



**HAL**  
open science

## Les signatures neurocognitives du changement de point de vue pour la reconnaissance des objets 3D

Héla Chouk, Michel-Ange M.-A. Amorim

### ► To cite this version:

Héla Chouk, Michel-Ange M.-A. Amorim. Les signatures neurocognitives du changement de point de vue pour la reconnaissance des objets 3D. Deuxième conférence française de Neurosciences Computationnelles, "Neurocomp08", Oct 2008, Marseille, France. hal-00331573

**HAL Id: hal-00331573**

**<https://hal.science/hal-00331573>**

Submitted on 17 Oct 2008

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

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

# Les signatures neurocognitives du changement de point de vue pour la reconnaissance des objets 3D

Héla Chouk

Université Paris-Sud, France

Michel-Ange Amorim

Université Paris-Sud, France

Laboratoire « Contrôle Moteur et Perception »

UPRES EA 4042 de l'Université Paris-Sud 11

91405 Orsay Cedex, France.

[Hela.chouk@u-psud.fr](mailto:Hela.chouk@u-psud.fr) / [michel-ange.amorim@u-psud.fr](mailto:michel-ange.amorim@u-psud.fr)

## RESUME

La littérature psychophysique sur la reconnaissance d'objets tridimensionnels complexes (de type arrangement de cubes), a montré qu'afin d'indiquer si deux objets présentés simultanément ou successivement sous des orientations différentes sont identiques ou non, l'observateur engageait un processus de rotation mentale de l'objet test afin de l'aligner à l'objet de référence. D'autre part, l'identification des postures humaines présentées sous différentes orientations active des structures cérébrales impliquées dans la perception sociale et semble être médiée par une connaissance implicite des contraintes biomécaniques du corps humain. De plus, notre connaissance du corps facilite sa rotation mentale autant que sa reconnaissance. Le thème général de notre étude est la compréhension du rôle fonctionnel (et des mécanismes cérébraux) du point de vue pour la reconnaissance des objets 3D et la spécificité du corps humain en tant qu'objet. A ce jour, aucune étude en imagerie cérébrale n'avait comparé la reconnaissance des postures humaines à celle de configurations spatiales équivalentes de cubes ou d'objets plus familiers (une lampe de bureau). Notre expérimentation met en évidence au moyen de la MEG l'organisation spatio-temporelle des activités corticales dans une tâche de reconnaissance visuelle impliquant l'anticipation des conséquences visuelles d'un changement de point de vue.

## MOTS-CLES

Objets tridimensionnels complexes/ Rotation mentale / Reconnaissance /Point de vue/ Neuro-imagerie / Magnétoencéphalographie.

## 1. Introduction

Dans la littérature psychophysique, on distingue deux familles de modèles pour la reconnaissance

des objets tridimensionnels. Les modèles dits « de gabarit » [1] dans lesquels la représentation des objets est faite sous forme de patterns visuels prototypiques dépendants du point de vue. Le cortex occipital latéral (LOC) et le gyrus fusiforme sont identifiés comme des structures dépendantes du point de vue pour la reconnaissance des objets [2, 3]. En outre, le gyrus fusiforme et plus précisément les portions du gyrus fusiforme qui se trouvent dans le LOC, sont sensibles aux manipulations sémantiques [4] et répondent plus fortement à des objets qui ont du sens qu'à des objets non réalistes [5, 6]. A l'opposé, pour les modèles dits de « représentation structurale » [7, 8], la reconnaissance des objets est un processus dans lequel l'image d'entrée est représentée sous forme d'arrangement géométriques simples (cylindre, cônes, briques etc...), que Biederman appelle « géons ». Selon lui, il existe une description structurale, un agencement de géons qui permet de lire les objets. Cette théorie se base sur la détection de certaines propriétés intrinsèques de l'objet qui sont généralement invariables, peu importe la qualité de l'image ou le point de vue.

Qu'en est-il alors du corps humain ? Est-il un objet comme les autres ?

Selon Marr & Nishihara [9], le corps humain est un objet 3D non spécifique, puisque il est considéré comme un assemblage particulier de cylindres ou de cônes tronqués comme dans le modèle biomécanique de Hanavan [10]. A contrario, les travaux de Parsons [11], indiquent que lorsqu'on doit juger la latéralité d'une partie du corps, certaines aires motrices cérébrales ou des régions connexes qui font appel à ces aires s'activent suggérant l'implication d'un mécanisme d'émulation [11] spécifique à la reconnaissance du corps. En effet, le temps de décision pour reconnaître une main droite d'une main gauche est

plus lent lorsque la difficulté biomécanique de la posture est difficile.

Par ailleurs, Christou et al. [12] ont montré que connaître son point de vue sur un objet facilite son identification. Si la géométrie des formes est codée sans avoir recours au contexte spatial environnant, alors aucune différence ne devrait être observée dans la capacité à identifier de nouvelles vues. Les résultats montrent que la présence d'un environnement visuel n'élimine pas la dépendance au point de vue mais elle réduit de manière significative les erreurs. Par ailleurs, il a été montré que si un avatar indique le futur point de vue sur une scène visuelle, la détection des changements dans cette scène est facilitée [13].

Le thème général de notre étude est la compréhension du rôle fonctionnel (et des mécanismes cérébraux) du point de vue pour la reconnaissance des objets 3D complexes et la spécificité du corps humain en tant qu'objet.

La magnétoencéphalographie (MEG) nous a permis de mettre en évidence l'organisation spatio-temporelle des activités cérébrales lors d'une tâche de reconnaissance visuelle.

## 2. Méthode et résultats

### 2.1 Tâche

Les participants comparent deux vues d'un objet présentées successivement, afin d'indiquer si c'est le même objet ou deux objets différents (un objet et sa version miroir). Nos stimuli inspirés des travaux d'Amorim et al. [14] représentent des configurations spatiales équivalentes de postures humaines (objet très familier), de lampe articulée de bureau (objet familier) et d'arrangements spatiaux de cubes façon Shepard & Metzler [15] (objet non familier).



Figure 1. Exemple de configurations spatiales équivalentes.

### 2.2 Procédure

Tout d'abord, après un point de fixation présenté 1s, un indice spatial (amorce) est présenté 2s afin d'indiquer la position virtuelle de l'observateur sur l'objet au moment de la présentation de la vue test. Il s'agit d'un personnage qui fait face à une sphère vide représentée sous forme filaire à l'intérieure de laquelle l'objet est censé figurer. Après 1s de point de fixation, la vue de référence sur l'objet est exposée 4s sous son orientation initiale. Le sujet doit transformer mentalement l'objet sur la base des informations initiales de l'indice spatial afin

d'anticiper les conséquences visuelles du changement de point de vue sur l'objet. Après 1s de point de fixation, la vue test est présentée pendant 3s. Le sujet indique aussi vite que possible, si l'objet est identique ou différent de celui à la vue de référence. Les vues tests varient par une rotation autour de l'axe vertical de 0°, 60°, 120°, dans le sens horaire ou anti-horaire, d'un essai à l'autre.

12 individus droitiers (moyenne d'âge = 27ans) ont pris part à cette expérimentation. Ils ont réalisés trois blocs de 96 essais correspondant chacun à un type de stimulus différent (posture, lampe, cubes) ; l'ordre des blocs étant contrebalancé au travers les sujets.

### 2.3 Matériel

L'expérience s'est déroulée au centre MEG-EEG de l'hôpital de la Pitié Salpêtrière à Paris. La MEG permet le recueil non invasif des champs magnétiques directement engendrés par les neurones qui sont orientés de façon tangentielle à la surface du scalp. Cet appareil est équipé de 151 capteurs magnétiques assurant une couverture totale de la tête. Les champs magnétiques enregistrés par le système MEG sont dus à un ensemble de courants intracellulaires qui circulent au niveau des neurones situés dans les sillons du cortex.

### 2.4 Résultats

#### 2.4.1 Données comportementales

Une analyse de variance à mesures répétées sur les temps de réaction (TRs) à la condition « objets identiques » a été effectuée sur la base de plan d'analyse suivant : 3 Types de stimuli (Posture, cube, Lampe) \* 3 Angles (0°, 60°, 120°). Les résultats indiquent un effet principal de l'Angle ( $F(2,20) = 143.23, p < .05$ ) à la vue test (vue 2). Toutefois, les TRs trouvés au cours de cette expérimentation, sont très rapides ( $< 1.2s$ ) et plus courts que ceux indiqués dans l'étude de rotation mentale d'Amorim et al. [14], pour la condition 0° (aux alentours de 1.5s). Ce dernier point suggère que tout le travail cognitif de transformation spatiale s'est réalisé au cours des 4s de présentation de la vue de référence (vue 1) et qu'au moment de la présentation de la vue test, les sujets n'ont pas fait de rotation mentale.

#### 2.4.2 Les données MEG

Toutes les analyses de données MEG ont porté sur les activités cérébrales aux réponses correctes. Une première analyse est réalisée sur les 500ms post onset de la vue 1 avec des fenêtres temporelles de 50ms se recouvrant toutes les 25ms (Facteur "temps"). Une Anova à mesures répétées a été réalisée sur les valeurs RMS (root mean square) d'intensité du champ magnétique en femto-tesla, en regroupant les capteurs MEG par région corticale (occipitale, pariétale, centrale, frontale et temporale). Les degrés de liberté des Anovas ont été corrigés (Huynh-Feldt) pour la non sphéricité

des données. Le plan d'analyse est : 3 Types de stimuli (Posture, Cube, Lampe) \* 3 Angles (0°, 60°, 120°) \* 2 Hémisphères \* 5 Régions \* 20 Temps, avec un facteur Objet test (Identique vs différent) supplémentaire pour les réponses magnétiques évoquées à la vue 2. De plus l'évolution de l'activité cérébrale au cours du temps a été comparée au niveau initial t1 par des tests post-hocs (Tukey).

Parmi la multitude de résultats observés, une interaction significative Régions \* Temps est retrouvée aussi bien à la vue 1 ( $F(5.175,46.58) = 4.58, p < .05$ ) qu'à la vue 2 pour l'objet test Identique ( $F(76,38.27) = 7.94$ ) et pour l'objet test Différent, ( $F(76,33.117) = 5.178, p < .05$ ). Cette interaction indique (Figure 2A), une activité occipito-temporale à partir des 75-125ms (t4 sur la Figure 2A), qui s'accompagne d'une activité pariétale à partir de 100-150ms (t5 sur la vue 2A).

Ce résultat est cohérent avec les travaux de Keehner et al. [16], montrant au cours d'une tâche de changement de point de vue, d'une part, une activité occipito-temporale qui coderait la forme 3D de l'objet et d'autre part, une activité pariétale qui calculerait les conséquences visuelles égocentrées du changement de point de vue.

Des analyses supplémentaires sur les 4s post-onset de la vue initiale par des fenêtres temporelles de 100ms, et dont le plan d'analyse est : 3 Types de stimuli (Posture, Cube, Lampe) \* 3 Angles (0°, 60°, 120°) \* 5 Régions \* 20 Temps, indiquent que pour les différences angulaires de 60° et 120°, la région pariétale reste sollicitée tout au long des 4sec de la vue de référence (Vue 1) alors que pour 0°, cette activité diminue rapidement (au bout d'1s). D'autre part, la région temporale est sollicitée pendant les 2 premières secondes pour les différences angulaires de point de vue 60° et 120°, alors qu'elle n'est sollicitée que très brièvement pour 0° (entre 200 - 500ms selon la nature de l'objet). On note également que les arrangements de cubes sont les seuls objets pour lesquels à 0° une activité occipitale augmente significativement à la dernière seconde de présentation de la vue de référence. Ceci peut traduire une mémoire visuelle 2D de la structure de l'objet [17], qui suggère que les cubes sont plus simples à coder visuellement que les autres objets.

En ce qui concerne les 500ms post-onset du stimulus, le niveau élevé de l'activité pariétale à la vue 1 (Figure 2A) est maintenu jusqu'à la vue 2 (pour 60° et 120°), mais il augmente encore à la vue 2 (Figure 2B) probablement pour vérifier le contenu du nouveau point de vue. Le même phénomène est observé pour l'activité temporale (Figure 2B).

D'autre part, les activités occipito-temporales et pariétales sont encore plus importantes à la vue 2 quand les stimuli sont différents que lorsqu'ils sont

identiques (une augmentation d'environ 30 femto-tesla par rapport au objet Identiques).

En outre, l'hémisphère gauche est beaucoup plus sollicitée que l'hémisphère droit dans ce type de tâche aussi bien à la vue de référence qu'à la vue test. Ce résultat est cohérent avec les travaux montrant que l'hémisphère gauche, plus particulièrement le LOC gauche est impliqué dans une analyse perceptive et sémantique de la reconnaissance d'objet [18].

### 3. Conclusion

En résumé, notre étude suggère que pour anticiper les conséquences visuelles d'un changement de point de vue, la première chose que fait le cerveau lorsqu'il perçoit la vue initiale c'est identifier la nature de l'objet via le LOC et en particulier de l'hémisphère gauche. Dans un deuxième temps, il commence la transformation spatiale de l'objet, sur la base des informations initiales de l'indice spatiale en sollicitant la région pariétale, pour modifier l'apparence de l'objet en fonction de l'amplitude du changement de point de vue dans un référentiel égocentré.

S'il s'agit de mémoriser l'apparence de l'objet (changement de point de vue de 0°) et que ce dernier est facilement codé de façon bidimensionnelle, alors l'imagerie mentale sollicitera de façon plus importante les régions occipitales plutôt que pariétales.

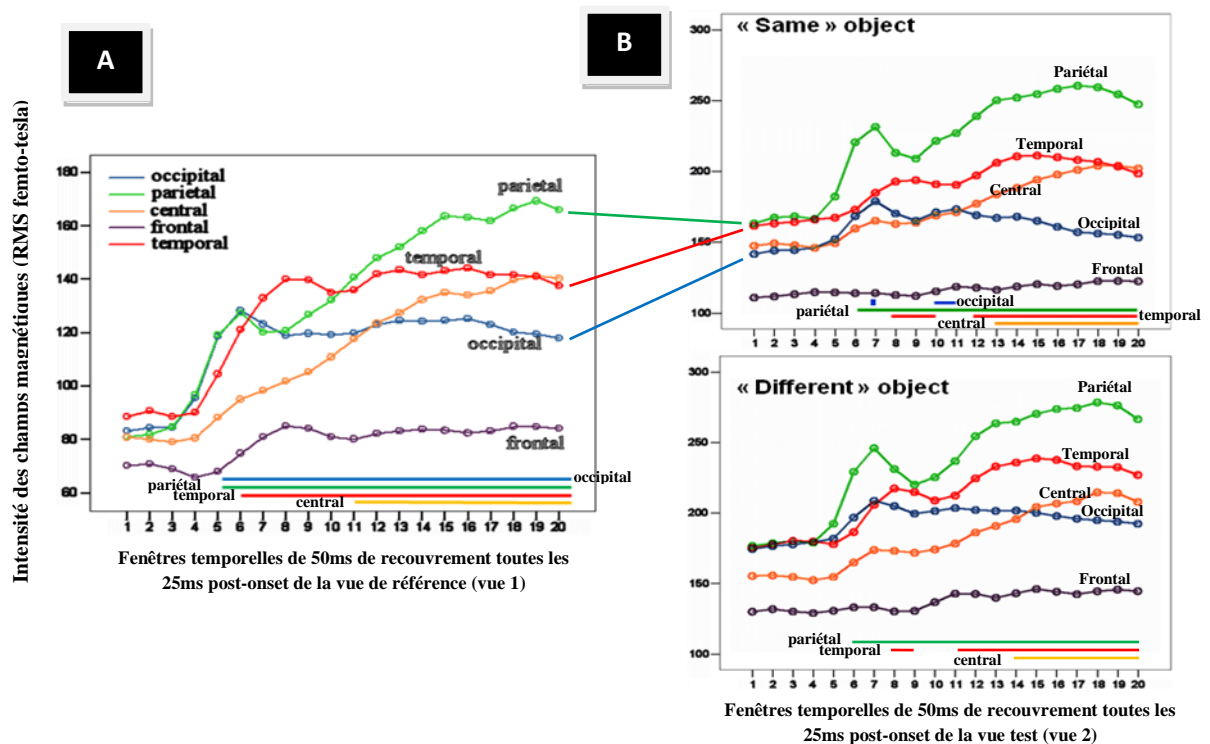
D'autre part, nos résultats suggèrent que pour anticiper les conséquences visuelles d'un changement de point de vue, il existe deux processus différents qui fonctionnent en parallèle : i) un processus de maintien de la structure de l'objet pris en charge par la région temporale et fonctionnant dans un référentiel spatial centré sur l'objet et ii), un processus qui sollicite la région pariétale pour calculer et/ou anticiper l'apparence visuelle de l'objet dans un référentiel égocentré.

### Remerciements

Le staff technique du centre MEG-EEG, en particulier A. Ducorps et D. Schwartz.

### References

- [1] T. Poggio & S. Edelman, A network that learns to recognize three dimensional objects, *Nature*, 343, 1990, 263-266.
- [2] F. Fang & S. He, Viewer-centred object representation in human visual system revealed by viewpoint aftereffects, *Neuron*, 45, 2005, 793-800.
- [3] I. Gauthier, W. G. Hayward, M. G. Tarr, A. W. Anderson, P. Skudlarski, & J. C. Gore, Bold activity during mental rotation and viewpoint-dependent object recognition, *Neuron*, 34, 2002, 161 - 171.



**Figure 2.** Organisation spatio-temporelle des activités cérébrales à la vue 1(A) et à la vue 2 (B). Les traits en couleurs indiquent quand une région donnée a une activité significativement plus élevée (Post-hoc Tukey) que son niveau initial (à t1).

[4] J. S. Simon, W. Koutstaal, S. Pince, A. D. Wagner, & D. L. Schacter, Neuron mechanisms of visual object priming: evidence for perceptual and semantic distinctions in fusiform cortex, *NeuroImage*, 19, 2003, 613 - 626.

[5] C. Gerlach, C. T. Aaside, G. W. Hymphreys, A. Gade, O. B. Paulson, & I. Law, Brain activity related to integrative processes in visual object recognition: bottom-up integration and the modulatory influence of stored knowledge, *Neuropsychologia*, 40, 2002, 1254-1267.

[6] P. Vuilleumier, R. N. Henson, J. Driver, & R. J. Dolan, Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming, *Nature Neuroscience*, 5, 2002, 491-499.

[7] I. Biederman, Recognition-by-components: a theory of human understanding, *Psychological Review*, 94, 1987, 115-147.

[8] I. Biederman, Viewpoint-dependent mechanisms in visual object recognition, *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1995, 1506-1514.

[9] D. Marr & H. K. Nishihara, Representation and recognition of the spatial organization of three dimensional structures, *Proceedings of the Royal Society of London B*, 200, 1978, 269-294.

[10] E. P. Hanavan, A mathematical model of human body. *AMRL Technical Report 64-102*, Wright Patterson Air Force Base, Ohio, 1964.

[11] L.M. Parsons, Temporal and kinematic properties of motor behavior reflected in mentally simulated action, *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1994, 709- 730.

[12] C.G. Christou, B.S. Tjan, & H.H. Bühlhoff, Extrinsic cues aid shape recognition from novel viewpoints, *Journal of Vision*, 3, 2003, 183 - 198.

[13] M. A. Amorim, What is my avatar seeing? The coordination of out-of-body and embodied perspectives for scene recognition across views, *Visual Cognition*, 10, 2003, 157 - 199.

[14] M. A. Amorim, B. Isableu, & M. Jarraya, Embodied spatial transformations: "body analogy" for the mental rotation of objects, *Journal of Experimental Psychology: General*, 135, 2006, 327-347.

[15] R.N. Shepard, & J. Metzler, Mental rotation of three dimensional object, *Science*, 171, 1971, 701-703.

[16] M. Keehner, S. A. Guerin, M. B. Miller, D. J. Turk, & M. Hegarty, Modulation of neural activity by angle of rotation during imagined spatial transformations, *NeuroImage* 33, 2006, 391-398.

[17] M. Kosslyn, & W. L. Thompson, When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, 129, 2003, 723 - 746.

[18] M. E. Large, A. Aldcroft, & T. Vilis, Task-related laterality effects in the lateral occipital complex, *Brain research*, 1128, 2007, 130 -138.