

ANISOTROPIE ET RECHERCHE VISUELLE : L'ORIENTATION CANONIQUE COMME DÉTERMINANT DE LA SAILLANCE PERCEPTIVE

Cécile Ballaz, Alan Chauvin, Christian Marendaz et Carole Peyrin

Laboratoire de Psychologie Expérimentale CNRS UMR 5105
BP. 47, 38 040 Grenoble Cedex 09
Mél : Cecile.Ballaz@upmf-grenoble.fr

Résumé

La recherche d'un segment de ligne 'cible' différant des segments de lignes 'distracteurs' par une orientation d'une vingtaine de degrés se marque d'une curieuse anisotropie : la recherche est plus facile (rapide) si la cible est oblique que si elle est verticale (ou horizontale). Dans des travaux récents, nous avons montré que l'anisotropie variait en fonction de la posture des sujets, signifiant que la recherche visuelle d'orientation, habituellement décrite et modélisée sous forme de traitements visuels de bas niveau (Foster et Ward, 1991 ; Bernard et al., 1997), se fait cependant après une étape d'intégration intersensorielle (Marendaz, 1998). Est-elle sensible aux informations mnésiques ? Pour répondre à cette question, nous avons substitué aux segments de ligne des formes représentant des animaux dont l'orientation canonique est soit verticale soit oblique. Les résultats montrent que c'est l'orientation canonique, et non l'orientation physique, qui détermine l'anisotropie. La recherche visuelle d'orientations est donc sensible aux informations mnésiques, lesquelles l'emportent sur le signal en cas de conflit alors qu'elles ne sont a priori d'aucune utilité pour résoudre la tâche. Ce que confirment deux contre-expériences, l'une réalisée chez le sujet cérébro-lésé (agnosie visuelle), l'autre manipulant le contenu fréquentiel des images.

Abstract

Visual search performance for a target line among distractors differing in orientation shows an asymmetrical pattern: Search for a tilted line (by approximately 20 degrees) among vertical (or horizontal) lines is faster and more accurate than search for a vertical line among tilted ones. In recent studies, we found that this search asymmetry is a function of the observer's posture (Marendaz, 1998). This suggests that orientation processing implies the integration of multi-sensorial information, rather than being a separate low-level visual processing stage (Foster et Ward, 1991; Bernard et al., 1997). In the present study we investigated whether the search asymmetry for orientation detection also occurs with objects. We replaced line segments with drawings of real-world objects (animals) whose intrinsic (canonical) orientation was either vertical or tilted. The results showed a search asymmetry for object orientation, and that this asymmetry depended on the canonical, rather than the physical, orientation of the target object. This suggests that visual search for orientation can be influenced by memory representations incorporating canonical object orientation although these representations are not a priori useful to resolve the task. This canonical orientation effect was confirmed in two counter-experiments one with a brain-injured participant (visual agnosia), and the other using spatial frequency manipulation on the same pictures.

Introduction

Si un consensus émerge en faveur d'un traitement visuel analytique, débutant par une étape massivement parallèle où sont extraites les caractéristiques élémentaires de la scène visuelle, la façon dont sont combinées ces caractéristiques en objets cohérents permettant une reconnaissance est beaucoup plus controversée. Ceci est dû au fait que la modélisation de cette intégration doit prendre en compte différentes propriétés de la reconnaissance humaine, dont la complexité conceptuelle est à la mesure de leur évidence phénoménologique : rapidité, résistance au bruit et invariance spatiale.

1. Rôle des axes d'élongation et des connaissances dans la reconnaissance d'objet

L'invariance spatiale est la propriété la plus difficile à modéliser. Elle renvoie au fait que nous sommes capables de reconnaître un objet quelles que soient sa localisation, sa taille, son orientation 2D (fronto-parallèle)

ou 3D (en profondeur). Si ces diverses possibilités d'images rétiniennes, pour un même objet, permettent d'aboutir à une reconnaissance, le coût cognitif peut varier selon le type de transformations. Les expériences montrent que ce coût, estimé par exemple en temps de traitement, est moindre pour un objet ayant subi une translation (changement de localisation) ou une dilatation (changement de taille) que pour un objet ayant subi une rotation 2D ou 3D (Biederman et Cooper, 1991a/b ; Biederman et Gerhardstein, 1993 ; Verfaillie et Boutsen, 1995 ; Corballis, 1988 ; Jolicoeur, 1985 ; Rock, 1973 ; voir Leek, 1998 et Vannucci et Viggiano, 2000, pour une discussion sur les objets ayant une orientation 2D préférentielle (canonique) ou non, et Hinton, 1981, pour une modélisation connexionniste simple de la reconnaissance d'objet ayant subi une rotation 2D). Pour rendre compte d'une manière économique des propriétés d'invariance spatiale et de rapidité de la reconnaissance humaine, les théories computationnelles proposent que chaque objet soit stocké en mémoire sous forme d'une représentation centrée sur l'objet lui-même. Ce type de représentations nécessite un référentiel spatial intra-objet. Il existe maintenant de nombreux arguments expérimentaux et théoriques

(simulations) permettant de penser que le ou les axes d'élongation et/ou de symétrie des objets constituent ce référentiel spatial sur la base duquel seraient codées topographiquement les différentes parties de l'objet (e.g., Marr et Nishihara, 1978 ; Marr, 1982 ; Biederman, 1987 ; Hummel et Biederman, 1992 ; Jolicoeur, 1992 ; Lawson *et al.*, 2000 ; Palmer, 1980 ; Palmer *et al.*, 1981 ; Quinlan et Humphreys, 1993 ; Srinivas, 1995). Dans ce type de modélisation, dite structurale, l'extraction des axes d'élongation et/ou de symétrie est donc fondamentale et doit, du point de vue du déroulement temporel du traitement, être extrêmement précoce (e.g., 2ème couche, dans le modèle connexionniste de Hummel et Biederman, 1992).

À l'heure actuelle, les modèles sur la perception visuelle sont principalement constitués de traitements ascendants. Ces modèles postulent que le traitement des propriétés physiques de l'objet soit terminé avant de pouvoir être modulé par des traitements descendants. C'est pourquoi, lorsque sont envisagés les traitements de haut niveau, telle que l'activité des représentations mnésiques, ils n'interviennent de manière automatique que sur les dernières étapes du traitement visuel (e.g., sur la construction de l'ébauche 3D pour Marr, 1982). L'intervention plus précoce de telles représentations n'est jusqu'alors envisagée que comme un processus volontaire. Par exemple, dans les modifications apportées au modèle d'intégration des caractéristiques (FIT), Treisman (1993) envisage la possibilité de traitements descendants pour rendre compte des effets liés aux attentes des sujets. Les attentes du sujet sur la cible à rechercher vont agir de manière descendante sur le contrôle attentionnel pour filtrer les informations non pertinentes (voir aussi Itti *et al.*, 1998 ; Wolfe, 1994). Cependant, de nombreux travaux mettent en évidence un accès automatique aux représentations mnésiques pouvant interférer sur le traitement visuel. La réponse ou le temps de réponse des sujets peuvent être modifiés par une activation involontaire des représentations mnésiques, bien qu'elles ne soient pas pertinentes pour la tâche. Cet effet s'observe sur de nombreuses tâches, comme l'interférence entre mots et images (Glaser et Glaser, 1989), la perception de scène (Biederman, 1981) ou encore l'amorçage négatif d'objets (Tipper, 1985). Boucart et Humphreys (1994, 1997) ont testé cette activation automatique des représentations mnésiques lorsque le sujet est explicitement informé qu'il doit traiter des propriétés physiques. Ils observent une interférence de la catégorie sémantique des items dans une tâche d'appariement qui nécessite que l'attention du sujet soit focalisée sur l'orientation physique principale de l'objet. Cette activation automatique des représentations mnésiques peut-elle intervenir dès les premières étapes du traitement ? Retrouve-t-on une telle interférence des représentations mnésiques dans les tâches plus 'basiques' de recherche visuelle ?

2. Recherche de cible et modélisation du codage de l'orientation

Lorsque l'on examine des scènes visuelles, certaines caractéristiques ou certaines formes « sautent aux yeux » (« pop out »). Cette détection immédiate a été interprétée en psychologie cognitive comme résultant de

traitements parallèles ascendants (e.g., Julesz et Schumer, 1981 ; Treisman, 1993). Cette extraction en parallèle des dimensions visuelles primaires peut être mise en évidence par la recherche visuelle d'une cible qui se distingue des distracteurs par une caractéristique élémentaire, comme l'orientation. Lorsque la cible est un segment de ligne qui diffère fortement en orientation des distracteurs, cette cible saute aux yeux et sera très rapidement détectée (cf. Treisman, 1985, 1993). Cependant, la perception de l'orientation opérationnalisée par une tâche de recherche de cible, se marque d'une anisotropie spécifique : lorsque l'orientation de la cible et celle des distracteurs diffèrent seulement d'une vingtaine de degrés, la recherche d'une oblique parmi des verticales ou des horizontales est immédiate, parallèle (le temps de réponse est indépendant du nombre de distracteurs), alors que la recherche d'une verticale ou d'une horizontale parmi des obliques est sérielle/attentionnelle (le temps de réponse augmente linéairement en fonction du nombre de distracteurs). Ce pattern asymétrique de recherche suggère un codage psychophysique 'grossier' de l'orientation en vision de bas niveau, basé sur un ensemble de filtres locaux accordés sur la verticale et sur l'horizontale (Foster et Ward, 1991 ; Marendaz, *et al.*, 1991). À partir de l'activité de ces filtres, il est possible d'établir une carte de saillance perceptive déterminant quelles sont, du point de vue du signal, les zones les plus 'saillantes' (les plus 'attractives') de l'image (voir figure 1 ; Chauvin *et al.* 2001 ; Itti et Koch, 2000). Si ce système de filtres rend bien compte de l'anisotropie, il est toutefois réducteur car il ne prend pas en compte un certain nombre d'informations qui sont déterminantes en terme de référentiel dans le traitement des orientations.

Dans quel référentiel spatial fonctionnent ces filtres psychophysiques ? En manipulant les informations gravito-inertielles (posture des sujets et centrifugation) et les informations visuo-contextuelles (cadre carré circonscrivant les segments de ligne et ayant diverses orientations possibles), de nombreux travaux ont montré que l'asymétrie de recherche est fonction des indices vestibulaires, proprioceptifs et visuo-contextuels disponibles pour résoudre la tâche. Le traitement de l'orientation en vision de bas niveau n'opère donc pas dans un référentiel rétinocentré, mais dans un référentiel spatial multisensoriel (Marendaz *et al.*, 1993 ; Marendaz, 1998). La prise en compte d'autres informations, comme ici les informations vestibulaires, ne s'intègre pas dans les modèles psychophysiques classiques (e.g., Foster et Ward, 1991). Certains modèles plus généraux de recherche visuelle prévoient que des informations descendantes parallèles puissent guider la recherche (e.g., Wolfe, 1994 ; Itti et Koch, 2000), mais ces informations ne sont que très rarement définies et sont souvent limitées aux attentes du sujet. Enfin, si les informations vestibulaires et proprioceptives sont impliquées dans le traitement des orientations, qu'en est-il des informations dites de plus haut niveau, comme notre connaissance sur l'orientation des objets ?

La recherche visuelle peut être utilisée à des fins diverses, parfois même, elle n'est pas mesurée *per se*. Par exemple, Butter et Goodale (2000) utilisent la recherche de cible comme condition expérimentale préalable à la reconnaissance, afin de montrer que la recherche d'une cible active sa représentation

mnésique. Dans ces travaux, la mesure porte sur une tâche d'identification indépendante de la recherche visuelle, les résultats montrent que l'activation initiée par la tâche de recherche persiste sur les essais d'identification et peut s'exprimer en dehors de la présence physique de la cible. La recherche visuelle induit une activation de la représentation mnésique de la cible à rechercher, cependant ces travaux ne nous renseignent pas directement sur le rôle d'une telle activation dans la recherche visuelle. L'activation des représentations mnésiques qui s'opère durant la recherche visuelle, peut-elle déterminer le traitement de l'orientation ? Pour répondre à cette question, nous avons substitué aux segments de ligne (sans signification) des images d'animaux ayant une orientation canonique verticale ou oblique.

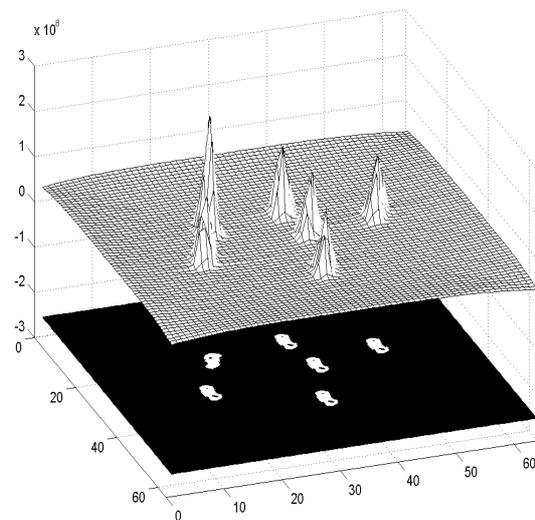


Figure 1 — Simulation d'une carte de saillance perceptive déterminée à partir du signal d'une configuration contenant six hippocampes.

L'examen de cette carte montre que l'hippocampe oblique est l'item le plus « saillant » de l'image. Pour plus de détails sur la modélisation des cartes de saillance, voir Chauvin et al. (2001)

3. Résultats expérimentaux

3.1. Expérience 1

Le but de l'expérience 1 était de voir si l'anisotropie observée dans les tâches de recherche visuelle d'orientations est modulée par les représentations mnésiques.

3.1.1. Méthode

Participants — Quarante-huit étudiants avec une vue normale ou corrigée ont passé l'expérience (douze pour chaque stimulus utilisé).

Matériel — Quatre stimuli représentant des animaux (catégorie sémantique dont la reconnaissance est la plus sensible aux effets de rotation 2D, Vannucci et Viggiano, 2000) ont été sélectionnés pour constituer deux conditions d'orientation canonique. Ces

représentations d'animaux ont été choisies le plus proches possibles de celles utilisées par Boucart et Humphreys (1997) dans leurs expériences montrant une interférence du sémantique sur le traitement des caractéristiques physiques. Les deux premiers stimuli étaient des items contrôles pour lesquels l'information mnésique et l'information physique prenaient la même orientation. Ils représentaient un hippocampe et un pingouin, tous deux ayant un axe d'élongation vertical dans leur orientation canonique (i.e. orientation canonique verticale). Ces items étaient présentés soit dans leur orientation canonique, soit avec une rotation de 23° par rapport à la verticale (leur axe d'élongation est alors physiquement oblique, voir figure 2).

La deuxième paire de stimuli correspondait aux items tests ; elle était constituée d'un kangourou et d'un pigeon, tous deux ayant un axe d'élongation orienté à 23° par rapport à la verticale (i.e. orientation canonique oblique). Comme les items contrôles, les items cibles pouvaient être présentés soit dans leur orientation canonique (l'axe d'élongation physique est orienté à 23°), soit avec une rotation de 23° (axe d'élongation physique vertical ; figure 2). La longueur de l'axe d'élongation de chaque animal était d'environ 1° de taille angulaire. Chaque essai consistait en la présentation simultanée de 3 ou 6 items représentant le même animal. Soit les items du pattern étaient tous les mêmes (la cible était absente), soit l'un différait des autres par son orientation (la cible était présente). La position de la cible et des distracteurs variait d'un essai à l'autre et la distance entre deux items était au minimum de 1,5° angulaire.

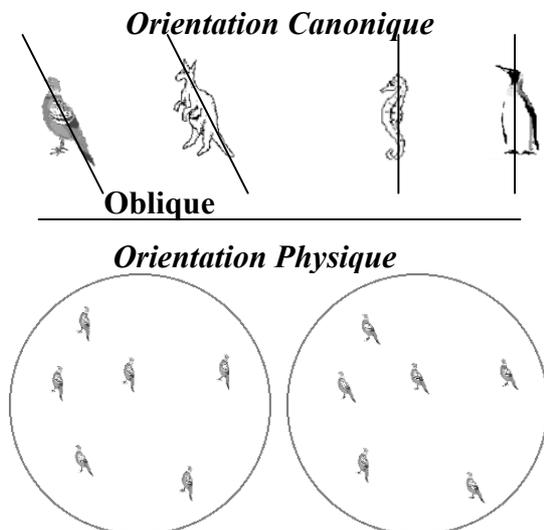


Figure 2 — Présentation des items cibles.

En haut, les différents exemplaires utilisés dans l'expérience 1 et 3 : à gauche, le Pigeon et le Kangourou ont une orientation canonique oblique et, à droite, l'Hippocampe et le Pingouin ont une orientation canonique verticale.

En bas, exemples de configuration : à gauche la cible est physiquement oblique, à droite elle est physiquement verticale

Procédure — Les items étaient présentés sur un écran 17 pouces. Les sujets étaient assis face à l'écran à une distance de 1,20 m. Ils plaçaient leur tête à l'intérieur

d'un cône noir supprimant tout autre indice visuel d'orientation. Chaque essai était constitué d'un point de fixation durant 1,5 secondes, d'une configuration de 3 ou 6 items qui restait visible jusqu'à la réponse du sujet, puis d'un écran blanc durant 2 secondes. Les sujets devaient déterminer le plus vite possible et le plus précisément possible si la cible, définie comme l'élément différant par son orientation, était présente ou non. Ils répondaient à l'aide de deux boutons poussoirs indépendants, l'un correspondant à la réponse absente et l'autre à la réponse présente. Les sujets étaient répartis de manière aléatoire dans une des quatre conditions définies par l'animal utilisé. Pour chaque exemplaire d'animal, l'expérience était constituée de deux blocs qui correspondaient aux deux orientations possibles de la cible : une recherche de cible oblique parmi des distracteurs verticaux, et une recherche de cible verticale parmi des distracteurs obliques. Dans chaque bloc, chaque combinaison de taille de pattern (3 et 6) et de cible présente ou absente était présentée 16 fois dans un ordre aléatoire, soit au total 64 essais pour chacune des orientations de la cible. L'ordre de présentation de chaque bloc a été contrebalancé entre les participants.

3.1.2. Résultats expérience 1

Une analyse de variance 4 x 2 x 2 a été réalisée à partir des temps de réponse pour détecter la cible présente. Les facteurs étaient le Type d'exemplaire animal (Kangourou vs. Pigeon vs. Hippocampe vs. Pingouin), l'orientation physique de l'axe d'élongation de la cible (verticale vs. oblique) et le nombre d'items (3 vs. 6). Le premier est un facteur inter-sujets, les deux autres sont des facteurs intra-sujets. Le facteur type d'exemplaire animal permet de distinguer chacun des items utilisés et comporte l'information d'orientation canonique lorsqu'ils sont comparés deux à deux (Kangourou + Pigeon vs. Hippocampe + Pingouin). Les temps de réponse moyens ne se distinguaient pas en fonction du facteur exemplaire ($F(3, 44) = 2.1$; $MSE = 249777$; $p > 0.10$). On constate qu'il n'y a pas d'effet principal de l'orientation physique des cibles ($F < 1$). Seul l'effet principal lié à la taille du pattern est significatif ($F(1, 44) = 112$; $MSE = 20850$; $p < 0.0001$). Les temps de réponses sont généralement plus courts lorsque le pattern est constitué de 3 items (782 ms) que lorsqu'il est constitué de 6 items (1002 ms). L'interaction simple entre les facteurs exemplaires et orientation physique de la cible est significative ($F(3, 44) = 12.8$; $MSE = 48215$; $p < 0.0001$), de même que l'interaction simple entre l'orientation canonique et l'orientation physique ($F(1, 44) = 35$; $MSE = 48215$; $p < 0.0001$). Les temps de réponses sont plus courts pour les exemplaires avec une orientation canonique verticale lorsque la cible est présentée physiquement oblique plutôt que verticale (725 ms vs. 922 ms, $F(1, 44) = 19.3$, $MSE = 48215$, $p < 0.0001$, voir figure 3). Cette différence se retrouve clairement pour l'Hippocampe et approche du seuil de significativité pour le Pingouin (668 ms vs. 941 ms pour l'Hippocampe, $F(1, 44) = 18.6$; $MSE = 48215$; $p < 0.0001$; et 781 ms vs. 901 ms pour le Pingouin ; $F(1, 44) = 3.6$; $MSE = 48215$; $p = 0.06$).

À l'inverse, pour les exemplaires ayant une orientation canonique oblique, c'est la cible physiquement verticale qui est plus rapidement détectée que la cible oblique (1051 ms vs. 871 ms, $F(1, 44) = 16$; $MSE = 48215$;

$p < 0.0007$). Cette différence se retrouve pour chacun des exemplaires (1041 ms vs. 941 ms pour le Kangourou, $F(1, 44) = 10$; $MSE = 48215$; $p < 0.003$ et 961 ms vs. 802 ms pour le Pigeon, $F(1, 44) = 6.3$; $MSE = 48215$; $p < 0.02$). Enfin, l'interaction double entre le facteur orientation canonique, l'orientation physique et la taille du pattern est significative ($F(1, 44) = 5.2$; $MSE = 14205$; $p < 0.03$).

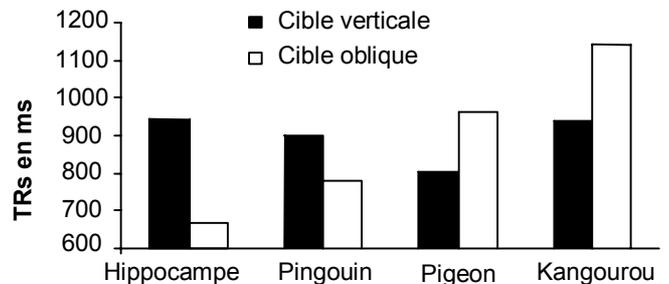


Figure 3 — Temps de réponse moyens pour détecter la cible présente en fonction de l'orientation physique de la cible (verticale vs. oblique) pour chacun des exemplaires.

3.1.3. Conclusion expérience 1

Les résultats obtenus sont en faveur d'une intervention des représentations mnésiques sur le traitement des orientations dans le cadre de la recherche visuelle (figure 3). Effectivement, on trouve, au niveau des temps de réponse moyens, une interaction simple entre l'orientation canonique des animaux et l'orientation physique de l'axe d'élongation. Dans les tâches classiques de recherche de cible avec des segments de ligne, la cible oblique parmi des distracteurs verticaux est plus rapidement détectée que la cible verticale parmi des distracteurs oblique. Ce résultat est ici reproduit avec les items contrôles dont l'orientation canonique est verticale. Au contraire pour les items tests dont l'orientation canonique est oblique, la détection est plus rapide lorsqu'il faut détecter une cible physiquement verticale parmi des obliques que l'inverse. Ce renversement du traitement de l'orientation physique en fonction de l'orientation canonique est ici systématique (il apparaît pour les deux items tests et pas pour les contrôles) indiquant une anisotropie déterminée par l'orientation canonique. L'ensemble des résultats va dans le sens d'un jugement perceptif basé sur l'orientation canonique. Pour conforter nos conclusions et s'assurer qu'il s'agit d'un effet de haut niveau lié aux représentations mnésiques, nous avons mené plusieurs contre-expériences.

3.2. Expérience 2

Pour diminuer le poids de certaines informations lors du traitement visuel, nous avons appliqué un filtre en fréquences spatiales sur chacune des configurations contenant les exemplaires Pigeon et Hippocampe (voir exemples figure 4). Sur chaque image d'origine, on effectue un filtrage passe-bas avec une fréquence de coupure de 1,19 cycles par degré et une fréquence de transition de 0,96 cycles par degré. Ce filtrage supprime l'information locale, ce qui a pour principale

conséquence de supprimer la possibilité d'une reconnaissance des animaux présentés tout en préservant l'information d'orientation globale. Sans reconnaissance, l'orientation canonique ne peut pas intervenir et seule l'orientation physique devrait diriger la recherche visuelle. Si tel est le cas, il reste toutefois à vérifier que les données obtenues sont bien dues à la perte des informations mnésiques, du fait du caractère passe-bas du filtrage, et non au bruit (à la dégradation de l'image : perte de la couleur, modification des contrastes locaux, etc.) lié au filtrage *per se*. C'est pourquoi, à titre de contrôle, nous avons effectué un filtrage passe-haut des images d'origine (fréquence de coupure de 6 cycles par degré et fréquence de transition de 2,4 cycles par degré). Ce filtrage passe-haut atténue les informations globales de l'image (telles que le contraste, la luminance...), ce qui rend l'orientation physique moins prégnante, mais préserve l'information locale (les contours) qui permet une reconnaissance. On s'attend donc, dans cette condition contrôle de filtrage, à retrouver une anisotropie déterminée par l'orientation canonique.

3.2.1. Méthode

Participants — Quarante-huit étudiants avec une vue normale ou corrigée ont passé l'expérience. Les sujets étaient répartis aléatoirement en quatre groupes de 12 sujets, un pour l'Hippocampe en basses fréquences, un pour le Pigeon en basses fréquences, un pour l'Hippocampe en hautes fréquences et un pour le Pigeon en hautes fréquences.

Procédure et matériel — La procédure était identique à celle de l'expérience 1. Les images comportant les Hippocampes et les Pigeons étaient filtrés en passe-haut et en passe-bas (voir figures 4 et 6)

3.2.2. Résultats expérience 2



Figure 4 — Exemple de configuration en filtrage passe-bas, comportant 3 Pigeons dont une cible physiquement verticale et deux distracteurs obliques.

L'interaction double entre le Type d'animal (i.e. l'orientation canonique), l'orientation physique et le type de filtrage spatial est significative ($F(1, 52) = 17.8$; $MSE = 46334$; $p < 0.0001$) de même que l'interaction triple qui inclut le nombre d'items ($F(1, 52) = 7.3$; $MSE = 7961$; $p < 0.01$).

Images avec un filtrage spatial passe-bas

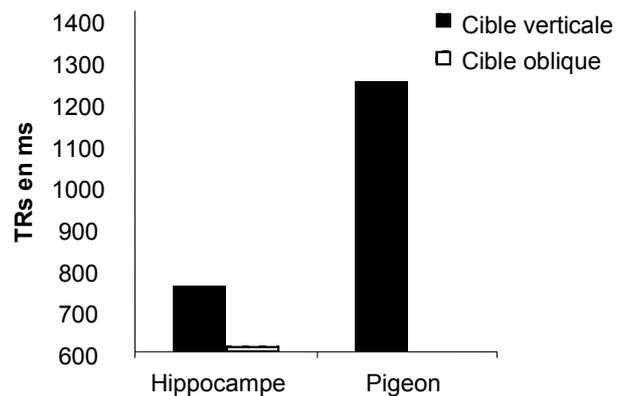


Figure 5 — Temps de réponse moyens pour détecter la cible présente en fonction de l'orientation physique de la cible (verticale vs. oblique) pour chacun des exemplaires en filtrage passe-bas.

En condition filtrage passe-bas, aucun sujet n'a reconnu ce qui lui a été présenté. Les temps de réponse moyens pour l'exemplaire Pigeon sont significativement supérieurs à ceux obtenus pour l'exemplaire Hippocampe (1074 ms vs. 684 ms, $F(1, 26) = 13.1$; $MSE = 325907$; $p < 0.002$). Cette différence peut s'expliquer par la forme des *blobs* qui, dans la condition Pigeon, ajoute de la difficulté à la tâche. En effet, alors que les *blobs* obtenus pour l'Hippocampe sont relativement rectilignes, ceux du Pigeon ont une forme de croissant (voir figure 4). Il y a un effet principal de l'orientation physique de la cible (la cible verticale est plus longue à détecter (995 ms) que la cible oblique (763 ms, $F(1, 26) = 26$; $MSE = 57945$; $p < 0.0001$), ainsi que de la taille du pattern (les temps sont plus courts pour 3 items (805 ms) que pour 6 (952 ms, $F(1, 26) = 42.4$; $MSE = 14245$; $p < 0.0001$). L'interaction simple entre l'orientation canonique des exemplaires et l'orientation physique des cibles est significative ($F(1, 26) = 4.4$; $MSE = 57945$; $p < 0.04$). L'avantage pour détecter la cible oblique par rapport à la cible verticale est plus marqué avec l'exemplaire « Pigeon » (la différence de TRs entre la cible verticale 1238 ms et la cible oblique 911 ms est de 327 ms, $F(1, 26) = 25.8$; $MSE = 57945$; $p < 0.0001$) qu'avec l'exemplaire « Hippocampe » (la différence de TRs entre la cible verticale 753 ms et la cible oblique 615 ms est de 138 ms, $F(1, 26) = 4.5$; $MSE = 57945$; $p < 0.05$, voir figure 5). L'interaction entre l'orientation canonique des exemplaires et la taille du pattern est significative ($F(1, 26) = 17.4$; $MSE = 14245$; $p < 0.0004$) de même que l'interaction simple entre l'orientation physique de la cible et la taille du pattern ($F(1, 26) = 9.7$; $MSE = 9769$; $p < 0.005$). L'augmentation des temps de réponse avec le nombre des distracteurs est plus importante lorsque la cible est physiquement verticale (893 ms pour 3 items vs. 1098 ms pour 6, $F(1, 26) = 31$; $MSE = 18983$; $p < 0.0001$) que lorsqu'elle est oblique (718 ms pour 3 items vs. 807 ms pour 6, $F(1, 26) = 31.9$; $MSE = 5030$; $p < 0.0001$). Ce profil correspond à celui de la littérature pour décrire l'anisotropie classique. Enfin, l'interaction double entre le Type d'animal, l'orientation de la cible et la taille du pattern n'est pas significative ($F(1, 26) = 2.93$; $MSE = 9769$; $p > 0.09$).

Avec un filtrage passe-haut, l'analyse des temps de réponse met en évidence l'intervention du facteur d'orientation canonique de l'exemplaire. Les temps de réponse sont plus longs pour l'exemplaire « Pigeon » que pour l'exemplaire « Hippocampe » (1193 ms vs. 819 ms, $F(1, 26) = 8.25$; $MSE = 475187$; $p < 0.008$). La différence entre les temps de détection de la cible physiquement oblique et ceux de la cible physiquement verticale s'approche de la significativité statistique (972 ms vs. 1041 ms, $F(1, 26) = 3.75$; $MSE = 34723$; $p = 0.06$). Enfin, il y a un effet principal de la taille du pattern, la cible est détectée plus rapidement parmi 3 items (840 ms) que parmi 6 (1173 ms, $F(1, 26) = 78.5$; $MSE = 39439$; $p < 0.0001$).

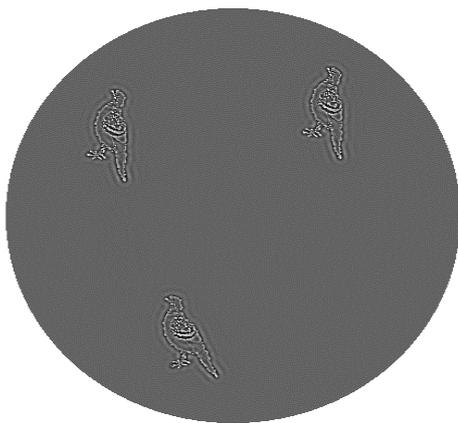


Figure 6 — Exemple de configuration en filtrage passe-haut, comportant 3 Pigeons dont une cible physiquement oblique et deux distracteurs physiquement verticaux.

Images avec un filtrage spatial passe-haut

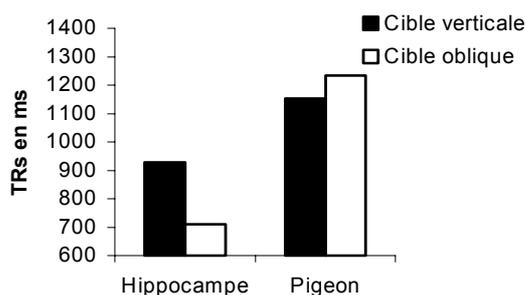


Figure 7 — Temps de réponse moyens pour détecter la cible présente en fonction de l'orientation physique de la cible (verticale vs. oblique) pour chacun des exemplaires en filtrage passe-haut.

Comme attendu dans le cas d'une incidence des représentations mnésiques, le Type d'exemplaire utilisé interagit avec l'orientation physique de la cible ($F(1, 26) = 17.7$; $MSE = 34723$; $p < 0.0003$, voir figure 7). La cible oblique est plus rapidement détectée que la cible verticale pour l'exemplaire « Hippocampe », et ce profil s'inverse pour l'exemplaire « Pigeon ».

L'interaction simple entre la taille du pattern et l'orientation physique de la cible n'est pas significative ($F(1, 26) = 3$; $MSE = 6124$; $p > 0.09$). L'orientation canonique (i.e. le Type d'exemplaire) interagit avec le nombre d'items ($F(1, 26) = 17.7$; $MSE = 39439$; $p < 0.0003$).

L'augmentation du temps de réponse avec le nombre d'items est plus marquée pour l'exemplaire « Pigeon » que pour l'exemplaire « Hippocampe ». Enfin, l'interaction double entre la taille du pattern, le Type d'exemplaire et l'orientation de la cible est significative ($F(1, 26) = 4.8$; $MSE = 6154$; $p < 0.04$).

3.2.3. Conclusion expérience 2

À partir d'images dégradées, on constate que la modification de l'anisotropie en fonction de l'orientation canonique des items dépend de leur reconnaissance effective. En condition filtrage passe-bas (résultats représentés figure 5), aucun sujet ne reconnaît ce qui lui a été présenté, et l'analyse des temps de réponse montre un profil classique d'anisotropie (i.e. un avantage pour détecter la cible physiquement oblique) quelle que soit l'orientation canonique de l'animal utilisé. La résolution de la tâche avec des images basses fréquences relève bien d'un traitement de l'orientation basé sur l'information physique. Par contre, en condition filtrage passe-haut, le profil classique d'anisotropie s'observe uniquement pour l'hippocampe dont l'orientation canonique est verticale. Pour le pigeon, dont l'orientation canonique est oblique, l'anisotropie est à nouveau renversée puisque c'est la cible physiquement verticale qui est le plus rapidement détectée. Les données obtenues en condition passe-haut suggèrent que ce n'est pas la dégradation *per se* des images qui permet d'observer un profil d'anisotropie classique en passe-bas, mais bien le fait que ce filtrage passe-bas ne permet plus une activation automatique des représentations mnésiques. Cette idée est actuellement confortée par le biais d'une tâche similaire avec des représentations non dégradées d'image à l'envers. Avec des animaux à l'envers, la reconnaissance est très largement ralentie et les données préliminaires montrent, dans cette situation, un profil d'anisotropie classique quelle que soit l'orientation canonique des items.

3.3. Expérience 3

Une autre façon d'étudier le rôle des représentations mnésiques est d'analyser le comportement de patients atteints de troubles de l'identification. Nous avons donc proposé la tâche de l'expérience 1 à un patient (M. H.J.A.) décrit par la littérature comme souffrant d'une forme particulièrement pure d'agnosie visuelle (e.g., Riddoch et Humphreys, 1987 ; Humphreys et Riddoch, 1987 ; Humphreys *et al.*, 1992). H.J.A. perçoit parfaitement bien les segments de ligne, les angles et les orientations (Humphreys *et al.*, 1992). Il a déjà participé à de nombreuses épreuves de recherche de cible et ses performances, bien qu'un peu ralenties, sont tout à fait normales lorsqu'il s'agit de rechercher un segment de ligne différant par son orientation. Cependant, H.J.A. ne reconnaît pas la plupart des objets qui lui sont présentés visuellement. H.J.A. n'ayant pas la possibilité d'accéder aux représentations mnésiques des items et donc à leur orientation canonique, on peut donc

s'attendre chez lui à une anisotropie de recherche classique, i.e. déterminée par l'orientation physique. Nous testons cette hypothèse en soumettant H.J.A. à l'expérience 1, restreinte aux items Hippocampe et Pigeon.

3.3.1. Méthode

Anamnèse d'H.J.A. — H.J.A., né en 1920, était âgé de 78 ans lors de la passation. Ses troubles font suite à un accident vasculaire cérébral post-opératoire en 1981, produisant une lésion bilatérale du cortex occipital étendue à une partie antérieure du cortex temporal. Son acuité visuelle est intacte sur une bonne partie du champ visuel, bien qu'il soit amputé dans sa partie supérieure. H.J.A. a des troubles dans la discrimination des couleurs et la reconnaissance. Ses performances sont relativement stables, il est, depuis 1981, incapable de reconnaître visuellement la plupart des objets même les plus communs, alors qu'il les reconnaît en exploration tactile et ne présente aucun trouble mnésique ou sémantique (il est parfaitement capable de définir les objets). Il peut nommer, si le temps de présentation n'est pas limité et s'il peut effectuer de nombreux déplacements du regard, 80% des objets réels qui lui sont présentés et environ 40% des dessins au trait. Certaines catégories d'objets telles que les outils, les voitures sont relativement préservées, mais il a beaucoup de difficulté pour les catégories vivantes comme les végétaux et les animaux.

Procédure et matériel — Nos expériences avec H.J.A. se sont déroulées au Centre des Sciences du Comportement et du Cerveau (Behavioural and Brain Centre) de l'université de Birmingham. Elles ont été réparties sur deux jours, au cours desquels H.J.A. a passé la tâche de recherche de cible de l'expérience 1 avec les exemplaires Hippocampe et Pigeon.

Après chacune des passations, nous lui demandions de nommer les exemplaires, mais il en était incapable. C'est pourquoi, à la fin des passations, nous lui avons remis une version imprimée de chacun des exemplaires pour qu'H.J.A. puisse procéder à une longue exploration visuelle. L'Hippocampe fut alors reconnu comme un oiseau, de même que le Pigeon. Suite à une demande de précision, H.J.A. déclare concernant le Pigeon « ne jamais en avoir vu et que ce doit être un oiseau rare, comme on en trouve dans les contrées lointaines ». En fait, H.J.A. se base uniquement sur les détails pour tenter de deviner ce que représentent ces exemplaires. C'est le long nez de l'Hippocampe qui, après une longue observation, lui a fait conclure à un bec.

3.3.2. Résultats expérience 3

L'analyse par item des temps de réponse d'H.J.A. montre un effet principal du Type d'animal utilisé ($F(1, 56) = 17.4$; $MSE = 76673$; $p < 0.0002$). H.J.A. répond plus vite pour l'exemplaire Pigeon (1048 ms) que pour l'exemplaire Hippocampe (1336 ms). Ce résultat est en accord avec les difficultés accentuées pour H.J.A. avec des dessins au trait comportant de nombreux détails internes au contour. On observe un effet principal de l'orientation physique de la cible ($F(1, 56) = 67$; $MSE = 76673$; $p < 0.0001$), puisqu'il détecte plus rapidement les cibles physiquement obliques (909 ms) que les cibles physiquement verticales (1475 ms). H.J.A. détecte plus rapidement la cible parmi 3 items (936 ms)

que parmi 6 (1449 ms, $F(1, 56) = 55$; $MSE = 76673$; $p < 0.0001$). Il n'y a pas d'interaction significative entre l'orientation canonique des exemplaires et l'orientation physique dans laquelle ils sont présentés ($F(1, 56) = 2.2$; $MSE = 76673$; $p > .10$), ni entre le Type d'exemplaire et le nombre d'items ($F(1, 56) = 2.2$; $MSE = 76673$; $p > .10$).

L'interaction entre l'orientation physique de la cible et le nombre d'items est significative ($F(1, 56) = 6.6$; $MSE = 76673$; $p < 0.02$). L'augmentation des temps de réponse en fonction du nombre d'items est plus marquée pour la cible verticale (1130 ms vs. 1821 ms $F(1, 56) = 50$; $MSE = 76673$; $p < 0.0001$) que pour la cible oblique (741 ms pour 3 items vs. 1077 ms pour 6 items, $F(1, 56) = 11.8$; $MSE = 76673$; $p < 0.002$). Enfin, l'interaction double entre l'orientation canonique des exemplaires, l'orientation physique dans laquelle ils sont présentés et le nombre d'items n'est pas significative ($F < 1$).

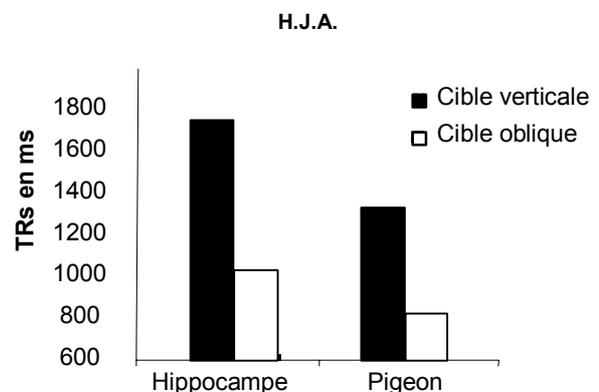


Figure 8 — Temps de réponse moyens d'H.J.A. pour détecter la cible présente en fonction de l'orientation physique de la cible (verticale vs. oblique).

3.3.3. Conclusion expérience 3

Les performances d'H.J.A. (figure 8) ne montrent pas de renversement de l'anisotropie en fonction de l'orientation canonique des animaux. En l'absence de reconnaissance on retrouve, quel que soit l'animal utilisé, une détermination physique de l'anisotropie (i.e. un avantage pour détecter la cible physiquement oblique).

4. Conclusion Générale

La recherche visuelle basée sur l'orientation physique d'objets révèle une interférence des informations mnésiques, bien que ces informations ne soient *a priori* d'aucune nécessité pour résoudre la tâche. L'expérience 1 montre que l'anisotropie dépend non pas de l'orientation physique des stimuli mais de leur orientation canonique. Les expériences 2 et 3 le confirment puisque sur la base des mêmes images, mais sans reconnaissance effective des items, l'anisotropie est à nouveau déterminée par l'orientation physique. Comme nous l'avons souligné en introduction, les influences de haut niveau sont peu représentées, voire parfois totalement absentes, des modélisations actuelles. L'ensemble des recherches présentées ici

suggère qu'une activation automatique et rapide des représentations mnésiques peut moduler la perception dès les premières étapes du traitement visuel. Pour renforcer cette conclusion, il serait intéressant d'envisager d'autres investigations des performances en vision de bas niveau (amorçage, ségrégation de textures...) confirmant l'intervention rapide des représentations mnésiques. Pour des modèles plus réalistes de la perception visuelle, capables de rendre compte des performances humaines, il faudrait alors envisager davantage de traitements descendants et donc préalablement mieux comprendre et identifier les facteurs pouvant intervenir. Soulignons pour conclure que l'activation automatique des représentations mnésiques est en accord avec les modèles parallèles distribués de la mémoire, où mémoire et traitement perceptif sont intimement mêlés. Mais l'incidence de cette activation sur le traitement perceptif ne peut être envisagée qu'au niveau théorique, car ce type de modèle se prête difficilement, actuellement, au parallélisme 'multiple objets' impliqué dans la recherche visuelle.

Références bibliographiques

- [Bernard et al., 1997] Bernard P., Guérin-Dugué A., Marendaz C. (1997). Modélisation d'une tâche de recherche visuelle par détecteurs d'orientation spatio-temporels en interaction. *JIOSC'97*. Lille, France.
- [Biederman, 1981] Biederman I. (1981). On the semantics of a glance at a scene. In M. Kubovy, & J.R. Pomeranz (Eds.), *Perceptual organisation*, Hillsdale : Lawrence Erlbaum. 213-254.
- [Biederman, 1987] Biederman I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*. 94, 115-145.
- [Biederman et Cooper, 1991a] Biederman I., Cooper E.E. (1991). Evidence for complete translational and reflectional invariance in visual object priming. *Perception*. 20(5), 585-593.
- [Biederman et Cooper, 1991b] Biederman I., Cooper E.E. (1991). Object recognition and laterality: Null effects. *Neuropsychologia*. 29(7), 685-694.
- [Biederman et Gerhardstein, 1993] Biederman I., Gerhardstein P.C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performances*. 19, 1162-1182.
- [Boucart et Humphreys, 1994] Boucart M., Humphreys G.W. (1994). Attention to orientation, size, luminance and color: attentional failure within the form domain. *Journal of Experimental Psychology: Human Perception and Performances*. 20, 61-80.
- [Boucart et Humphreys, 1997] Boucart M., Humphreys G.W. (1997). Integration of physical and semantic information in object processing. *Perception*. 26, 1197-1209.
- [Butter et Goodale, 2000] Butter C.M., Goodale M.A. (2000). Visual search selectively enhances recognition of the search target. *Visual Cognition*. 7, 769-784.
- [Chauvin et al., 2001] Chauvin A., Marendaz C., Hérault J. (2001). Modèle de carte de saillance attentionnelle et attracteurs visuels. *CJC IV*, Lyon, France.
- [Corballis, 1988] Corballis M.C. (1988). Recognition of disoriented shapes. *Psychological Review*. 95, 115-123.
- [Foster et Ward, 1991] Foster, D. H. Ward, P. A. (1991). Asymmetries in oriented-line detection indicate two orthogonal filters in early vision. *Proceedings of the Royal Society of London*. 243, 75-81.
- [Glaser et Glaser, 1989] Glaser, M.O. Glaser, W.F. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*. 118, 13-42.
- [Hinton, 1981] Hinton G.E. (1981). A parallel computation that assigns canonical object-based frames of reference. *Proceeding of the International Joint Conference on Artificial Intelligence*. Vancouver, Canada.
- [Hummel et Biederman, 1992] Hummel J.E., Biederman I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*. 99, 480-517.
- [Humphreys et Riddoch, 1987] Humphreys G.W., Riddoch M.J. (1987). *To see but not to see: A case study of visual agnosia*. London : Erlbaum.
- [Humphreys et al., 1992] Humphreys G.W., Riddoch M.J., Donnelly N., Freeman T., Boucart M., Muller H.M. (1992). Intermediate visual processing and visual agnosia. In M. Farah & G. Ratcliff (Ed.), *Higher level cortical vision*. London : Erlbaum.
- [Itti et Koch, 2000] Itti L., Koch C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*. 40, 1489-1506.
- [Itti et al., 1998] Itti L., Koch C., Neibur E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transaction on Pattern Analysis and Machine Intelligence*. 20, 1254-1259.
- [Jolicoeur, 1985] Jolicoeur P. (1985). The time to name disoriented natural objects. *Memory and cognition*. 13 (4), 289-303.
- [Jolicoeur, 1992] Jolicoeur P. (1992). Identification of disoriented objects: a dual-system theory. In G.W. Humphreys (Ed.), *Understanding vision*. Oxford : Blackwell. 180-198.
- [Julesz et Schumer 1981] Julesz B., Schumer R.A. (1981). Early visual perception. *Annual Review of Psychology*. 32, 575-627.
- [Lawson et al., 2000] Lawson R., Humphreys G.W., Jolicoeur P. (2000). The combined effects of plane disorientation and foreshortening on picture naming: One manipulation or two? *Journal of Experimental Psychology: Human Perception and Performance*. 26, 568-581.
- [Leek, 1998] Leek E.C. (1998). Effects of stimulus orientation on the identification of common polyoriented objects. *Psychonomic Bulletin & Review*. 5, 650-658.
- [Marr, 1982] Marr D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. New York : W.H. Freeman and Company.
- [Marr et Nishihara, 1978] Marr D., Nishihara, H.K. (1978). Representation and recognition of the spatial organisation of three-dimensional shapes. *Proceedings of the Royal Society of London*. B207, 187-217.
- [Marendaz, 1998] Marendaz C. (1998). Nature and Dynamics of reference frames in visual search for orientation. Implications for early visual processing. *Psychological Science*. 9, 27-32.
- [Marendaz et al., 1993] Marendaz C., Stivalet P., Barraclough L., Walkowiak P. (1993). Effect of gravitational cues on visual search for orientation. *Journal of Experimental Psychology: Human Perception and Performance*. 19, 1266-1277.
- [Marendaz et al., 1991] Marendaz C., Stivalet P., Genon D. (1991). Coding of orientation in early vision: search asymmetry revisited. *European Bulletin of Cognitive Psychology*. 11(4), 427-440.

[Palmer, 1980] Palmer S.E. (1980). What makes triangles point: Local and global effects in configurations of ambiguous triangles. *Cognitive Psychology*. 12, 285-305.

[Palmer et al., 1981] Palmer S.E., Rosch E., Hase P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention and Performance IX*. Hillsdale : Lawrence Erlbaum. 135-151.

[Quinlan et Humphreys, 1993] Quinlan P.T., Humphreys G.W. (1993). Perceptual frames of reference and two-dimensional shape recognition: Further examination of internal axes. *Perception*. 22, 1343-1364.

[Riddoch et Humphreys, 1987] Riddoch M.J., Humphreys G.W. (1987). A case of integrative visual agnosia. *Brain*. 110, 1431-1462.

[Rock, 1973] Rock I. (1973). *Orientation and form*. New York : Academic Press.

[Srinivas, 1995] Srinivas K. (1995). Representation of rotated objects in implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 21, 1019-1036.

[Tipper, 1985] Tipper S.P. (1985). The negative priming effect: inhibitory priming by ignored objects. *Quarterly Journal of experimental Psychology*. 37, 571-590.

[Treisman, 1985] Treisman A. (1985). Preattentive processing in vision. *Computer vision, graphics, and image processing*. 31, 156-177.

[Treisman, 1993] Treisman A. (1993). The perception of features and objects. In D.A. Baddeley et L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control*. England: Oxford University Press.

[Vannucci et Viggiano, 2000] Vannucci M. Viggiano M.P. (2000). Category effects on the processing of plane rotated objects. *Perception*, 29, 287-302.

[Verfaillie et Boutsen, 1995] Verfaillie K., Boutsen L. (1995). A corpus of 714 full-color images of depth-rotated objects. *Perception and Psychophysics*. 57(7), 925-961.

[Wolfe, 1994] Wolfe J. M. (1994). Guided search 2.0 a revisited model of visual search. *Psychonomic Bulletin & Review*. 1, 202-238.

Les auteurs

Après un doctorat sur l'intégration des informations visuo-posturales et mnésiques en vision de bas niveau, préparé à Grenoble, Cécile Ballaz est actuellement en stage post-doctoral au *Behavioural and Brain Centre* (Birmingham, Royaume-Uni). Alan Chauvin prépare un doctorat d'informatique et Carole Peyrin un doctorat de Psychologie Cognitive, tous deux à Grenoble. Christian Marendaz est professeur des universités (UPMF Grenoble).

