | 44               | -66 | -2 | <b>59</b> | <b>6.30</b> |

Tableau 2.3 : Coordonnées de Talaraich de l'aire du mouvement MT / V5 différenciée pour les sessions de localisation, les postures du corps humain en mouvement – action et les objets d'action.

La figure 2.9 représente l'effet de l'amplitude maximale dans l'aire MT / V5 pour la session de localisation et les deux sessions expérimentales : corps humain et objets manufacturés. Il n'y a pas de différence significative pour l'amplitude maximale en fonction de l'hémisphère (LH ou RH). Pour chaque session, l'activation dans les deux hémisphères est équivalente ( $F(2,1) < 1$ ). Une augmentation significative de l'amplitude maximale est observée pour la session de localisation comparée aux deux sessions expérimentales. (Pour la session de localisation, l'amplitude maximale est significativement augmentée comparativement à celle observée pour les deux autres sessions expérimentales). Par ailleurs, la taille de l'effet ne diffère pas entre les sessions corps humain et objets manufacturés.

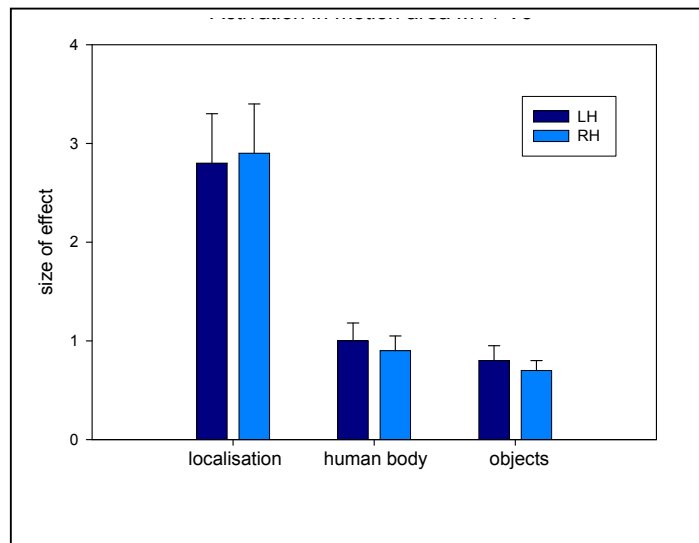


Figure 2.9 : Histogrammes représentant la taille de l'effet dans l'aire MT / V5 dans chaque hémisphère (LH, RH) pour la session de localisation, corps humain et objets manufacturés.

## Discussion

Les résultats de cette expérience ont permis de mettre en évidence une activation bilatérale de l'aire du mouvement MT / V5. Cette activation a été délimitée, pour le groupe expérimental, lors de la session de localisation, et évoquée par la catégorisation de postures du corps humain en mouvement et d'objets manufacturés en objets d'action. Par ailleurs, ces résultats montrent que l'introduction d'une tâche de catégorisation explicite sur le caractère de mouvement (implicite) a permis d'évoquer des représentations de mouvement lors de la présentation de postures du corps humain et d'objets manufacturés. Cette activation bilatérale de l'aire MT / V5 est partiellement superposée dans l'hémisphère gauche et apparaît discrète dans l'hémisphère droit (cf. figure 2.8). L'étendue des clusters et la taille de l'amplitude de l'activation est comparable pour les postures du corps humain et les objets, et cela dans les deux hémisphères.

Dans notre première expérience en IRMf, seules les différentes postures du corps humain ont évoqué des représentations de mouvement, et cela indépendamment de leurs postures : mouvement implicite et au repos. Une explication possible à ce résultat est que les sujets auraient fait comme seule distinction sémantique des stimuli le caractère animé et inanimé. Une des principales caractéristiques intrinsèques d'un objet animé est sa capacité à se déplacer dans l'espace, d'où l'évocation de représentations de mouvement spécifiques du corps humain. Dans l'expérience présente, les sujets avaient pour consigne de catégoriser les postures du corps humain et les objets manufacturés en mouvement implicite et statique. La prise en compte de ces deux résultats suggère que l'activation bilatérale de la région MT / MST observée par Kourtzi & Kanwisher (2000) est probablement due à l'interprétation de l'action perçue sur la photographie (i.e. un athlète en train de lancer un disque), plutôt qu'à l'évocation d'une représentation de mouvement. En revanche, la comparaison des résultats de nos deux expériences a d'abord montré une absence d'activation de l'aire MT / V5 lors de la première expérience, mais l'activation bilatérale de l'aire MT / V5 lorsque les sujets reçoivent pour consigne de catégoriser les objets manufacturés en objets d'action et en objets statiques. Une explication de ce résultat est que lors de la première expérience les objets manufacturés ont bien été distingués par rapport aux stimuli du corps humain, mais qu'il n'y a pas eu de différenciation à un niveau sémantique plus élevé, par exemple sur l'utilisation pratique des objets. Ainsi, nos résultats suggèrent que l'activation dans l'aire V5, induite par la présentation de photographies d'objets en mouvement implicite (cf. figure 2.1), et observée par Senior et al. (2000), est probablement due à l'interprétation de l'ensemble de la scène visuelle (c'est-à-dire une tasse en train de tomber d'une table et une main essayant de la rattraper), plutôt qu'au mouvement implicite de la tasse représentée dans l'air.

Par conséquent, ces données suggèrent qu'il est important de distinguer, lors de la présentation d'images contenant du mouvement implicite, la description d'une scène visuelle complexe décrivant un mouvement et / ou une action, de la présentation d'un objet seul, animé et / ou inanimé mais aussi de distinguer des représentations de mouvements évoquées par des images contenant un mouvement implicite, de celles contenant des actions implicites.

## **Expériences comportementales sur la spécificité du corps humain:**

Pour tester des hypothèses concernant le contenu des représentations de mouvement spécifique du corps humain, parallèlement à des expériences de neuroimagerie, plusieurs expériences utilisant des mesures psychophysiques ont été menées.

Dans un premier temps, ces expériences, qui utilisent des mesures psychophysiques, se focalisent plus particulièrement sur la question des représentations de mouvements du corps humain à partir d'images statiques. Dans un deuxième temps, ces expériences portent sur des représentations de mouvements décrivant une action à partir d'une séquence d'images du corps humains utilisant le paradigme de mouvement apparent de Ternus (1926). La première série d'expériences, présentant des images statiques du corps humain, vise à évaluer dans quelle mesure les sujets possèdent des représentations implicites sur les propriétés biomécaniques du corps humain. Pour ce faire, les sujets devaient répondre si le corps humain présenté dans différentes postures était capable de se maintenir en équilibre ou non. La deuxième série d'expériences, présentant des images dotées d'un mouvement apparent, teste dans quelle mesure ce phénomène perceptif, qui résulte de mécanismes de bas niveau, peut être modulé par des mécanismes plus tardifs tels que l'activation de représentations sémantiques ; plus particulièrement lorsque les images présentées illustrent différentes postures du corps humain en train d'exécuter une action.

### **I - Représentations visuelles d'actions à partir d'images statiques**

La compréhension et l'interprétation de différentes postures du corps humain sont des activités importantes réalisées en permanence et sans aucune difficulté apparente chez l'homme. Cette expérience vise à tester la reconnaissance des conséquences d'une posture d'un corps humain, en l'occurrence une chute ou un maintien en équilibre, grâce à la présentation d'une image statique. Toute tâche utilisant des images est fortement marquée culturellement (e.g. Deregowski, 1999), et nécessite le recours à des représentations mentales. Afin de préciser la spécificité des représentations du corps humain, différentes images d'objets articulés (animés et inanimés) seront utilisées. En particulier, les expériences décrites

ci-dessous visent à vérifier dans quelle mesure des images statiques d'un objet articulé, i.e. le corps humain, suffisent à anticiper les conséquences d'une action à partir de différentes postures de ce dernier.

L'expérience comportait trois types d'objets : des images d'un corps humain, d'un mannequin de dessin et d'un squelette. Ces deux derniers objets ont été introduits afin de déterminer si les représentations biomécaniques spécifiques du corps humain sont restreintes aux images d'un corps humain « normal » ou bien si elles se généralisaient à des objets anthropomorphes.

Dans une première série d'expériences, trois objets en couleur : un corps humain, un mannequin en bois et un squelette, reposant sur un socle horizontal de couleur orange, étaient présentés sur un fond blanc (cf. Figure 1.1). Tous les objets présentaient différents degrés d'inclinaison vers l'avant ou vers l'arrière ( $0^\circ$ ,  $20^\circ$ ,  $40^\circ$ ,  $60^\circ$ ) de la partie supérieure du corps. L'inclinaison se faisait à partir de la hanche, une jambe étant droite et l'autre repliée d'un angle de  $90^\circ$  au niveau du genou pour faire contrepoids (cf. Figure 1.1). Ils étaient présentés latéralement de profil gauche ou droit avec une inclinaison vers l'avant ou l'arrière. Dans toutes les conditions, la tâche des sujets était de juger, en temps de réaction, si l'objet allait chuter ou non (Bonnet, Paulos & Nithart, 2005). Les temps de réaction étaient mesurés avec la consigne de répondre le plus rapidement possible sans faire d'erreur. Cette tâche ne présentait aucune difficulté apparente pour les sujets. Les temps de réaction ont été analysés en détail dans Bonnet & Paulos (2004a, 2004b).



Figure 1.1 : Deux exemplaires de chaque stimulus : le corps humain, le mannequin de dessin et le squelette, présentés de profil droit et selon une inclinaison avant (F) ou arrière (B) de 40°.

### Résultats 1

La figure 1.2 présente les fréquences de réponses « chute » en fonction des différents degrés d'inclinaison des trois objets présentés : le corps humain, le mannequin en bois et le squelette humain. Une analyse de variance (ANOVA) sur le facteur latéral (profil gauche, profil droit) ne montre pas d'effet significatif de ce facteur ( $F(1,9) = 2.95$  ;  $p < 0.12$ ) ni aucune interaction avec les autres facteurs. Par conséquent, pour les analyses suivantes, les résultats ont été moyennés pour le facteur latéral et une ANOVA à 3 facteurs (direction de la chute x degré d'inclinaison x type d'objets) a été réalisée.

La fréquence des réponses « chute » est plus élevée pour les chutes arrière comparativement aux chutes avant ( $F(1,9) = 12.81$  ;  $p < 0.006$ ). La fréquence des réponses « chute » augmente significativement avec le degré d'inclinaison ( $F(3,27) = 65.19$  ;  $p < 0.0001$ ). De même, une différence significative est observée en fonction des trois objets présentés ( $F(2,18) = 2.93$  ;  $p < 0.08$ ).

Il n'y a pas de différence significative en fonction du type d'objet présenté et la direction de la chute avant – arrière ( $F(2,18) = 1.29$  ; n.s.). Néanmoins, il y a une interaction significative entre le facteur objet et la direction avant – arrière de la chute ( $F(6.54) = 3.55$  ;  $p < 0.005$ ). De même, une interaction significative est observée entre le facteur direction de la chute et le degré d'inclinaison ( $F(3,27) = 7.42$  ;  $p < 0.0009$ ).

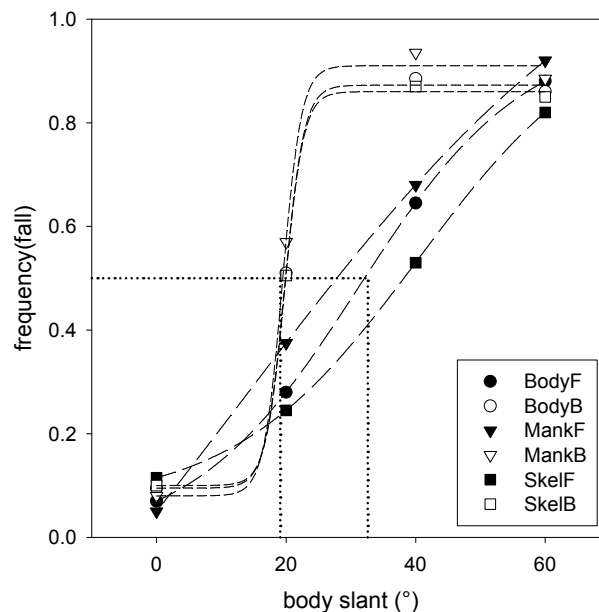


Figure 1.2 : Les fréquences de réponses « chute » avant (F) et arrière (B) sont présentées en fonction du degré d'inclinaison 0°, 20°, 40° et 60° pour le corps humain (Body), le mannequin en bois (Man) et le squelette (Skel).

Des fonctions psychométriques logistiques ont été ajustées aux résultats afin de déterminer les différents points d'équilibre subjectif (PES) qui correspondent au degré d'inclinaison donnant 50% de réponses « chute ». Les PES ne diffèrent pas significativement entre les objets pour les chutes en arrière (PES=19.1°) et donnent pour les chutes avant : 32.1° pour le corps humain, 27.9° pour le mannequin de dessin et 32.6° pour le squelette humain.

### Discussion 1

Les résultats montrent en premier lieu que les différences de réponses « chute » données pour les différents objets sont minimales (non significatives). L'effet le plus important est la présence d'une asymétrie des réponses chute avant – arrière, et cela pour les trois objets présentés. Le point de chute réel d'un corps humain a été difficile à mesurer empiriquement. Néanmoins, un contrôle a été réalisé en demandant à un sujet d'adopter les différents degrés d'inclinaison de postures de nos stimuli. Le sujet a été sécurisé pour les chutes avant – arrière et des photographies ont été prises afin d'estimer le point de chute du corps humain. Une asymétrie des chutes avant – arrière a été retrouvée lors de cette mesure empirique correspondant aux résultats expérimentaux observés. Cette comparaison suggère que, pour réaliser la tâche expérimentale, les sujets ont implicitement utilisé leurs connaissances spécifiques du corps humain en respectant ses contraintes biomécaniques. De plus, ces représentations spécifiques du corps humain sont transposées implicitement à des objets inanimés anthropomorphes, à savoir au mannequin de dessin et au squelette.

Néanmoins, ce premier résultat pourrait s'expliquer par un effet de contexte (e.g. Boucart & Bonnet, 1991) dû à la présence, dans l'expérience, d'images réalistes du corps humain parmi les images d'objets inanimés anthropomorphes, tels le mannequin de dessin et le squelette.

Afin de tester le caractère implicite d'utilisation de ces représentations dynamiques spécifiques du corps humain pour différents types d'objets inanimés (ne possédant pas les mêmes caractéristiques physiques et fonctionnelles du corps humain) un contexte, qui ne présente que des images d'un mannequin de dessin, a été créé.

Une deuxième expérience a été réalisée en présentant uniquement comme stimulus des photographies d'un mannequin de dessin en bois. La variabilité des différentes postures de ce mannequin a été augmentée (série a et b) dans le but de tester à la fois l'hypothèse de l'effet de contexte expérimental et l'hypothèse de la généralisation des représentations spécifiques du corps humain à des objets inanimés anthropomorphes. Comme dans l'expérience précédente le stimulus était présenté avec différents degrés d'inclinaison 0°, 20°, 40° (Figure 1.3) et 60°.



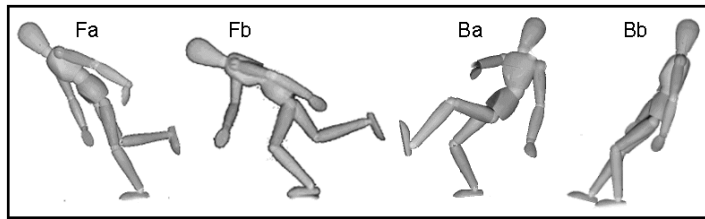


Figure 1.3 : Mannequin en bois présenté dans deux types d'inclinaisons avant (Fa et Fb) de 40° et deux types d'inclinaisons arrière (Ba et Bb) de 40°.

## Résultats 2

Sur la figure 1.4 les fréquences de réponses « chute » sont présentées en fonction du degré d'inclinaison pour les chutes avant (Fa et Fb) et arrière (Ba et Bb). Le facteur de variabilité de la chute (série a et b) n'a pas d'effet significatif ( $F(1,19) = 1.11$  ; n.s.). Aucun effet significatif n'a été obtenu pour le facteur latéral profil gauche – droite et les résultats ont été analysés comme précédemment en réalisant une ANOVA à 3 facteurs (direction de la chute x degré d'inclinaison x type d'objets).

La fréquence des réponses « chute » est plus grande pour les inclinaisons en direction arrière que les inclinaisons avant ( $F(1,19) = 55.16$  ;  $p < 0.0001$ ). Néanmoins, il y a une différence significative pour ce facteur avec la direction de la chute ( $F(1,19) = 10.27$  ;  $p < 0.005$ ). Cette différence est seulement visible pour les réponses chute avant. Une différence significative est également obtenue entre la direction de la chute et le degré d'inclinaison ( $F(3,57) = 8.76$  ;  $p < 0.0001$ ).

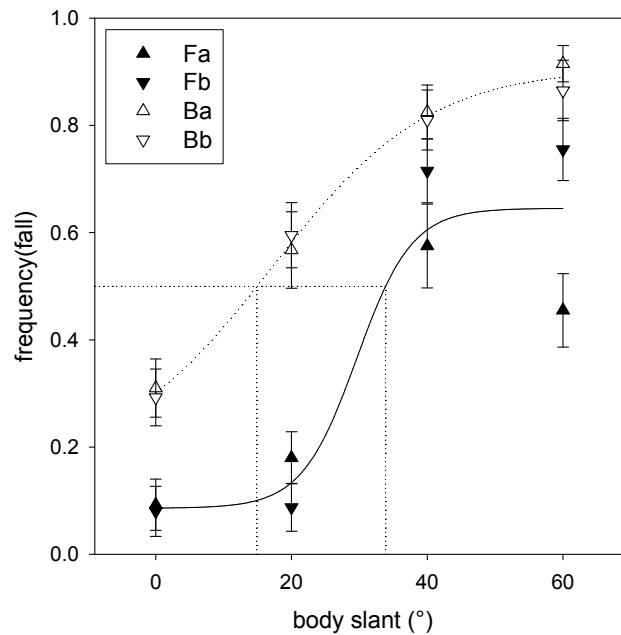


Figure 1.4 : Fréquences des réponses « chute » en fonction des différents degrés d'inclinaison pour les chutes avant et arrière et la variabilité de la chute (série a et b).

Des fonctions psychométriques ont été ajustées aux résultats afin d'estimer les différents points d'équilibre subjectif (PES). Le PES moyen arrière ( $14.8^\circ$ ) est moins élevé que celui vers l'avant ( $33.8^\circ$ ).

## Discussion 2

Les résultats de l'expérience 2 sont en accord avec ceux obtenus lors de la première expérience. Tout d'abord on peut observer que les fréquences des réponses « chute » augmentent avec le degré d'inclinaison. Nous retrouvons également une différence selon la direction de la chute avant – arrière. Le point de chute avant est plus tardif que celui de la chute arrière. Ce résultat, totalement en accord avec celui de l'expérience précédente, suggère que les sujets utilisent de nouveau des représentations spécifiques du corps humain en respectant ses contraintes biomécaniques et physiques, bien que l'objet présenté soit inanimé. Ces résultats suggèrent que l'apparence anthropomorphe du mannequin de dessin a suffi à activer les représentations spécifiques du corps humain.

En conclusion, cette expérience montre que les sujets ont recours spontanément à leurs représentations spécifiques du corps humain. Ils transposent leurs connaissances implicitement au mannequin de dessin sans y avoir été sollicité ni par la présence d'images du corps humain (cf. effet de contexte), ni par les consignes expérimentales.

Une troisième expérience a été menée afin de vérifier que dans le cas d'un objet inanimé non anthropomorphe, seules les contraintes physiques de la gravité rendraient compte des jugements de chute. Ainsi, selon notre hypothèse, avec un tel objet, aucune asymétrie entre les fréquences des réponses chute avant – arrière ne devrait être observée.

Dans cette expérience, des images d'un corps humain et d'un nouvel objet ont été présentées lors d'une même session afin de contrôler simultanément la transposition des représentations spécifiques du corps humain et l'effet de contexte. Le choix de l'objet inanimé est critique. Néanmoins, un objet de couleur, d'aspect structural et proportionnel similaire à celui du corps humain a été fabriqué (Figure 1.5). Les caractéristiques physiques (matière, poids approximatif, etc.) de l'objet ont été précisées lors des consignes.

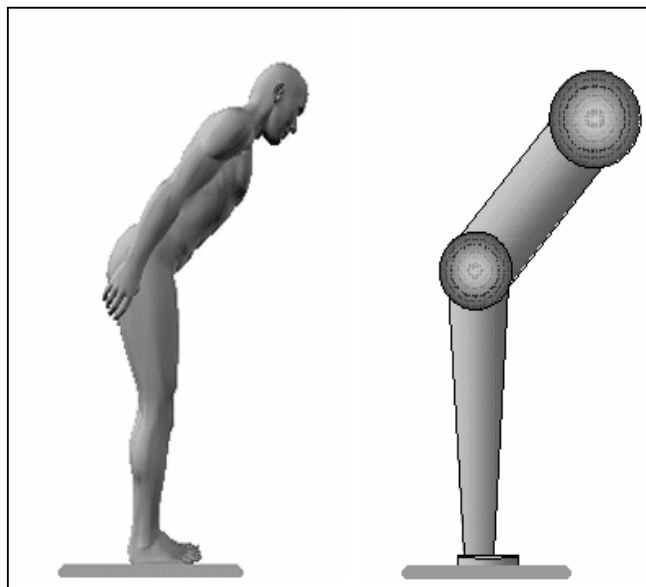


Figure 1.5 : Un exemplaire de profil droit du corps humain et de l'objet manufacturé avec la même inclinaison avant de 40°.

### Résultats 3

La figure 1.6 présente les fréquences des réponses « chute » du corps humain et de l'objet manufacturé en fonction de la direction de la chute et des différents degrés d'inclinaison. Les résultats ont été analysés de la même manière que précédemment.

Les fréquences de réponses « chute » ne diffèrent pas significativement entre les deux objets ( $F(1,9) = 0.25$  ; n.s.). La différence de direction de la chute est significative ( $F(1,9) = 13.48$  ;  $p < 0.005$ ) et la fréquence des réponses « chute » augmente avec le degré d'inclinaison ( $F(3,27) = 60.19$  ;  $p < 0.00001$ ). L'interaction entre les facteurs objet et la direction de la chute est significative ( $F(1,9) = 12.44$  ;  $p < 0.006$ ). De même, l'interaction entre la direction de la chute et le degré d'inclinaison est significative ( $F(3,27) = 7.81$  ;  $p < 0.0006$ ). Par contre, il n'y a pas d'interaction entre les facteurs objets et le degré d'inclinaison ( $F(3,27) = 1.32$  ; n.s.).

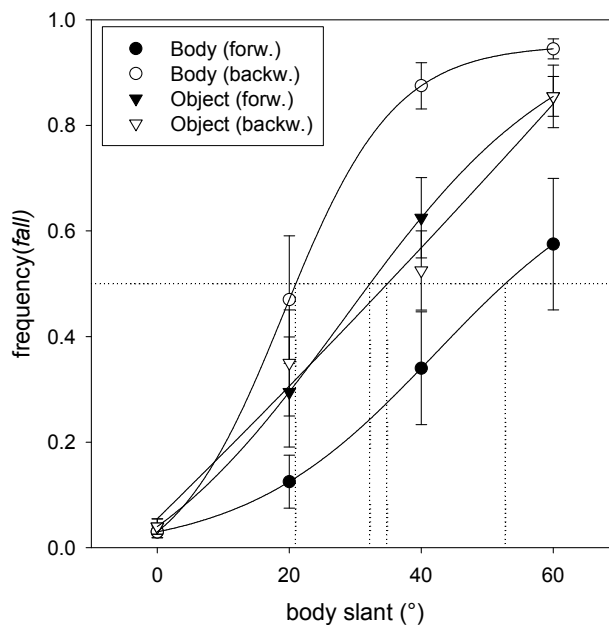


Figure 1.6 : Fréquences des réponses chute pour le corps humain (B) et l'objet manufacturé (O) en fonction du degré d'inclinaison et de la direction de la chute avant (forw.) et arrière (backw.).

A nouveau, les différents PES ont été estimés en ajustant des fonctions psychométriques aux résultats. En ce qui concerne l'objet, les PES (avant – arrière) sont quasiment identiques : avant (32.2°) et arrière (34.8°), la différence n'est pas significative. Par contre, comme dans la première expérience le PES avant pour le corps humain est plus élevé (52.7°) que le PES arrière (20.9°).

### Discussion 3

Dans cette troisième expérience, les réponses des sujets diffèrent en fonction de la nature sémantique des stimuli. Lorsque le stimulus représente un corps humain, les résultats suggèrent à nouveau que des représentations spécifiques du corps humain sont utilisées. Dans toutes les expériences, les fréquences des réponses « chute » augmentent avec le degré d'inclinaison, et cela indépendamment du type d'objet présenté. Particulièrement, l'asymétrie des réponses chute en fonction de la direction avant – arrière est retrouvée uniquement pour les images du corps humain.

Par contre lorsque le stimulus est un nouvel objet (inanimé) manufacturé, les sujets ne semblent pas spontanément transposer les représentations spécifiques du corps humain et cela en dépit du contexte de présentation mixte puisque d'un essai à l'autre, le sujet peut voir, au hasard, un corps humain ou un objet inanimé.

Cependant les résultats montrent que les réponses « chute » augmentent également en fonction du degré d'inclinaison de l'objet manufacturé. Néanmoins, aucun effet du facteur de direction de la chute n'est observé pour celui-ci. Leurs réponses reflètent plutôt l'utilisation de caractéristiques purement physiques d'un objet inanimé, ne lui attribuant pas de force intrinsèque. L'utilisation de ces caractéristiques purement physiques se traduit par des réponses « chute » dès qu'il y a inclinaison de l'objet. Ces résultats confirment à nouveau que lorsque les images représentent un corps humain, les sujets utilisent spontanément des représentations spécifiques de celui-ci.

### Discussion générale et conclusions

Les sujets ont jugé de manière consistante si les objets présentés étaient en train de chuter ou bien s'ils étaient dans une posture d'équilibre. Pour réaliser la tâche, lorsque l'objet représente un corps humain, les sujets utilisent implicitement leurs connaissances du corps humain telles que les contraintes biomécaniques. Cette interprétation est soutenue à la fois par l'augmentation des fréquences des réponses « chute » avec le degré d'inclinaison, et par l'asymétrie observée entre les fréquences des réponses « chute » avant et arrière. Pour ce dernier critère, les chutes arrière apparaissent avec un degré d'inclinaison plus petit du corps humain, ce qui correspond tout à fait aux mesures empiriques réalisées sur un corps humain réel.

De plus, lorsque des objets inanimés anthropomorphes sont présentés, les sujets transposent implicitement ces connaissances spécifiques du corps humain à ceux-ci, tel qu'un mannequin de dessin ou un squelette. On observe alors, comme pour le corps humain, une augmentation des réponses « chute » en fonction du degré d'inclinaison et l'asymétrie des réponses « chute » avant – arrière (cf. expérience 1). Ces résultats persistent lorsque l'objet inanimé (mannequin de dessin) est présenté dans un contexte pur (cf. expérience 2). La comparaison des expériences 1 et 2 permet d'écarter l'hypothèse d'un effet de contexte sémantique dû à une présentation mixte car dans l'expérience 2 la présentation d'un objet inanimé anthropomorphe engendre spontanément des réponses reflétant l'utilisation de représentations spécifiques du corps humain.

Lors d'une autre série d'expériences (Nithart, 2004), des images du corps humain et de l'objet manufacturé, en utilisant une rotation au niveau de la cheville, ont été présentées (cf. Figure 1.7). Différents degrés d'inclinaison des deux objets étaient présentés aux sujets suivant la même procédure expérimentale que celle décrite pour les expériences exposées ci-dessus.

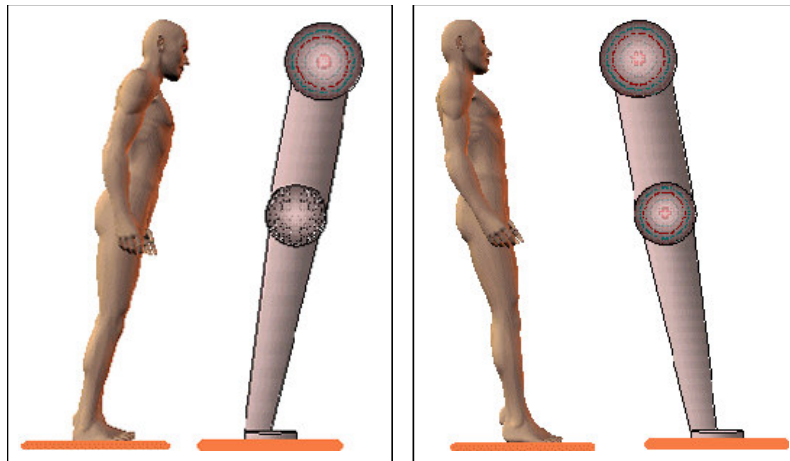


Figure 1.7 : Deux exemplaires de chaque stimulus profil droite (en couleur dans l'expérience) présenté au niveau d'inclinaison avant (F) et arrière (B) de 5° à partir de la cheville : le corps humain et l'objet manufacturé.

De manière générale, les expériences avec la rotation au niveau de la cheville montrent à nouveau une augmentation des fréquences des réponses chute en fonction du degré d'inclinaison indépendamment de la direction avant – arrière, et cela pour le corps humain et l'objet manufacturé. De plus, l'asymétrie des fréquences réponses chute avant – arrière n'est pas retrouvée, ni pour le corps humain, ni pour l'objet manufacturé. L'absence de l'asymétrie des réponses chute pour le corps humain, correspond dans ces expériences également à l'utilisation de représentations spécifiques du corps humain, respectant les contraintes biomécaniques de celui-ci. Effectivement, le déplacement du site de rotation du corps humain de la hanche vers la cheville correspond à une modification du polygone de sustentation du corps entier. Dans ces conditions, le point de chute physique pour l'homme se situe au moment où le centre de gravité sort du polygone de sustentation impliquant que la différence du degré d'inclinaison des réponses chute avant – arrière devient statistiquement non significative. Cette dernière expérience montre que l'utilisation de représentations spécifiques au corps humain n'est active et efficace, lorsque celles-ci conduisent à des prédictions différentes des contraintes de la force de la gravité.

En conclusion, ces expériences apportent des arguments supplémentaires en faveur de l'hypothèse de traitements spécifiques lorsque les images représentent un corps humain contrairement à d'autres objets inanimés non anthropomorphes. D'autres études récentes ont également montré des indices en faveur de traitement spécifique du corps humain : une préférence visuelle pour l'image d'un corps humain entier par rapport à mélangé (Slaughter et

al., 2002), l'effet d'inversion au corps humain (Reed et al., 2003), une différence de traitements en fonction des parties du corps humain (Reed et al., 2004) et le cas de troubles sélectifs chez des patients cérébro-lésés (i.e. Campbell & Manning, 1996 ; Manning & Warrington, 1996). Ces expériences mettent également en évidence que les caractéristiques spécifiques du corps humain sont spontanément généralisées à des objets inanimés anthropomorphes.

Enfin, des expériences récentes en neuroimagerie montrent que la présentation de différentes postures du corps humain active des représentations visuelles de mouvement (Kourtzi & Kanwisher, 2000 ; Paulos et al., 2003). Différentes postures du corps humain considérées comme étant au repos n'activent pas de telles représentations (Kourtzi & Kanwisher, 2000 ; Paulos et al., 2004). Ces processus sont utilisés évidemment au quotidien, mais également dans des productions picturales, comme par exemple dans la peinture ou dans les bandes dessinées. Cutting (2002) met en évidence les différents indices utilisés dans la sculpture et la peinture afin de représenter un mouvement d'un animal ou d'un personnage. Nous avons implicitement ou explicitement recours à ces connaissances à travers notre expérience personnelle et l'utilisation de représentations spécifiques du corps humain. Celles-ci nous aident également dans toute interaction sociale avec des congénères en préparant le corps à agir dans son environnement (e.g. Chaminade et al., 2002 ; Decety et al., 2003).



## **II - La spécificité du corps humain dans un paradigme de mouvement apparent**

La partie ci-dessous présente une synthèse d'une série d'expériences dont le détail figure en annexe (cf. Bonnet & Paulos ; a ; b, soumis). D'autres expériences complémentaires peuvent être consultées dans les travaux de fin d'étude de maîtrise suivants : Benoliel (2003), Pirra (2004), Seereekissoon (2004) et Van Assche (2005).

### **A. Facteurs structuraux et sémantique intervenant dans le paradigme de Ternus**

L'objectif général de ces expériences est d'examiner dans quelle mesure des facteurs sémantiques concernant la représentation d'un corps humain en action peuvent moduler des effets perceptifs résultant de mécanismes de bas niveau, et en particulier dans le cadre expérimental du mouvement apparent en utilisant le paradigme de Ternus (1926).

Les expériences de Shiffrar et Freyd (1990, 1993) sur le mouvement apparent bêta montraient que le prolongement de la durée de traitement, prenant en compte les contraintes biomécaniques, avait un effet sur la perception de la forme de la trajectoire du mouvement. Cependant, du point de vue méthodologique, la discrimination d'une trajectoire directe ou courbe, représente un phénomène moins robuste (plus labile) que le phénomène de Ternus que nous avons utilisé.

L'intérêt du paradigme de Ternus dans notre problématique est que, contrairement aux expériences de Shiffrar et Freyd (1990, 1993), la tâche des sujets ne porte pas directement sur les caractéristiques phénoménales qui nous intéressent. C'est à partir des variations des fréquences des réponses *élément* et *groupe* que nous évaluerons l'effet des différents facteurs mis en jeu.

Un mouvement apparent peut être créé par la présentation répétée et successive de deux images illustrant une même scène tenant compte d'un décalage spatiale d'un ou de plusieurs objets. Ce phénomène est largement utilisé dans différentes techniques de dessins animés comme par exemple les mangas (dessins animés japonais). Le phénomène de mouvement apparent de Ternus utilise deux images représentant chacune trois objets identiques. Le mouvement apparent est créé, ici par l'alternance pendant 2 cycles et demi de deux images

comportant chacune 3 objets régulièrement espacés en position A, B, C pour la première image et B, C, D pour la seconde. La durée d'exposition de chaque image est de 83,3 ms. L'intervalle temporel entre deux images varie d'essai en essai allant de 16.6 ms à 166 ms. En fonction de la durée de l'intervalle inter stimulus (ISI) deux types de mouvement sont perçus : mouvement élément pour les ISI les plus courts et mouvement groupe pour les ISI les plus longs. Le mouvement élément est perçu comme un saut de l'objet entre les 2 positions extrêmes, tandis que le mouvement groupe est perçu comme un déplacement spatial des 3 objets. La tâche des sujets est, avec une consigne de rapidité, d'indiquer s'il perçoit un mouvement **élément** du seul objet en position extrême (de la position A vers la position D) ou bien un mouvement **groupe** de 3 objets (des positions ABC vers les positions BCD).

Toutes les expériences sont précédées d'une phase de familiarisation à la tâche expérimentale, où les sujets doivent rapporter en temps de réaction la perception du mouvement de l'item élément ou du groupe à partir de la présentation d'images de disques noirs.

Dans nos expériences les stimuli présentés sont de différente nature sémantique: soit des images du corps humain, soit des images d'une chaise. Les résultats présentés ci-dessous sont les fréquences de réponses « groupe » pour les différentes conditions respectives en fonction de la durée de l'intervalle inter stimulus (ISI), mesuré en nombre de trames de balayage de l'écran de l'ordinateur (60Hz ; 1 trame = 16.67ms). Toutes les expériences ont été générées à partir du logiciel Inquisit V1.32 de Millisecond. Chaque expérience a été menée auprès d'un différent groupe composé de 16 sujets (étudiants en psychologie), tous naïfs par rapport aux hypothèses sous investigation et avec une vue normale ou corrigée à la normale.

Classiquement, la fréquence des réponses « groupe » croît avec l'augmentation de la durée de l'ISI (fonction psychométrique logistique) lorsque les stimuli d'une séquence sont identiques. Pour l'essentiel, les résultats de la littérature, depuis l'expérience princeps de Ternus (1926), montrent que, à ISI constant, la fréquence des réponses *groupe* est d'autant plus grande que les différents items sont structurellement identiques. En particulier, des différences d'orientation entre éléments réduisent considérablement cette fréquence (e.g. Allais & Lorenceau, 2002). Même si ces mécanismes responsables de la transition entre des réponses *élément* et des réponses *groupe* sont encore l'objet de discussions et d'interrogations, tous les auteurs s'accordent à penser qu'il s'agit de mécanismes de « bas niveau », mis en jeu essentiellement par les caractéristiques structurales (spatiales et temporelles) des stimuli.

Sans rentrer dans les détails des différentes expériences qui utilisent toutes le même paradigme, nous en mentionnerons seulement ici les principaux résultats. Tout d’abord et quel que soit le type d’objets utiliser (un personnage humain ou une chaise), nous vérifions que les effets classiques des facteurs structuraux apparaissent bien dans nos conditions. Dans un second temps, nous chercherons à voir si ces effets structuraux peuvent être modulés par des facteurs d’ordre sémantique, i.e. des représentations d’une action, en nous centrant en particulier sur des situations dans lesquelles les objets dans les différentes positions sont quatre postures séquentielles tirées d’une action.

Selon le résultat classique du paradigme de Ternus, la fréquence des réponses « groupe » augmente avec la durée de l’ISI. La pente de cette fonction psychométrique est maximale lorsque les 4 items sont (structuralement et sémantiquement) identiques (Figure 2.2).

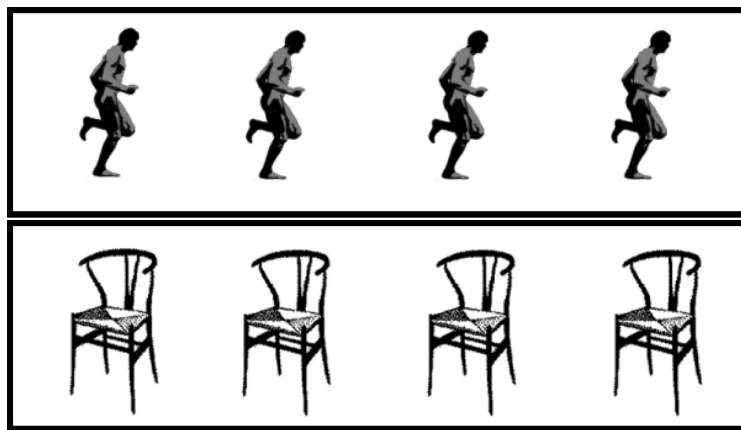


Figure 2.1 : Exemple des stimuli utilisés dans les premières expériences : images d’un personnage humain (H) et images d’une chaise (C).

Dans le cas où les items sont identiques, les fréquences des réponses « groupe » évoluent de la même manière que les items représentent un personnage humain ou une chaise.

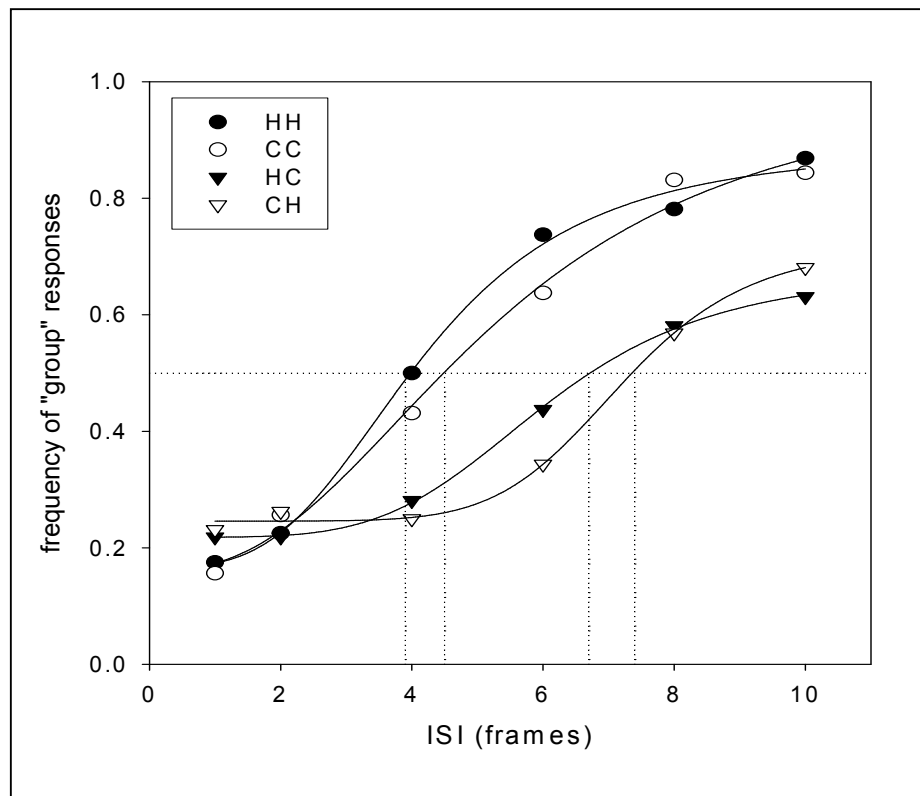


Figure 2.2 : Fréquences des réponses « groupe » lors de différentes conditions expérimentales. HH, CC les 4 éléments sont identiques : soit un personnage humain, soit une chaise. L'élément situé aux positions extrêmes A et D est différent : soit un personnage humain (HC), soit une chaise (CH).

Ce premier résultat montre l'augmentation de la fréquence des réponses « groupe » en fonction de la durée de l'intervalle inter stimulus (ISI). La pente de cette fonction est maximale lorsque les éléments sont strictement identiques (HH et CC). Dans les conditions HC et CH la pente des fréquences des réponses « groupe » est plus faible. Ainsi, suivant un changement structural (i.e. forme de l'objet) et un changement sémantique (i.e. personnage humain – chaise) perturbent considérablement le déploiement de la fonction de réponse groupe.

Dans une deuxième expérience, une séquence mixte a été créée en utilisant uniquement des chaises. Celles-ci étaient de différentes formes et orientation (Figure 2.3). Notre hypothèse était de comparer à l'intérieur de la catégorie sémantique « chaise » l'effet sur les fréquences

des réponses « groupe » de changements structuraux au niveau de la forme et de l'orientation de la chaise.

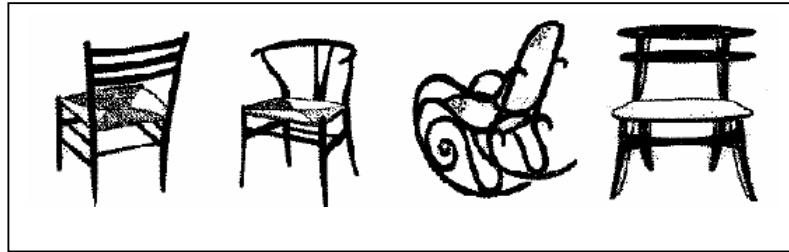


Figure 2.3 : Séquence mixte de la catégorie sémantique « chaise » incluant des changements structuraux au niveau de la forme et de l'orientation.

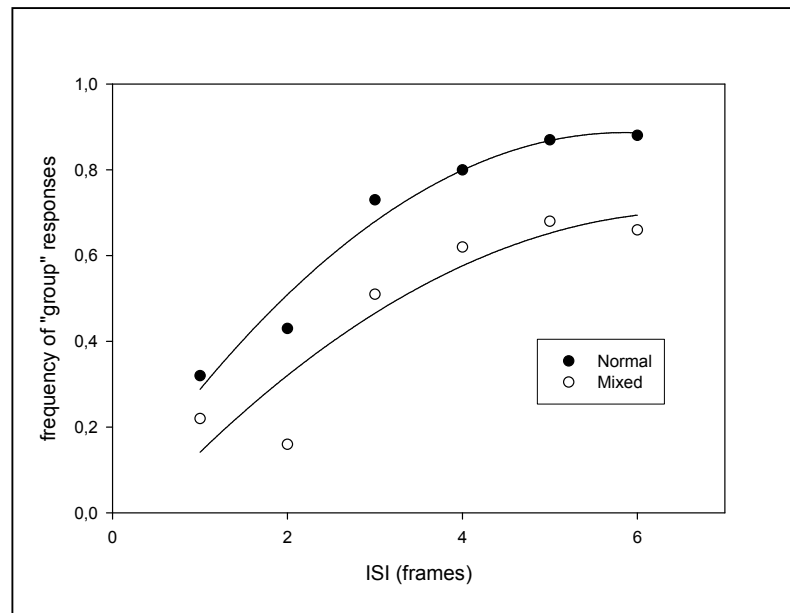


Figure 2.4 : Fréquences des réponses « groupe » pour la séquence mixte (i.e. changements de forme et d'orientation) et la séquence normale (i.e. chaise identique).

La figure 2.4 montre une augmentation des fréquences des réponses « groupe » en fonction de la durée de l'ISI, et cela pour la séquence normale et la séquence mixte. Cependant, l'évolution des réponses « groupe » pour la séquence mixte est considérablement perturbée n'atteignant pas 70% de réponses « groupe » pour les ISI les plus longs.

Dans la même perspective, une troisième expérience se focalise sur des changements structuraux de forme et d'orientation à l'intérieur de la catégorie sémantique corps humain. Deux séquences supplémentaires (Figure 2.5) ont été introduites en réalisant un changement de la forme (i.e. différentes postures du corps humain) incluant en dernière position soit un changement de direction (i.e. *séquence miroir*), soit un changement d'orientation (i.e. *séquence inverse*).

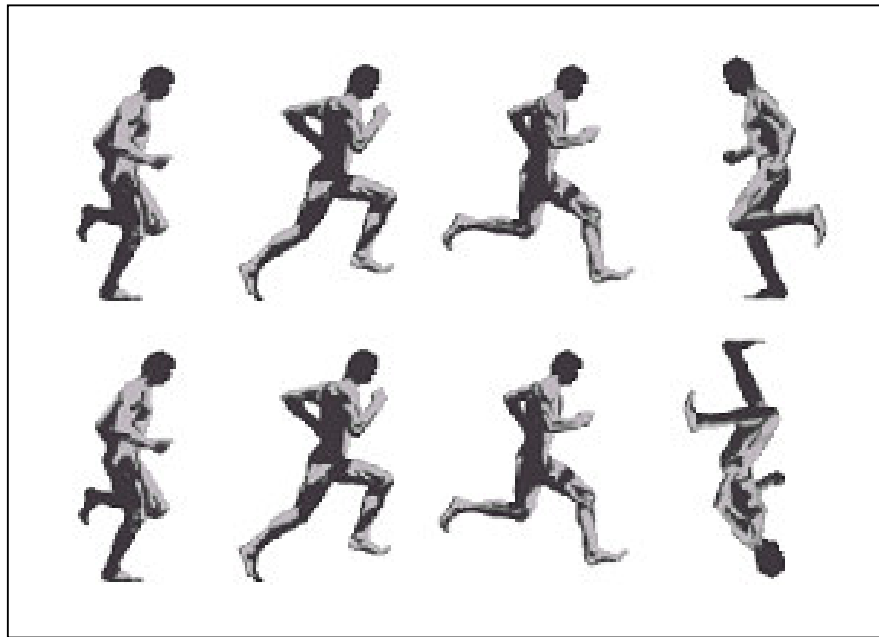


Figure 2.5 : Séquences du corps humain incluant des changements structuraux de la forme et de l'orientation : séquence miroir (en haut), séquence inverse (en bas).

Dans l'expérience précédente, les changements d'orientation étaient référencés à la verticale. Mais l'effet des changements structuraux d'orientation se manifeste aussi lorsque l'on inverse le haut et le bas des objets (cf. séquence inverse). Une telle inversion haut bas de l'objet en dernière position par rapport aux précédents réduit notablement la fréquence des réponses « groupe » (figure 2.6) que les stimuli soient des chaises (39% vs. 62%) ou des humains (53% vs. 60%).

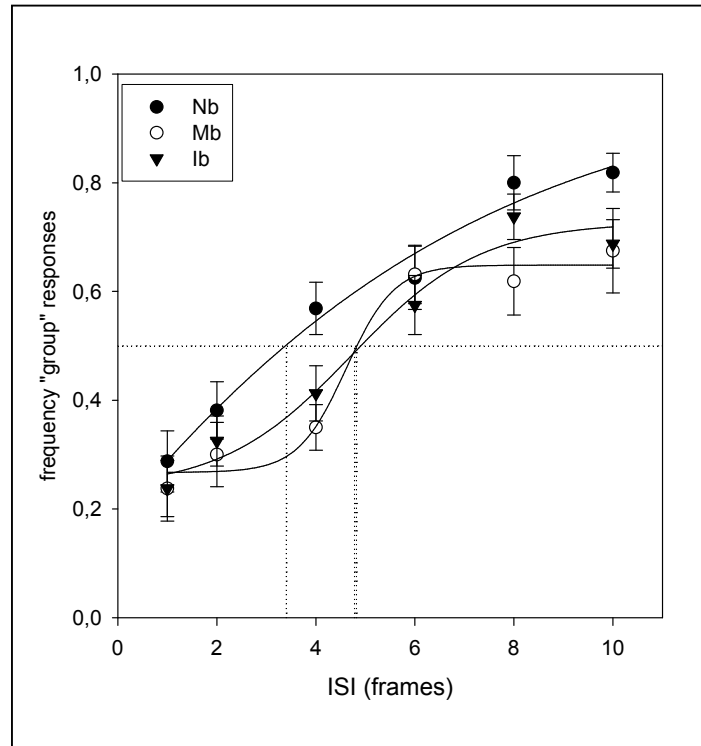


Figure 2.6 : Fréquences des réponses « groupe » pour deux séquences du corps humain incluant des changements structuraux de forme et d'orientation (cf. séquence miroir et séquence inverse) en fonction de la durée de l'intervalle inter stimulus (ISI).

Enfin, l'effet d'une hétérogénéité structurale entre les items a été testé en utilisant des objets différents mais appartenant à la même catégorie sémantique (quatre formes de chaises). Cette hétérogénéité structurale réduit la fréquence des réponses « groupe », par rapport à une condition de 4 chaises identiques (47% vs. 67%). De même, un effet analogue est retrouvé lorsque les objets appartiennent à la même catégorie sémantique (chaise) et présentent la même orientation de haut en bas (figure 2.3) que s'ils appartiennent à des catégories sémantiques distinctes (chaise et corps humain, 37% vs. 52%).

Dans une dernière expérience présentant un corps humain deux séquences supplémentaires ont été comparées incluant des changements structuraux uniquement au niveau de l'orientation. La première a été fabriquée à partir d'un corps humain rigide réalisant une

rotation de 90° à chaque position (figure 2.7 en haut). La deuxième a été réalisée en utilisant un corps humain en posture de poirier (figure 2.7 en bas).

En comparant une condition où les quatre items sont une posture en poirier d'un personnage avec une condition dans laquelle l'orientation d'un corps humain rigide change à chaque position, Van Assche (2005) observe en moyenne 62% de réponses « groupe » pour la première (orientation constante du corps), mais seulement 32% de ces réponses lorsque l'orientation est variable. Dans les deux conditions, cette fréquence augmente avec l'ISI, mais plus fortement pour l'orientation constante.

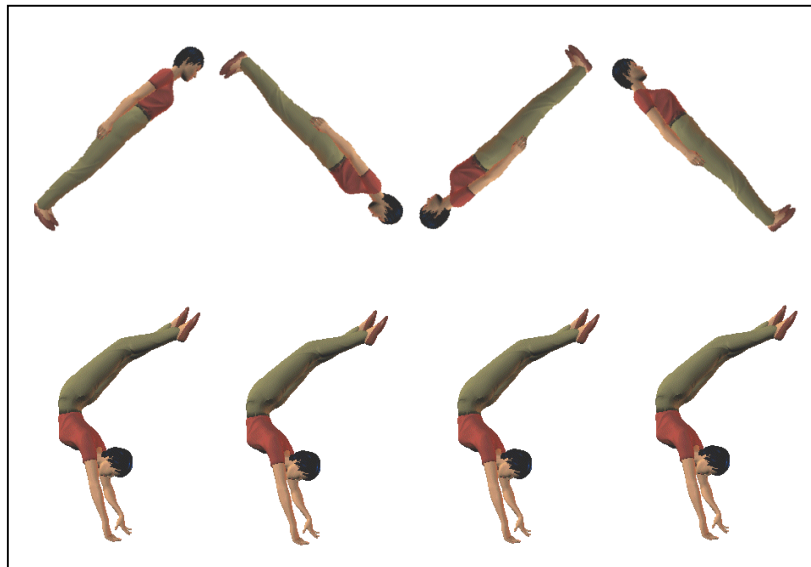


Figure 2.7 : Séquences de corps humain incluant des changements structuraux de l'orientation en maintenant la forme constante.

En résumé, les résultats de ces travaux montrent que l'augmentation des réponses « groupe » en fonction de la durée de l'ISI est retrouvée de manière générale dans toutes les différentes conditions. Néanmoins, le déploiement de la fonction des réponses « groupe » est nettement perturbé lorsque des changements structuraux interviennent entre les différents items d'une séquence.



## **B. Représentations visuelles du corps humain à partir d'une séquence d'action**

A partir des résultats mentionnés ci-dessus, nous avons cherché à voir dans quelle mesure les perturbations du facteur d'hétérogénéité structurale pouvaient être atténuées par un facteur sémantique. Nous nous sommes alors concentrés sur l'étude de séquences d'action normales ou perturbées avec l'hypothèse que la présentation d'une séquence normale, malgré d'éventuelles différences structurales entre les items (i.e. différentes postures du corps humain), favoriseraient les réponses « groupe ». Ce qui caractérise généralement un mouvement d'un être animé est non seulement le déplacement spatial, mais aussi les déformations d'un corps non rigide. Ainsi, des changements de l'orientation des items par rapport à la verticale dans une séquence « naturelle » d'action pourraient réduire l'effet négatif de ces changements structuraux d'orientation sur la fréquence des réponses « groupe ».

Nous avons utilisé 4 items extraits d'une séquence d'action d'un personnage. Trois actions (figure 2.8) ont été testées correspondant à trois grands types de mouvement et impliquant des différences structurales plus ou moins importantes entre les différentes postures :

- **course** : seuls les bras et les jambes changent d'un item à l'autre. La translation spatiale est cohérente avec une représentation de course.
- **pirouette** : la figure implique une rotation du corps autour de l'axe vertical du corps et donc des changements notables des positions relatives des bras et des jambes. Ici, la translation spatiale n'est pas cohérente avec la figure représentée qui s'effectue normalement « sur place ».
- **saut carpié** : la figure implique une rotation autour d'un axe horizontal et donc des changements d'orientation par rapport à la verticale ainsi que de la forme même du corps (non rigide). La translation spatiale est cohérente avec une représentation de cette figure de gymnastique.

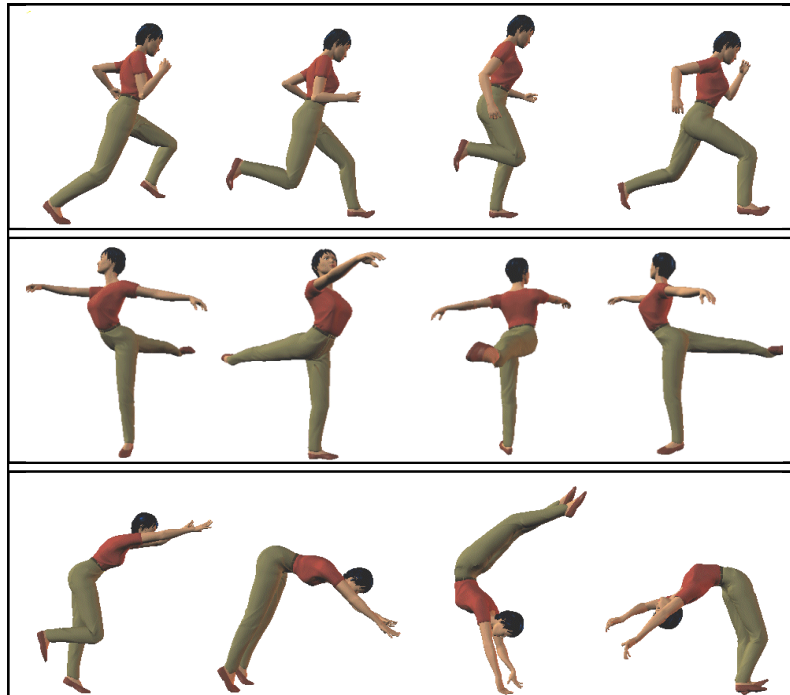


Figure 2.8 : Séquences d'un personnage en train de réaliser une action : la course (en haut), une pirouette (au milieu) et un saut carpé (en bas).

Testée sur des groupes indépendants de sujets (i.e. étudiants en psychologie), la variation de la fréquence des réponses « groupe » avec l'ISI (figure 2.9) est très voisine pour les trois types d'actions et voisine de ce qu'elle est lorsque l'item est identique dans les 4 positions. La fréquence moyenne des réponses groupe est respectivement de 50% pour la course, de 47% pour la pirouette et de 45% pour le saut carpé. On observe donc un léger effet des changements d'orientation du corps avec un peu plus de réponses « élément » pour la pirouette que pour la course et encore un peu plus pour le saut carpé. Mais les différences restent minimes, ce qui est compatible avec un effet modulateur de la nature sémantique des stimuli sur les effets des changements structuraux. Notons de plus que les TR moyens augmentent avec la complexité de la séquence d'action, passant respectivement de 711 ms pour la course, à 1102 ms pour la pirouette et à 1154 ms pour le saut carpé.

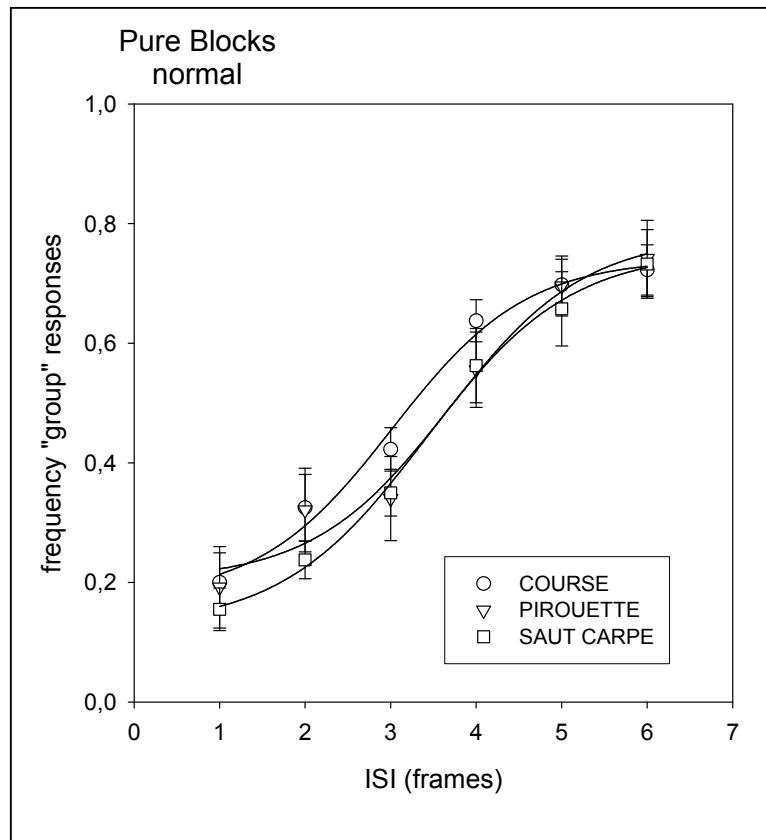


Figure 2.9 : Fréquences de réponses « groupe » en fonction de la durée de l'intervalle inter stimulus (ISI) pour les séquences d'un personnage en action : course, pirouette et saut carpe.

La régularité de ces trois actions a été perturbée donnant lieu à deux séquences supplémentaires. La figure 2.10 présente la séquence régulière et les deux séquences perturbées pour l'exemple du saut carpe. En utilisant les mêmes conditions expérimentales l'effet de l'augmentation de l'ISI sur les réponses « groupe » peut devenir très atypique et se traduire par un plateau ne différenciant pas les réponses « groupe » des réponses « éléments ». Cependant, il y a de fortes raisons de penser que cet effet ne résulte pas de facteurs sémantiques, mais de facteurs de bas niveau. En effet, les changements utilisés conduisaient à la perception d'un mouvement apparent de type bêta des éléments, entraînant alors une forte

augmentation de la fréquence des réponses « groupe ». Celles-ci traduisent alors, non un renforcement de la cohérence perceptive du groupe des items, mais un substitut à une réponse de plusieurs éléments en mouvement.

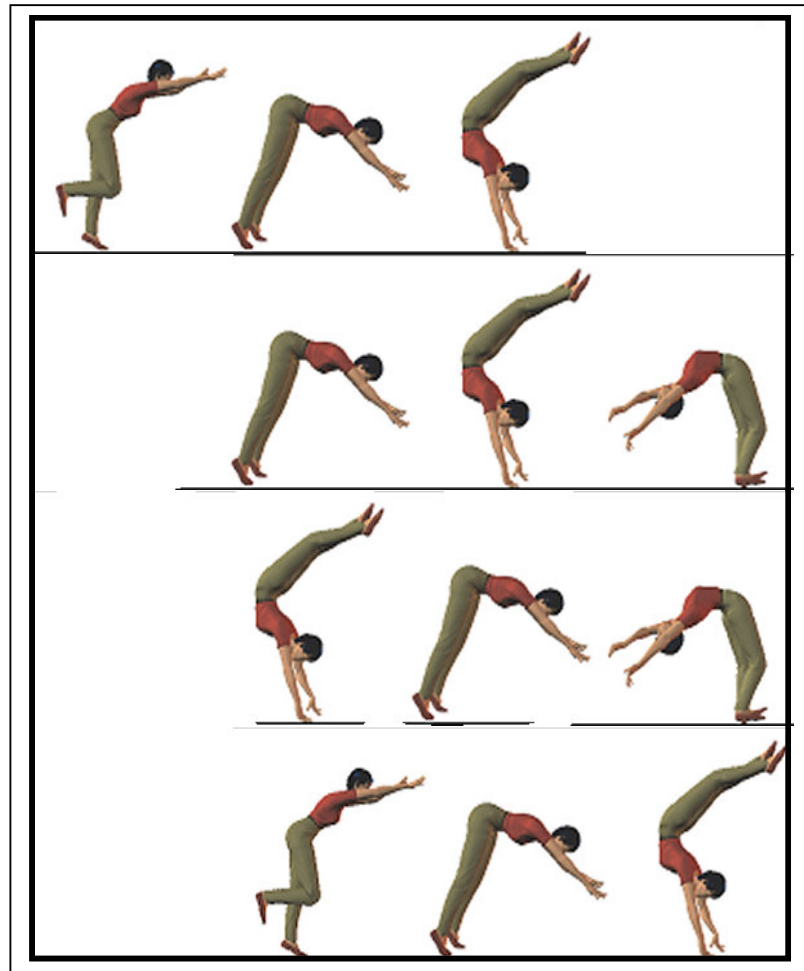


Figure 2.10 : Exemple de la séquence d'action « saut carpé » présentant la première image de la séquence (en haut), et la deuxième image représentant la séquence régulière et deux séquences perturbées : inversion des deux postures centrales et décalage des 3 postures initiales des positions ABC vers les positions BCD (cf. translatée).

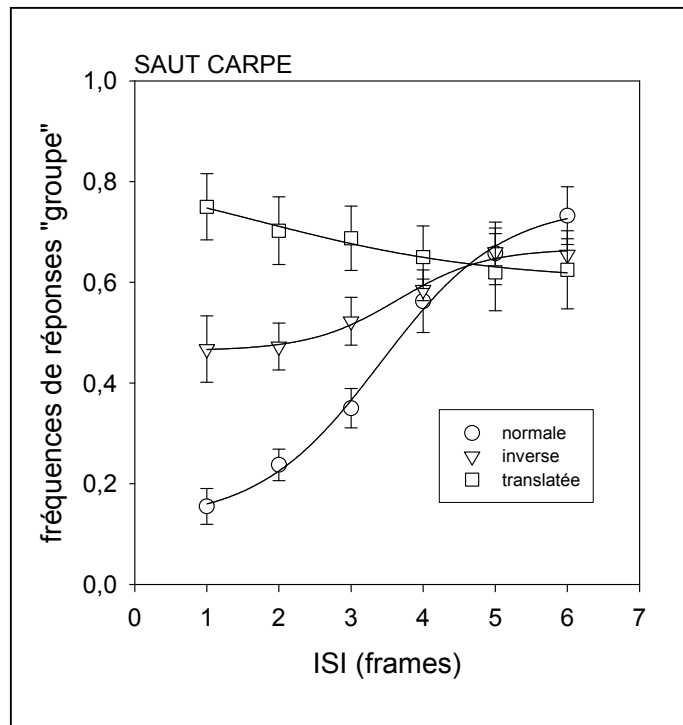


Figure 2.11 : Fréquences des réponses « groupe » pour trois séquences du saut carpé présentant une séquence régulière (i.e. normale), une inverse et une translatée en fonction de la durée de l'intervalle inter stimulus (ISI).

Ces expériences montrent la robustesse des effets des facteurs structuraux sur le phénomène de Ternus, laissant peu de place à une modulation de ces effets par des facteurs sémantiques, tout au moins de ceux que nous avons testés, en particulier une séquence cohérente de différentes postures d'un corps humain représentant une action. Cependant, cette modulation n'est pas absente comme nous l'avons vu dans plusieurs conditions. Récemment, Jarraya et al. (2005) montrent que les sujets anticipent correctement la suite d'une séquence de mouvements de gymnastique.

L'utilisation d'une consigne de rapidité et la mesure des temps de réaction permet d'observer que dans les conditions dans lesquelles les facteurs sémantiques semblent jouer un rôle modulateur, les TR sont nettement plus longs que dans les conditions dans lesquelles où seul le rôle des facteurs structuraux est suspecté. Cet allongement des TR peut s'interpréter comme un temps de traitement plus long de l'information et donc comme une plus grande chance que les caractéristiques sémantiques soient prises en compte dans les réponses. C'est un point de vue analogue qui est adopté par Shiffrar et Freyd (1990, 1993) pour rendre compte du passage des réponses directes (biomécaniquement impossibles, i.e. un bras passe à travers une jambe) à des réponses biomécaniquement possibles (i.e. un bras contourne une jambe).

### **C. Discussion générale et perspectives**

Les travaux expérimentaux présentés ci-dessus visent à vérifier dans quelle mesure les mécanismes des traitements des informations relatives au corps humain sont spécifiques.

Dans une première expérience en IRMf, deux catégories de stimuli ont été présentées à un groupe de sujets qui avaient pour seule consigne de maintenir leur regard au centre du champ visuel (cf. croix de fixation). Dans cette expérience les stimuli utilisés représentaient d'une part différentes postures du corps humain et d'autre part différents objets manufacturés. Le premier objectif de cette expérience était de vérifier dans quelle mesure les régions cérébrales impliquées dans les traitements des informations pour chacune de ces catégories seraient différenciables. En particulier, deux sous-catégories ont été créées a priori (Paulos, 2001) présentant différentes postures d'un personnage en train de réaliser une action (e.g. un mouvement) par rapport à différentes postures d'un personnage en repos. Les objets manufacturés avaient été classés en objets manipulables dans une action (e.g. un mouvement) par rapport à des objets ne suggérant pas au premier abord une action ou un déplacement spatial. Le deuxième objectif de cette expérience était de vérifier dans quelle mesure les sous-catégories de personnages et d'objets de « mouvement » évoqueraient des représentations de mouvement localisables dans les mêmes structures cérébrales impliquées dans les processus de mouvement visuel, et plus précisément l'aire MT / V5 (Kourtzi & Kanwisher (2000), Senior et al. (2000). Les résultats de cette première expérience en IRMf montrent principalement une différenciation des structures cérébrales impliquées au niveau du gyrus fusiforme latéral (GFL) et médian (GFM), respectivement pour les personnages et les objets (e.g. Beauchamps et al. (2002, 2003), Haxby et al. 2001, Ishai et al. 2000, 1999 ; Bartels & Zeki, 2004). Des activations cérébrales dans l'aire MT / V5 ont été localisées. Contrairement à nos attentes, celles-ci n'étaient pas présentes pour les sous-catégories de personnages et d'objets en « mouvement », mais spécifiquement pour la catégorie des personnages et cela indépendamment de leurs postures suggérant la réalisation d'une action ou un état de repos. Une explication de ce résultat serait que les sujets, en absence de tâche explicite de catégorisation, aient implicitement procédé à une tâche de catégorisation en *objets animés* – *inanimés*. Comparant ainsi les personnages aux objets manufacturés, une principale caractéristique des personnages est leur capacité à exécuter des actions et à se déplacer spatialement.

Dans une deuxième expérience en IRMf les mêmes stimuli ont été repris en introduisant une tâche explicite de catégorisation. Les sujets avaient pour consigne de catégoriser les différentes postures du corps humain en personnages en train de réaliser une action (cf. un mouvement) ou au repos, et les objets manufacturés en objets facilement utilisables dans une action (cf. mouvement) ou « statique » n'impliquant pas directement une action. Cette tâche ne représentait aucune difficulté pour les sujets. L'objectif de cette expérience était de comparer les activations cérébrales lors d'une tâche de catégorisation en « mouvement » versus « statique » pour des personnages et des objets manufacturés. En particulier, des activations cérébrales dans la région de l'aire MT / V5 étaient attendues pour les catégorisations « mouvement » pour les personnages et les objets. Les résultats de cette deuxième expérience en IRMf montrent des activations cérébrales de l'aire MT / V5 en fonction de la catégorisation « mouvement » et cela indépendamment du fait que les stimuli étaient des personnages ou des objets. Ici, le contexte sémantique de la tâche de catégorisation est suffisant à évoquer des représentations de mouvement localisables dans la région de l'aire MT / V5. Néanmoins, en comparant ces activations dans la région de l'aire MT / V5 pour les personnages et les objets, celles-ci ne se recoupent que partiellement. Ce résultat indiquerait d'une part des représentations de mouvement communes aux deux catégories (i.e. l'action utilisant l'objet), et d'autre part des représentations de mouvement spécifiques au corps humain (i.e. son mouvement propre), et spécifiques à des objets (i.e. déplacement spatiale de l'objet).

D'un point de vue de traitement de l'information, les résultats de ces deux expériences pourraient refléter d'une part des projections corticales de la voie ventrale, en particulier du gyrus fusiforme vers l'aire du mouvement MT / V5, classiquement attribuées à la voie dorsale. Ainsi, l'activation de MT / V5 pourrait constituer une étape de traitement préalable à la préparation de l'activation du sillon temporal supérieur (STS) à l'intérieur d'un réseau neuronal distribué (e.g. pour une revue Allison et al., 2000). Cette région cérébrale paraît également jouer un rôle crucial dans les processus visuels se référant de manière plus générale à la communication sociale entre congénères, comme par exemple dans des comportements d'anticipation, d'imitation et d'échange de regard (e.g. Chaminade et al., 2002, Decety et al., 2004, 2003 ; Perrett et al. 2003 ; Schuller & Rossion, 2001).

La différence de résultats de ces deux expériences, concernant les représentations visuelles du mouvement et l'activité cérébrale en MT / V5, souligne le caractère spécifique du corps humain à évoquer implicitement des représentations de mouvement par rapport aux objets manufacturés. Néanmoins, cette spécificité n'apparaît plus lorsque la tâche des sujets était de



catégoriser explicitement les différents stimuli. Ainsi, se pose également la question d'extensibilité des représentations de mouvement implicite à d'autres objets vivants comme des animaux. Kourtzi et Kanwisher (2000) présentent dans leur expérience IRMf à côté des photographies d'un athlète réalisant un exercice physique, par exemple des photographies d'un dauphin effectuant un saut en dehors de l'eau. Pour cette condition de mouvement implicite elles observent également de l'activation bilatérale dans les régions MT / MST. Une expérience intéressante serait de mesurer l'activation cérébrale dans l'aire MT / V5 à partir de la présentation visuelle de dessins au trait d'animaux en *mouvement implicite* et en *repos*. Ce type de stimuli a déjà été créé et testé en comportement (Paulos, 2001), donnant des résultats très similaires, en pourcentage de catégorisation et de temps de réaction, à ceux obtenus avec les dessins au trait de différentes postures du corps humain. De même, il serait très intéressant de vérifier avec un paradigme événementiel les effets sur l'activation de l'aire MT / V5 en fonction de la tâche des sujets : *passive viewing* et *tâche de catégorisation*, pour ce type de matériel.

Dans les deux dernières expériences du travail de recherche présenté ci-dessus, l'objectif général était de préciser le contenu de ces représentations de mouvement, et de préciser en quoi celles-ci étaient spécifiques au corps humain. Lors d'une première expérience, les sujets avaient pour consigne de juger différents états d'équilibre, respectivement de déséquilibre à partir d'images d'un corps humain comparé à différents objets inanimés (i.e. un mannequin en bois, un squelette). Les caractéristiques biomécaniques du corps humain permettent que son équilibre ne résulte pas seulement de la force de la gravité. Lorsque celles-ci interviennent, les résultats montrent de manière générale une asymétrie des réponses « chute » en faveur des inclinaisons « arrière ». Cette asymétrie correspond à un point d'« équilibre » subjectif plus tardif pour les chutes vers l'avant, que pour les chutes vers l'arrière. Celui-ci a été vérifié empiriquement et correspond au point d'« équilibre » réel du corps humain, c'est-à-dire qu'un corps humain chute plus rapidement vers l'arrière que vers l'avant. Ces connaissances spécifiques au corps humain ont été spontanément transposées à des stimuli anthropomorphes, lorsque ceux-ci représentaient un mannequin en bois ou un squelette. Néanmoins, cette asymétrie des réponses « chute » correspondant à l'utilisation de connaissances spécifiques au corps humain n'est plus retrouvée pour un objet artificiel. Pour ce dernier, seules les contraintes de la force de la gravité interviennent. Ces résultats soulignent l'existence et l'utilisation de connaissances sur les propriétés biomécaniques et physiques du corps humain, et leurs attributions spontanées à des objets inanimés anthropomorphes.

Dans notre dernière série d'expériences, un paradigme de mouvement apparent présentait des séquences d'action à partir d'images montrant différentes postures du corps humain. L'objectif général de ces expériences était de vérifier dans quelle mesure des représentations de mouvement (i.e. une séquence d'action) pouvaient moduler des effets de plus bas niveau résultant des mécanismes perceptifs du mouvement apparent utilisé. Les résultats de ces expériences montrent dans certaines conditions, en particulier lorsqu'une séquence de différentes postures du corps humain représente différents moments d'une action cohérente, un effet modulateur de groupement sur les mécanismes perceptifs. Une explication de cet effet modulateur serait l'activation des représentations de mouvement évoquée par les différentes postures du corps humain constituant une séquence d'action cohérente.

Quelques arguments supplémentaires présents dans littérature suggèrent également des processus spécifiques au traitement de l'information du corps humain entier. Récemment, Slaughter et al. (2002) montrent dans une approche développementale des différences de traitements pour des configurations du corps humain présentées d'une manière cohérente par comparaison à incohérente (scrambled). Ces auteurs suggèrent que chez les enfants âgés de 18 mois seulement, commencerait la mise en place de processus spécifiques montrant une préférence visuelle pour des configurations canoniques du corps humain. De même, Reed et al. (2003) présentent des caractéristiques d'un effet d'inversion pour le corps humain entier très similaires à celles observées pour les visages humains. Néanmoins, ces processus de traitements de l'information spécifiques à ces deux catégories de stimuli semblent se différencier par la relation entre des représentations du corps humain et des représentations de mouvement ou d'action (e.g. Astafiev et al., 2004, Slaughter et al., 2004, Downing et al., 2006). La mise en commun de ces différents résultats suggère la mise en place de processus cognitifs spécifiques au traitement visuel du corps humain entier, présents très tôt dans le développement.

## **D. Bibliographie**

Allison T., Puce A., & McCarthy G. (2000) Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, 4, 267–278.

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.

Batelli L., Cavanagh P., Thornton I. M. (2003) Perception of biological motion in parietal patients. *Neuropsychologia*, 41, 13, 1808 – 1816.

Bartels A. & Zeki S. (2004) Functional brain mapping during free viewing natural scenes. *Human Brain Mapping*, 21, 75 – 85.

Beauchamp M., Lee K., Haxby J. & Martin A., (2002) Parallel visual motion processing streams for manipulable objects and human movements. *Neuron*, 34, 149-159.

Beauchamp M., Lee K., Haxby J. & Martin A., (2003) fMRI Responses to video and point-light displays of moving humans and manipulable objects. *J. Cogn. Neurosc.*, 15, 7, 991 – 1001.

Benoliel N. (2003) Influence de la similarité structurale sur la perception du mouvement apparent d'éléments figuratifs : paradigme de Ternus. Mémoire de la Maîtrise de Psychologie Cognitive, Université Louis Pasteur, Strasbourg.

Bonnet C. & Lestienne F. (2003) Percevoir et produire le mouvement. Armand Colin, collection U.

Bonnet C. & Paulos C. (2004a) A wooden mannequin perceived as an animated object? In A. Oliveira et al. *Fechner Day 2004. Proceedings of the ISP conference*, Coimbra (315-319).

Bonnet C. & Paulos C. (2004b) Reaction Times as a measure of uncertainty. In A. Oliveira et al. *Fechner Day 2004. Proceedings of the ISP conference*, Coimbra (320-324).

Bonnet C., Paulos C. & Nithart C. (2005) Visual representations of dynamic actions from static pictures. *Perception*, 34, 835 – 846.

Bonnet C. & Paulos C. (soumis a) Structural and semantic factors in the Ternus paradigm. *Visual Cognition*.

Bonnet C. & Paulos C. (soumis b) Human body actions in the Ternus paradigm. *Visual Cognition*.

Boucart, M., & Bonnet, C. (1991) A study of the effect of structural information and familiarity in form perception. *Quarterly Journal of Experimental Psychology*, 43A, 223-248.

Bruce V., Cowey A., Ellis A. & Perrette D. (1992) Processing the facial image. Oxford University Press.

Campbell R. & Manning L. (1996) Optic aphasia: A case with spared action naming and associated disorders. *Brain and Language*, 53, 183 – 221.

Chaminade T., Meltzoff A. N. & Decety J. (2002) Does the end justify the means? A PET exploration of the mechanisms involved in human imitation. *NeuroImage*, 15, 318 – 328.

Chaminade, T. & Decety, J. (2002) Leader or follower? Involvement of the inferior parietal lobule in agency. *NeuroReport*, 15, 1975-1978.

Chatterjee S. H., Freyd J. J., Shiffrar M. (1996) Configural processing in the perception of apparent biological motion. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 916 - 929.

Cutting J. E. (2002) Representing motion in a static image: constraints and parallels in arts, science, and popular culture. *Perception*, 31, 1165 - 1193.

- Decety J. & Sommerville J. A. (2003) Shared representations between self and other: a social cognitive neuroscience view. *Trends in Cognitive Sciences*, 7, 12, 527 – 533.
- Decety J. & Grezes J. (1999) Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Sciences*, 3, 172 – 178.
- Derrington A. M., Allen H. A., Delicato L. S. (2004) Visual mechanisms of motion analysis and motion perception. *Annual Review of Psychology*, 55, 181 - 205.
- Deregowski J. B. (1999) Pictorial perception: individual and group differences within the human species. *Current Psychology of Cognition*, 18, 1031 - 1063.
- Downing P. E., Peelen M. V., Wigget A. & Tew B. (2006) The role of the extrastriate body area in action perception. *Social Neuroscience*, 1, 52 – 62.
- Downing P. E., Jiang Y., Shuman M. & Kanwisher N. (2001) The extrastriate body area (EBA). *Science*, 293, 2470 – 2473.
- Dukelow S.P., DeSouza J. F. X., Culham J.C., van den Berg A. V., Menon R. S. and Vilis T. , (2001) Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *J. Neurophysiol.*, 86, 1991-2000.
- Gauthier I., Anderson A.W., Tarr M.J., Skudlarski P., Gore J.C. (1997). Levels of categorization in visual recognition studied with functional MRI. *Current Biology*, 7, 645-651.
- Gauthier I., Skudlarski P., Gore J. C. and Anderson A.W. (2000) Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3, 2, 191-197.
- Gauthier I., Tarr M.J. (2002) Unraveling mechanisms for expert object recognition: bridging brain activity and behaviour. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 431-446.

- Grèzes J. & Decety J. (2002) Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia*, 40, 212 – 222.
- Grèzes J., Fonlupt P., Bertenthal B., Delon-Martin C., Segebarth C & Decety J. (2001) Does Perception of Biological Motion Rely on Specific Brain Regions? *NeuroImage*, 13, 775 – 785.
- Grèzes J. & Decety J. (2001) Functional anatomy of execution, mental simulation, observation and verb generation of actions: A meta-Analysis. *Human Brain Mapping*, 12, 1 – 19.
- Farah M. J. (2000) The Cognitive Neuroscience of Vision. Fundamentals of Cognitive Neuroscience. Blackwell Publishers.
- Farah M. J. (2004) Visual Agnosia. Cambridge, MIT Press.
- Freyd J. J. (1983) The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics*, 33, 6, 575 – 581.
- Freyd J. J. & Pantzer T. (1995) Static patterns moving in mind. In S. M. Smith, T. B. Ward & R. A. Finke (Eds.), *The Creative Cognition Approach*. Cambridge, MIT Press, 181 – 204.
- Friston K. J., Holmes A., Worsley K. et al. (1995) Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2, 189-201.
- Friston K.J., Fletcher P., Josephs O., Holmes A., Rugg M.D., Turner R. (1998) Event-related fMRI: characterizing differential responses. *NeuroImage*, 7, 30 – 40.
- Hancock P. J. B., Burton A. M. & Bruce V. (1996) Face processing: Human perception and principal components analysis. *Memory & Cognition*, 24, 26-40.
- Haxby J., Gobbini M., Furey M., Ishai A., Schouten J. & Pietrini P. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293, 2425 – 2430.

Huk, A.C. & Heeger, D.J. (2001) Pattern–motion responses in human visual cortex. *Nature Neuroscience*, 5, 72–75.

Huk, A.C., Dougherty, R.F. & Heeger, D.J. (2002) Retinotopy and functional subdivision of human areas MT and MST. *J. Neurosci.*, 22, 7195–7205.

Ishai A., Ungerleider L., Martin A., Schouten J. & Haxby J. (1999) Distributed Representation of Objects in the Human Ventral Visual Pathway. *Proc. Natl. Acad. Sci.*, 96 (16), 9379 – 9384.

Ishai A., Ungerleider L., Martin A., Schouten J. & Haxby J. (2000) The Representation of Objects in the Human Occipital and Temporal Cortex. *Journal of Cognitive Neuroscience*, 12 (2), 35 – 51.

Jarraya M., Amorim M.-A., Bardy B.G. (2005) Optical flow and viewpoint change modulate the perception and memorization of complex motion. *Perception & Psychophysics*, 67, 951 – 961.

Jeannerod M. (2004) Visual and action cues contribute to the self-other distinction. *Nature Neuroscience*, 7, 5, 422 – 423.

Johansson G. (1973) Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14, 201 – 211.

Kanwisher N., McDermott J., Chun M.M. (1997) The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J. Neuroscience*, 17, 4302 – 4311.

Kourtzi Z. & Kanwisher N. (2000) Activation in human MT / MST by static images with implied motion. *Journal of Cognitive Neuroscience*, 12, 48 – 55.

Kourtzi Z. & Kanwisher N. (2000) Implied motion activates extrastriate motion-processing areas. Response to David and Senior (2000). *Trends in Cognitive Science*, 4, 8, 295 – 296.

- Kourtzi Z., Shiffrar M. (1999) Dynamic representations of human body movement. *Perception*, 28, 49 - 62.
- Kozlowski L. & Cutting J. E. (1977) Recognizing the sex of a walker from dynamic point-light displays. *Perception & Psychophysics*, 21, 575-580.
- Manning L. (2006) *La neuropsychologie clinique : approche cognitive*. Paris, Armand Colin.
- Manning L., Warrington E. K. (1996) Two routes to naming: a case study. *Neuropsychologia*, 34, 809 – 817.
- Maurer D., Le Grand R. & Mondloch C. J. (2002) The many faces of configural processing. *Trends in Cognitive Science*, 6, 255 – 260.
- Milner A. D. & Goodale M. A. (1995) *The Visual Brain in Action*. Oxford Science Publications.
- Moscovitch M., Winocur G. & Behrmann M. (1997) What is special about face recognition ? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience*, 9, 555 – 604.
- Moscovitch M. & Moscovitch D.A. (2000) Super face-inversion effects for isolated internal or external features, and for fractured faces. *Cognitive Neuropsychology*, 17, 201-219.
- Newsome W. T. & Pare E. B. (1988) A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience*, 8, 2201 – 2211.
- Nithart C. (2004) *Le statut perceptif particulier du corps humain. Mémoire de la Maîtrise de Psychologie Cognitive*, Université Louis Pasteur, Strasbourg.
- Orban G. A., Fize D., Peuskens H., Denys K., Nelissen K. Sunaert S., Todd J. & Vanduffel W. (2003) Similarities and differences in motion processing between the human and macaque brain: evidence from fMRI. *Neuropsychologia*, 41, 1757 – 1768.



Paradis A., Cornilleau-Pérès V., Droulez J., Van de Moortele P., Lobel E., Berthoz A., Le Bihan D. & Pioline J. B. (2000) Visual perception of motion and 3D structure from motion: an fMRI study. *Cerebral Cortex*, 10, 772 – 783.

Pascalis O., de Haan M. & Nelson C.A. (2002) Is face processing species-specific during the first year of life? *Science*, 14, 199 – 209.

Pascalis O., Demont E., de Haan M. & Campbell R. (2001) Recognition of faces of different species: a developmental study between 5-8 years of age. *Infant and Child Development*. 10, 39 – 45.

Pascalis O. & Bachevalier J. (1998) Face recognition in Primates: a cross species study. *Behavioural Processes*, 43, 87 – 96.

Paulos C. (2001) Représentations des objets et des actions: études préparatoires à une recherche en IRMf. Mémoire de DEA de Psychologie Cognitive sous la direction du Pr. Claude Bonnet. Université Lumière – Lyon 2.

Paulos C., Scheiber C., Manning L. & Bonnet C. (2001) Visual representations of object and actions: an fMRI study. Neurex Annual Meeting, Strasbourg.

Paulos C., Scheiber C., Manning L. & Bonnet C. (2003) Distributed representations in the visual cortex to human body and object line drawings. In *NeuroImage*, 19, 2, 1, S1-S101. Annual meeting of Cognitive Neuroscience Society, New York (CNS 2003).

Paulos C., Scheiber C., Manning L. & Bonnet C. (2004) To move or Not to move. Visual motion representations in motion area MT / V5. Neurex Annual Meeting, Freiburg.

Pavlova, M., Lutzenberger, W., Sokolov, A., & Birbaumer, N. (2004) Dissociable cortical processing for recognizable and non-recognizable biological movement: Analysing gamma MEG activity. *Cerebral Cortex*, 14, 181 – 188.

Perrett, D.I., Harries, M.H., Bevan, R., Thomas, S., Benson, P.J., Mistlin, A.J., Chitty, A.J., Hietanen, J.K. & Ortega, J.E. (1989) Frameworks of analysis for the neural representation of animate objects and actions. *J. Exp. Biol.*, 146, 87– 113.

Pinto, J. & Shiffrar M. (1999) Subconfigurations of the human form in the perception of biological motion displays. *Acta Psychologica*, 102, 293-318.

Pinto J. & Shiffrar M. (2003) Subconfigurations of the human form in the perception of biological motion displays. *Cognitive Neuroscience Society*, New York.

Pirra E. (2004) Influences de facteurs sémantiques sur un mouvement apparent d'images du corps humain : paradigme de Ternus. Mémoire de la Maîtrise de Psychologie Cognitive, Université Louis Pasteur, Strasbourg.

Reed, C.L., McGoldrick, J.E., Shackelford, R., & Fidopiastis, C. (2004). Are human bodies represented differently from other animate and inanimate objects? *Visual Cognition*, 11, 523-550.

Reed C. L., Stone V. E., Bozova S. & Tanaka J. (2003) The body-inversion effect. *Psychological Science*, 14, 4, 302 – 308.

Rossion B., Gauthier I., Tarr M. J., Despland P., Bruyer R., Linotte S. & Crommelinck M. (2000) The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *NeuroReport*, 11, 1, 69 – 74.

Schenk, T., Mai, N., Ditterich, J. & Zihl, J. (2000). Can a motion-blind patient reach for moving objects? *European Journal of Neuroscience*, 12, 3351-3360.

Schuller A-M & Rossion B. (2001) Spatial attention triggered by eye gaze increases and speeds up early visual activity. *Neuroreport*, 12, 2381-2386.

Shiffrar M. & Freyd J. J. (1993) Timing and apparent motion path choice with the human body photographs. *Psychological Science*, 4, 6, 379 – 384.

Seereekissoon R. (2004) L'apport des connaissances sémantiques dans la perception du mouvement apparent. Mémoire de la Maîtrise de Psychologie Cognitive, Université Louis Pasteur, Strasbourg.

Senior C. & David A. S. (2000) Implicit motion and the brain. *Trends in Cognitive Science*, 4, 8, 293 – 295.

Senior C., Ward J. & David A. S. (2002) Representational momentum and the brain: An investigation into the functional necessity of V5 / MT. *Visual Cognition*, 9 (1/2), 81 – 92.

Slaughter V., Heron M. & Sim S. (2002) Development of preferences for the human body shape in infancy. *Cognition*, 85, B71 – B81.

Slaughter V., Stone V. E. & Reed C. (2004) Perception of faces and bodies. Similar or Different? *Current Directions in Psychological Science*, 13, 219 – 223.

Snodgrass J. G. & Vanderwart M., (1980) A Standardized Set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity and Visual Complexity. *J. Exp. Psychology: Human Learning and Memory*, 6 (2), 174-215.

Snowden R.J., Treue S., Erickson R.G. et Andersen R.A. (1991) The response of area MT and V1 neurons to transparent motion. *Journal of Neuroscience*, 11, 2768-2785.

Sunaert S. (2001) Functional magnetic resonance imaging. Studies of visual motion processing in the human brain. Thèse de doctorat

Sunaert S., Van Hecke P., Marchal G. & Orban GA. (1999) Motion-responsive regions of the human brain. *Experimental Brain Research*, 127, 355-370.

Talairach J., Tournoux P. (1988) Co-planar Stereotaxic Atlas of the Human Brain. Thieme, New York.

Ternus J. (1926) Experimentelle Untersuchungen über phänomenale Identität. *Psychologische Forschung*, 7, 81 – 136.

Vaina L. M., Solomon J. Chowdhury S., Sinha P. & Belliveau J. W. (2001) Functional neuroanatomy of biological motion perception in humans. *PNAS*, 98, 20, 11656 – 11661.

Valentine T. (1988) Upside-down faces : A review of the effect of inversion upon face recognition. *British Journal of Psychology*, 79, 471 – 491.

Van Assche M. (2005) Effets de facteurs sémantiques sur des facteurs structuraux dans le paradigme de Ternus : images de corps humains. Mémoire de la Maîtrise de Psychologie Cognitive, Université Louis Pasteur, Strasbourg.

Watson J. D., Myers R., Frackowiak R. S. J., Hajnal J. V., Woods R. P., Mazziotta J. C., Shipp S. & Zeki S. (1993) Area V5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral cortex*, 3, 79 – 94.

Yin R. K. (1969) Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141 – 145.

Zeki S. (1993) A Vision of the Brain. Blackwell Scientific Publications.

Zihl J., von Cramon D., & Mai N. (1983) Selective disturbance of movement vision after bilateral brain damage. *Brain*, 106, 313 – 340.

Zihl J, von Cramon D, Mai N, Schmid C. (1991) Disturbance of movement vision after bilateral posterior brain damage. Further evidence and follow up observations. *Brain*, 114, 5, 2235 – 2252.

## **E. Annexes**

Paulos C., Scheiber C., Manning L., Bonnet C. (\*\*\*\*) Visual representations of human bodies and objects: an fMRI study. (submitted)

Bonnet C., Paulos C., Nithart C. (2005) Visual representations of dynamic actions from static pictures. *Perception*, 34, 835 – 846.

Bonnet C. & Paulos C. (\*\*\*\*) Human body actions in the Ternus paradigm. (submitted).

Bonnet C. & Paulos C. (\*\*\*\*) Structural and semantic factors in the Ternus paradigm. (submitted).

## **Visual representations of human bodies and objects: an fMRI study.**

Carlos Paulos<sup>1</sup>, Christian Scheiber<sup>1</sup>, Lilianne Manning<sup>2</sup> and Claude Bonnet<sup>1</sup>

<sup>1</sup> Institut de Physique Biologique (CNRS UMR 7004)

<sup>2</sup> Laboratoire de Neurosciences Comportementales et Cognitives (CNRS UMR 7521)

Université Louis Pasteur, Strasbourg, France.

Keywords: fMRI, inter-individual variability, motion representation, semantic category, ventral and dorsal streams

25 pages

3 Figures, 3 Tables

Correspondence to Professor C. Bonnet,

Faculté de Psychologie,

Université Louis Pasteur,

12 rue Goethe,

67000 Strasbourg (France).

Tel. +33 3 90 24 19 40

e-mail : [claud.bonnet@psycho-ulp.u-strasbg.fr](mailto:claud.bonnet@psycho-ulp.u-strasbg.fr)

#### Acknowledgments:

This work was partially supported by the Hôpitaux Universitaires de Strasbourg. The financial contribution of the Conseil Régional d'Alsace to the upgrading of the MRI device is gratefully acknowledged. We would like to thank the Centre d'Investigation Clinique des Hôpitaux Universitaires de Strasbourg, for the medical examination of healthy volunteers prior to their inclusion in the study. We would also like to thank Mrs Corinne Marrer for her hard-working contribution in the acquisition of fMRI data, and Ms Nathalie Heider for correcting our manuscript. Carlos Paulos is supported by a grant from the Ministère de la Culture, de l'Enseignement supérieur et de la Recherche, Grand Duchy of Luxembourg.

## **Abstract**

The aim of the present study was to investigate the posterior cortical organization with respect to the visual processing of various semantic categories in order to understand why areas defined as having the same function are localized with such great variability in the literature. We relied on the contrast between living and non-living categories to activate discrete areas in the ventral stream. A major characteristic of living items is their motion aptitude. By presenting different whole-body postures, activation foci were also expected in the motion complex MT / V5, as part of the dorsal stream, since the stimuli could easily evoke motion representations. The size of the effect and spatial extent of the clusters were compared for both hemispheres. After standard group analysis and localization of the ROIs, inter-individual maps evidenced great localization variability across subjects and suggested a more consistent view of brain organization across studies.



## **Introduction**

Neuroimaging studies have shown that stimuli/items/pictures belonging to different semantic categories are processed by different visual areas [19, 21, 39]. Several studies comparing living and non-living items have demonstrated activation foci in different separable areas within the ventral temporal lobes [6, 30]. All these studies evidenced that living items activate the more lateral portions of the ventral temporal lobes, while non-living items involve the more medial portions [2, 3, 6, 7, 19]. The dissociation between living and non-living categories has also been studied in clinical neuropsychology [5, 20], in particular defective performance regarding a living sub-category, i.e. familiar faces, (also known as prosopagnosia) following circumscribed lesions of the lateral portion of the fusiform gyrus [9, 42].

Facial processing was also analyzed in healthy subjects using neuroimaging techniques. Kanwisher et al. [24] provided an excellent illustration of the architecture of visual representations in the ventral temporal lobes. They put forward a modular approach suggesting the existence of the fusiform face area (FFA): an area located in the lateral fusiform gyrus, specialized in face processing. Using the same approach, another brain region located near the FFA was suggested to be activated selectively by human bodies and human body parts but not by faces [11] and was labeled the extrastriate body area (EBA). Within the same modular conception, Gauthier et al. [14, 15, 16] argued instead for a functionally broader “expert recognition” processing area, which would include the recognition of human faces among other stimuli belonging to different semantic categories. From this point of view, the involvement of the lateral fusiform gyrus (or FFA) in visual processing depends not so much on the living vs. non-living feature as on the subjects’ expertise level [15].

Without rejecting altogether the modular hypothesis, an alternative notion about / theory on brain organization was suggested: distributed representations within the ventral temporal lobes. Ishai et al. [21, 22], working with faces and man-made objects, and Haxby et al. [19], presenting faces, cats, and five categories of man-made objects, have shown several areas activated by different categories and compared their respective percent signal change. These authors have reported that besides one brain region which responds maximally to a specific semantic category, e as the lateral fusiform gyrus does

to faces, for instance, several other cerebral areas also contribute, to a lesser degree, to the visual processing of the semantic category under consideration.

From a perceptual point of view, the main characteristic of a living stimulus is motion ability whether actual or implicit. As far as visual motion processing as such is concerned, neurophysiological studies in monkeys have identified a population of neurons, in the medial temporal regions (MT / V5), particularly sensitive to visual motion patterns [28]. Further to these findings the existence of a similar region within the dorsal stream in humans was hypothesized [33]. Anatomico-clinical studies are most enlightening. Indeed, Zihl et al. [46] reported the well-known case of a patient (LM) in whom focal bilateral lesions of the medial temporal region (MT / V5), provoked a selective motion blindness. The implication of area MT / V5 in visual motion processing in healthy subjects has been confirmed in many neuroimaging studies [4, 33, 37, 44, 45]. These studies also demonstrated that, surrounding the motion area MT / V5, several other areas participated/contributed/were involved in visual motion processing. Not surprisingly, it has been demonstrated that the medial temporal area (MT), together with the superior temporal sulcus (STS), are also activated by the visual processing of biological motion [17, 18, 43]. Focusing on the medial temporal region in the absence of actual motion, Kourtzi and Kanwisher [25] reported activation elicited by static pictures of athletes performing an exercise as opposed to pictures of people at rest. The authors described activation foci in the medial temporal area (MT) and in the neighboring medial superior temporal (MST) area, which were thought to reflect the process of implied motion, according to the hypothesis that a mental representation of a motion and / or action is formed while viewing “frozen-action” pictures [13, 25, 34, 35]. It is quite interesting to note that the localization of area MT / V5 and the previously mentioned EBA [11], activated by human bodies and human body parts, partially overlap.

Many studies have concentrated on the organization and on the degree of selectivity of differential and separable areas within the ventral temporal lobes activated by living and non-living items [6, 11, 19, 21, 22, 24]. Few studies have focused only on the possible implication of the MT complex, considered as part of the dorsal stream, in which activation is elicited by frozen-action pictures representing items of the living category [25, 35]. The first aim of the present study was to relate these two sets of data.

To this end, we used a box-car paradigm to present stimuli depicting still line drawings of living (human bodies) and non-living (man-made objects) items. . Whole human body postures were considered as good candidates to evoke motion representations Greater attention was paid either to lateral or medial activation foci in the ventral temporal lobes, in particular in the fusiform gyrus, depending on whether they were elicited by representations of human bodies, or by man-made objects, respectively. Human body representations were also expected to produce activation in the MT / V5 complex, as the different whole body postures presented could easily induce motion representation (see stimuli set). Our second purpose was to ascertain whether or not there are differences in terms of degree of selectivity and activation amplitude between homologous regions in the two hemispheres, for the two categories. Thirdly, the hypothesis of areas selectively activated by different semantic categories of pictures should be supported by consistent localization between studies. However, areas defined as having the same function have been localized with great variability in different studies. Assuming that such variability was to a great extent due to sampling fluctuations, we intended to compare our results with others, using a two-step analysis. After standard group analysis, volumetric analyses on the individual maps were processed. Based on the group-defined ROIs, cubic volumes of interest (VOIs) were reconstructed taking into account the inter-individual localization variability and the stimulus—category-specific. These VOIs being larger than the ROIs, we expected them to encompass to a degree the localization variability observed in the literature for similar kinds of stimuli.

## **Material and methods**

### *Subjects and task*

Ten right-handed subjects (5 females and 5 males, mean age 25; range 23 – 28) were included in the study, after having given prior informed consent. All the subjects had normal or corrected-to-normal vision (contact lenses), and reported no history of neurological or psychiatric disorder. The research protocol was carried out in compliance with the local ethical regulations, and approved by the French Ethical Committee for the Protection of People taking part in Biomedical Research (CCPPRB Alsace n° 1). The subjects were requested to maintain their gaze on the center of the visual field (passive viewing) while they were shown series of pictures.

### *Stimuli set*

Line drawings of human bodies (contour silhouettes of human bodies; authors' stimuli set) and of man-made objects (re-drawn from Snodgrass and Vanderwart's stimuli set [36]) were used (Figure 1). A box-car paradigm was used to present the stimuli belonging to two different semantic categories. The subjects were informed about the presence of the two categories. Each experimental block contained 10 different drawings belonging to the same category, each one being presented, randomly, twice in the same block. Two different sets of stimuli for each category were used in different blocks.

Figure 1 about here

Pictures were shown binocularly as white line drawings on a black background through a Silent Vision fMRI visual system (Avotec Inc., Florida, USA). The approximate size of each drawing was  $10^{\circ} \times 6^{\circ}$  and the luminance of the white lines was about  $10 \text{ cd/m}^2$ . The visualization system was computer-driven. The appearance of the stimuli was triggered by the fMRI scans.

Each picture was shown for 167 ms, and followed by a blank interval of about one second. The entire block sequence lasted for 25 seconds and was followed by the presentation of a fixation cross (rest period) for the same duration. Blocks of the two different semantic categories alternated. Each set was repeated 5 times for a single session.

### *fMRI method and data analysis*

Local variations in blood oxygenation level-dependent (BOLD) responses were measured on a 2 Tesla S200 (Bruker, Karlsruhe, Germany) MRI system using susceptibility-based functional magnetic resonance imaging (fMRI), applying gradient recall echo-planar imaging (EPI) sequences. Thirty-two parallel 4 mm thick planes of the whole brain were imaged using the following acquisition parameters:  $64 \times 64$  (voxel size  $4 \times 4 \times 4 \text{ mm}^3$ ),  $RT = 5000 \text{ ms}$ ,  $TE = 43 \text{ ms}$ ,  $FA = 90^{\circ}$ . Anatomical images required for the localization of functional responses were obtained using a turbo-RARE  $T_2$ -weighted

sequence using the following parameters: 128 x 128 x 64 (voxel size 4 x 4 x 4 mm<sup>3</sup>; RT = 15000 ms, TE = 73.8 ms, Turbofactor = 10).

## Results

The fMRI data were processed with the software *Statistical Parametric Mapping* (SPM 99, Wellcome Department of Cognitive Neurology, London). The fMRI time series was pre-processed in three steps: realignment (all the images were realigned on the first reference image from the scan), normalization (non-linear transformation of the brain into a reference template in Talairach's anatomical space) and smoothing (Gaussian filter FWHM: 4 mm).

SPM maps were then obtained using a box-car model for the two conditions, human bodies and man-made objects (*ON* period), contrasted with the fixation cross (*OFF* period) through a single session. In order to find out regions of interest (ROIs) selectively activated by the images of one semantic category, the following procedure was used. A group analysis using multiple comparisons was performed to obtain activation maps at  $p < 0.001$  (uncorrected) for each condition, and then submitted to an exclusive masking procedure at  $p < 0.05$  (uncorrected) with an extent threshold  $k \geq 30$  voxels: human bodies masked by objects and objects masked by human bodies. Both contrasts show the regions exclusively activated by each condition: human bodies displayed in red color and objects displayed in green (Figure 2). Three-dimensional projections of these functional maps of group data are presented in Figure 2, exhibiting a sagittal and coronal view. There are clearly regions selectively activated by either one of the two categories.

Figure 2 about here

This global analysis highlighted several separate regions, almost all located in the posterior part of the brain and responding exclusively to one of the two semantic categories. Some activation foci appearing in the anterior part of the brain ( $y > 0$ ) for the object contrast were disregarded for subsequent analyses as they appeared inconsistent across subjects. To begin with, two bilateral clusters were observed (see Figure 2 and Table 1) in the lower part of the occipito-temporal region. The Talairach coordinates of the clusters are given in Table 1, together with the spatial extent of the clusters ( $K_E$ ) and

the size of the effect ( $E$ ). The more lateral cluster, identified as the Lateral Fusiform Gyrus (LFG) responds exclusively to human bodies. The more medial one, identified as the Medial Fusiform Gyrus (MFG) responds exclusively to objects. A third pair of bilateral clusters can further be seen above the LFG, all of which respond exclusively to human bodies. Its localization corresponds to area MT / V5 (see Table 3). A fourth bilateral cluster can be seen in the dorsal bank, at the border between BA19 and the Superior Parietal Lobule (SPL), which responds exclusively to objects. These four (pairs of) clusters are our main ROIs. One can see in Table 1 that the clusters of voxels activated by human bodies on the one hand, and objects on the other, are larger in the right and in the left hemisphere, respectively.

Two unilateral clusters with  $k > 30$  are left, which respond exclusively to objects. One is located in the lower part of the left occipito-temporal cortex; its coordinates correspond to the superior border of the fusiform gyrus (SFG). The last cluster is situated in the right parietal cortex; its coordinates correspond to the inferior parietal lobule (IPL).

Table 1 about here

Given that each subject may contribute differently to the localization of the ROIs, to the cluster extent and to the size of the effect, further analysis was run on individual data. Using the coordinates of these ROIs as references, we studied the distributions of individual SPM maps, after exclusive masking in these regions, in order to determine inter-individual Volumes of Interest (VOIs). For subsequent analysis on individual data, only the local maxima having a height threshold of  $t \geq 3.13$  at  $p < 0.001$  (uncorrected) with an extent threshold  $k \geq 30$  voxels were considered. The larger VOIs thus defined remained category selective, with one exception at the border of the fusiform gyrus in the left hemisphere (see below). Activation foci in the anterior part of the brain ( $y > 0$ ) were again disregarded as our analyses focused on the posterior part of the brain. In the posterior cortex data falling below  $z < -30$  were considered as located in the cerebellum and were also disregarded.

Individual coordinates reaching the local maxima and surrounding each ROI defined from the group data were retained to reconstruct a cubic volume for each region (one local maximum per subject for a given region). The results of this analysis for the four bilateral VOIs are shown in Table 2. For each hemisphere, the mean coordinates of

each volume (with half the range of variation to the extreme values) and the number of subjects showing a significant response within the volume are provided.

Areas activated in response to human bodies exclusively were found in the LFG and MT, while activation foci in response to objects exclusively were found in the MFG and in BA19-SPL. However, 3 subjects also showed significant responses to objects in the VOI defined for human bodies (LFG) in the left hemisphere. These responses account for the above-mentioned cluster, in the group data, whose coordinates were said to correspond to the SFG (see Figure 2b and Table 1). The right hemisphere seems to be more selective than the left hemisphere.

Table 2 about here

Once individual data were thus sorted out, the percent signal change (PSC) of the BOLD responses for each category in each location was obtained from SPM99. Histograms of these results are shown in Figure 3. The means were computed from all the subjects ( $n = 10$ ). The picture would have been very similar if only the subjects showing significant local maxima had been used. The subjects who did not have significant local maxima were included in the analysis with value '0'. The distributions of mean individual data were submitted to an analysis of variance ( $n = 10$ ).

Figure 3 about here

Globally, the percent signal change does not vary significantly as a function of the regions ( $F(3,27) = 1.03$ ;  $p = \text{n.s.}$ ), hemispheres ( $F(1,9) = 1.08$ ;  $p = \text{n.s.}$ ) or categories ( $F < 1$ ). However, a significant interaction is shown between regions and categories ( $F(3,27) = 24.13$ ;  $p < 0.0001$ ). In fact, as expected, the percent signal change observed for human bodies exceeds the percent signal change observed for objects in the LFG and in MT, while the reverse phenomenon appears in the MFG and in BA19/SPL. The mean percent signal change observed for human bodies is larger in the right hemisphere, in the LFG (LH: 0.99; RH: 1.32) and MT (LH: 1.05; RH: 1.3) regions. The percent signal change observed for objects exhibits a similar trend, but only in the BA19/SPL region (LH: 0.48; RH: 1.04).

A similar procedure was applied for the last ROI observed in the group analysis, previously referred to as the inferior parietal lobule (IPL). This region appeared on group data only in the right hemisphere. Since a number of the local maxima remained visible on the individual data in this dorsal region, they were analyzed. On the basis of the

individual data for this region, two VOIs appeared, situated in each hemisphere, corresponding to activation foci arising from two subjects only. The mean Talairach coordinates for this region were, in the left hemisphere: (-47(11), 47(11), 44 (14)) and in the right hemisphere (40(10), -56(6), 48(8)). This region is essentially located bilaterally in the Inferior Parietal Lobule (BA 40). However, the mean percent signal change observed for human bodies and for objects are statistically equivalent (LH: PSC(H) = 0.32; PSC(O) = 0.49 and RH: PSC(H) = 0.59; PSC(O) = 0.54). In the group analysis, no voxel survived the selection criteria in the left hemisphere (height threshold of  $t \geq 3.09$  with  $p < .001$  and an extent threshold  $k \geq 30$  voxels).

## Discussion

A large portion of the primary and secondary visual cortices is approximately equally activated by both semantic categories. Presumably, these activation foci, mainly observed in the most occipital part of the posterior cortex, correspond to the early structural processing of the stimulus features [32]. These commonly activated areas were no longer visible on the contrasts (Figure 2), showing only the areas selectively activated by each semantic category. Owing to the exclusive masking procedure used here, the present results could not exhibit much overlapping activity, as described by some authors [19, 21, 22].

Before discussing the selective role of each region in the processing of our two stimulus categories, let us concentrate on the definition of each of these regions. Each selective region was first identified and labeled from the group data (ROIs) by using Talairach and Tournoux coordinates [38]. We subsequently compared our data with a large set of references given in the literature (cf. Table 3), whenever available. These comparisons evidenced more or less important divergences regarding the localization of the ROIs, even in cases where rather similar stimuli were used. However, if for each specific region we consider the coordinates of the VOIs determined from the individual data, the ROIs are in every cases (in the present study as well as in the literature) included within the VOIs. Besides, the selectivity of each region with respect to the semantic category remains significant. This holds true for the selectivity of the LFG vs. the MFG [2, 3, 6, 7, 21, 22], as well as for MT / V5 (see below). Of course, the size of the VOIs is much larger than that of the ROIs. This result suggests that the variability of coordinates



reported by different authors in different studies could reflect, at least partially, sampling fluctuations and provides a more consistent view of brain organization. Another source of variability arises from the different tasks used in different paradigms. In the present study, line drawings and passive viewing were used: two conditions known to reduce the amplitude of the percent signal change [21]. Moreover, passive viewing is likely to increase inter-individual variability of the data owing to differences in subject performance during the trials. It would be interesting to compare the variability of activation in a study involving the same stimuli set and an explicit task such as, for example, a categorization task.

Table 3 about here

Line drawings of human bodies and man-made objects activate separable areas in both the ventral and the dorsal streams [29, 41]. Whether on the basis of ROI analysis or VOI analysis, two category-selective regions have been identified in the ventral stream: the lateral fusiform gyrus (LFG) and the medial fusiform gyrus (MFG). The medial fusiform gyrus (MFG) was shown to be involved in processing of non-living items or man-made objects (including tools and houses) [2, 6, 22]. The lateral fusiform gyrus (LFG) was suggested to be selectively engaged in face recognition [6, 21, 22, 24]. But the LFG appears to be also activated by animals [6, 15] and human body images [2, 3 and the present experiment]. Taking into account Gauthier's work [14, 15, 16] on the expert recognition of 'greebles' or birds, it may be suggested that this region is involved in the recognition of living object categories on which subjects are experts. 'Greebles' could be considered as having 'pseudo-living' characteristics. However this interpretation should be qualified: indeed, the LFG is also activated by images of cars when subjects are car-experts [15]. Further work is (obviously) called for to clarify this dilemma about the specialization of the LFG: living object categories or expert recognition. As it happens, in our study the human bodies have both characteristics: they pertain to the living category and we are all very familiar with representations of human bodies [10, 23].

Category-selective activity is not limited to the ventral temporal lobes. Indeed, a second region shows a strong selectivity regarding human bodies that we identified as area MT / V5, known to be involved in visual motion processing and usually attributed to the dorsal stream. Our coordinates included the MT / V5 coordinates that have been obtained in other studies involving different stimuli, such as real visual motion patterns

[12, 26, 27, 31, 37], biological motion [1, 18] and ‘motion – action’ representations evoked from static images [25, 34, 35]. Our VOI coordinates for MT / V5 encompassed also the extrastriate body area (EBA) as described by Downing et al. [11]. These authors suggested that the EBA was exclusively activated by human bodies and body parts. Interestingly, the localization of this region partially overlaps the posterior superior temporal sulcus (pSTS) described by Beauchamp et al. [2, 3] who reported activation for human motion but not for tool motion. All these results confirm at least that the MT complex is activated by human body motion, whether real or implied.

A further two regions were identified in the dorsal stream. Although, these activation foci were less consistent across subjects (cf. Table 2). Nevertheless, a selective region responding to objects was found bilaterally at the border between the precuneus (BA19) and the superior parietal lobule (SPL). This region is slightly posterior to the unilateral region reported for tools by Chao and Martin [7], (see Table 3) and is in agreement with the coordinates of a parietal region, which responds to moving tools [2]. Another region activated in our sample was located within the inferior parietal lobule. This region presented a consistent selectivity regarding objects in the right hemisphere, but was equally activated by both categories in the left hemisphere. These parietal activation foci may reflect evoked representations of action, spatial knowledge/awareness and object use [40]. This region also included the coordinates of a parietal region described by Culham et al. [8] as being involved in object-grasping behavior.

After the selective regions in the posterior cortex were identified, the activation foci observed in the VOIs were compared for both hemispheres. Remarkably, the selectivity of the two areas in the ventral stream responding to each category is well delimited in the right hemisphere. The selectivity with respect to each semantic category is less definite in the left hemisphere (LFG), although still significant. Interestingly, the size of the effect for both regions (see Figure 3) is larger in the right hemisphere. The size of the effect in area MT / V5 appears also larger in the right hemisphere. Regarding the spatial extent of clusters in area MT / V5, more voxels were also activated in the right hemisphere. Conversely, the object-selective superior parietal lobule (SPL) is more responsive in the right hemisphere but has more voxels activated in the left hemisphere. Globally, human bodies activate more voxels in the right hemisphere while man-made

objects activate more voxels in the left hemisphere. This difference may be explained by the larger pictorial variability of the stimuli set used for the objects and their probable denomination.

In conclusion, the present results confirm that different areas within the ventral temporal lobes are involved in the processing of different semantic categories of pictures. Not only, the categorization of the picture elements as living or non-living items has been observed, but it is confirmed that static pictures belonging to the living category (namely human body images) also generate activation in an area (MT / V5) known to be involved in the processing of complex visual motion. This result is consistent with the general hypothesis that mental representations are governed by brain structures also involved in perception [7, 13, 23, 25, 34, 35, 40].

## References

- [1] E. Bonda, M. Petrides, D. Ostry, A. Evans, Specific involvement of human-parietal systems, and the amygdala, in the perception of biological motion. *The Journal of Neuroscience* 16 (1996) 3737-3744.
- [2] M. Beauchamp, K. Lee, J. Haxby and A. Martin, Parallel visual motion processing streams for manipulable objects and human movements. *Neuron* 34 (2002) 149-159.
- [3] M. Beauchamp, K. Lee, J. Haxby and A. Martin, fMRI Responses to video and point-light displays of moving humans and manipulable objects. *J. Cogn. Neurosci.* 15 (2003) 991-1001.
- [4] O. J. Braddick, J. M. O'Brien, J. Wattam-Bell, J. Atkinson, T. Hartley and R. Turner, Brain areas sensitive to coherent visual motion. *Perception* 30 (2001) 61-72.
- [5] A. Caramazza and J. R. Shelton, Domain specific knowledge systems in the brain: the animate-inanimate distinction. *J. Cogn. Neurosci.* 10 (1998) 1-34.
- [6] L. Chao, J. Haxby and A. Martin, Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neurosci.* 2 (1999) 913-919.
- [7] L. Chao and A. Martin, Representations of manipulable man-made objects in the dorsal stream. *Neuroimage* 12 (2000) 478-484.
- [8] J.C. Culham, S.L. Dankaert, J.F.X. De Souza, J.S. Gati, R.S. Menon, M.A. Goodale, Visually-guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Exp. Brain Res.* 153 (2003) 180-189.
- [9] A.R. Damasio, D. Tranel, H. Damasio, Face agnosia and the neural substrates of memory. *Annu. Rev. Neurosci.* 13 (1990) 89-109.

- [10] J. Decety, T. Chaminade, J. Grèzes and A. N. Meltzoff, A PET exploration of the neural mechanisms involved in reciprocal imitation. *Neuroimage* 15 (2002) 265-272.
- [11] P. E. Downing, Y. Jiang, M. Shuman and N. Kanwisher, A cortical area selective for visual processing of the human body. *Science* 293 (2001) 2470-2473.
- [12] S.P. Dukelow, J. F. X. De Souza, J.C. Culham, A. V. van den Berg, R. S. Menon and T. Vilis, Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *J. Neurophysiol.* 86 (2001) 1991-2000.
- [13] J. J. Freyd The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics* 33 (1983) 575-581.
- [14] I. Gauthier, A.W. Anderson, M.J. Tarr, P. Skudlarski and J.C. Gore, Levels of categorization in visual recognition studied with functional MRI. *Current Biology* 7 (1997) 645-651.
- [15] I. Gauthier, P. Skudlarski, J. C. Gore and A.W. Anderson, Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience* 3 (2000) 191-197.
- [16] I. Gauthier, M.J. Tarr, Unraveling mechanisms for expert object recognition: bridging brain activity and behaviour. *Journal of Experimental Psychology: Human Perception and Performance* 28 (2002) 431-446.
- [17] J. Grèzes, P. Fonlupt, B. Bertenthal, C. Delon-Martin, C. Segebarth, J. Decety, Does perception of biological motion rely on specific brain regions? *Neuroimage* 13 (2001) 775-785.
- [18] E. D. Grossman and R. Blake, Brain areas active during visual perception of biological motion. *Neuron* 35 (2002) 1167-1175.

- [19] J.V. Haxby, M. I. Gobbini, M. L. Furey, A. Ishai, J. L. Schouten and P. Pietrini Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293 (2001) 2425-2430.
- [20] G. W. Humphreys, E.M.E. Forde, Hierarchies, similarity and interactivity in object recognition: On the multiplicity of 'category-specific' deficits in neuropsychological populations. *Behavioral and Brain Sciences* 24 (2000) 453-509.
- [21] A. Ishai, L. G. Ungerleider, A. Martin, J. L. Schouten and J. V. Haxby, Distributed representation of objects in the human ventral visual pathway. *Proc. Natl. Acad. Sci.* 96 (1999) 9379-9384.
- [22] A. Ishai, L. G. Ungerleider, A. Martin and J. V. Haxby, The representation of objects in the human occipital and temporal cortex. *J. Cogn. Neurosci.* 12 (2000) 35-51.
- [23] M. Jeannerod, Neural mechanisms underlying representations for action. In: Grafman J., Boller, F. (Eds.), *Handbook of Neuropsychology* Elsevier Amsterdam vol. 11 167-183, 1997.
- [24] N. Kanwisher, J. McDermott, M. M. Chun, The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17 (1997) 4302-4311.
- [25] Z. Kourtzi and N. Kanwisher, Activation in human MT/MST by static images with implied motion. *J. Cogn. Neuroscience* 12 (2000) 48-55.
- [26] Z. Kourtzi, H. H. Bühlhoff, M. Erb and W. Grodd, Object-selective responses in the human motion area MT/MST. *Nature Neuroscience* 5 (2002) 17-18.

- [27] N. Kriegeskorte, B. Sorger, M. Naumer, J. Schwarzbach, E. van den Boogert, W. Hussy, and R. Goebel, Human cortical object recognition from a visual motion flow field. *J. Neurosci.* 23 (2003) 1451-1463.
- [28] V.L. Macar, J. Zihl, A. Cowey, Comparing the visual deficits of a motion blind patient with the visual deficits of monkeys with area MT removed. *Neuropsychologia* 35 (1997) 1459-1465.
- [29] D. Milner and A. Goodale (Eds.), *The Visual Brain in Action*. Oxford Psychology Series, 1995.
- [30] C.J. Moore and C.J. Price, A functional neuroimaging study of the variables that generates category-specific object processing differences. *Brain* 122 (1999) 943-962.
- [31] A. L. Paradis, V. Cornilleau-Pérès, J. Droulez, P. F. Van de Moortele, E. Lobel, A. Berthoz, D. Le Bihan and J. B. Poline, Visual perception of motion and 3-D structure from motion: an fMRI study. *Cereb. Cortex* 10 (2000) 772-783.
- [32] C. Pernet, S. Basan, B. Doyon, D. Cardebat, J.-F. Démonet, P. Celsis, Neural timing of visual implicit categorization. *Cogn. Brain Res.* 17 (2003) 327-338.
- [33] G. Rees, K. Friston and C. Koch, A direct quantitative relationship between the functional properties of human and macaque V5. *Nature Neurosci.* 3 (2000) 716-723.
- [34] C. Senior, J. Barnes, V. Giampietro, M. Brammer, E. Bullmore, A. Simmons and A.S. David, The functional neuroanatomy of implicit motion perception or 'representational momentum'. *Current Biology* 10 (2000) 16-22.
- [35] C. Senior, J. Ward and A. S. David, Representational momentum and the brain: An investigation into the functional necessity of V5/MT. *Visual Cognition* 9 (2002) 81-92.

- [36] J. G. Snodgrass and M. Vanderwart, A Standardized Set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity and Visual Complexity. *J. Exp. Psychology: Human Learning and Memory* 6 (1980) 174-215.
- [37] S. Sunaert, P. Van Hecke, G. Marchal and G. A. Orban, Motion-responsive regions of the human brain. *Exp. Brain Research* 127 (1999) 355-370.
- [38] J. Talairach, P. Tournoux (Eds.), *Co-planar Stereotaxic Atlas of the Human Brain*. Thieme, New York, 1988.
- [39] L.K. Tyler, H.E. Moss, Towards a distributed account of conceptual knowledge. *Trends Cogn. Sci.* 5 (2001) 244-252.
- [40] L.K. Tyler, E.A. Stamatakis, E. Dick, P. Bright, P. Fletcher and H. Moss, Objects and their actions: evidence for a neurally distributed semantic system. *Neuroimage* 18 (2003) 542-557.
- [41] L.G. Ungerleider, M. Mishkin, Two cortical pathways. In D.D. Ingle, M.A. Goodale, R.J.M. Mansfield (Eds.) *Analysis of visual behavior*. Cambridge MIT Press, pp 549-586, 1982.
- [42] I. Uttner, H. Bliem and A. Danek, Prosopagnosia after unilateral right cerebral infarction. *J. Neurology* 249 (2002) 933-935.
- [43] L. M. Vaina, Solomon J., Chowdhury S., Sinha P., Belliveau J. W., Functional neuroanatomy of biological motion perception in humans. *Proc. Natl. Acad. Sci.* 98 (2001) 11656-11661.
- [44] S. Zeki, J. D. Watson, C. J. Lueck, K. J. Friston, C. Kennard and R. S. Frackowiak, A direct demonstration of functional specialization in human visual cortex. *J. Neurosc.*, 11 (1991) 641-649.



[45] S. Zeki (Eds.), *A Vision of the Brain*. Blackwell Scientific Publications, 1993.

[46] J. Zihl, D. von Cramon, N. Mai, Selective disturbance of movement vision after bilateral brain damage. *Brain* 106 (1983) 2313-2340.

### **Figures captions**

**Figure 1** One example of each subset of the two semantic categories presented: human bodies and man-made objects.

**Figure 2** Statistical maps of the group data ( $n = 10$ ) displayed on a 3D structural representation of the brain (sagittal and coronal view). In red, areas exclusively activated by human bodies; in green, areas exclusively activated by objects. The maps were obtained after applying an exclusive masking procedure (extent threshold  $k \geq 30$  voxels,  $p < 0.001$ ). The areas activated: 1 – Lateral fusiform gyrus (LFG), 2 – medial fusiform gyrus (MFG), 3 – motion (processing) area MT / V5, 4 – superior parietal lobule (SPL), 5 – inferior parietal lobule (IPL), 6 – superior fusiform gyrus (SFG).

**Figure 3** Mean percent signal change (PSC) of the BOLD response with the inter-individual standard errors for each stimulus category, each hemisphere and each VOI.

**Table 1****Talairach coordinates of the ROIs defined from the group data (n = 10)**

| Region    | Left Hemisphere |     |     |            |             | Right Hemisphere |     |     |            |             |
|-----------|-----------------|-----|-----|------------|-------------|------------------|-----|-----|------------|-------------|
|           | X               | Y   | Z   | $K_E$      | $E$         | X                | Y   | Z   | $K_E$      | $E$         |
| LFG (H)   | -42             | -48 | -22 | <b>49</b>  | <b>2.9</b>  | 42               | -54 | -18 | <b>92</b>  | <b>4.3</b>  |
| MFG(O)    | -24             | -48 | -24 | <b>318</b> | <b>-4.9</b> | 26               | -36 | -24 | <b>216</b> | <b>-4.3</b> |
| MT (H)    | -52             | -72 | 2   | <b>215</b> | <b>4.0</b>  | 52               | -70 | 6   | <b>628</b> | <b>4.2</b>  |
| 19 SPL(O) | -22             | -78 | 34  | <b>92</b>  | <b>-2.3</b> | 34               | -72 | 46  | <b>59</b>  | <b>-3.0</b> |
| SFG(O)    | -48             | -60 | -14 | <b>53</b>  | <b>-2.7</b> |                  |     |     |            |             |
| IPL(O)    |                 |     |     |            |             | 48               | -52 | 52  | <b>34</b>  | <b>-2.7</b> |

Extent cluster size and mean differences of the size of the effect for human bodies

(H) – objects (O) are shown in columns  $K_E$  and  $E$ .

**Table 2**  
**Central coordinates of the 4 VOIs in both hemispheres**  
**Means of individual data**

| Region   | Left Hemisphere |             |             |      |      | Right Hemisphere |             |        |      |      |
|----------|-----------------|-------------|-------------|------|------|------------------|-------------|--------|------|------|
|          | X               | Y           | Z           | N(M) | N(O) | X                | Y           | Z      | N(M) | N(O) |
| LFG (H)  | -45(5)          | -59(2<br>3) | -21(13<br>) | 7    | 3    | 47(7)            | -61(15<br>) | -17(5) | 7    | 0    |
| MFG (O)  | -30(8)          | -49(1<br>9) | -18(6)      | 0    | 8    | 30(10<br>)       | -61(29<br>) | -21(9) | 0    | 7    |
| MT (H)   | -47(5)          | -74(1<br>0) | 4(12)       | 5    | 0    | 42(12<br>)       | -68(14<br>) | 10(12) | 8    | 0    |
| 19SPL(O) | -26(4)          | -74(6)      | 42(12)      | 0    | 3    | 33(13<br>)       | -78(6)      | 41(9)  | 0    | 5    |

The numbers in parentheses for each hemisphere in the first columns correspond to the half-range of variation. The number of subjects showing a significant response within the volume for each stimulus category is indicated in columns N (H for human bodies) and N (O for objects).

**Table 3****Coordinates of the ROIs from the literature used for comparison purposes**

| <b>Chao et al. (1999), Chao Martin (2000)</b> | x   | y   | z   |                                                                    |
|-----------------------------------------------|-----|-----|-----|--------------------------------------------------------------------|
| LFG                                           | 38  | -56 | -12 | Animals (passive viewing)                                          |
|                                               | -40 | -59 | -10 |                                                                    |
| MFG                                           | 26  | -48 | -9  | Tools (passive viewing)                                            |
|                                               | -26 | -47 | -5  |                                                                    |
| <b>Ishai et al. (1999,2000)</b>               |     |     |     |                                                                    |
| LFG                                           | 40  | -52 | -19 | Faces (p.v. and delayed matching)                                  |
|                                               | -36 | -55 | -20 |                                                                    |
| MFG                                           | 28  | -57 | -13 | Houses (p.v. and d.m.)                                             |
|                                               | -26 | -57 | -14 |                                                                    |
|                                               | 30  | -49 | -16 | Chairs (p.v. and d.m.)                                             |
|                                               | -27 | -51 | -19 |                                                                    |
| LFG                                           | 41  | -60 | -20 | Faces (delayed matching)                                           |
|                                               | -39 | -64 | -28 |                                                                    |
| MFG                                           | 28  | -63 | -18 | Houses (d.m.)                                                      |
|                                               | -25 | -58 | -16 |                                                                    |
|                                               | 30  | -61 | -19 | Chairs (d.m.)                                                      |
|                                               | -29 | -56 | -22 |                                                                    |
| <b>Beauchamp et al. (2002)</b>                |     |     |     |                                                                    |
| LFG                                           | 39  | -46 | -11 | Humans (forced choice categorization task: Moving human vs. tools) |
| MFG                                           |     |     |     | Tools (idem)                                                       |
|                                               | -28 | -66 | -11 |                                                                    |
| <b>Beauchamp et al. (2003)</b>                |     |     |     |                                                                    |
| LFG                                           | 44  | -42 | -15 | Humans (idem)                                                      |
|                                               | -36 | -42 | -11 |                                                                    |
| MFG                                           | 28  | -42 | -24 | Tools (idem)                                                       |
|                                               | -21 | -40 | -16 |                                                                    |
| <b>Paradis et al. (2000)</b>                  |     |     |     |                                                                    |
| V5+ (MT)                                      | 45  | -54 | 6   | Random motion (passive viewing)                                    |
|                                               | -51 | -72 | 3   |                                                                    |

|                            |       |       |      |                                                            |
|----------------------------|-------|-------|------|------------------------------------------------------------|
| Dukelow et al. (2001)      |       |       |      |                                                            |
| MT                         | 44    | -64   | 5    | Optic flow stimuli (visual pursuit task)                   |
| Kourtzi et al. (2002)      |       |       |      |                                                            |
| MT/MST                     | 43.3  | -71.2 | 5.7  | Moving Objects (intact vs. scrambled)                      |
|                            | -41.3 | -74.4 | 4.7  |                                                            |
| Kriegeskorte et al. (2003) |       |       |      |                                                            |
| hMT+                       | 46    | -62   | 0    | Moving dots (structure from motion categorization)         |
|                            | -43   | -68   | -3   |                                                            |
| Sunaert et al. (1999)      |       |       |      |                                                            |
| MT                         | 42    | -62   | 6    | Visual motion                                              |
|                            | -42   | -66   | 2    |                                                            |
| Bonda et al. (1996)        |       |       |      |                                                            |
| MT                         | 48    | -69   | 8    | Random motion – random static                              |
|                            | -52   | -64   | 7.5  |                                                            |
| Senior et al. (2002)       |       |       |      |                                                            |
| MT                         | 48    | -67   | 12   | Implied motion static picture                              |
|                            | -42   | -77   | 6    |                                                            |
| Downing et al. (2001)      |       |       |      |                                                            |
| EBA                        | 51    | -71   | 1    | Human body parts and bodies (1-back repetition task)       |
|                            | -51   | -72   | 8    |                                                            |
| Grossman & Blake (2002)    |       |       |      |                                                            |
| EBA                        | 40.6  | -65.7 | 10.6 | Stationary pictures of headless bodies – household objects |
|                            | -39.3 | -70.1 | 13.5 |                                                            |
| Beauchamp et al. (2002)    |       |       |      |                                                            |
| SPL                        | 22    | -50   | 42   | Human and tools motions                                    |
|                            | -22   | -57   | 44   |                                                            |
| Culham et al. (2003)       |       |       |      |                                                            |
| AIP (IPL)                  | 40    | -50   | 50   | Grasping - reaching objects                                |
|                            | -38   | -48   | 52   |                                                            |

## **Structural and semantic factors in the Ternus paradigm**

**Claude Bonnet & Carlos Paulos**

### **Abstract**

Three experiments are reported in order to investigate the relative role of structural and semantic factors in the report of “group” responses in the Ternus visual apparent movement. In the first experiment, the picture elements of each image were either a running man body or a chair. Within a session the picture elements were either homogeneous or heterogeneous. In homogeneous conditions, the frequency of “group” responses increases in a steeper way with the ISI than in heterogeneous conditions. The second experiment concentrates on human body images. Changing the orientation of the picture in the last position reduces the slope of the function relating the frequency of “group” responses to ISI. In the third experiment, a similar design was used with chair pictures and raises similar results. It is concluded that under the present conditions, the effects of structural factors of similarity and orientation prevailed on the semantic factors. However, there are some suggestions of the influence of human body representations.

### **Introduction**

Apparent or stroboscopic visual motion has been largely investigated since Wertheimer’s seminal paper (see Sekuler, 1996). Among the apparent motion phenomena, the Ternus paradigm has received some attention. The Ternus paradigm (Ternus, 1926; Kolers, 1972) is an apparent motion phenomenon in which two frames (or images) alternate (separated by a blank interval). Each frame shows three equally spaced pictures arranged so that two pictures occupy the same position in both frames (see Figure 1). These two alternating frames are separated by a blank interstimulus interval (ISI). Two different percepts are seen depending on the duration of this ISI

(e.g. Kramer & Yantis, 1997). When the ISI is short (or null), subjects only report the motion of the extreme element ( $A \rightarrow D$ , call *element* response), while for longer ISI, they report more and more the motion of a group of three pictures ( $ABC \rightarrow BCD$ , call *group* response). Increasing the duration of each frame increases the probability of a *group* report (e.g. Dawson & Wright, 1989). The more spaced are the pictures within and between frames, the lower is the probability of *group* responses (e.g. Burt & Sperling, 1981; Dawson & Di Lollo, 1990; Kolers, 1972).

Most research on the Ternus phenomenon has concerned the effects of low-level factors. Different explanations have been proposed for the switch between *element* and *group* responses, while no definitive consensus appears. Among those explanations, some authors (e.g. Breitmeyer & Ritter, 1986 a, b) argue that *group* motion is prevented in conditions in which the first image persists while the second image is shown. However, persistence does appear neither necessary nor sufficient to explain the phenomenon (e.g. Kramer & Rudd, 1999). More frequently, the switch between *element* to *group* responses is suggested to reflect the switch between two motion processes proposed by Braddick (1980; Braddick & Adlard, 1978): a short-and a long-range process. It is suggested that *element* motion is perceived via the activation of a short-range process (i.e. for short spatial and temporal intervals), while the *group* motion would activate a long-range process. Another complementary version of the hypothesis refers to the parallel processing in the parvo-and magnocellular visual system (e.g. Bradley & Born, 2003). In neither case, the results were totally conclusive (e.g. Skottun, 2001).

The similarity between the pictures of the display plays a major role in the appearance of *group* responses (e.g. He & Ooi, 1999; Oyama et al., 1999). The similarity factors which have been studied are contrast, colour, orientation, spatial frequency, and shape. The association field hypothesis (Alais & Lorenceau, 2002) combines this factor of similarity with the short-and long-range explanation. These authors proposed that neighbouring pictures of the Ternus display group together as much as they share similar orientation. The grouping effect would be greater for longer ISI, i.e. when the long-range process dominates.

Many experiments were dealing with the study of similarity in the Ternus phenomenon combining simple geometric figures. Few are known about the influence



of semantic factors combined with structural factors. Yu (2000) has shown a weak effect of a semantic context induced by a first meaningful image used to induce a semantic context: A cartoon of a person reduces slightly the frequency of *group* responses. In this respect, it is worth to mention the particular effect of human body images in apparent beta motion shown by Shiffrar and her colleagues. The perceived trajectory of the apparent motion is in general the shortest track (i.e. a straight linear trajectory). These authors (Shiffrar & Freyd, 1990, 1993; Chatterjee, Freyd & Shiffrar, 1996; Stevens et al., 2000) have shown that the perceived trajectory of the apparent motion can be seen as curved when the motion of a body segment is biomechanically impossible. This happens only for large ISI allowing the complete processing of semantic factors. For instance a hand is perceived as moving around the head, while for shorter ISI the shortest trajectory is perceived (i.e. through the head).

The present experiments try to evaluate in which extent structural and semantic factors influence the perception of apparent *group* motion in a Ternus paradigm. Semantic factor was manipulated through the content of the images which were either pictures of human body or pictures of chair. The choice of animate and inanimate stimuli was done in considering the many arguments demonstrating a different processing of these two classes of visual stimuli including the above reported results. Neuroimaging studies demonstrate that different brain structures are activated by (inanimate) object images than by human body images (Kourtzi & Kanwisher, 2000; Senior et al., 2000, 2002; Paulos et al., submitted). Neuropsychological studies stressed differences between recognizing objects and recognizing (mostly naming) actions from pictures suggesting a double dissociation between two different processes (e.g. Campbell & Manning, 1996; Manning & Warrington, 1996).

The attempt to disentangle the structural and semantic factors in the Ternus phenomenon was done here in applying parallel changes in similarity of the pictures between and within two semantic categories: human body pictures and chair pictures.

## **Experiment 1**

In a first experiment pictures of a human body and a chair were presented. They were combined in four conditions: two homogeneous conditions and two heterogeneous ones. Each condition was run in pure blocks including different levels of

ISI of trials with a change in the ISI within each block. In homogeneous conditions we want to know whether or not these two semantic categories of pictures raise similar change in the frequency of “group” responses with the increase in ISI. Tentatively, it was expected that human body pictures would raise more *group* responses because of their implicit dynamic characteristics. In heterogeneous conditions, the two central pictures (in **B** and **C**) belonged to the other semantic category than those presented at the extreme positions (in **A** and **D**). This change affects both structural and semantic factors and should reduce the frequency of *group* responses. If the body pictures raise more *group* responses in the homogeneous conditions, they should also raise relatively more *group* responses in the heterogeneous condition when they are the moving element (from **A** to **D**).

### Material and method

The experiment was run on a high resolution (1024 x 768 px) computer display (60 Hz) controlled with an Inquisit program V1.32 by Millisecond Software. At every trial, two images (**A-B-C** and **B-C-D**) alternate for 2.5 cycles. Each image is shown for 5 frames (83.3 ms) separated from the next one by an ISI.

Each experimental session started with 20 practice trials using black disk pictures as shown in Figure 1. In these practice trials, only two extreme values of the ISI were used (1 and 10 frames). The subject was requested to indicate in pressing one of the two keys of the keyboard whether they have perceived the motion of an *element* (**A** → **D**) or if they perceived the motion of a *group* (**ABC** → **BCD**). They were encouraged to respond as fast as possible, even before the end of the presentation. Responses were corrected, the word “error” appearing on the monitor if they respond *group* for the shorter ISI, or if they respond *element* for the longer one.

Figure 1 about here

In the main experiment, the duration of this ISI varies randomly from trial to trial. Its durations were 1, 2, 4, 6, 8 or 10 frames (about 16.6 – 33.2 – 66.4 – 99.9 – 133.3 – 166.6 ms). In each condition (see below), each ISI was shown 10 times. No corrections of the responses were given. Depending on the ISI, the total durations of the stimulus were: 483, 549, 682, 816, 950, 1083 ms.

Four conditions were used in separate trial blocks. They differ by which pictures were presented. In two conditions, thereafter called homogeneous, the three figures (**ABC** and **BCD**) were identical being in one condition, images of a human body (condition HH) and in the other condition, images of a chair (condition CC). The four pictures were identical (see Figure 2). In the two other conditions, called heterogeneous, the extreme figures (in positions **A** and **D**) were different from the two central ones: a human body with a chair in positions **B** and **C** (condition HC) and a chair with a human body in positions **B** and **C** (condition CH).

Figure 2 about here

Sixteen subjects took part in the experiment. They were psychology students and voluntary and give their consent to participate. Each subject was run successively in the four conditions during the same session. The first condition was always a homogeneous one. Homogeneous and heterogeneous conditions alternated and were separated by a short rest pause.

Subjects sat at 120 cm from the monitor with his/her head on a chin and head rest. The background luminance of the monitor was about 10 cd/m<sup>2</sup>. It was the only source of light in the experiment room. Figures were well contrasted and shown in shades of grey (grey scaled). Each one has 1.7° height and a group of three figures covered 5.6°. Centre to centre, the distance between figures was 1.9°.

At every trial, subject's task was to indicate as fast as possible (reaction times were measured) whether the apparent motion of a single picture (*element* response) or the apparent motion of a group of three figures (*group* response) is seen. They were encouraged to respond, if possible, before the end of the sequence. Reaction times were measured from the onset of the second image.

## Results

Figure 3 shows the frequencies of *group* responses as a function of the ISI duration. The effect of the later factor is significant ( $F(5,75) = 38.37$ ;  $p < 0.00001$ ): The frequency of *group* responses increases with the duration of the ISI ( $F(1,15) = 53.14$ ;  $p < 0.0001$ ). The frequency of the *group* responses differs between the four conditions ( $F(3,45) = 9.04$ ;  $p < 0.0001$ ). A significant interaction was observed between the two factors (conditions x ISI) ( $F(15,225) = 5.37$ ;  $p < 0.0001$ ).

Figure 3 about here

The main effect of the conditions results essentially from the difference between the homogeneous conditions (HH and CC) and the heterogeneous ones (HC and CH); the former leading to more *group* responses (0.53) than the later (0.39) ( $F(1,15) = 19.27$ ;  $p < 0.0005$ ). In fact, the quadratic regression between these two subgroups (homogeneous and heterogeneous) is significant ( $F(1,15) = 27.77$ ;  $p < 0.0009$ ). None of the differences between the conditions of each subgroup were significant ( $F < 1$ ).

The increase in the frequency of *group* responses with the ISI reflects an increase in the probability of such a response. For this reason, as seen in Figure 3, psychometric sigmoid functions were fitted to the results. This was done in order to estimate the duration of the liminal ISI leading to the maximum uncertainty of the responses (50%, a point of subjective equality or PSE). This PSE is an index of the strength of the *group* responses: the lower it is the stronger are *group* responses. The estimates are for each condition: HH = 3.9 frames; CC = 4.5 frames; HC = 6.7 frames; CH = 7.4 frames. These results correspond to the fact that the mean frequency of *group* responses is higher in condition HH (0.54) than in condition HC (0.39) and in condition CC (0.52) than in condition CH (0.38).

Figure 4 about here

The first analysis on reaction times mixed *element* and *group* responses. These RT shown in Figure 4 are the mean of the individual median for each experimental condition. The mean RT in the four conditions are rather similar (HH = 626.8 ms; CC = 594.5 ms; HC = 643.2; CH = 666.3). In the global ANOVA, neither the conditions, nor the ISI, and nor their interaction are significant ( $F < 1$ ). In fact, RT are longer in the medium range of the ISI in such a way that a quadratic regression is significant for the effect of ISI ( $F(1,15) = 3.08$ ;  $p < 0.09$ ).

A descriptive analysis of the RT data was then done separately for the *group* responses and for the *element* response. Since the number of responses for each factor is variable, it was chosen to use the interindividual medians. The averaged RT is longer for *group* responses (596.5 ms) than for *element* responses (549.5 ms). Moreover, it is observed that the RTs for *group* responses are shorter than RT for *element* responses when the ISI is short, but longer when the ISI is long.

## Discussion

The present results show that the frequency of the *group* responses in the Ternus paradigm is a growing probability function of the ISI duration (Dawson & Wright, 1989; Kramer & Yantis, 1997; Allais & Lorenceau, 2001). The identity of the shown pictures in the four positions (*A–B–C–D*) is more favourable to *group* responses: an earlier switch from *element* responses to *group* responses appears for shorter ISI when the four figures are identical than when they are not. The main observed effect is the one of the structural (and semantic) factor (homogeneous vs. heterogeneous). The semantic content of the images has a negligible effect: As expected, the switch between *element* to *group* responses appears a little earlier when the extreme (‘moving’) picture is a human body (HH and HC) than when it is a chair (CC and CH). In fact, the difference appears significant only for ISI= 6 frames ( $t(15)=2$ ,  $p=0.06$ ), i.e. when the uncertainty between *group* and *element* responses is maximum.

Reaction time data indicate that *group* and *element* responses are not used symmetrically by the subjects. Their variation seems to reflect essentially changes in the certainty of the subject’s responses.

The semantic and structural context was constant in each block of 60 trials during which the same pair of images was presented. From a methodological point of view, it may be thought that this procedure is more favourable to the production of *group* responses in heterogeneous conditions since the subjects may not neglect to use *group* responses. In the two next experiments, the conditions (homogeneous and heterogeneous) will be mixed within experimental sessions.

## **Experiment 2 (chair)**

The aim of the second experiment was to explore the effects of structural factors within the same semantic category (chair). The different conditions were mixed within the session. The structural factor was manipulated in two ways: by the type of the chair and by its orientation. Chair exemplars were taken from Kolars (1972, see Figure 5). It is expected, as classically found, that reducing the structural similarity between the pictures will decrease the frequency of “group” responses. The very question is to compare the effects of a complete heterogeneity between the pictures (see below *mixed* condition) with a change in the orientation of the extreme picture between positions *A*

and **D**. Such change in orientation has frequently been reported as very detrimental to “group” perception (e.g. Alais & Lorenceau, 2002).

### Stimuli and method

Unless mentioned, the experimental conditions, including the ISI, were the same as in Experiment 1. The instructions ask the subject to concentrate on the motion aspect (*element* vs. *group*) without paying much attention to the nature of the single pictures. Three conditions were used varying randomly from trial to trial.

Figure 5 about here

The three conditions were defined by which pair of images was used:

-In the *normal* condition (Nc) four identical chairs were shown in positions **A**, **B**, **C**, and **D**, as in condition CC of Exp. 1. The chair picture was the one seen in position **B** of Figure 5.

-In the *inverse* condition (Ic), the two central chairs (**B** and **C**) were the same as in condition Nc (Chair **B** of Figure 5). The chairs in position **A** and **D** were identical (chair **A** of Figure 5); they were different from the ones in positions **B** and **C**. The chair shown in position **D** was upside down.

-In the *mixed* condition (Mc) the four chairs (see Figure 5) were different with a similar upside orientation but seen from different point of views (around the vertical axis).

Another group of 16 psychology students participated as voluntary subjects in the experiment. Three hundred eighty trials (6 ISI x 3 images conditions x 10 replications) were given to each subject. Their order was randomized.

### Results

Figure 6 about here

Figure 6 shows the frequency of *group* responses as a function of the ISI duration for each of the three conditions. The mean frequency of *group* responses was respectively for each condition: Nc = 0.67, Ic = 0.39, Mxc = 0.47 ( $F(2,30) = 13.67$ ;  $p < 0.0006$ ). The frequency of *group* response varies with the duration of the ISI ( $F(5,75) = 23.48$ ;  $p < 0.0001$ ). In fact, the increase in the *group* responses with the increase in the ISI was significant ( $F(1,15) = 31.6$ ;  $p < 0.0001$ ). A significant interaction was observed

between the two factors (conditions  $\times$  ISI) ( $F(10,150) = 1.61$ ;  $p=0.10$ ). The homogeneous condition (Nc) gives significantly more *group* responses than the two heterogeneous ones (Ic,Mxc) ( $F(1,15) = 13.98$ ;  $p<0.002$ ). The *mixed* condition (Mxc) gives more *group* responses than the *inverse* (Ic) one ( $F(1,15) = 9.79$ ;  $p<0.01$ ).

As for experiment 1, psychometric sigmoid functions were fitted to the results which allow estimating the duration of the ISI leading to the maximum uncertainty of the response (0.50). The estimates are for each condition: Nc = 2.2 frames; Ic = 6.5 frames and Mc = 4.1 frames.

Figure 7 about here

The first analysis on reaction times mixed *element* and *group* responses. The mean RT is 612 ms. These RT shown in Figure 7 are the mean of the individual median for each experimental condition. The mean RT are respectively: Nc = 525.2 ms; Ic = 658.4 ms; Mc = 653.8 ms ( $F(2,30) = 23.11$ ;  $p<0.0001$ ). The ISI has a significant effect ( $F(5,75) = 6.38$ ;  $p<0.0001$ ). Globally, RT decline when the ISI increases ( $F(1,15) = 11.89$ ;  $p<0.005$ ). A significant interaction was observed between the conditions and the ISI ( $F(10,150) = 3.51$ ;  $p<0.001$ ). Such an interaction is due to the fact that RT declines with ISI only in the condition Nc ( $F(1,15)=14.19$ ;  $p=0.001$ ). The difference between the two heterogeneous conditions is not significant ( $F<1$ ).

As before, RT for *group* responses and RT for *element* responses were analyzed. On average, RT for *group* responses are longer ( $RT_g = 651.1$  ms) than RT for *element* responses ( $RT_e = 617.8$  ms). For both responses, RT declines when ISI increases.

## Discussion

Again, in all three conditions, the probability of *group* responses increases when the ISI increases. Sigmoid probability functions were satisfactorily fitted to the data. As expected, the heterogeneity of the stimuli (conditions Ic and Mxc) reduces the number of *group* responses and increases the RT. Likely due to the difference in the experimental procedure (pure blocks in Exp.1 and mixed blocks in Exp. 2), the difference between homogeneous conditions and heterogeneous conditions is much larger in the second experiment than in the first one (24% vs. 15%). Among the structural factor of heterogeneity, the orientation (upside down) has a larger effect in reducing the *group* responses than the heterogeneity in the shapes of the chair (see

Allais & Lorenceau, 2002). The question is then to know whether this orientation effect is structural and/or semantic. Since the vertical axis of the object is preserved, it may be thought that the incongruity of seeing a chair standing upside down is responsible for more uncertainty in the *group* responses.

### **Experiment 3 (Body)**

Within the chair category it was possible to present very different chair exemplar as in Exp. 2. With images of human bodies, some structural diversity could be obtained by the posture of each exemplar. In the present, the four pictures in position **A**, **B**, **C**, **D** show different postures and were extracted from a film showing a man running (see Figure 8). They form a regular sequence of an action. Two conditions introduced a perturbation of this regular sequence for the picture in position **D**; in one the human body is facing left while the three others are facing right, in the other one, the human body is upside down and facing left. The aim of the present experiment is to compare the detrimental effect of these two changes in orientation of the last picture. A second aim is to compare the effect of a regular action sequence as used here which introduces some dissimilarity in the posture between pictures with the constant posture as used in Experiment 1 (condition HH).

### **Stimuli and method**

Unless mentioned, the experimental conditions, including the ISI, were the same as in Experiment 1. The instructions ask the subject to concentrate on the motion aspect (*element* vs. *group*) without paying much attention to the nature of the figures. Three conditions were used varying randomly from trial to trial.

Figure 8 about here

Each image in this experiment contains three successive postures of a running man in positions **A**, **B** and **C** for image 1 and **B**, **C**, **D** for image 2. Three conditions were used characterized by the posture of the body picture in **D**:

-In the *normal* condition (Nb), the posture of the man in **D** follows ‘naturally’ his posture in position **C** (see Figure 8). All figures were facing rightward.

-In the condition *mirror* (Mb), the body posture in **D** is facing right so as to appear as a mirror image of the man in **A**.



-In condition *inverse* (Ib), the body posture in **D** is upside down and facing left.

Another group of 16 psychology students participated as voluntary subjects in the experiment. Three hundred eighty trials (6 ISI x 3 images conditions x 10 replications) were given to each subject. Their order was randomized.

## Results

The mean frequencies of *group* responses are shown in Figure 9. The frequencies of *group* responses differ between conditions (Nb = 0.58; Mb = 0.47; Ib = 0.49) ( $F(2,30) = 8.30$ ;  $p=0.001$ ). The frequency of *group* response varies with the duration of the ISI ( $F(5,75) = 48.78$ ;  $p<0.0001$ ). In fact, the increase in the *group* responses with the increase in the ISI was significant ( $F(1,15) = 82.35$ ;  $p<0.0001$ ). The interaction between conditions and ISI is not significant.

Figure 9 about here

The homogeneous condition (Nb) gives significantly more *group* responses than the two heterogeneous ones (Mb,Ib) ( $F(1,15) = 30.69$ ;  $p<0.001$ ). The two heterogeneous conditions (Mb vs. Ib) do not differ significantly ( $F<1$ ).

As for the two previous experiments, psychometric sigmoid functions were fitted to the results which allow estimating the duration of the ISI leading to the maximum uncertainty of the response (0.50). The estimates are for each condition: Nb = 3.4 frames; Mb = 4.8 frames and Ib = 4.7 frames.

Figure 10 about here

The first analysis on reaction times mixed *element* and *group* responses. The mean RT is 652 ms. These RT shown in Figure 10 are the mean of the individual median for each experimental condition. The mean RT are respectively: Nb = 612.7 ms; Mb = 674.3 ms; Ib = 669.8 ms ( $F(2,30) = 6.96$ ;  $p<0.01$ ). The ISI has not globally a significant effect ( $F(5,75) = 1.86$ ;  $p>0.1$ ). However, RT decline significantly when the ISI increases ( $F(1,15) = 3.38$ ;  $p<0.05$ ). A significant interaction was observed between the conditions and the ISI ( $F(10,150) = 2.11$ ;  $p<0.05$ ).

RT are shorter and decrease faster with the increase in ISI in condition Nb than in the two other conditions ( $F(1,15) = 8.04$ ;  $p<0.05$ ). The difference between the two heterogeneous conditions is not significant ( $F<1$ ).

As before, RT for *group* responses and RT for *element* responses were analyzed.

On average, RT for *group* responses and for *element* responses are similar ( $RT_g = 737.7$  ms;  $RT_e = 731.7$  ms). However, for *group* responses RT declines when ISI increases in all three conditions, while for *element* responses, this happens only for condition Nb. **Discussion**

Again, in all three conditions, the probability of *group* responses increases when the ISI increases. Sigmoid probability functions were satisfactorily fitted to the data. As expected, the heterogeneity of the stimuli (conditions *mirror* and *inverse*) reduces the number of *group* responses and increases the RT. However, it cannot be concluded whether the effects of the introduced perturbations result from a structural factor (change in the orientation of the picture) or a semantic one (disruption of a regular action sequence).

Comparison of similar conditions between Experiment 1 (condition HH) and the present experiment (condition *normal*) shows no significant difference in the frequency of *group* responses ( $F(1,30) < 1$ ; ns); that frequency increases with the increase in ISI ( $F(5,150) = 61.29$ ;  $p < 0.0001$ ) and more interesting, there is an interaction of the two factors ( $F(5,150) = 2.43$ ;  $p < 0.04$ ). That interaction results from the fact that the psychometric function is steeper in the present experiment than in experiment 1. Correlatively, the PSE is lower in the present experiment (3.4 frames) than in experiment 1 (3.96). This difference is consistent with the hypothesis of using pictures of a regular human body action sequence as a facilitating factor for grouping perception. However, the difference in the experimental conditions (pure blocks in experiment 1, mixed blocks in experiment 3) should be taken into consideration.

### **General discussion**

In the three experiments, the classical increase in the frequency of *group* responses with the increase in ISI was observed (Dawson & Wright, 1989; Kramer & Yantis, 1997; Allais & Lorenceau, 2001). The present results show that the grouping effect, due to the increase in ISI, resists globally even for large structural dissimilarities.

However, it may be suggested that *group* responses obtained in heterogeneous conditions do not reflect simply subjects' perception but are also constrained by the experimental procedure. This can be the case in Exp. 1 in which pure blocks of conditions are presented. Being instructed that two responses are possible, subjects may

be reluctant to use only one of the two (for instance the *element* response). If that tendency was present to a large extent, almost no *group* response would be found in Exp. 2 and 3 in the heterogeneous conditions. A careful comparison of the procedure effect is then necessary between Exp. 1 (pure blocks) on the one hand and Exp. 2 and 3 (mixed blocks) on the other hand.

In order to achieve these comparisons, we have compared sigmoid probability functions fitted to the averaged data. We first compare the mean frequencies of *group* responses for the homogeneous conditions between the two sets of experiments (HH,CC vs. Nc,Nb). It is shown that more *group* responses are produced in the mixed block experiments (62.5%) than in the pure block ones. This difference appears in the lower range of ISI values. For heterogeneous conditions, *group* responses grow later in the pure block experiment than in the mixed block ones. These differences in the frequency of the responses were accompanied by differences in RT: on average, RTs are shorter in the mixed block experiments than in the pure block experiment. From these comparisons, it can be concluded that the effect of the change in procedure (presentation mode), from pure to mixed blocks, does not affect the results in a critical way. In fact, the mean difference between homogeneous and heterogeneous conditions is very similar with the two procedures.

Since the differences between homogeneous and heterogeneous conditions cannot mainly result from response bias due to the experimental procedure, we can now examine the effect of the different structural and semantic factors. Dissimilarities between the pictures of each image reduce the frequency of the *group* responses. In the homogeneous conditions (HH, CC, Nc and Nb) about 58% of the responses are *group* responses while only 43% of *group* responses are observed in heterogeneous conditions. Consequently the PSE is lower in homogeneous (mean = 3.4) than it is in heterogeneous conditions (mean = 6.1). Correlatively, overall RTs are shorter in homogeneous conditions than in heterogeneous ones (homogeneous = 589.8 ms; heterogeneous = 642.9 ms).

Two aspects of the dissimilarity effect are examined which are confounded with the effect of the procedure (pure vs. mixed blocks). This difference appears essentially for the middle range of the ISI. For the shortest and for the longest ISI, no difference is demonstrated on average. We would then suggest that heterogeneity between

categories is more detrimental to *group* responses than heterogeneity within a category. Firstly, is the case of the structural and semantic heterogeneity within each image. In that case the ‘moving’ picture (in positions A and D) is identical between the two frames. This is the case of condition HC and CH in Exp.1 (favouring “element” responses). Other cases combine heterogeneity within and between the two frames: the ‘moving’ picture is different in position A and D. This is the case in conditions Ic, Mxc, Ib and Mb. These effects are clearly demonstrated if the frequency of “group” response in each heterogeneous condition is subtracted from the frequency in the corresponding homogeneous condition.

Table 1 about here

Exp. 3 using only human body images exhibit the smallest effect of dissimilarities (between frames) on the grouping phenomenon. This means that in spite of the fact that different postures are used, these changes have only a small influence on the grouping process. The heterogeneity within figures as in Exp.1 reduces the frequency of *group* responses. In this experiment, it is not possible to know whether the more important factor is structural dissimilarity or semantic dissimilarity. Finally, the largest effects of dissimilarities are found in Exp. 2. While we have no metric of structural dissimilarity, it is obvious that our chairs were structurally more dissimilar than our human bodies. This is enough to explain why the difference between homogeneous and heterogeneous conditions is larger in Exp. 2 than in Exp. 3.

Some of the results suggest that the semantic identity of the pictures may modulate the effects of structural factors. In experiment 1, a smaller PSE is observed when the extreme (‘moving’) picture is a human body (HH and HC) than when it is a chair (CC and CH). A second suggestion comes from the comparison of two similar conditions of experiment 1 (HH) and experiment 3 (*normal*). Tentatively, it was seen that an action sequence (running man) leads to a steeper psychometric function than a identical posture of the body picture. It would be premature to conclude to a privilege of human body representations in modulating the Ternus phenomenon. Further experiments will concentrate on this point.

More experiments are certainly needed to clarify completely the question. However, the present results demonstrate first the robustness of the grouping effect as dependent of the duration of the ISI. Secondly, they confirmed the effect of structural

similarity in modulating the Ternus effect (Alais & Lorenceau, 2002; He & Ooi, 1999; Oyama et al., 1999). And finally, they suggest that if semantic factors may have an effect, this effect will be very small at least in the present experimental conditions.

### References

Allais, D. & Lorenceau, J. (2002). Perceptual grouping in the Ternus display: evidence for an “association field” in apparent motion. *Vision Research*, 42, 1005-1016.

Bradley, D. C. & Born, R. (2003). Structure and Function of Area MT *Annual Review of Neuroscience* (in press)

Braddick, O.J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London, Series B*, B290, 137-151.

Braddick, O.J. & Adlar, A. (1978). Apparent motion and the motion detector. In J.C. Armington, J. Krauskopf, B.R. Wooten (Eds.) *Visual Psychophysics and Psychology*. New York, Academic Press.

Breitmeyer, B. G. & Ritter, A. (1986a). The role of visual pattern persistence in bistable stroboscopic motion, *Vision Research*, 26, 1801-1806.

Breitmeyer, B. G. & Ritter, A. (1986b). Visual persistence and the effect of eccentric viewing, element size and frame duration on bistable stroboscopic motion percepts. *Perception and Psychophysics*, 39, 275-280.

Burt, P. & Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review*, 88, 171-195.

Campbell, R. & Manning, L. (1996). Optic aphasia: A case with spared action naming and associated disorders. *Brain and Language*, 53, 183 – 221.

Chatterjee, S. H., Freyd, J. J. & Shiffrar M. (1996). Configural processing in the perception of apparent biological motion. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 916-929.

Dawson, M.R.W. (1991). The How and Why of What Went Where in Apparent Motion: Modeling Solutions to the Motion Correspondence Problem. *Psychological Review*, 98, 569-603.

Dawson, M. & Di Lollo, V. (1990). Effects of adapting luminance and stimulus contrast on the temporal and spatial limits of short-range motion. *Vision Research*, 30, 415-429.

Dawson, M.R.W. & Wright, R.D. (1989). The consistency of element motion affects the visibility but not the direction of apparent movement. *Spatial Vision*, 4, 17-29.

He, Z.J. & Ooi, T.L (1999). Perceptual organization of apparent motion in the Ternus display. *Perception*, 28, 877-892.

Kolers, P.A. (1972). *Aspects of motion perception*. Oxford, Pergamon Press.

Kourtzi, Z. & Kanwisher, N. (2000). Activation in Human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience*, 12, 48 - 55.

Kramer, P. & Rudd, M. (1999). Visible persistence and form correspondence in Ternus apparent motion. *Perception & Psychophysics*, 61, 952-962.

Kramer, P. & Yantis, S. (1997). Perceptual grouping in space and time: Evidence from the Ternus Display. *Perception & Psychophysics*. 59, 87-99.

Manning, L. & Warrington, E. K. (1996). Two routes to naming: a case study. *Neuropsychologia*, 34, 809 – 817.

Oyama, T., Simizu, M. & Tozawa, J. (1999). Effects of similarity on apparent motion and perceptual grouping. *Perception*, 28, 739 – 748.

Paulos, C., Scheiber, C., Manning, L. & Bonnet, C. Visual representations of human bodies and objects: an fMRI study. (submitted).

Petersik, J. T. (1989). The two-process distinction in apparent motion. *Psychological Bulletin*, 106, 107-127.

Sekuler, R. (1996). Motion perception: a modern view of Wertheimer's 1912 monograph. *Perception*, 25, 1243-1258.

Senior, C., Ward, J. & David, A. S. (2002). Representational momentum and the brain: An investigation into the functional necessity of V5/MT. *Visual Cognition*, 9, 81 - 92.

Senior, C., Barnes, J., Giampietro, V., Brammer, M., Bullmore, E., Simmons, A. & David, A. S. (2000). The functional neuroanatomy of implicit motion perception or 'representational momentum. *Current Biology*, 10, 16-22.

Shiffrar, M. & Freyd, J. J. (1990). Apparent motion of the human body. *Psychological Science*, 1, 257-264.

Shiffrar, M. & Freyd, J. J. (1993). Timing and apparent motion path choice with human body photographs. *Psychological Science*, 4, 379-384.

Skottun, B.C. (2001). On the use of the Ternus test to assess magnocellular function.

*Perception*, 30, 1449-1457.

Stevens, J.A., Fonlupt, P., Shiffrar, M. & Decety, J. (2000). New aspects of motion perception: selective neural encoding of apparent human movements. *NeuroReport*, 11, 109-115.

Ternus, J. (1926). Experimentelle untersuchungen über phänomenale identität. *Psychologische Forschung*, 7, 81-136.

Yu, K. (2000). Can semantic knowkledge influence motion correspondence? *Perception*, 29, 693 – 707.



### Figure captions

Figure 1: Schema of the Ternus phenomenon. Pictures are here black disks. Their names and positions are indicated by capital letters.

Figure 2: Pictures of the stimuli used in experiment 1 shown in their four positions used in condition HH (top row) and CC (bottom row). In the heterogeneous conditions, the two central pictures were different from the two extreme ones.

Figure 3: Frequency of “group” responses as a function of the ISI (in frames) for the 4 conditions of experiment 1. For sake of clarity, standard error bars are omitted. Sigmoid psychometric functions are fitted and the PSE values shown with dotted lines (see text).

Figure 4: Reaction times of all responses as a function of the ISI for the 4 conditions of experiment 1.

Figure 5: Stimuli used in experiment 2. The chair picture in position B is used in condition *normal*, the chair pictures in positions A and B are used in condition *mirror* and the 4 chair pictures are used in condition *mixed*.

Figure 6: Frequency of “group” responses as a function of the ISI (in frames) for the 3 conditions of experiment 2. Standard error bars are shown. Sigmoid psychometric functions are fitted and the PSE values shown with dotted lines (see text). Nc for *normal* condition, Ic for *inverse* condition, Mxc for *mixed* condition.

Figure 7: Reaction times of all responses as a function of the ISI for the 3 conditions of experiment 2.

Figure 8: Running man pictures used in experiment 3. Pictures are in positions **A**, **B**, **C** and **D** as in the normal condition.

Figure 9: Frequency of “group” responses as a function of the ISI (in frames) for the 3 conditions of experiment 3. Standard error bars are shown. Sigmoid psychometric functions are fitted and the PSE values shown with dotted lines (see text).

Figure 10: Reaction times of all responses as a function of the ISI for the 3 conditions of experiment 3.

### Acknowledgements

Acknowledgements are due to Emilie Pirra, to Nathalie Benoliel and to Raaj Seereekissoon for carefully running the experiments.

Figure 1:

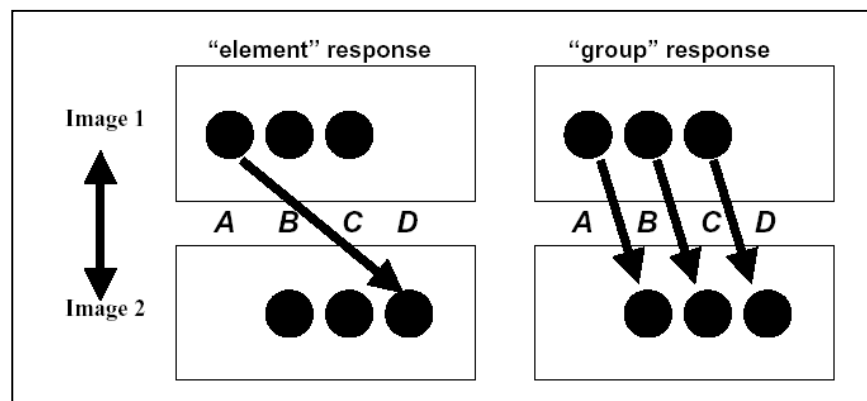


Figure 2:

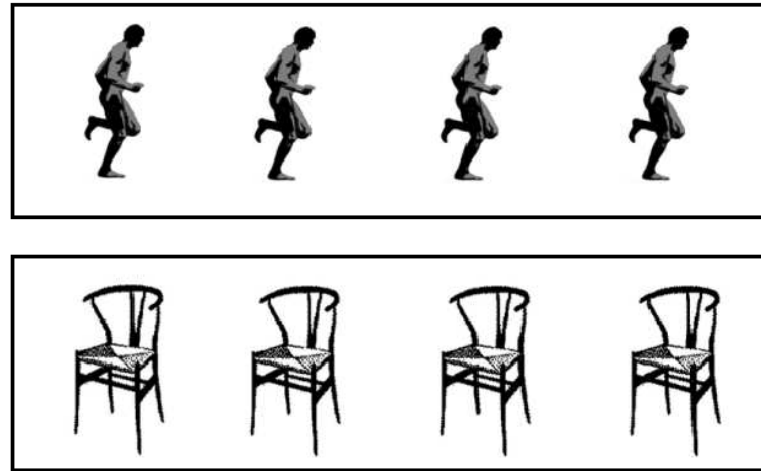


Figure 3:

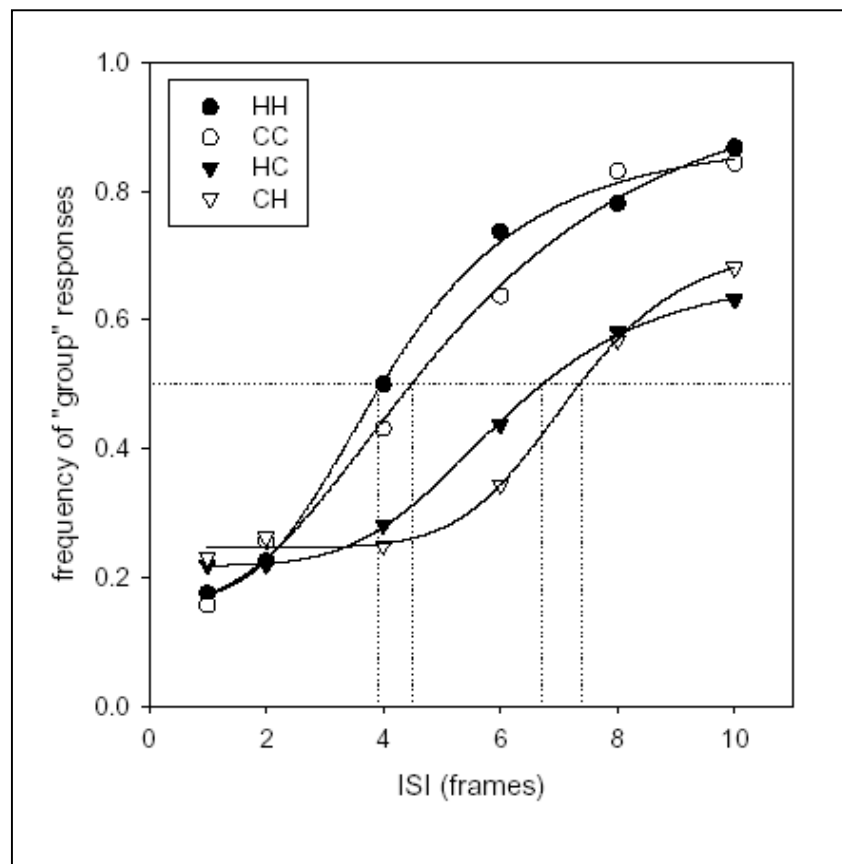


Figure 4:

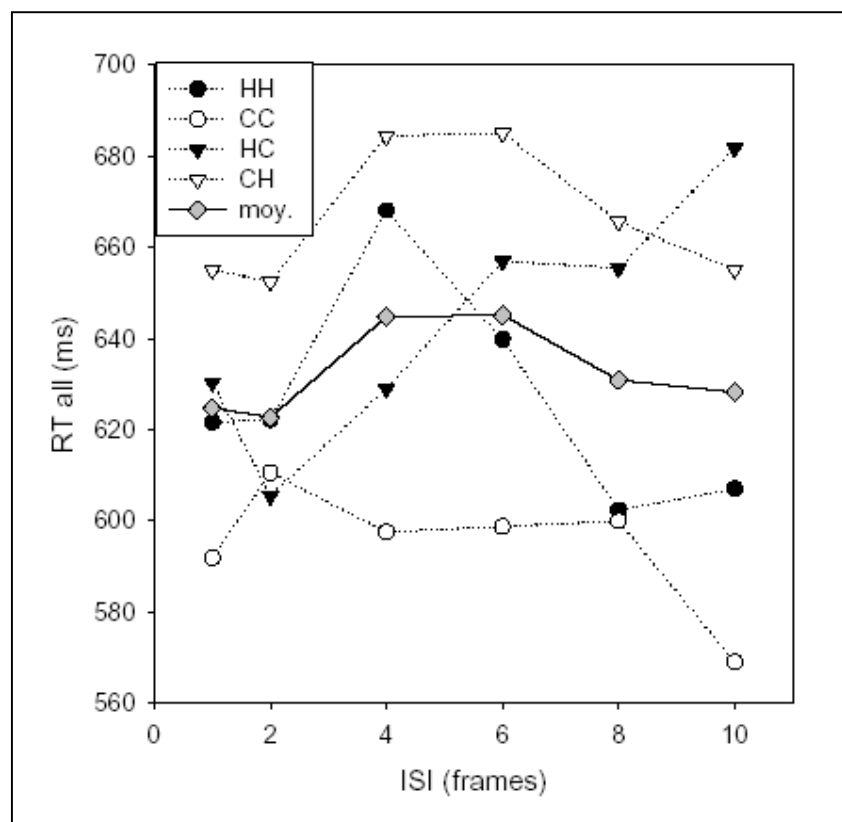


Figure 5:

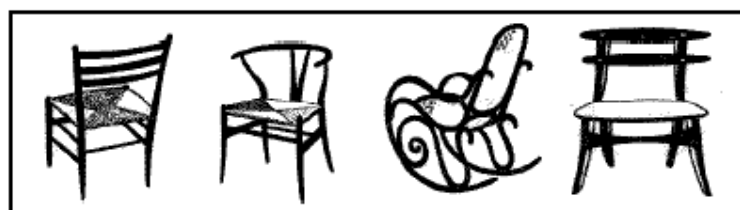


Figure 6:

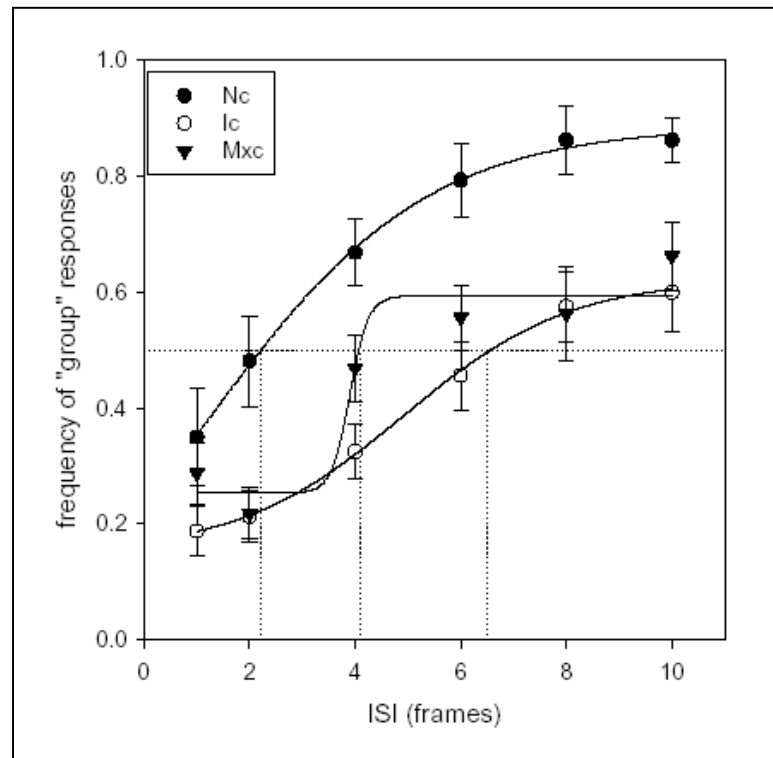


Figure 7:

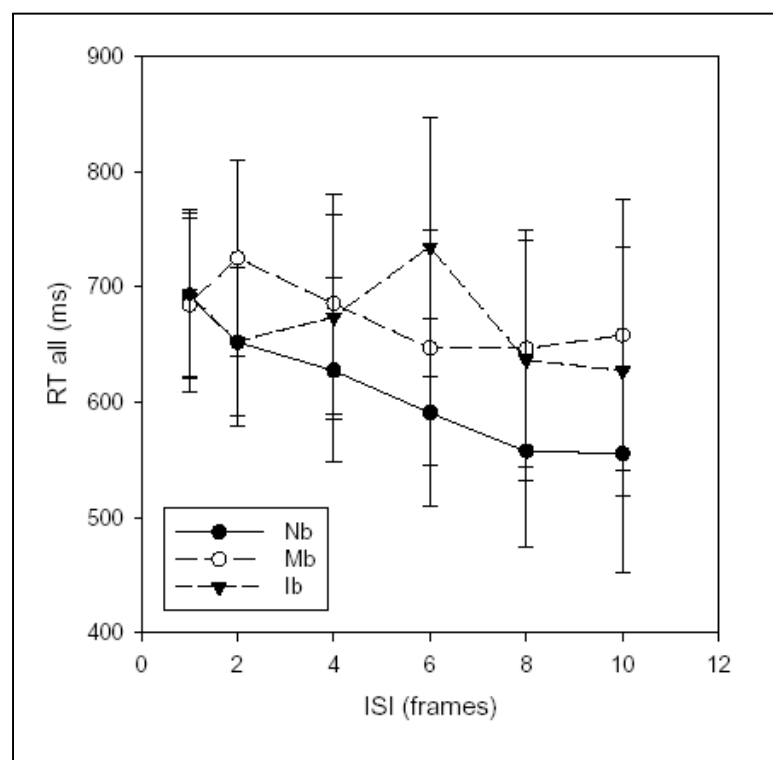


Table 1

Difference in the frequency of “group” responses of each heterogeneous condition  
with respect to the corresponding homogeneous condition

| HC   | CH   | lc   | Mxc  | Mb   | lb   |
|------|------|------|------|------|------|
| 0.17 | 0.14 | 0.28 | 0.20 | 0.11 | 0.09 |

## **Human body actions in the Ternus paradigm**

**Claude Bonnet & Carlos Paulos**

### **Abstract**

The present experiments explored the possibility that human body action representations may modulate the effects of structural factors in the Ternus apparent movement. The first experiment compares the frequency of “group” responses for increasing ISI when the pictures elements represent a sequence of body action: running, pirouette and somersault. Results are similar when each action is presented in pure blocks of trials; strong interactions appear when they are presented in mixed blocks. The second experiment examines the effect of perturbing the regularity of a sequence of pirouette; perturbations increase the frequency of “group” responses. The third experiment examines the effect of perturbing the regularity of a sequence of somersault with similar results as experiment 2. All together, the results suggest some semantic influence of body representations, but the comparison between pure block designs and mixed block design indicate the fragility of these semantic influences.

### **Introduction**

Effects of structural and semantic factors in the Ternus paradigm have been examined in a previous paper (Bonnet & Paulos, *submitted*). Structural similarity between the pictures was confirmed to be a major factor favouring the frequency of “group” responses. However, some semantic effects were reported in the cases of pictures of human body and more likely when the pictures represent an action. The present paper continues in the same framework, in concentrating on human body actions in order to try to reveal the possible effect of semantic factors.

The Ternus paradigm (Ternus, 1926; Kolers, 1972) is an apparent motion phenomenon in which two frames alternate separated by a blank interval. Each frame or image shows three equally spaced pictures arranged so that two pictures occupy the same position in the two frames (see Figure 1). These two alternating frames are separated by a blank interstimulus interval (ISI). Two different percepts are seen depending on the duration of this ISI (*e.g.* Kramer & Yantis, 1997). When the ISI is short (or null), subjects only report the motion (called “*element*”) of the extreme picture ( $A \rightarrow D$ ),

while for longer ISI, they report more and more the motion of a “group” of three pictures ( $ABC \rightarrow BCD$ ). Increasing the presentation duration of each frame also increases the probability of a “group” report (e.g. Dawson & Wright, 1989). The more spaced are the pictures within and between frames, the lower the probability of “group” responses (e.g. Burt & Sperling, 1981; Dawson & Di Lollo, 1990; Kolers, 1972). Most research on the Ternus phenomenon has concerned the effects of low-level factors. Different explanations have been proposed for the switch between “element” and “group” responses, while no definitive consensus appears. The structural similarity between the pictures of the display plays a major role in the appearance of “group” responses (e.g. He & Ooi, 1999; Oyama et al., 1999; Bonnet & Paulos, *submitted*). The similarity factors which have been studied are contrast, color, orientation, spatial frequency, and shape (e.g. Alais & Lorenceau, 2002).

In our previous experiments (Bonnet & Paulos, *submitted*), comparing the frequency of “group” responses when pictures of human body or chair were presented in homogeneously or heterogeneously structural conditions lead to the conclusion that structural factors prevailed. The semantic factor was manipulated between object categories (human vs. chair). Also noticeable was the effect of changing the orientation of picture in position **D** compared to **A**. In these cases, the differently oriented fourth picture (in **D**) clearly disrupted an implicit sequence of action and reduces the frequency of group responses.

The present paper explores the possibility that semantic factors may prevail on structural ones when the four pictures are made of a sequence of human action sequence. The reason for this choice is that human body images showing an action are likely processed by specific mechanisms in which semantic representations are of importance. First, human body images are likely to be processed by different mechanisms than inanimate objects. Different arguments suggested that it is the case. Like for faces, infants of about 1 year show a preference for normal human body shape as opposed to scrambled ones (Slaughter, Heron & Sim, 2002). The inversion effect frequently considered as face-specific also appears for body images (Reed et al., 2003). The relations among parts of human bodies are different than other objects (Reed et al., 2004). Neuroimaging studies demonstrate that different brain structures are activated by (inanimate) object images than by human body images (Kourtzi & Kanwisher, 2000;



Senior et al., 2000, 2002; Paulos et al., submitted).

Second, neuropsychological studies stressed differences between recognizing objects and recognizing (mostly naming) actions from pictures. For instance, Campbell and Manning (1996) reported the case of a patient with an optic aphasia who had a strong deficit in naming object pictures but a preserved capacity for naming actions. On the other hand, Manning and Warrington (1996) reported another case in which the patient is impaired for naming actions but the naming of objects is preserved.

Third is the fact that known biological constraints may sometimes prevail on low level perceptual constraints in the case of human body images. An example of such an effect was presented by Shiffrar and her colleagues (Shiffrar & Freyd, 1990, 1993; Chatterjee, Freyd & Shiffrar, 1996; Stevens et al., 2000) who compared the perceived trajectory of the apparent beta motion when the stimulus was a member of a body image or an abstract figure. The perceived trajectory of the apparent motion is generally the shortest track (i.e. a straight linear trajectory). These authors have shown that the perceived trajectory of the apparent motion can be seen as curved when biomechanical constraints of a moving body segment are taken into account. For instance, a hand is perceived as moving around the head until ISI gets too short. With the decreasing ISI, the shortest trajectory is perceived (i.e. through the head). The biomechanical representations have an effect only for large ISI allowing the complete processing of semantic factors.

In the present experiments, the question is to know whether or not a change in the orientation of human body during an action sequence would have an effect on the “group” response within a Ternus paradigm. These changes in the orientation of the human body between the four pictures will form in some conditions a “natural” sequence (later called *normal* mode) of an action such as running, a dance pirouette or a running forward flip into a hand stand, shortly called later somersault. In these conditions (Exp.1) it was expected that if the semantic factor (the action sequence) prevails on the structural factor of body orientation, the change in the frequency of “group” responses with the increase in ISI will be the same for the three actions in spite of the large differences in the orientations of the body. The experiment will be run in two designs. In a pure block design, only one action sequence is presented to one group of subjects for different ISI. In a mixed block design, the 3 action sequences will be presented within the same experimental session, each one for different ISI. The mixed

block design is used in order to evaluate the strength of a semantic effect, when some response biases are favoured.

The second and third experiments investigate the effect of a perturbation in the regularity of the sequence of action on the frequency of *group* responses as a function of the ISI. This was done for the pirouette and for the somersault. Two types of perturbations were used: an *inversion* of the two central pictures between the two images and a *translation* of the three pictures from their positions **A**, **B**, **C** to the position **B**, **C**, **D**. The *translate* mode will clearly favoured “group” responses likely to the point that the ISI could have no effect. In comparison, the *inversion* mode should favour more the element responses and at least reveals an effect of the ISI. Again, in the same logic as above, each experiment was run both in pure blocks (independent groups) and in mixed blocks.

### Stimuli and method

Images of the same trouser dressed woman seen laterally in full colours were drawn through the Poser 3 software by MetaCreations. Three actions were used: running, dance pirouette and somersault. Running is a simple translatory sequence. Pirouette is a rotation of the body around its vertical axis. Somersault is combination of a rotation in space around a horizontal axis and of a translation. Four pictures of each of these actions in a normal sequence (see Figure 1) were selected from an animation so that when sequentially presented in the animation a smooth movement of the body in action was perceived. The three first pictures (in position **A**, **B**, and **C**) formed image 1, the three last pictures (in position **B**, **C** and **D**) formed image 2.

FIGURE 1 ABOUT HERE

The experiments were run on a high resolution (1024 x 768 px) computer display (60 Hz non interlaced) controlled with an Inquisit program V1.32 by Millisecond Software. At every trial, two images (each one composed of three pictures in position **A-B-C** then **B-C-D**) alternate for 2.5 cycles. Each image is shown for 10 frames (166 ms) separated from the next one by an ISI. Six ISI were used: 1, 2, 3, 4, 5 and 6 frames (1 frame=16.6ms). In each experiment they were presented in a random order. The luminance background of the monitor was 10 cd/m<sup>2</sup>. It was the only source of light in the experiment room. Subjects have their head on a head and chin rest placed at 120 cm

from the monitor. The total centre to centre length of the four pictures (from position **A** to position **D**) was about  $13^\circ$  and the height of each picture is  $2.9^\circ$

when the body was vertical. The positions of the four pictures were regular; about  $4.3^\circ$  separate two neighbouring pictures. These dimensions were selected so as to allow a clear recognition of the action.

On each trial, subjects reported, following a two alternative choice procedure, whether they saw a motion of the extreme “element” ( $A \rightarrow D$ ) or a motion of the “group” ( $ABC \rightarrow BCD$ ). Reaction times were measured from the beginning of the second image. A pause of 1300 ms followed the subject’s response.

A training period preceded the main experiment. The Ternus phenomenon was first illustrated for the two extreme ISI with black closely spaced disks, then with the pictures of a woman standing upright with the same interposition distance and height as in the main experiments. In the training session, only the two extreme ISI (1 and 6 frames) were used. The responses were corrected at every trial (“element” for ISI=1; “group” for ISI=6). The training ended when subjects had at least 80% correct responses in 10 successive trials. Every subject reached the criterion in 2 to 4 blocks of 10 trials.

## **Experiment 1**

The first experiment compared the performances for the three selected actions: running, pirouette and somersault in pure block and in mixed block designs. The pictures appearing in positions **A**, **B**, **C** and **D** formed a regular sequence (mode *normal*). In pure block conditions, it was expected that the effect of the ISI on the frequency of “group” responses will be similar for the three actions in spite of the large differences in the orientation of the body for the pirouette and the somersault. Since the posture of the body, particularly its orientation in space, changes to a different extent between the three actions, the use of mixed blocks was introduced in order to check whether this structural factor can modulate a more semantic effect in introducing response biases.

## **Method**

Two experiments were run. In the first (Exp. 1p), the three actions were shown

in pure blocks of trials to three independent groups of 10 subjects. For each action (and each group of 10 subjects), the 6 ISI were randomly presented and each ISI appeared 20 times in a random order. In the second experiment (Exp. 1m), call mixed, the three actions were shown to the same 10 subjects. The 3 actions and the 6 ISI vary randomly from trial to trial. There were 10 trials for each of the 18 experimental conditions. As for the following experiments, subjects were naïve and none participates to more than one experiment.

## Results

The results of Exp. 1p (pure blocks) are shown in Figure 2a. The mean frequency of the “group” responses is respectively 0.50 for running, 0.47 for pirouette and 0.45 for somersault. However, these differences are not significant ( $F(2,27)<1$ ; ns). The frequency of “group” responses increases with the length of ISI ( $F(5,135)=78.2$ ;  $p<0.0001$ ). There is no significant interaction between the two factors ( $F(10,135)<1$ ; ns). In order to estimate the critical ISI (PSE for point of subjective equality) which is the limit between “element” and “group” responses (i.e. corresponding to 50% “group” responses), sigmoid psychometric functions were fitted to the results and the PSE estimated from this continuous function. The mean PSE for the three conditions is  $PSE=59.3$  ms.

### FIGURE 2 ABOUT HERE

The results of Exp 1m (mixed blocks) are shown in Figure 2b. The mean number of “group” responses for all the conditions is 0.40. It is larger for running (0.61) than for the pirouette (0.52) and for the somersault (0.31) ( $F(2,18)=130.5$ ;  $p<0.0001$ ). The frequency of the “group” responses increases with the ISI length ( $F(5,45)=104.7$ ;  $p<0.0001$ ). The two factors interacted ( $F(10,90)=5.41$ ;  $p<0.0001$ ). The PSE differ considerably for the three conditions: for the running condition  $PSE=32.8$ ms, for the pirouette condition  $PSE=43.2$ ms and for the somersault condition  $PSE>99.6$ ms (i.e. out of the range).

Concerning reactions times, in the pure block conditions, the mean RT is 1036ms; RT are respectively 852ms for running, 1102 for pirouette and 1153 for somersault ( $F(2,27)=2.58$ ;  $p<0.10$ ). On average, they increase with ISI ( $F(5,135)=19.19$ ;  $p<0.0001$ ). The interaction of the two factors is not significant ( $F(10,135)=1.38$ ; ns). In the mixed

block conditions, the mean RT is 704ms; RT are respectively 667ms for running, 721ms for pirouette and 723ms for somersault ( $F(2,18)=14.09$ ;  $p<0.0002$ ). On average, RT increases with ISI ( $F(5,45)=13.31$ ;  $p<0.0001$ ). The two factors interact ( $F(10,90)=2.67$ ;  $p<0.007$ ).

FIGURE 3 ABOUT HERE

## **Discussion**

In the pure block conditions, the three actions show similar results as far as the frequency of “group” responses is concerned. This is the main argument for concluding that the semantic factor (regularity of the action sequence) prevails on a structural factor such as the heterogeneity of the body orientation. However, subjects respond faster in the running condition than in the two other ones ( $F(1,27)=5.04$ ;  $p<0.04$ ). The increase in RT may reflect the necessity for more processing in structurally heterogeneous conditions. The mixed block conditions discriminate more the 3 actions with more “group” responses and shorter RT for running. “Group” responses are smaller for the pirouette than for the running and the smallest. In the later, the frequency of “group” responses does not reach 0.50 within the range of the used ISI. However, in the three conditions the frequency of “group” responses increases with the ISI. The mean RT is about the same in the two conditions pirouette and somersault. In fact, the large difference between the mean RT of the pure block and mixed block conditions suggest that subjects are not performing the same task in the two conditions. The discrimination between “element” and “group” is a difficult task in the present experimental conditions. In the mixed block conditions, most subjects tend to reduce (or to escape) this difficulty in responding much more frequently “group” in the running condition and much more frequently “element” in the somersault condition. As a result, the slope of the relationship between ISI and frequency of “group” responses is smaller in the mixed conditions than in the pure conditions. It can be suggested that in the mixed block conditions, the structural factor of the orientation of the body within and between the two images reduces the effect of the semantic factor.

## **Experiment 2 (pirouette)**

The second set of experiments concentrates on the pirouette sequence. We investigate the effect of perturbing the regular sequence on the frequency of “group”

responses. This was done again in pure block conditions (Exp. 2p) and in mixed block conditions (Exp. 2m). The first perturbation was a mode of *inversion* of the pictures in position **B** and **C** between the first and second image. The second perturbation was a mode *translation* of the same three pictures from positions **ABC** to position **BCD**. The *inversion* condition was expected to generate more “element” responses than the normal condition. However, it should be mentioned that this inversion generates a local apparent movement since the two central pictures are not identical from one frame to the next one (see Figure 1). The *translate* condition was expected to generate more “group” responses because the three same pictures slip for one position between images 1 and 2. In fact, such a procedure produces an apparent beta movement between the two images. In the pure block conditions, two groups of 10 new subjects were run, one with the *inverse* mode, the other one for the *translate* mode. Each subject of each group sees the 6 ISI conditions 20 times. The results of the *normal* mode (see Exp 1p) are included for the ANOVA. The experiment 2m was run in mixed block conditions on another group of 10 new subjects. The three modes of presentation of the action were the *normal* one (as in experiment 1), the *inverse* and the *translate* ones. In a single experimental session, each subject was shown 18 conditions mixing randomly 6 ISI and the 3 modes. Each of these conditions was shown 10 times.

## Results

The results are shown in Figure 4. In the pure conditions (Figure 4a) the mean frequency of “group” responses is 0.52. It does not vary significantly between the three modes ( $F(2,27)=1.47$ ; ns). ISI have a significant effect  $F(5,135)=28.49$ ;  $p<0.0001$ ). Frequency of “group” responses increases significantly for each mode with the ISI;  $F(1,9)$  is respectively 31.33 in normal mode, 7.44 in the inverse mode, and 5.89 in the translate mode. There is a significant interaction between the two factors ( $F(10,135)=3.72$ ;  $p<0.0002$ ). As a result of this interaction, the PSE are different for the 3 modes; for the *normal* mode  $PSE=61.9ms$ , for the *inverse* mode  $PSE=54.3ms$  and for the *translate* mode  $PSE=27.8$ .

FIGURE 4 ABOUT HERE

The mean frequency of “group” responses in the mixed conditions (Figure 4b) is 0.49. The 3 modes give different results; on average that frequency was 0.29 in the *normal*

mode, 0.64 in the *inverse* mode and 0.83 in the *translate* mode ( $F(2,18)=22.24$ ;  $p<0.0001$ ). There is a significant effect of ISI ( $F(5,45)=5.87$ ;  $p<0.0003$ ). Frequency of “group” responses increases with the ISI, but the regression is not significant in every mode; normal ( $F(1,9)=7.00$ ;  $p<0.03$ ), inverse ( $F(1,9)=4.39$ ;  $p<0.07$ ), translate ( $F(1,9)=2.58$ ; ns). There is no significant interaction between the two factors ( $F(10,90)<1$ ; ns). The slopes of the functions relating the ISI to the frequency of “group” responses is very small. In fact, the PSE could not be estimated since they were out of the range of used ISI.

The mean RT in pure block conditions is 1056 ms. The modes are not significantly different ( $F(2,27)<1$ ). There is a significant effect of the ISI ( $F(5,35)=6.12$ ;  $p<0.00004$ ). The regression effect is not significant in every mode; normal ( $F(1,9)=22.07$ ;  $p<0.001$ ), inverse ( $F(1,9)=0.6$ ; ns), translate ( $F(1,9)=3.72$ ;  $p<0.09$ ). The interaction between mode and ISI is marginally significant ( $F(10,135)=1.71$ ;  $p<0.09$ ).

#### FIGURE 5 ABOUT HERE

The mean RT in mixed block conditions is 659 ms. The modes have a marginally significant effect ( $F(2,18)=3.02$ ;  $p<0.08$ ) as well as the interaction between modes and ISI ( $F(10,90)=1.82$ ;  $p<0.07$ ). In fact, only in the normal mode, RT increases in a slightly significant way with ISI ( $F(1,9)=4.34$ ;  $p<0.07$ ).

## Discussion

In the pure block conditions, the mean frequency of “group” responses is equivalent for the 3 modes of presentation. However, while in the *normal* mode, the slope of the relationship between ISI and frequency is steeper, that variation is much smaller for the two other modes (*inverse* and *translate*). In fact, in the *translate* mode, only the shortest ISI generates more “element responses; beyond that ISI, that factor has nearly no effect. The interaction between ISI and modes together with the values of the PSE suggested a hierarchy between these three modes. In the low range of ISI, subjects produce more “group” responses for the *translate* mode than for the *inverse* mode, and more in the later than in the *normal* mode. That hierarchy is strongly amplified in the mixed conditions and applies to the entire range of ISI. Clearly, the two modes *inverse* and *translate* perturb gravely the Ternus phenomenon. As predicted, the translation of the three pictures increases the frequency of the “group” responses up

to the point that ISI has practically no effect. Contrary to our expectation, the *inverse* mode produces more “group” responses than the *normal* one. The local apparent movement of the two central pictures may be responsible for this result: all three pictures of one image seem to move from that image to the next one.

As in the first experiment, strong response biases are observed in the mixed block design as compared to the pure block design. Likely, subjects make a different task in the mixed and in the pure conditions as attested by the large difference in RT. In the mixed block conditions, their responses appeared more influenced by the mode of presentation than by the ISI. In fact, their mean results suggested that they are close to a (uninstructed) categorisation task between the modes decreasing the effect of the ISI.

### **Experiment 3 (somersault)**

This third experiment concentrates on the somersault sequence. We investigate the effect of perturbing the regular sequence of a somersault on the frequency of “group” responses. As in experiment 2, two modes of perturbation are used. In the *inverse* condition, the position of the pictures between positions **B** and **C** is inverted between the two images. In the *translate* condition, the same three pictures are drifted from position **ABC** to position **BCD** between the two images. Since the changes in the posture orientations of the human body are larger for this action than for the pirouette, we expected larger effects of the experimental conditions. The experiments were run both in a pure block design and in a mixed block design.

### **Results**

The results are shown in Figure 6. The mean frequency of “group” responses in the pure conditions (Figure 6a) is 0.56. The mode of presentation has a significant effect ( $F(2,27)=9.25$ ;  $p<0.0009$ ) with a mean frequency of “group” responses of 0.45 in *normal* mode, 0.56 in the *inverse* mode and 0.67 in the *translate* mode. There is a significant effect of the ISI ( $F(5,135)=11.73$ ;  $p<0.0001$ ), and a significant effect of the interaction of the two factors ( $F(10,135)=9.67$ ;  $p<0.0001$ ). The frequency of “group” responses increases significantly with the ISI in the normal mode ( $F(1,9)=47.13$ ;  $p<0.00008$ ) and in the inverse mode ( $F(1,9)=5.19$ ;  $p<0.05$ ), but is not significant in the translate mode. The PSE are different; for the *normal* mode PSE=62ms, for the *inverse*



mode PSE=45ms. It cannot be estimated for the *translate* mode, since at each ISI the frequency of “group” responses is larger than 0.50.

#### FIGURE 6 ABOUT HERE

The mean frequency of “group” responses in the mixed condition is 0.57 (see Figure 6b). The mode of presentation has a significant effect ( $F(2,18)=14.68$ ;  $p<0.0002$ ) with a smaller frequency in the *normal* mode (0.20), a medium one in the *inverse* mode (0.65) and a higher one in the *translate* mode (0.87). Globally, there is no significant effect of the ISI ( $F(5,45)=1.89$ ; ns). However, the increase in the frequency of “group” responses is significant in the normal mode ( $F(1,9)=5.16$ ;  $p<0.05$ ). No significant interaction between the two factors ( $F(10,90)<1$ ; ns) is observed.

Reaction times are shown in Figure 7. The mean RT is 1040ms in the pure conditions (Figure 7a). The mode of presentation has no significant effect ( $F(2,27)=1.76$ ; ns). RT varies with ISI ( $F(5,135)=1379$ ;  $p<0.0001$ ) and shows an interaction with the modes ( $F(10,135)=2.22$ ;  $p<0.02$ ). RT increases with ISI in the normal mode ( $F(1,9)=22.95$ ;  $p<0.001$ ) and in the inverse condition ( $F(1,9)=6.20$ ;  $p<0.04$ ). The mean RT is 598ms in the mixed conditions (Figure 7b). Only the mode of presentation has a marginally significant effect ( $F(2,18)=3.25$ ;  $p<0.06$ ). However, RT increases with ISI only in the normal mode ( $F(1,9)=4.82$ ;  $p<0.06$ ).

#### FIGURE 7 ABOUT HERE

### Discussion

In the pure block conditions, as in Experiment 2, the functions relating the frequency of the “group” responses and the ISI differ according to the mode of presentations. The function is steeper in the *normal* mode than in the *inverse* one. The ISI has no significant effect in the *translate* mode. The hierarchy in the frequency of “group” responses in the low range of ISI is the same in pure block conditions and in mixed block one: less “group” responses are observed in the *normal* mode than in the *inverse* one, and less in the later than in the *translate* mode. Then, as predicted, the *translate* mode increases the frequency of “group” responses which is insensitive to the effect of ISI. The inversion of the two central pictures has an intermediate effect: it increases the frequency of “group” responses as compared to the *normal* mode, but it shows some sensitivity to the ISI. Again, strong differences are observed between the pure block condition and the mixed block one. These differences are obtained both in

the frequency of “group” responses and in the RT. The later are much shorter in the mixed condition, suggesting a change in the task done by the subjects.

### **General discussion**

In a pure block design, the three action sequences (running, pirouette and somersault) lead to similar results: the frequency of “group” responses increases with the ISI with a PSE = 59.3 ms. This result happens in spite of the large structural differences in the body postures between neighbouring positions for the three actions. This result should be considered in comparison with the results reported in a previous paper (Bonnet & Paulos, *submitted*) showing large effects of the structural differences between the pictures. In the present experiments, it may be concluded that the semantic factors have prevailed on the structural ones. In pure block conditions, RT are long, suggesting a deeper processing of the stimuli and presumably of their semantic content. In this respect, since the pictures are structurally more heterogeneous for the pirouette and the somersault than for the running, it is not surprising that RT are longer in the two former than in the later.

Perturbating the action sequence increases the frequency of “group” responses for both the pirouette and the somersault. Comparing the *inversion* mode for the pirouette and the somersault in pure block conditions leads neither to a significant difference nor to an interaction with the ISI ( $F < 1$ ). However, in every case, the frequency of the “group” responses increases with ISI. The same conclusion holds for that comparison in mixed block conditions. The absence of difference between pirouette and somersault in pure as well as in mixed block conditions for the *inversion* mode suggests that the semantic content plays a major role in these conditions, as it is the case for the *normal* mode. Likely, the local apparent movement disrupts the sequence of the depicted action. The fact that the semantic content plays a major role for this action is also suggested by the fact the RT are, on average, longer in this mode than in the *normal* mode since semantic processing requires longer processing time.

If the above interpretations are correct, with the *translation* mode we must find differences between the two actions, since presumably, structural factors are dominant in this condition. First the fact that “group” responses are dominant in this mode, with almost no effect of ISI, is not surprising. In fact, such a mode can be considered as an

apparent beta movement of three pictures. The fact that subjects were instructed to discriminate between “element” and “group” responses is sufficient to explain why, in pure block condition, the frequency of “group” responses is not at the maximum. It reaches practically that maximum in mixed block conditions. We then expect that structural factors play a larger role in this mode. This is the case since the frequency of “group” responses is significantly larger with the somersault than with the pirouette in the pure block conditions ( $F(1,18)=4.39$ ;  $p<0.05$ ), and these two modes show an interaction with the effect of the ISI ( $F(5,90)=4.53$ ;  $p<0.001$ ). These differences between the two actions in that mode are consistent with an effect of structural factors in comparison of what happens for the two other modes. In mixed blocks, the two actions did not differ significantly ( $F<1$ ). They show a moderate effect of ISI ( $F(5,45)=2.93$ ;  $p<0.03$ ) with the frequency of “group” responses being on average 0.80 for the three lowest ISI and 0.89 for the three largest ones. No interaction between the action and the ISI is observed.

In each of the 3 experiments, the mean frequency of “group” responses is about the same between the pure blocks and the mixed block. But their distribution among the factors is different. In the mixed block experiments, all the effects observed in pure blocks are amplified, except the effect of the ISI. Mixing the conditions make subjects to be more influenced by the category of the stimuli. In the first experiment, they give more “group” response for the running than for the pirouette and finally few “group” responses for the somersault. Their shorter RTs in the mixed condition suggest a strategy associating “group” response to running and “element” responses to somersault. In the second and third experiments, that type of strategy is drastically amplified to the point that the ISI has no significant effect for the somersault. Schematically, it is suggested that associating “group” responses to the *translate* condition and “element” responses to the *normal* condition facilitate their task and reduces their RTs but make them to neglect the demanded main task to discriminate “element” motion from “group” motion.

Comparing the perturbation conditions used in the present experiment with the heterogeneous conditions of running used in the previous experiment (Bonnet & Paulos, submitted) leads to suggesting several conclusions. In the previous experiment the perturbation was introduced only in the picture in position **D** which either changes its

facing (mirror condition) or changes both its facing and orientation (up-side-down down / inverted). These two transformations disrupted the grouping perception and increase the frequency of “element” responses. In the present experiments, the perturbations were done on pictures on position **B** and **C**. This type of change was chosen in order to maintain the structural identity of the pictures between the two images. They induced larger frequencies of “group” responses in the lower range of the ISI. Such a result was expected for the *translate* condition in which the ISI has nearly no effect. A rigid apparent motion of the three pictures is perceived. Such a phenomenon does not appear when the four picture elements are identical as we verified in the training sessions. In the *inverse* condition, the frequencies of “group” responses increase with ISI, but even at the shortest one this frequency is close to 0.5. The *inverse* perturbation induces an apparent motion in each picture which is likely responsible for the increase in “group” responses. Furthermore, the inversion does not really break to action sequence as the perturbations introduced in the previous experiment. In their post experimental comments, subjects never report such an irregularity of the action.

While in the previous study, structural factors prevailed clearly on semantic ones on the frequency of “group” responses, in the present one, some of the results are in favour of a semantic effect. The differences in the exposure times of the figure (166 ms vs. 83 ms) and in their spatial dimensions were introduced in order to allow more processing of the semantic characteristics of the stimuli. In spite of the structural differences in the orientation of the body between the three actions, they generate similar frequency of responses in pure block condition (Exp. 1p). Likely, the regularity of the sequence of the posture is a basic factor for this result as seen in the results of experiment 2p and 3p. As a general conclusion, the present experiments show that in some restricted conditions, semantic factors can influence a perceptual phenomenon, here the Ternus effect, known to be very dependent on structural factors. While, in the apparent motion experiments of Shiffrar and Freyd (1990, 1993), Chatterjee, Freyd and Shiffrar (1996) and Stevens et al. (2000), subjects have to directly report their phenomenal experience, we choose here to look for indirect evidence. This was done in order to reduce the risk of more or less conscious cognitive strategies. The results obtained in mixed blocks conditions show that such strategies may occur. Now, the present experiments left open

many questions which should be investigated further. Are the semantic effects specific of human body images or do they apply to any animate ones? And more, are these semantic effects restricted to images of animate stimuli or do they extend to dynamic objects such as vehicles?

### References

Allais D., Lorenceau J. (2002) Perceptual grouping in the Ternus display: evidence for an “association field” in apparent motion. *Vision Research*, 42, 1005-1016.

Bonnet C, Paulos C (\*\*\*\*) Structural and semantic factors in the Ternus paradigm (submitted)

Braddick O.J. (1980) Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London, Series B*, B290, 137-151.

Braddick O.J., Adlar A. (1978) Apparent motion and the motion detector. In J.C. Armington, J. Krauskopf, B.R. Wooten (Eds.) *Visual Psychophysics and Psychology*. New York, Academic Press.

Breitmeyer B. G., Ritter A. (1986a). The role of visual pattern persistence in bistable stroboscopic motion, *Vision Research*, 26, 1801-1806.

Breitmeyer B. G., Ritter A. (1986b). Visual persistence and the effect of eccentric viewing, “element” size and frame duration on bistable stroboscopic motion percepts. *Perception and Psychophysics*, 39, 275-280.

Burt P.; Sperling, G. (1981) Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review*, 88, 171-195.

Chatterjee S. H., Freyd J. J., Shiffrar M. (1996). Configural processing in the perception of apparent biological motion. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 916-929.

Dawson M.R.W. (1991). The How and Why of What Went Where in Apparent Motion: Modeling Solutions to the Motion Correspondence Problem. *Psychological Review*, 98, 569-603.

Dawson M, Di Lollo V. (1990) Effects of adapting luminance and stimulus contrast on the temporal and spatial limits of short-range motion. *Vision Research*, 30, 415-429.

Dawson M.R.W., Wright R.D. (1989). The consistency of “element” motion affects the visibility but not the direction of apparent movement. *Spatial Vision*, 4, 17-29.

He Z.J., Ooi T.L (1999). Perceptual organization of apparent motion in the Ternus display. *Perception*, 28, 877-892.

Kolers P.A. (1972). *Aspects of motion perception*. Oxford, Pergamon Press.

Kramer P., Rudd M. (1999). Visible persistence and form correspondence in Ternus apparent motion. *Perception & Psychophysics*, 61, 952-962.

Kramer P., Yantis S. (1997). Perceptual “group”ing in space and time: Evidence from the Ternus Display. *Perception & Psychophysics*. 59, 87-99.

Oyama T., Simizu M., Tozawa J. (1999) Effects of similarity on apparent motion and perceptual “group”ing . *Perception*, 28, 739 – 748.

Petersik, J. T. (1989). The two-process distinction in apparent motion. *Psychological Bulletin*, 106, 107-127.

Sekuler R. (1996) Motion perception: a modern view of Wertheimer's 1912 monograph.

*Perception*, 25, 1243-1258.

Shiffrar M, Freyd J J, (1990) Apparent motion of the human body. *Psychological Science*, 1, 257-264.

Shiffrar M, Freyd J J, (1993) Timing and apparent motion path choice with human body photographs. *Psychological Science*, 4, 379-384.

Skottun B.C. (2001) On the use of the Ternus test to assess magnocellular function. *Perception*, 30, 1449-1457.

Stevens J.A., Fonlupt P., Shiffrar M., Decety J. (2000). New aspects of motion perception: selective neural encoding of apparent human movements. *NeuroReport*, 11, 109-115.

Ternus J. (1926). Experimentelle untersuchungen über phänomenale identität. *Psychologische Forschung*, 7, 81-136.

### **Figures captions**

Figure 1: The four pictures of the three action sequences used in the experiments in their positions **A**, **B**, **C** and **D** (from left to right).

Figure 2: Frequency of “group” responses as a function of the ISI (in frames) in the pure block conditions (left, open symbols) and in the mixed block conditions (right, black symbols) in the *normal* mode for the three actions. Vertical bars are the standard errors. Psychometric sigmoid functions were fitted to the results in order to estimate points of subjective equality (PES, see text).

Figure 3: Reaction times (“group” and “element” responses mixed) as a function of the ISI (in frames) in the pure block conditions (left, open symbols) and in the mixed block conditions (right, black symbols) in the *normal* mode for the three actions. The ordinate scales are different between pure and mixed blocks.

Figure 4: Frequency of “group” responses as a function of the ISI (in frames) in the pure block conditions (left, open symbols) and in the mixed block conditions (right, grey symbols) in the *normal*, the *inverse* and the *translate* modes for the pirouette. Vertical bars are the standard errors. Psychometric sigmoid functions were fitted to the results in order to estimate points of subjective equality (PES, see text).

Figure 5: Reaction times (“group” and “element” responses mixed) as a function of the ISI (in frames) in the pure block conditions (left, open symbols) and in the mixed block conditions (right, grey symbols) in the *normal*, the *inverse* and the *translate* modes for the pirouette. The ordinate scales are different between pure and mixed blocks.

Figure 6: Frequency of “group” responses as a function of the ISI (in frames) in the pure block conditions (left, open symbols) and in the mixed block conditions (right, grey symbols) in the *normal*, the *inverse* and the *translate* modes for the somersault. Vertical bars are the standard errors. Psychometric sigmoid functions were fitted to the results in order to estimate points of subjective equality (PES, see text).

Figure 7: Reaction times (“group” and “element” responses mixed) as a function of the ISI (in frames) in the pure block conditions (left, open symbols) and in the mixed block conditions (right, grey symbols) in the *normal*, the *inverse* and the *translate* modes for the somersault. The ordinate scales are different between pure and mixed blocks.

### **Acknowledgements**

Acknowledgements are due to Emilie Pirra and to Raaj Seereekissoon for carefully running the experiments.



Figure 1:

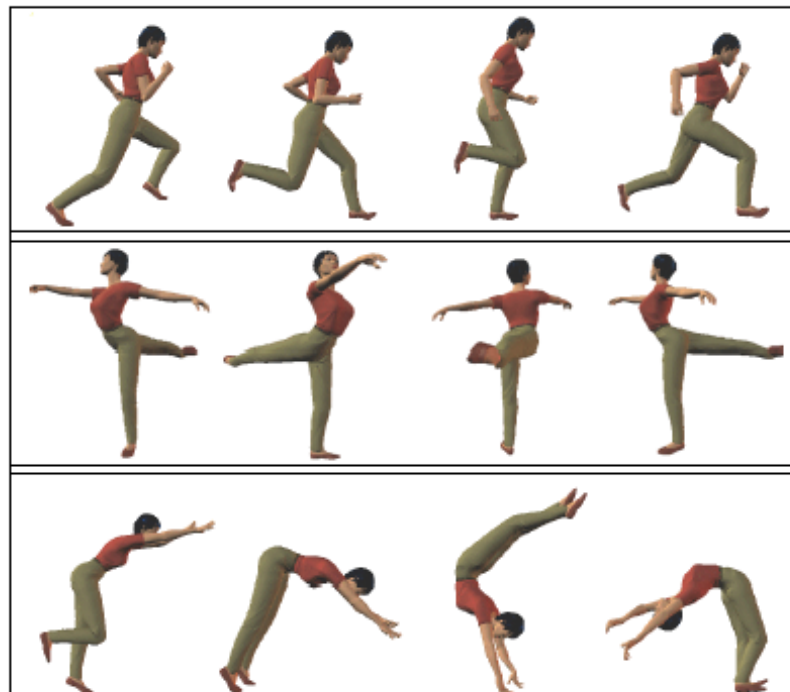


Figure 2:

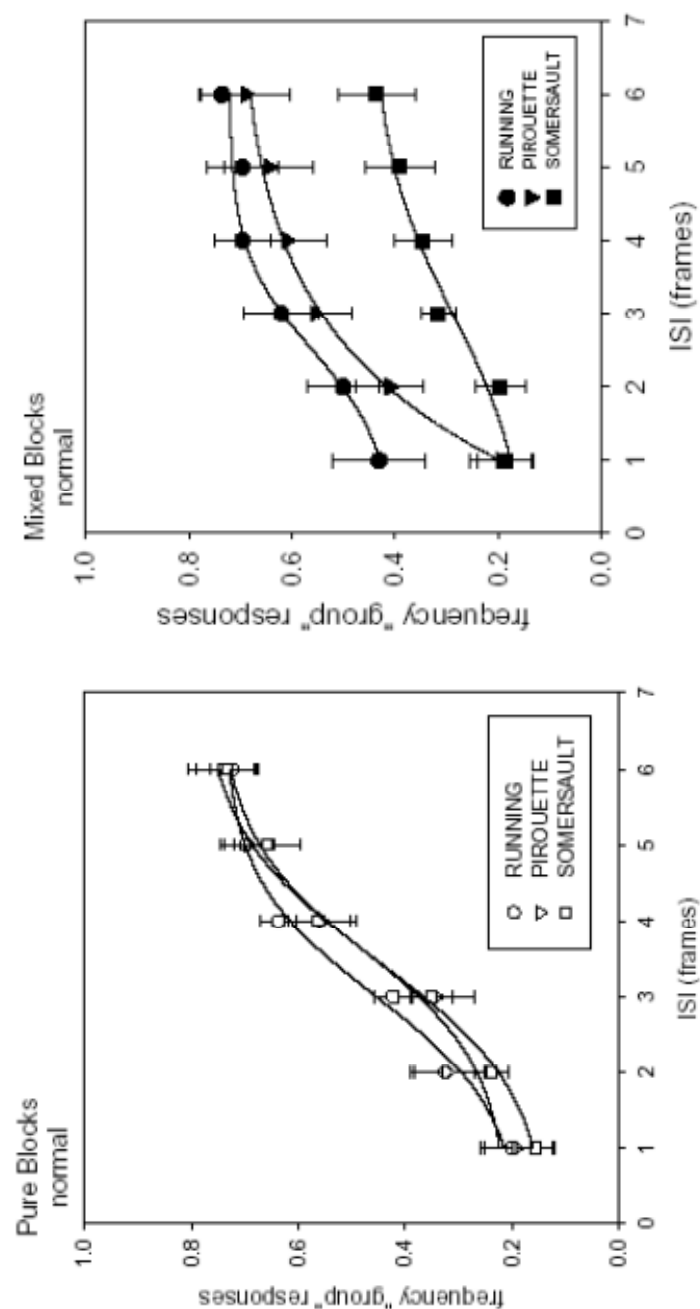


Figure 3:

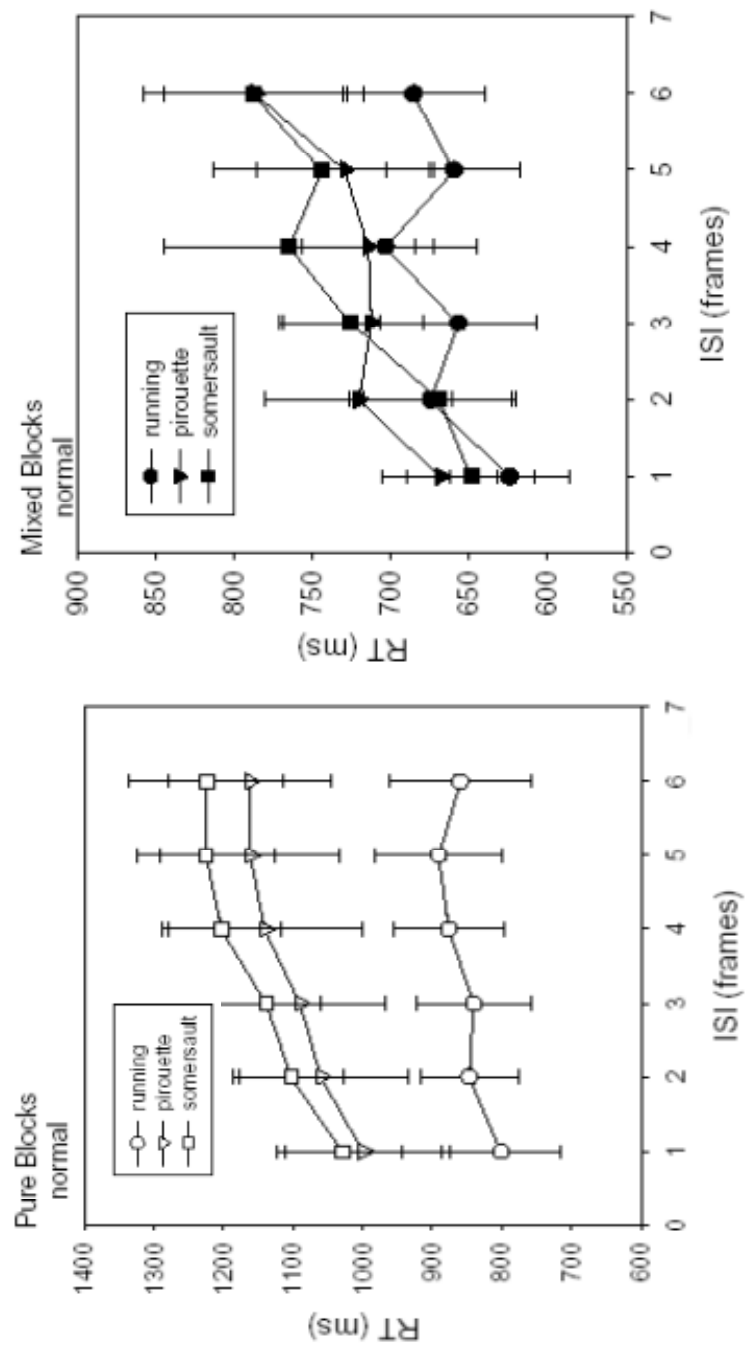


Figure 4:

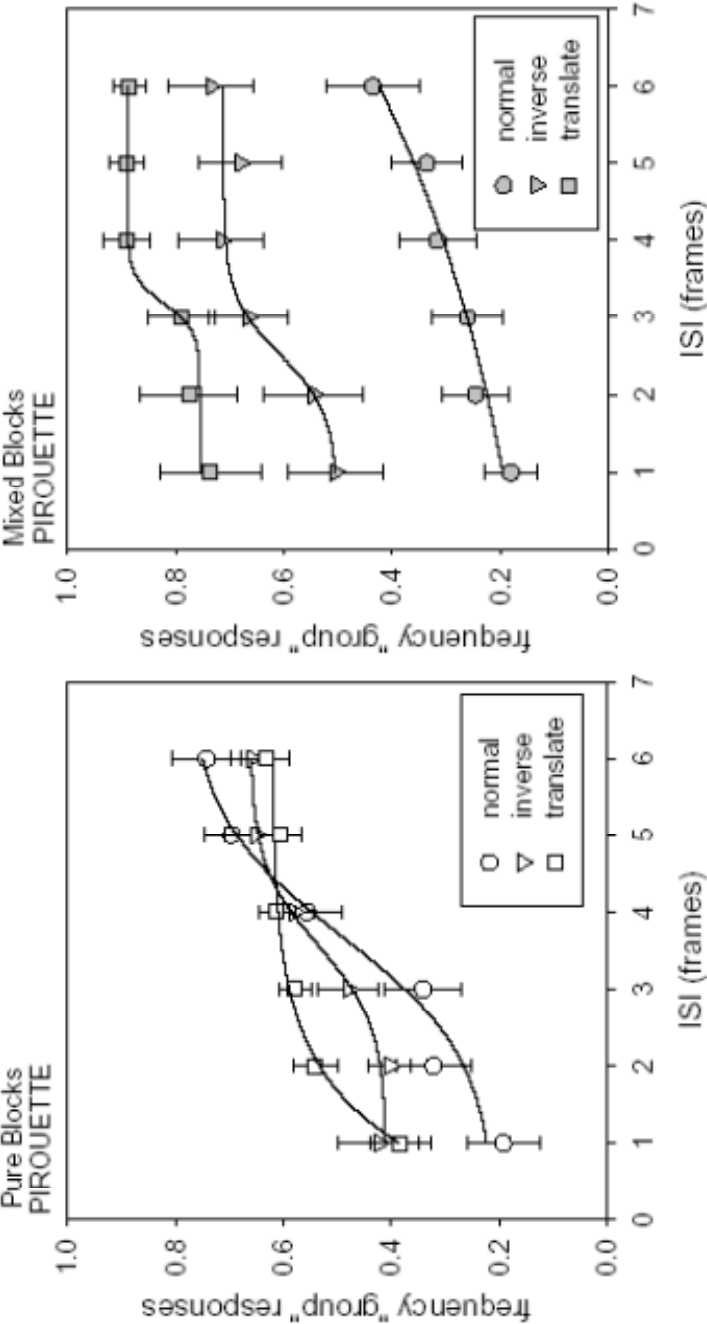


Figure 5:

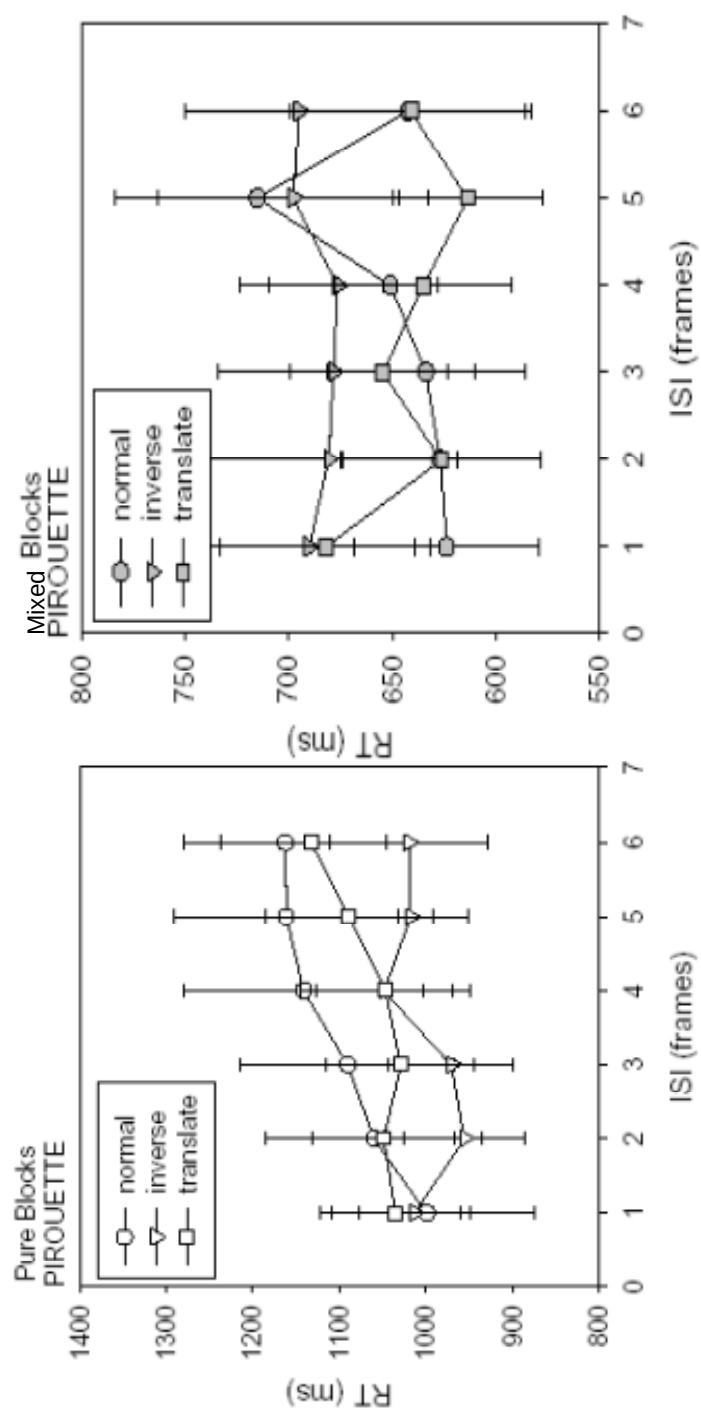


Figure 6:

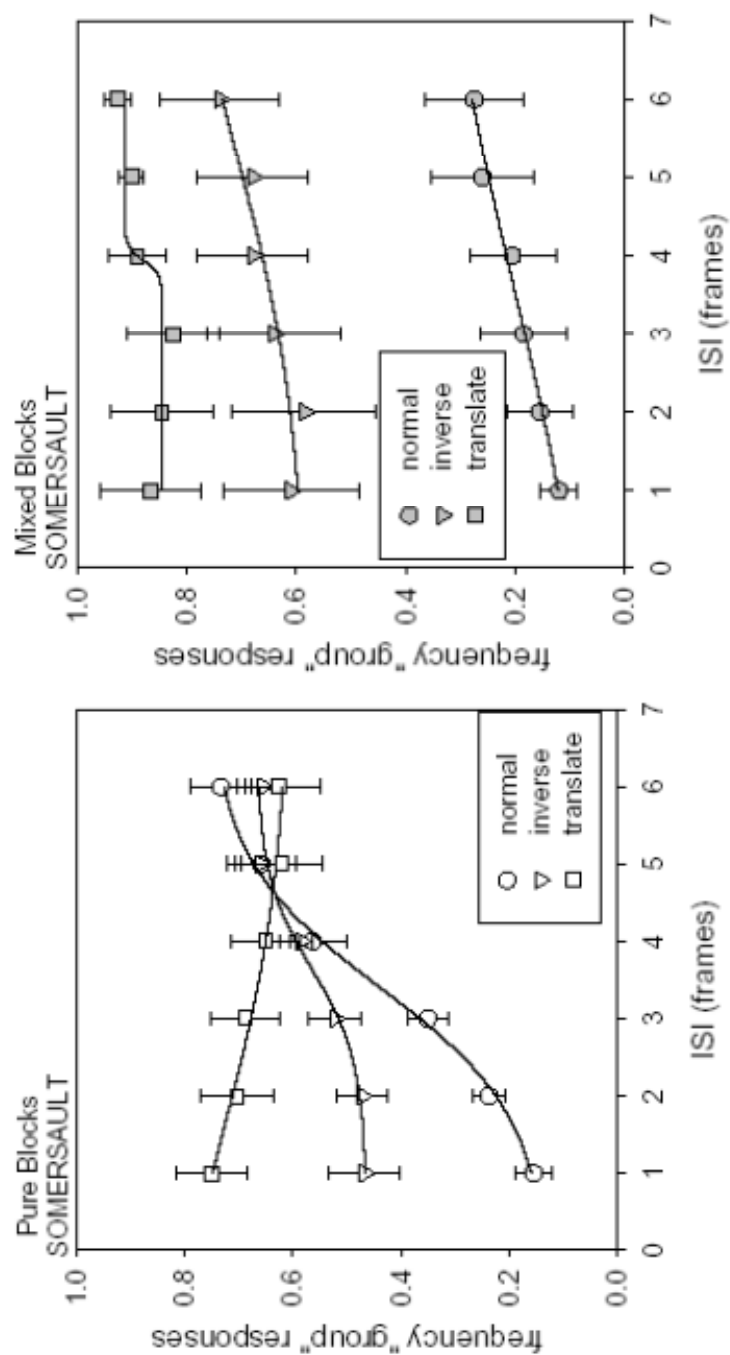


Figure 7:

