

La perception visuelle des objets

Muriel BOUCART

Résumé

Cet article résume les connaissances actuelles sur la perception visuelle des objets. Les données de la neurophysiologie, de la psychophysique, de la psychologie cognitive chez le sujet normal, de la pathologie cérébrale et des modèles computationnels sont prises en compte. L'article est divisé en quatre parties. La première partie traite des régions cérébrales impliquées dans la perception des objets. La seconde partie décrit les processus d'intégration et de segmentation des formes et les mécanismes impliqués dans le traitement des contours et de la surface des formes. La troisième partie présente les modèles d'organisation des représentations d'objets en mémoire et des données sur le rôle des informations de forme et de surface dans la reconnaissance des objets. Les effets de contexte sont traités dans la dernière partie qui présente les connaissances actuelles sur le traitement des scènes visuelles.

Mots clés : perception visuelle, reconnaissance d'objets.

Key words: visual perception, object recognition.

Adresse : Hôpitaux Universitaires de Strasbourg, INSERM U 405,
Département de Psychiatrie I, 67091 Strasbourg Cedex, France (e-mail:
boucart@alsace.u-strasbg.fr).

INTRODUCTION

Ces dernières années ont vu se développer un intérêt croissant pour la compréhension des processus cognitifs impliqués dans la reconnaissance des objets. Comme pour la mémoire ou le langage, il ressort clairement des données de la psychologie et de la neuropsychologie cognitive que la reconnaissance des objets ne relève pas d'un processus unitaire mais implique une multitude de traitements. L'objectif de cet article est de présenter de la manière la plus exhaustive possible les connaissances actuelles sur la perception des objets. J'ai choisi d'employer le terme "perception" plutôt que la terminologie classique, "reconnaissance" des objets (object recognition), parce que le terme reconnaissance fait référence à la comparaison de la forme physique d'un objet aux représentations des objets stockées en mémoire afin d'attribuer une identité à cette forme. Le traitement visuel des objets ne commence pas à ce niveau, loin s'en faut. Une multitude d'opérations a eu lieu sur les caractéristiques physiques de la forme bien avant l'accès aux représentations en mémoire et ce sont ces traitements, trop souvent ignorés, qui conditionnent la reconnaissance des objets. La présentation prendra donc en compte les données de la neurophysiologie, de la psychophysique, de la psychologie expérimentale chez le sujet normal et de la neuropsychologie. Ces différentes approches ne seront pas développées de manière équivalente dans la mesure où plusieurs articles dans ce numéro traitent spécifiquement des données provenant des études sur l'animal (voir les articles de Vighetto et de Thorpe) et de la pathologie cérébrale (voir les articles de Bruyer, de Charalet et Carbonnel et de Sartori).

ASPECTS PHYSIOLOGIQUES

Les données électrophysiologiques chez le singe montrent que l'information visuelle est traitée dans des voies qui impliquent des structures anatomiques indépendantes spécialisées dans l'analyse d'informations différentes (Desimone et Ungerleider, 1989 ; Merigan et Maunsell, 1993). Ainsi, l'aire V1 (le cortex strié) est composé de régions qui comportent des cellules sélectives à l'orientation, aux fré-

quences spatiales, aux longueurs d'onde, aux directions de mouvement ... (Desimone et Ungerleider, 1989 ; Hubel et Livingstone, 1987 ; Maunsell et Newsome, 1987 ; Zeki, 1990). Cette subdivision se retrouve dans l'aire V2 avec un degré d'intégration plus élevé puisque les régions anatomiques comportent des neurones qui intègrent des signaux provenant de différentes classes de cellules (Peterhans et von der Heydt, 1989). Néanmoins l'aire V2 reste spécialisée au sens où elle comporte, comme l'aire V1, des neurones répondant sélectivement à différentes caractéristiques physiques élémentaires. Les neurones des différentes régions de l'aire V2 se projettent dans différentes aires corticales. A ce niveau, les données physiologiques divergent sur le degré de spécialisation de ces aires. Zeki (1990) suggère une spécialisation de l'aire V4 dans le traitement des informations de couleur tandis que l'aire V3 comporte des cellules répondant à des caractéristiques de forme (orientation, taille) et l'aire MT traite les informations de mouvement. Le traitement indépendant des informations de forme et de couleur est discuté par d'autres auteurs (Desimone et Schein, 1987 ; Heywood et Cowey, 1987 ; Schein et Desimone, 1990) qui rapportent l'existence de cellules répondant à des informations combinées de forme et de couleur dans l'aire V4. Il se peut néanmoins que la réponse de ces cellules soit influencée par les connexions latérales entre les aires V3 et V4 (voir Vighetto dans ce numéro). Par ailleurs, des données récentes obtenues chez le sujet humain avec la technique de la caméra à position semblent conforter l'hypothèse d'une sélectivité pour la couleur dans l'aire V4 (le gyrus fusiforme chez le sujet humain). Dans cette étude Zeki et collègues (1991) ont comparé l'activité du cortex dans trois conditions expérimentales : (1) pendant la présentation de stimuli composés de rectangles colorés, (2) pendant la présentation des mêmes stimuli en nuances de gris et (3) dans une condition contrôle où les sujets avaient les yeux fermés. La soustraction des images cérébrales obtenues dans ces trois conditions met en évidence une activation du gyrus fusiforme pour les stimuli colorés. Un traitement des informations de forme et de couleur dans cette région aurait donné lieu à une activation pour les stimuli en nuances de gris qui ne différaient des figures colorées que par la présence d'informations chromatiques. Les neurones des régions V3 et V4 se projettent dans le cortex inféro-temporal (IT) qui peut être défini comme la région où la perception ren-

contre la mémoire (Miyashita, 1993). A ce niveau les cellules répondent à des informations combinées de couleur, de texture, de taille, d'orientation (Tanaka et al., 1991) et à des informations structurées telles que des formes (Fujita et al., 1992) ou des visages (Baylis et al., 1987 ; Hasselmo et al., 1989). Les études neurophysiologiques montrent que le traitement de l'information visuelle suit une progression d'un traitement très local des caractéristiques élémentaires du signal à une représentation intégrée plus globale.

LA STRUCTURATION DES FORMES

Comment le système visuel intègre-t-il des caractéristiques locales codées séparément en une configuration globale qui servira de base pour la comparaison aux représentations d'objets stockés en mémoire ? Considérons d'abord le traitement des contours. Dans une scène visuelle, la délimitation des surfaces des formes constitue une étape importante de la perception des objets puisque ce traitement permet la ségrégation d'une figure du fond, la segmentation des figures entre elles et le codage des relations spatiales entre les formes et entre les parties des formes. Dans l'environnement naturel, les limites des surfaces peuvent être définies par des différences de luminance, de texture, de couleur, de disparité ou de direction de mouvement. Malgré cette hétérogénéité, le système visuel est capable de généraliser et de traiter ces contours de manière équivalente même si les bords présentent des discontinuités. Les données de l'électrophysiologie et de la psychophysique visuelle montrent que deux principales classes de cellules interviennent dans le traitement des informations de contour : (1) les détecteurs d'orientation qui codent les alignements locaux de transitions de couleur, de luminance, et (2) les détecteurs de terminaison (end-stopped cells) qui codent les discontinuités telles que les fins de lignes, les coins ou les jonctions. Ces deux classes de cellules ont des fonctions différentes. Les détecteurs d'orientation interviennent de façon prévalente dans les processus d'intégration des contours locaux en une configuration globale tandis que les détecteurs de terminaison interviennent plutôt dans les processus de segmentation, c'est-à-dire la délimitation des par-

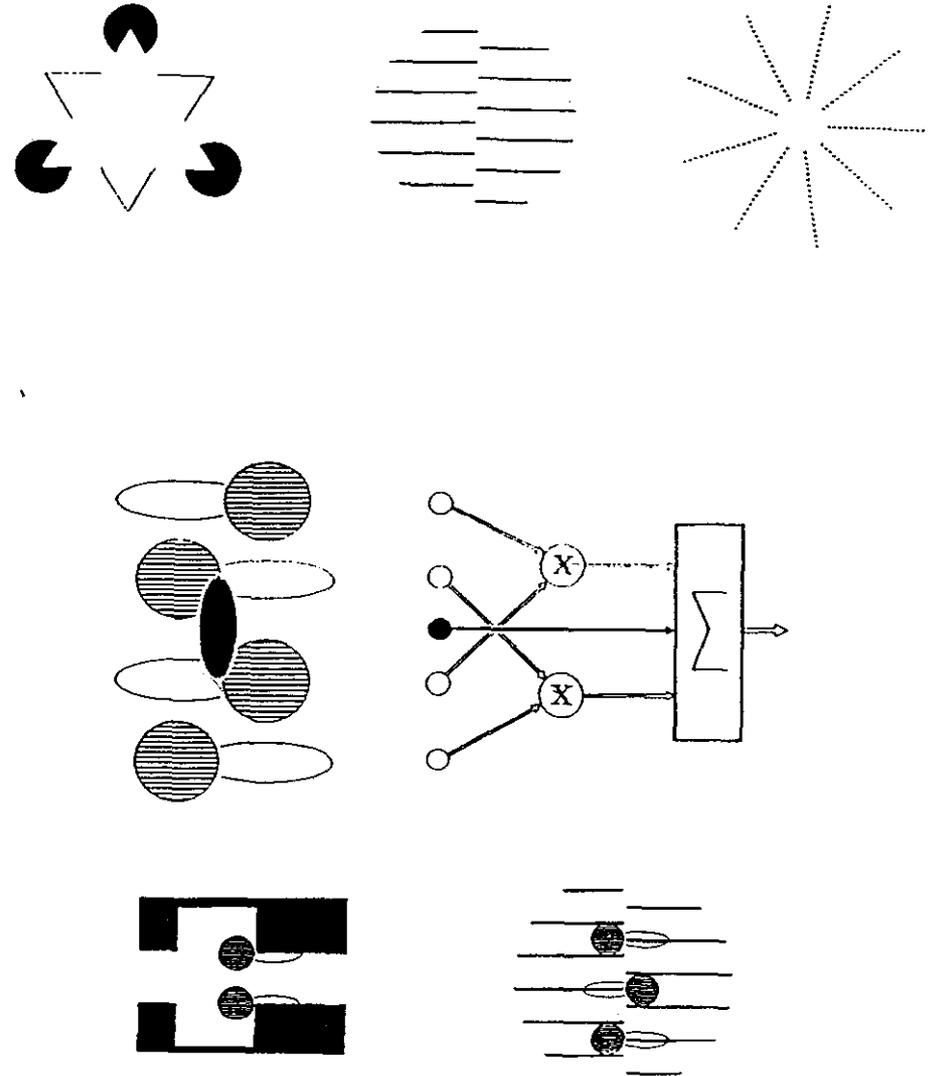
ties des formes ou des formes entre elles (Boucart, Delord et Giersch, 1994 ; Lorenceau et Shiffrar, 1991).

Peterhans et von der Heydt (1989, 1991) ont étudié les mécanismes impliqués dans le traitement précoce des contours en enregistrant la réponse des cellules des aires V1, V2 et V3 du cortex du singe pour des bords continus (des barres) et des bords discontinus composés de points alignés ou de barres alignées produisant un contour illusoire (voir Figure 1). Ils ont montré que les cellules sélectives à l'orientation dans l'aire V1 répondent à la présentation de barres continues mais ne répondent pas à des barres composées d'éléments discontinus. Une petite brèche de 2 à 8 minutes d'arc est suffisante pour abolir totalement la réponse des cellules. Ces données suggèrent que les brèches sont codées très précocement dans le cortex strié. En revanche, un tiers des cellules enregistrées dans l'aire V2 répondent de façon équivalente pour des bords continus et des bords discontinus composés d'éléments alignés. Ces résultats montrent que les cellules de l'aire V2 sont capables d'intégrer des contours locaux. Le mécanisme proposé par Peterhans et von der Heydt (1989, 1991 ; von der Heydt et Peterhans, 1989) pour expliquer ce processus d'intégration est basé sur un mécanisme de sommation de signaux provenant des détecteurs d'orientation (les champs récepteurs des cellules simples ou complexes de l'aire V1) et des détecteurs de terminaison dont les champs récepteurs sont alignés aux extrémités des axes des détecteurs d'orientation (voir Figure 1). L'alignement (la colinéarité) constitue une propriété physique importante dans l'intégration des contours locaux. L'amplitude de la réponse des neurones qui intègrent des caractéristiques locales (les "contour neurons", Peterhans et von der Heydt, 1989) est d'autant plus élevée que le nombre de coins ou d'extrémités de barres tombant dans la région excitatrice des champs récepteurs des détecteurs de terminaison est important. Une démonstration est présentée Figure 2. Le graphique montre qu'un écart aussi petit que 2 minutes d'arc dans l'alignement des points suffit pour réduire de moitié l'amplitude de la réponse de la cellule enregistrée et un écartement plus important élimine totalement la réponse du neurone (Peterhans et von der Heydt, 1991). Cet effet se produit à la fois pour des bords stationnaires et des bords en mouvement.

L'équivalence perceptive des bords continus et des bords discontinus composés d'éléments alignés et proches a également été démontrée chez le sujet humain dans différents paradigmes. Par exemple, Zucker et Davis (1988) ont testé l'existence d'une distance critique en deçà de laquelle des éléments discrets sont intégrés et traités de façon similaire à une ligne continue. L'illusion du soleil était utilisée comme stimulus (voir Figure 1). Lorsque les extrémités des barres forment un cercle, un rehaussement de luminosité apparaît au centre de la figure. Le centre paraît plus brillant que le fond. Zucker et Davis ont utilisé des lignes composées de points dont l'espacement était manipulé. Pour chacun des espacements, le sujet devait dire s'il percevait le rehaussement de luminosité au centre de la figure. Les résultats montrent que le rehausse-

Figure 1. Exemples de stimuli à contours illusoires. L'alignement des extrémités des barres induit la perception d'un contour illusoire. Le mécanisme de traitement des contours proposé par Peterhans et von der Heydt (1989, 1991) est présenté dans la partie inférieure. Les barres orientées activent les champs récepteurs des cellules sensibles à l'orientation. Les extrémités des barres activent les champs récepteurs des détecteurs de terminaison (end-stopped cells). Les "contour neurons" de l'aire V2 somment l'activité des signaux de ces deux classes de cellules. Plus le nombre de fins de lignes tombant dans la région excitatrice (ellipse blanche) des détecteurs de terminaison est important plus le "contour neuron" est activé. La réponse est réduite lorsque les fins de ligne tombent dans la région inhibitrice des détecteurs de terminaison (parties grisées).

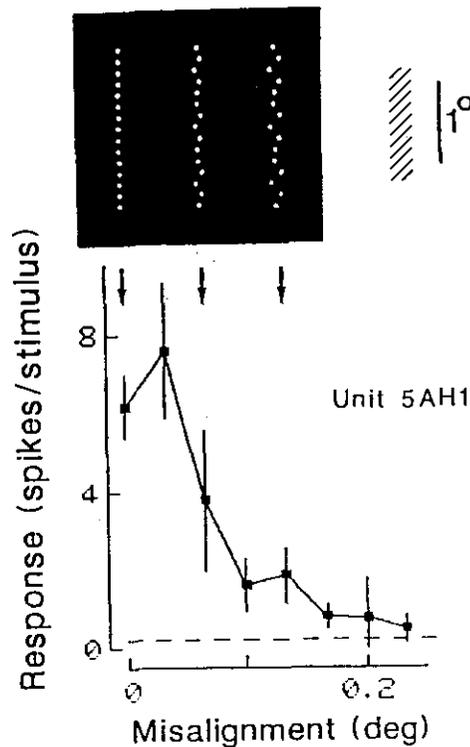
Figure 1. Examples of stimuli with illusory contours. The alignment of the line ends induces the perception of an illusory contour. The mechanism proposed by Peterhans and von der Heydt (1989, 1991) is displayed in the lower part of the figure. Oriented bars activate the receptive field of cells sensitive to orientation. The line ends activate the receptive field of end-stopped cells. "Contour neurons" in area V2 sum signals from these two types of cells. The degree of activation of the "contour neuron" is determined by the number of line ends that fall in the excitatory part (white ellipses) of the receptive field of end-stopped cells. The response is reduced when line ends fall in the inhibitory part (grey circles) of the receptive field.



ment de luminosité est perçu pour des distances entre les points inférieures à 6 minutes d'arc. Au-delà de cette distance critique, les points ne sont pas traités de la même façon qu'une ligne continue par le système visuel. L'explication des auteurs est la suivante : en deçà d'un espacement critique, les points tombent dans le champ récepteur des neurones sélectifs à l'orientation des lignes. Le champ récepteur du détecteur d'orientation est activé. L'activation est d'autant plus importante que les points sont proches et, donc, que la densité est élevée. Si la distance est trop grande, les points tombent aux extrémités ou en dehors du champ récepteur. L'activation est alors réduite ou éliminée. Cette distance critique ne vaut vraisemblablement que pour des présentations fovéales puisque la taille des champs récepteurs augmente en fonction de l'excentricité rétinienne.

Figure 2. Illustration du rôle de la colinéarité dans les processus d'intégration des contours. La réponse du neurone de l'aire V2 diminue en fonction de l'écartement des points.

Figure 2. Illustration of the role of colinearity in integration processes of local contour information. The response of cells in area V2 decreases with the increase in misalignment of the dots.



Boucart et al. (1994) ont montré qu'une distance critique pour l'intégration de contours locaux pouvait aussi s'appliquer à des stimuli plus complexes. Dans cette étude les stimuli étaient des dessins d'objets incomplets dont le contour était composé de segments de ligne. L'espacement et l'alignement des contours locaux étaient manipulés. Les sujets devaient appairer une figure de référence présentée centralement à deux figures présentées latéralement. La performance, en termes de temps de réponse et de taux d'erreurs, était équivalente pour les dessins dont le contour était complet et pour des dessins dont les éléments du contour étaient espacés de moins de 10 minutes d'arc. La performance se dégradait brutalement lorsque l'espacement passait de 10 à 20 minutes d'arc. Les stimuli et les résultats sont présentés Figure 3. Toutefois, les auteurs ont montré que, pour des objets, la perturbation des processus d'intégration précoces due à un espacement trop grand peut être compensée par des processus de haut niveau. Ainsi la présentation de l'objet complet, aisément identifiable, avant la présentation de sa version incomplète améliore la précision de la performance, même si la tâche de comparaison n'implique pas directement l'identification des objets. Ces données suggèrent que des stratégies de compensation peuvent être mises en oeuvre en cas de déficit des processus d'intégration précoces. Une démonstration empirique a récemment été rapportée par Giersch et al. (1995) dans une étude pharmacologique sur une substance affectant les inhibitions latérales entre des cellules sensibles à l'orientation.

Le mécanisme proposé par Peterhans et von der Heydt (1989) permet de rendre compte du traitement des contours locaux sur une petite région du champ visuel. Comment le système visuel intègre-t-il les contours locaux sur des grandes parties du champ visuel ? Plusieurs types de processus d'intégration (binding processes) ont été proposés. Gilbert et Wiesel (1989) et, plus récemment, Field et al. (1993) ont proposé la notion de "champ d'association" (association field). Le cortex visuel est composé de colonnes comportant des cellules sensibles à différentes orientations. Un champ d'association se définit par des connexions latérales entre des colonnes de cellules sensibles à la même orientation ou à des orientations voisines. Le champ d'association couvre une région plus grande que la taille des champs récepteurs des détecteurs d'orientation qui codent les contours locaux de la forme. Le

Figure 3. Exemples de stimuli utilisés par Boucart et al. (1994). Les contours locaux des objets varient sur leur colinéarité (NC = non colinéaire) et sur la proximité spatiale des éléments (S2 = 4.2 minutes d'arc, S5 = 10.8 minutes d'arc et S10 = 22.2 minutes d'arc). Les résultats montrent que la performance se dégrade brusquement pour les stimuli ayant un espacement supérieur à 11 minutes d'arc et pour les figures dont les éléments ne sont pas alignés.

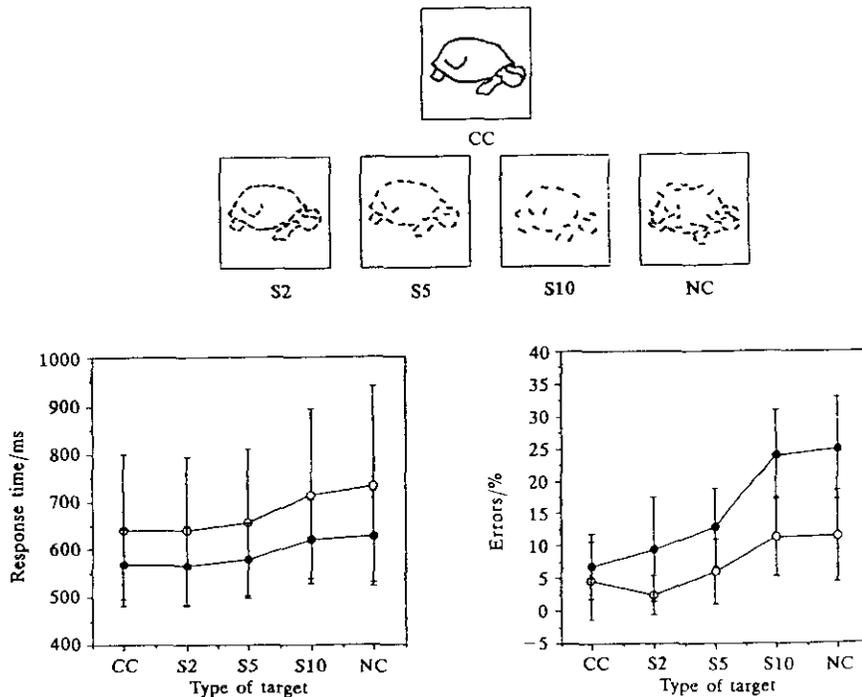


Figure 3. Examples of the stimuli used by Boucart et al. (1994). Local contours of objects vary on colinearity (NC = non colinear) and spacing between elements (S2 = 4 minutes of arc, S5 = 10.8 minutes of arc, and S10 = 22.2 minutes of arc). The results show that performance is degraded for stimuli with a large spacing between elements (more than 11 minutes of arc) and for stimuli lacking colinearity).

groupement par un mécanisme d'excitation latérale entre des cellules voisines permet de rendre compte de l'intégration de lignes ou de courbes sur des grandes parties du champ visuel (Field et al., 1993). Une démonstration empirique présentée par Kovacs et Julesz (1993, 1994) illustre la notion de champ d'association. Le groupement de segments d'orientation voisine facilite la discrimination d'un cercle parmi des distracteurs (voir Figure 4).

Figure 4. Illustration de la notion de champ d'association. Les connexions latérales entre des cellules sensibles à des orientations voisines permettent de grouper les éléments et de discriminer la figure du fond (tiré de Kovacs & Julesz, 1993).

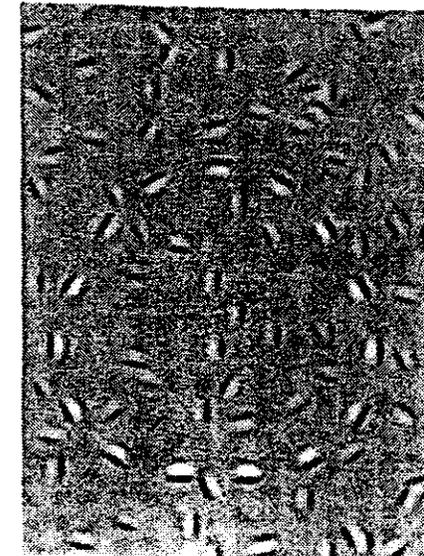


Figure 4. Illustration of the notion of association field. Lateral connections between cells sensitive to similar orientations allow a grouping of local elements and figure/ground segregation (from Kovacs & Julesz, 1993).

Un autre mécanisme, qui est complémentaire de la notion de champ d'association, a été proposé par Gray et collègues (1989, 1991 ; voir également Eckhorn et al., 1988 ; Eckhorn et Obermuller, 1993). Il est basé sur la synchronisation des oscillations des réponses des neurones sélectifs à la même orientation. Ces études réalisées chez le chat montrent que, lorsqu'une barre orientée traverse le champ visuel de l'animal, les cellules sensibles à l'orientation de cette barre synchronisent leur réponse. La synchronisation de l'activité des cellules se produit à la fois entre des colonnes de cellules proches avec recouvrement des champs récepteurs mais aussi pour des régions éloignées avec des espacements de l'ordre de 7 mm dans l'aire 17 du cortex visuel du chat. Lorsque deux barres perpendiculaires sont présentées, deux groupes distincts de cellules synchronisent leur activité pour chacune des orientations. Ce mécanisme pourrait donc constituer le processus de mise en relation des contours locaux dans différentes régions du champ visuel.

Un mécanisme de synchronisation de l'activité des unités a été introduit dans le modèle connexionniste de reconnaissance d'objets de Hummel et Biederman (1992). Dans ce modèle, les unités oscillent en phase lorsqu'elles représentent les caractéristiques locales (bords et angles) d'une même partie de l'objet et elles se désynchronisent lorsqu'elles représentent les caractéristiques physiques de différentes parties de l'objet. Ce mécanisme permet donc d'isoler les différentes parties de l'objet.

D'autres mécanismes de groupement faisant intervenir des processus de plus haut niveau ont été suggérés. Par exemple Ullman (1984) et plus récemment Jolicoeur (1986), Jolicoeur et al. (1991), et McCormick et Jolicoeur (1992) ont proposé l'existence d'une "routine de traçage" (boundary tracing operation) qui se réalise en déplaçant un opérateur de taille variable le long des lignes et des courbes du contour de la forme. L'opérateur est une fenêtre attentionnelle qui se déplace d'une région à une autre du contour. La routine de traçage, telle qu'elle est présentée par les auteurs, apparaît constituer une opération séquentielle et coûteuse en temps. Des études sur la perception de scènes visuelles montrent que la plupart des objets ainsi que leurs relations spatiales et sémantiques sont identifiés avec des durées de présentation très brèves (150 msec ou moins ; Biederman, 1981 ; Boyce et Pollatsek, 1989 ; voir Thorpe dans ce numéro). Pour cette raison un

traitement en parallèle des limites des surfaces sur différentes régions du champ visuel semble plus vraisemblable.

Les objets et les scènes visuelles ne sont pas composés que de lignes et de courbes. Ils comportent d'autres caractéristiques physiques qui contribuent à l'organisation perceptive des éléments de la forme et à l'identification des objets. Ces caractéristiques sont les jonctions et les informations de surface. Des études sur le traitement des jonctions suggèrent que ces caractéristiques sont extraites très précocement et qu'elles sont utilisées pour délimiter les parties des formes (Enns et Rensink, 1991 ; Humphreys et al., 1994). Enns et Rensink (1991) ont montré que différents types de jonctions fournissent des informations différentes dans l'analyse de l'image. Les jonctions en forme de L signalent les coins des surfaces. Les jonctions en forme de T correspondent le plus souvent à des occlusions ; elles signalent que deux surfaces diffèrent dans leur profondeur relative. Les jonctions en forme de Y ou de flèches fournissent des informations sur la profondeur ; elles signalent les coins de deux ou trois surfaces visibles et elles donnent des informations sur l'orientation convexe ou concave d'une surface. Une illustration est présentée Figure 5. Les jonctions sont traitées dans la couche 2 du modèle connexionniste de Hummel et Biederman (1992). Elles ont pour fonction de segmenter la forme en composantes (les parties de l'objet) et aussi de spécifier la nature des composantes (volumétriques ou plates) et les relations entre les parties de l'objet (leur position relative les unes par rapport aux autres et la présence d'occlusions).

Dans une version antérieure d'un modèle de reconnaissance d'objets, Biederman (1987) avait démontré empiriquement que ces régions du contour jouent un rôle important dans l'identification des objets. La suppression des jonctions dans un dessin au trait affecte plus la précision et le temps de dénomination que la suppression des segments de ligne ou de courbe dans des proportions équivalentes. Cet effet a été confirmé dans une étude récente utilisant un paradigme d'amorçage (Boucart, Grainger et Ferrand, 1995). Dans cette étude, un objet incomplet dont le contour était composé soit de segments de lignes soit de coins ou de jonctions (Y, T et flèche) était présenté en amorce. Des exemples des stimuli sont présentés Figure 6. Il était suivi d'un masque puis d'un objet qui était la version complète de la figure présentée

comme amorce. La tâche des sujets était de dénommer l'objet au contour complet. Les résultats montrent que la présentation préalable d'un contour composé de jonctions facilite l'identification de la version complète de l'objet présenté ultérieurement. La facilitation était plus prononcée pour les jonctions tridimensionnelles (Y, T et flèche) que pour les coins. Ce résultat suggère (1) que les objets pourraient être représentés en mémoire sous la forme d'entités volumétriques, et (2) que les caractéristiques tridimensionnelles facilitent l'identification en réduisant le nombre de candidats objets activés avant l'apparition de l'objet complet.

Figure 5. Exemples de jonctions dans une scène visuelle. Les jonctions en L signalent les extrémités des objets ou des parties des objets. Les jonctions en T délimitent les objets ou les parties d'un même objet. Les jonctions en Y ou en flèche signalent les coins de deux ou trois surfaces qui diffèrent sur leur profondeur.

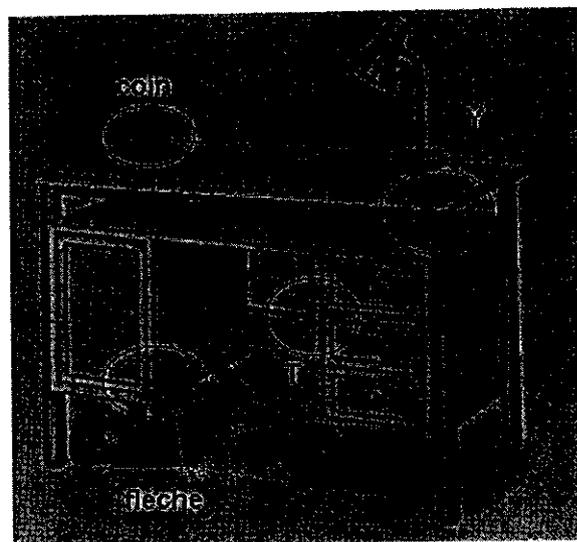


Figure 5. Examples of junctions in a visual scene. L junctions signal the corners of objects or of object parts. T junctions separate objects or object parts. Y and arrow junctions signal the corner between two or three visible surfaces differing in their depth.

Figure 6. Stimuli utilisés dans l'étude de Boucart et al. (1995). (A) formes comportant des coins et des segments alignés sur le contour, (B) formes comportant uniquement des coins, (C) formes comportant uniquement des segments de lignes, (D) formes comportant des éléments non alignés sur le contour et (E) formes comportant des coins et des jonctions tridimensionnelles (Y, T et flèches).

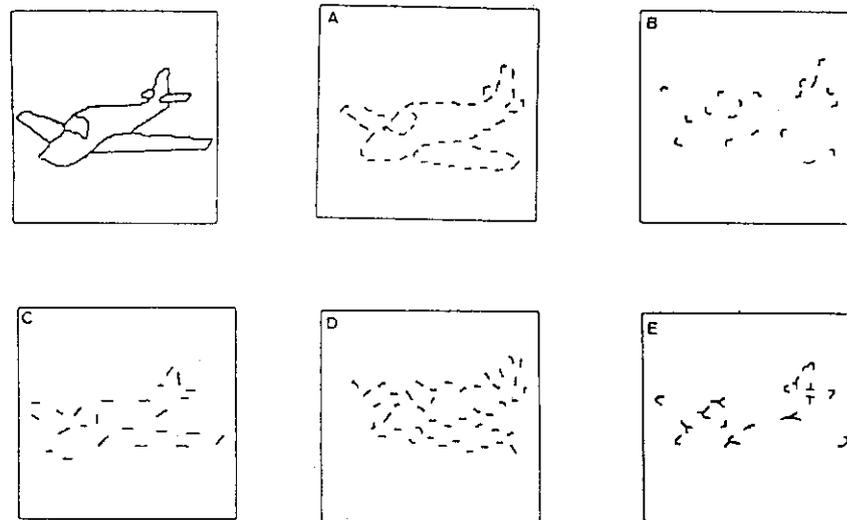


Figure 6. Stimuli used in Boucart et al. (1994). A: forms with corners and aligned elements; B: forms containing only corners; C: forms containing only mid-segments; D: forms with non-collinear elements; E: forms with corners and three-dimensional features (Y, T, and arrow junctions).

Où sont codées les jonctions ? Pour certains auteurs, les jonctions sont extraites très précocement. Elles sont codées par les champs récepteurs asymétriques des détecteurs de terminaison (Peterhans et von der Heydt, 1991). Un traitement plus tardif a été suggéré par Fujita et al. (1992) qui ont rapporté l'existence de colonnes de cellules répondant à des formes en L, T, Y ou flèche dans le cortex inféro-temporal du singe.

Les études sur la perception des formes distinguent le traitement des caractéristiques de forme (contour, taille, orientation) et le traitement des caractéristiques de surface (luminance, couleur, texture). Le traitement précoce des informations de surface a principalement été étudié sur des figures à contour illusoire (voir Figure 1) qui se caractérisent par la perception d'un contour et d'un rehaussement de luminosité au centre de la figure.

Dans une approche computationnelle, Grossberg et Mingolla (1985, 1987 ; Grossberg, 1992) ont proposé l'existence de deux systèmes pour le traitement des contours et des surfaces. Les deux systèmes opèrent en parallèle et coopèrent. Le système de traitement du contour (Boundary Contour System) est sensible à l'orientation mais pas à la polarité du contraste ; l'activité des unités qui codent les bords se propage vers l'intérieur, dans le prolongement des axes des détecteurs d'orientation. Le système de traitement de la surface (Feature Contour System) est sensible à la polarité du contraste mais pas à l'orientation ; l'activité des unités qui codent la surface se propage vers l'extérieur. Les mécanismes de propagation des signaux des deux systèmes sont illustrés Figure 7.

Sur le plan neurophysiologique la perception d'une surface plus lumineuse que le fond est expliquée de la façon suivante : près d'un bord qui présente une inversion de contraste (des disques noirs sur un fond blanc) l'activité des cellules à champ récepteur concentrique qui codent les informations clair/sombre s'accroît et s'étend. Le signal se diffuse vers l'extérieur des bords du disque. L'accroissement et la diffusion d'activité sont d'autant plus importants que le contraste entre le disque et le fond est élevé et que la taille du disque est grande (Dresp et Bonnet, 1991). La limitation du rehaussement de luminosité à l'intérieur de la figure est expliquée par Grossberg et Mingolla (1985) par l'intervention du système de traitement des contours. Le système envoie des "signaux d'arrêt" qui contiennent la diffusion d'activité du système de traitement des surfaces à l'intérieur de la figure. La diffusion s'arrête lorsqu'elle se heurte à un contour. Le rôle des informations de surface pour les stimuli plus complexes (les objets) sera développé dans la partie concernant la reconnaissance des objets.

Figure 7. Illustration du fonctionnement des systèmes impliqués dans le traitement des informations de contour et de surface. Le "boundary contour system" code l'orientation, quelle que soit la polarité du contraste (noir ou blanc). Les signaux se propagent vers l'intérieur (flèches). Le "feature contour system" code la luminance ou la polarité du contraste, quelle que soit l'orientation. Une inversion brutale de contraste déclenche la réponse du neurone. Le signal se diffuse vers l'extérieur (tiré de Grossberg, 1992).

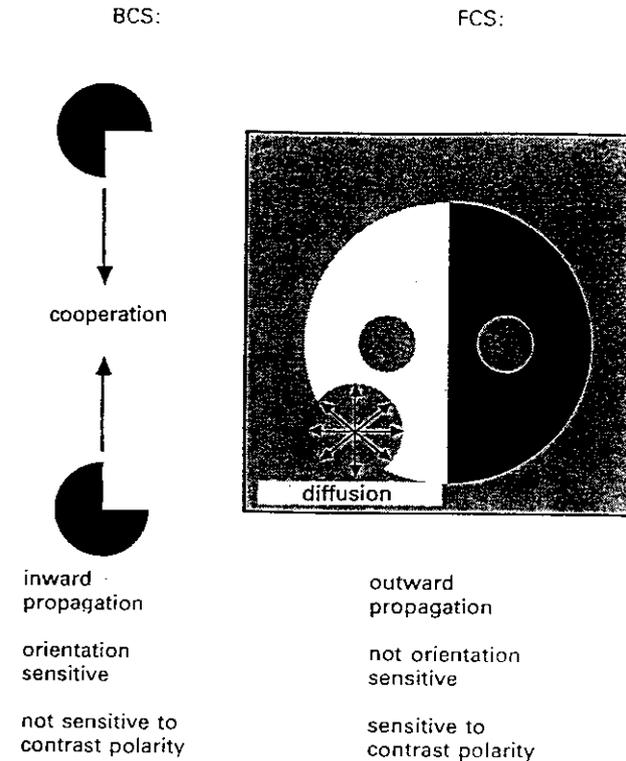


Figure 7. Illustration of the systems involved in the processing of contour and surface information. The "Boundary Contour System" codes orientation independently of the direction of contrast (black or white). Signals propagate inwardly. The "Feature Contour System" codes luminance and contrast polarity but is insensitive to orientation. An inversion of contrast induces the response of the neuron. Signals propagate outwardly (from Grossberg, 1992).

Les processus impliqués dans le traitement des informations de contour et de surface des formes sont pour tous les auteurs dont les travaux ont été résumés ci-dessus des processus précoces, automatiques et cognitivement impénétrables. Dans les scènes visuelles naturelles, les discontinuités dues à des occlusions de parties du contour d'un objet par d'autres objets sont fréquentes. Un exemple est présenté Figure 8. Par ailleurs, la taille des brèches est souvent bien supérieure aux espaces entre éléments utilisés dans les investigations portant sur les mécanismes d'intégration des contours. Dans ce cas, d'autres mécanismes faisant appel à des processus de plus haut niveau dans la séquence de traitement des objets interviennent pour la complétion des parties manquantes.

Des études réalisées sur des figures incomplètes montrent que l'identification est facilitée lorsque des indications sur la classe sémantique des objets sont fournies au sujet avant l'apparition de la figure incomplète (Gregory, 1972 ; Reynolds, 1985). Cet effet suggère que l'activation de la représentation en mémoire de l'objet aide la complétion du contour. Néanmoins, des manipulations systématiques des propriétés physiques qui interviennent dans les processus d'intégration précoces des informations de contour (la colinéarité et la proximité spatiale des éléments) montrent que les figures incomplètes dont les éléments sont distants et non alignés restent difficiles à identifier, même lorsque l'objet complet dont la figure fragmentée est dérivée est présenté dans les millisecondes qui précèdent l'apparition de la figure incomplète (Boucart et al., 1994).

De même, un déficit des processus d'intégration des contours peut avoir des répercussions sur tous les traitements ultérieurs de l'objet. Par exemple, Boucart et Humphreys (1992 a et b) ont étudié les performances d'un patient (HJA) présentant une agnosie intégrative dans des tâches de comparaison de formes incomplètes contrôlées sur les propriétés physiques d'alignement et de proximité spatiale. Tandis que les sujets normaux groupent sans difficulté des contours locaux proches et alignés, HJA présente une performance similaire pour les figures structurées et pour les figures non structurées sur les propriétés de colinéarité et de proximité spatiale. La performance n'est pas améliorée par la présentation de la version complète de la figure fragmentée. Ces données montrent que les processus précoces d'intégration et de seg-

mentation des formes jouent un rôle important pour l'identification des objets.

Figure 8. Des parties des contours des objets sont fréquemment masquées par les contours d'un autre objet : ici le contour du château est interrompu régulièrement par les barreaux du portail. Notre connaissance des objets nous permet de compléter les parties manquantes du contour et de traiter le château et les barreaux comme des objets indépendants.

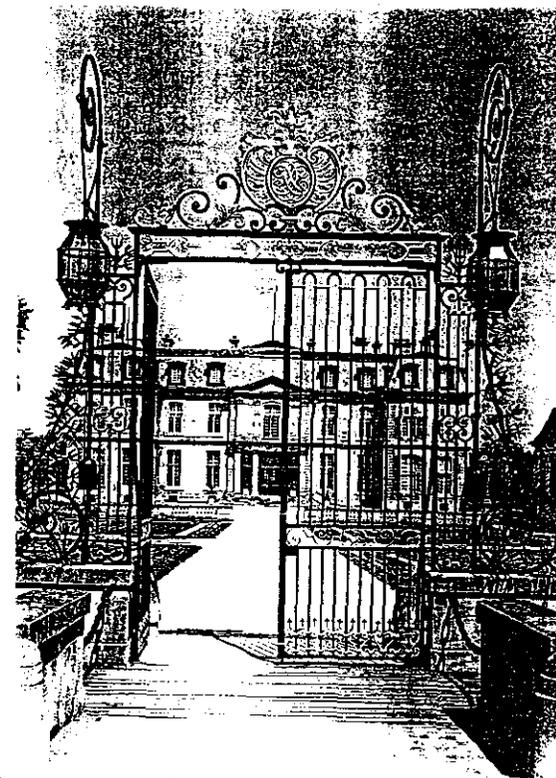


Figure 8. Parts of objects' contour are often masked or occluded by other objects. In this example the contour of the castle is periodically interrupted by the bars of the gate. Our knowledge of the structure of objects allows us to complete the missing parts and to process the castle and the gate as independent objects.

LA RECONNAISSANCE DES OBJETS

Au cours du traitement perceptif, un objet est une forme ; il n'a pas de signification. L'attribution d'une signification repose sur la comparaison de la forme physique aux représentations des objets stockées en mémoire ; c'est le niveau de l'interprétation des données perceptives. Comment les objets sont-ils stockés en mémoire et quelles sont les propriétés physiques efficaces pour l'activation des représentations d'objets ?

Les modèles théoriques divergent sur ces questions. Si tous les auteurs s'accordent pour reconnaître à l'information de forme un rôle principal dans l'activation des représentations d'objets, deux hypothèses s'opposent sur le type de caractéristiques de forme efficace pour activer ce niveau d'entrée. Pour Cavanagh (1991), Sanocki (1993), ou Watt (1994), l'information physique qui est à la base de l'activation des représentations est une forme bidimensionnelle grossière. Watt (1994) a utilisé des techniques de filtrage spatial de l'image. Il montre que quelques taches orientées sont suffisantes pour identifier un objet. Cette démonstration suggère qu'au niveau d'entrée les représentations en mémoire pourraient contenir le minimum d'informations nécessaires pour activer une classe d'objets. Par exemple, la Figure 9 montre que quelques taches horizontales permettent de reconnaître un visage. De la même façon, Sanocki (1993) propose que la reconnaissance commence sur la base d'une représentation bidimensionnelle comportant quelques éléments de contour. Cette représentation permet de délimiter une première classe de candidats objets et sert de guide pour l'analyse des détails intégrés plus tardivement.

Cette hypothèse a été testée dans une expérience d'amorçage qui montre que l'identification d'un objet cible est facilitée par la présentation préalable du simple contour externe d'un objet. Cependant, le nombre d'objets utilisés était très limité (6 stimuli) et les deux catégories sémantiques représentées étaient aisément discriminables sur leur configuration globale (maisons versus véhicules). Le contour externe n'aurait probablement pas constitué une information suffisante pour des stimuli très similaires sur leur configuration globale.

Figure 9. Au moyen de techniques de filtrage d'image, Watt (1994) montre que quelques taches orientées sont suffisantes pour identifier un objet.

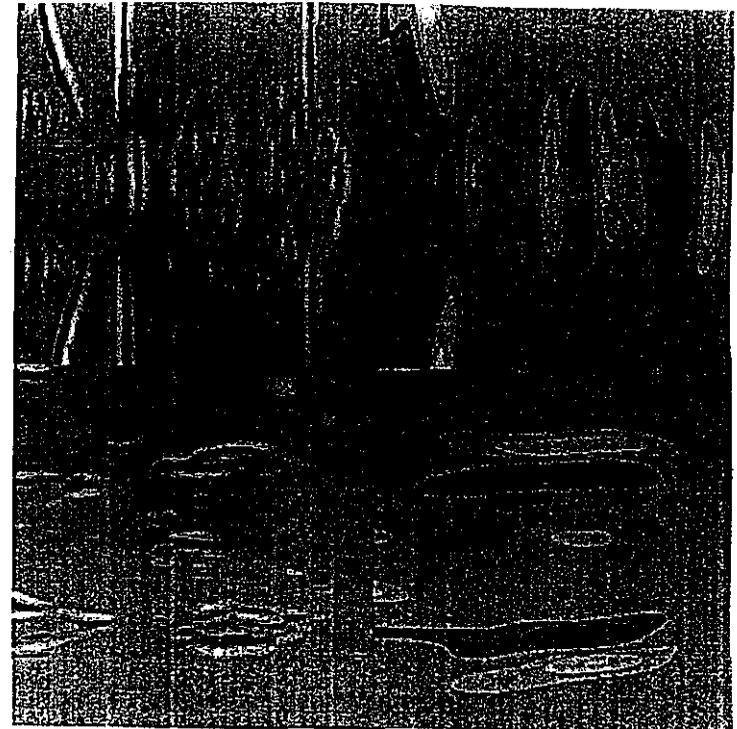


Figure 9. Watt (1994) has shown, with methods of image filtering, that a few oriented blobs are sufficient to identify an object.

Un point de vue plus "localiste" est proposé par d'autres auteurs (Marr, 1982 ; Pentland, 1986 ; Biederman, 1987). Dans ces modèles, la forme est d'abord segmentée en parties. Ces composantes sont comparées à un premier niveau de représentations qui comporte un ensemble limité d'entités volumétriques simples (des blocks, des cônes, des cylindres) communes à tous les objets. L'identification des composantes ainsi que leur arrangement spatial activent les représentations de

la forme des objets et permet d'identifier très rapidement les objets. La comparaison des parties de la forme au stock de composantes volumétriques (les "géons", Biederman, 1987) se déroule en parallèle. La dénomination d'un objet n'est pas affectée par le nombre de parties qui le composent (Biederman, 1987). De plus, le caractère tridimensionnel des géons garantit une propriété d'invariance à l'orientation, permettant ainsi d'identifier les objets sous différents angles pour autant que toutes les parties de l'objet restent visibles (Biederman et Gerardstein, 1993).

Comment sont organisées les représentations d'objets ? Dans les années 70, Rosch (1975) et Rosch et al. (1976) ont proposé une organisation hiérarchique de la mémoire sémantique en trois niveaux de représentations : la catégorie de base (e.g., chien), la supercatégorie (e.g., animal) et la sous-catégorie (e.g., dalmatien). La catégorie de base est définie comme le niveau comportant les objets qui ont le plus grand nombre de caractéristiques physiques communes avec les autres membres de leur classe sémantique. Elle constitue le point d'entrée des représentations sémantiques et sert de base pour l'activation de la classe sémantique. Jolicoeur, Gluck et Kosslyn (1984) ont montré que ce point de vue ne vaut que pour les exemplaires les plus typiques de la classe sémantique. Ainsi, l'image d'une mouette peut activer rapidement la classe sémantique oiseau tandis qu'un pingouin est plus difficile à catégoriser dans cette classe sémantique.

Cette flexibilité du système sémantique, liée aux interactions entre la forme de l'objet et le concept qu'il évoque, a conduit à la suggestion d'autres types d'organisation des représentations d'objets. Flores d'Arcais et Schreuder (1987) proposent l'existence de deux voies pour l'accès à l'identité de l'objet : (1) une voie directe déterminée par l'activation des caractéristiques physiques (P-elements) qui activent les caractéristiques fonctionnelles (F-elements) connectées aux représentations de la forme activées par les caractéristiques physiques, et (2) une voie indirecte qui ne dépend pas des caractéristiques physiques de l'objet mais qui est déterminée par le résultat de l'activation d'un concept. Par exemple, pour la voie directe, l'objet "ballon" active les représentations d'objets ayant une forme similaire (e.g., orange, melon, balle de tennis ...). Les représentations des objets "orange" et "melon" activent le concept "fruit". Les représentations des objets "ballon" et "balle de tennis" activent le concept "sport". Pour la voie indirecte, le

concept "sport" activé par le ballon et la balle de tennis active la représentation de l'objet "raquette", dont la forme n'avait pas été activée par la présentation d'un ballon.

Une démonstration empirique avec un paradigme d'amorçage met en évidence un effet des relations physiques et sémantiques entre la cible et l'amorce mais l'amplitude de la facilitation varie en fonction des caractéristiques physiques et fonctionnelles communes à l'objet cible et à l'amorce. Par rapport à une condition neutre, la facilitation la plus importante est observée pour des objets qui partagent des informations physiques et fonctionnelles (amorce : guitare – cible : violon). Elle diminue pour des objets similaires par leurs caractéristiques physiques mais différents par leurs caractéristiques fonctionnelles (amorce : guitare – cible : raquette de tennis) et elle diminue encore pour des objets ayant uniquement des caractéristiques fonctionnelles communes (amorce : guitare – cible : accordéon). Une inhibition est observée pour les objets n'ayant aucune caractéristique physique et fonctionnelle commune (amorce : guitare – cible : chaise).

Une organisation voisine a été proposée par Riddoch et Humphreys (1987) mais ce modèle ne comporte qu'une seule voie. Selon Riddoch et Humphreys, l'identification d'un objet repose sur l'accès à trois niveaux de représentations organisés en cascade. La forme construite au cours du traitement perceptif est d'abord comparée à un stock de représentations comportant des informations sur les caractéristiques physiques des objets (les représentations structurales). Ce niveau de description n'a pas été clairement défini par les auteurs. Il pourrait s'agir de représentations de la forme prototypique des objets dans leur orientation la plus familière, c'est-à-dire, celle qui comporte le maximum d'informations physiques pour réduire le nombre de candidats-objets activés. L'activation des représentations de la forme active automatiquement les représentations sémantiques connectées aux formes activées au niveau des représentations structurales. Les représentations sémantiques comportent des informations sur la fonction de l'objet (son usage, l'endroit où on peut le trouver) et sur les propriétés associatives de l'objet (à quel objet il est fréquemment relié).

Dans un modèle plus récent (Humphreys et al., sous presse), les auteurs ont divisé les représentations sémantiques en deux niveaux. Le premier niveau se réfère aux propriétés fonctionnelles et associatives

des objets. Le second comporte des informations sur la classe sémantique de l'objet (la super-catégorie). Le dernier niveau, appelé phonologique, correspond au lexique interne où sont stockés les noms des objets.

L'organisation en cascade, plutôt que sérielle, de ces trois niveaux de représentations est suggérée par le fait que l'activation des représentations à un niveau donné commence avant la fin du traitement au niveau précédent. Par exemple, Vitkovitch, Humphreys et Lloyd-Jones (1993) ont réalisé une expérience dans laquelle les sujets avaient pour instruction de dénommer le plus rapidement possible (speeded naming condition) les objets présentés sous forme d'images. La performance était comparée pour deux groupes d'objets : (1) des objets ayant une grande similarité physique intra-catégorie (e.g., pomme, orange, melon ...) et (2) des catégories composées d'objets physiquement différents (e.g., chaise, table, lit ...). Les résultats montrent que le nombre de confusions physiques et/ou sémantiques est plus important pour les catégories d'objets physiquement similaires que pour les catégories d'objets physiquement différents. Une organisation sérielle des représentations d'objets ne peut pas rendre compte de ce résultat. Dans une organisation sérielle, le traitement à un niveau commence sur la base des résultats du traitement réalisé au niveau précédent qui doit être complètement terminé.

L'existence des trois niveaux de représentations (structurales, sémantiques et lexicales) est corroborée par des données neuropsychologiques qui montrent que des patients peuvent présenter des déficits sélectifs des représentations de la forme, de la fonction ou du nom des objets ainsi que des connexions entre ces niveaux de représentations (Humphreys et Riddoch, 1987 a et b ; Sartori et Job, 1988 ; Sartori, Job et Coltheart, 1992 ; Rumiati et al., 1994 ; voir Bruyer dans ce numéro pour une revue).

Une organisation en trois niveaux de représentations est-elle suffisante ? Un nombre croissant de données neuropsychologiques montre que des patients agnosiques peuvent présenter des troubles sélectifs de l'identification de certaines classes d'objets. La perturbation la plus fréquemment observée concerne les catégories d'objets naturels (animaux, fruits, végétaux) par rapport aux objets manufacturés (voir Charnalet et Carbonnel et Sartori dans ce numéro). Ces données ont conduit à pro-

poser l'existence d'une multitude de sous-systèmes sémantiques (Warrington et Shallice, 1984 ; Shallice, 1988 ; Farah, McMullen et Meyer, 1991) correspondant à la classe sémantique, au type de stimulus (mot ou image) et à la modalité d'entrée du signal (visuelle, auditive, tactile). Cette hypothèse est contestée par Riddoch et al. (1988) qui suggèrent qu'un déficit sélectif pour l'identification des objets naturels peut s'expliquer par des perturbations à différents niveaux de la séquence de traitement des objets. Il peut résulter de difficultés au niveau du traitement perceptif induites par la similarité physique importante des objets provenant de catégories naturelles (Sheridan et Humphreys, 1993). Il peut aussi provenir d'un déficit au niveau des représentations de la forme des objets, également induit par la similarité physique des membres de catégories naturelles, puisque l'activation de nombreux candidats introduit du bruit dans le système. Par l'organisation en cascade des niveaux de représentations, ce bruit se répercute aux niveaux suivants.

Certaines propriétés physiques sont-elles plus efficaces que d'autres pour l'activation des représentations d'objets ? Les modèles théoriques traitant des informations d'entrée des représentations mettent l'accent sur le rôle des informations de forme, qu'elles soient caractérisées en termes de descriptions locales (Biederman, 1987 ; Biederman et Cooper, 1991) ou plus globales (Sanocki, 1993). L'influence des caractéristiques de surface sur l'accès à l'identité de l'objet a été examinée dans quelques études. Par exemple, Biederman et Ju (1988) ont comparé la performance pour des dessins au trait d'objets et des photos en couleur des mêmes objets dans une tâche de vérification. L'objet était précédé d'un mot afin de générer des attentes sur les propriétés physiques de l'objet qui suivait. Les sujets avaient pour instruction de décider si l'objet présenté correspondait au mot ou non. Les résultats montrent une performance équivalente pour les dessins au trait et les photos d'objets en couleur, suggérant que les informations de couleur ne sont pas indispensables pour l'identification des objets.

Price et Humphreys (1989) montrent que l'utilisation des informations de couleur peut être modulée par les caractéristiques de formes de l'objet. Dans cette étude, les temps de dénomination ou de catégorisation étaient comparés pour des objets très similaires par leur forme (e.g., pomme, orange, melon ...) et pour des objets différents par leur

forme. Les résultats montrent que l'identification des objets est facilitée par leur présentation en couleur correcte, par rapport à une couleur incorrecte ou des nuances de gris. Cette facilitation ne se produit que pour les objets similaires par leur forme. Ce résultat suggère que les informations de couleur sont secondaires. Elles n'interviennent que lorsque les informations de forme sont insuffisantes pour discriminer efficacement les objets. Des difficultés à traiter efficacement les informations de formes pourraient expliquer la performance classiquement meilleure des patients agnosiques pour les photos d'objets en couleur par rapport aux dessins au trait. L'utilisation des informations de couleur dans des conditions particulières a également été mise en évidence par Tanaka et Bunoski (1993) dans un paradigme de recherche visuelle d'une cible parmi des distracteurs. Dans cette étude, un mot précédait l'apparition de six objets disposés en cercle. Les objets étaient présentés en couleur ou en nuances de gris. La couleur facilitait l'identification de la cible uniquement pour les objets ayant naturellement une couleur spécifique (citron, banane, tomate, voiture de pompier...). L'ensemble des données suggère que l'information de forme est plus efficace que la couleur pour l'identification des objets.

L'OBJET DANS SON CONTEXTE

Quelle est l'influence de l'environnement spatial et sémantique sur l'identification des objets ? L'interprétation correcte d'une scène visuelle implique l'identification des objets qui la composent et des relations spatiales et sémantiques entre les objets. Biederman et collaborateurs (1981, 1982) ont proposé l'existence de cinq règles pour caractériser une scène bien organisée. Deux de ces règles sont définies comme des contraintes physiques. Il s'agit de la règle de support, selon laquelle un objet doit reposer sur une surface, et de la règle d'interposition qui se réfère à l'occlusion : un objet opaque cache les objets qui se trouvent derrière. Les trois autres règles sont définies comme des contraintes sémantiques. La première se réfère à la probabilité de la présence d'un objet dans une scène : un objet est plus probable dans une scène que dans une autre. La seconde concerne la localisation spatiale : un objet probable dans une scène doit se trouver à un endroit spécifique. La

troisième est une contrainte de taille : un objet probable dans une scène doit avoir une taille spécifique par rapport aux autres objets de la scène.

Biederman et ses collègues ont testé l'impact de la transgression de ces règles sur l'identification des objets d'une scène. Le paradigme expérimental est présenté Figure 10. La performance était comparée dans trois conditions expérimentales : (1) l'objet cible transgressait une, deux ou trois règles, (2) un distracteur (un autre objet de la scène) transgressait une ou plusieurs règles et (3) aucune règle n'était transgressée. Les résultats montrent que le temps de réponse et le taux d'erreurs augmentent avec le nombre de règles transgressées pour l'objet cible. En revanche la transgression d'une règle par un distracteur n'a aucune influence sur la performance. La transgression de

Figure 10. Paradigme utilisé par Biederman (1981) pour les investigations sur le traitement des scènes visuelles. Un mot désignant un objet cible est suivi d'une scène visuelle. La scène est suivie d'un masque comportant un indice (point). Le sujet a pour instruction de décider si l'objet présenté à l'emplacement de ce point dans la scène correspond au mot.

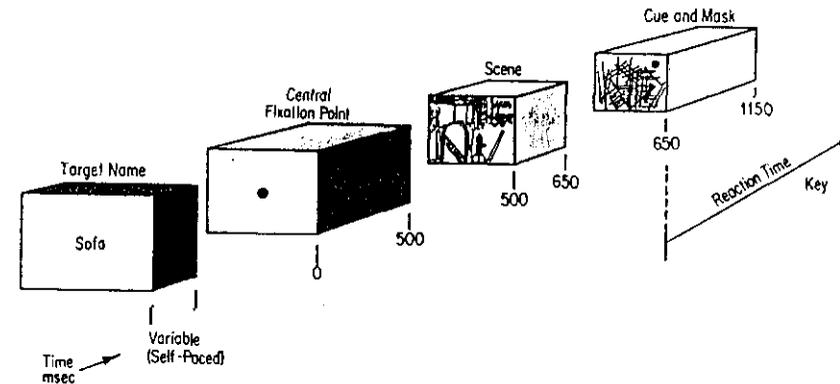


Figure 10. A paradigm used by Biederman (1981) to investigate the processing of visual scenes. A word corresponding to a target object is followed by a drawing of a scene. The scene is followed by a mask containing a cue (dot). Subjects are asked to decide whether the object located where the dot was displayed corresponds to the word.

certaines règles affecte l'identification des objets plus que d'autres. Le temps de réponse est plus long pour un objet peu probable dans une scène, ainsi que pour un objet qui n'a pas une taille cohérente par rapport aux objets qui l'entourent, que pour un objet qui transgresse les règles de localisation spatiale, d'interposition ou de support. Ces effets ont été répliqués dans des études sur les mouvements oculaires (de Graef et al., 1990, 1992) qui montrent que la durée de fixation est plus longue pour les objets non compatibles avec la scène (un caddy dans une cour de ferme ou des poules aux caisses d'un supermarché).

Deux grandes classes d'interprétations ont été proposées pour rendre compte des effets de contexte sur l'identification des objets. L'une repose sur l'activation d'un schéma mental. Elle est défendue par Biederman (1981), Antes et Penland (1981), Friedman (1979) et Boyce et al. (1989). Selon cette hypothèse, des représentations prototypiques de scènes rencontrées par le sujet ont été stockées en mémoire. Le schéma comporte quelques objets spécifiques d'une scène. Il est activé très rapidement et il guide l'interprétation des composantes de la scène (les objets). L'interprétation se déroule en trois étapes : (1) l'identification du fond (la présence de quelques éléments spécifiques) active la représentation d'un schéma qui, à son tour, active les candidats objets compatibles avec ce schéma. Plusieurs arguments étayaient cette hypothèse. Par exemple, Palmer (1975) rapporte des temps d'identification plus rapides pour un objet précédé d'une scène compatible (un pain précédé par une cuisine) par rapport à un contexte non approprié (une boîte à lettres précédée par une cuisine). L'identification d'un objet est facilitée lorsqu'il est présenté sur une scène organisée par rapport à un fond désorganisé (des parties de scènes disposées aléatoirement) (Biederman, 1972, 1973, 1974). Plus récemment, Boyce et al. (1989) ont montré qu'un objet cible est identifié plus précisément dans une scène comportant des objets compatibles avec l'objet cible que lorsqu'il est présenté sur un fond non significatif de complexité équivalente ou sur une scène non compatible avec l'objet cible. Ces données suggèrent que l'information globale fournie par le fond facilite l'identification des éléments de la scène.

Ce point de vue a été contesté par Henderson (1992) qui propose une interprétation en termes de traitement local (local processing hypothesis). Selon Henderson, l'hypothèse de l'activation d'un schéma de base

n'est pas nécessaire pour rendre compte des effets de contexte. L'identification d'un seul objet de la scène active les représentations des objets appartenant à la même classe sémantique par un mécanisme de propagation de l'activation des objets connectés. Ainsi, un objet peu probable dans une scène est plus difficile à identifier qu'un objet probable, parce que sa représentation n'a pas été activée par un autre objet présent dans la scène.

Cette hypothèse a fait l'objet de nombreuses critiques, la plus importante concernant le caractère épisodique plus que sémantique des relations entre les objets d'une scène visuelle. Boyce et Pollatsek (1992) ont manipulé les relations sémantiques entre un objet cible et les objets distracteurs d'une scène. Les résultats montrent que les relations sémantiques entre l'objet cible et les distracteurs n'affectent pas l'identification de l'objet cible. En revanche, un fond non compatible perturbe l'identification de la cible (des palmes devant un immeuble par rapport à des palmes au bord d'une piscine). L'hypothèse de l'activation d'un schéma de base semble mieux argumentée. Elle n'a cependant pas échappé à certaines critiques. Par exemple, dans le paradigme (Figure 10) utilisé par Biederman et al. (1982) et Boyce et al. (1989), la présentation d'un mot avant la scène peut générer des attentes sur les objets qui vont être présentés. Dans ce cas l'activation d'un schéma de base n'est peut-être pas nécessaire pour expliquer la facilitation induite par un fond cohérent. Par ailleurs, un objet est plus plausible qu'il n'est probable dans une scène visuelle. Enfin, ces études ne prennent pas en compte le traitement perceptif des propriétés physiques de la scène. Des objets qui transgressent certaines règles peuvent être plus difficiles à identifier simplement pour des raisons perceptives, telles que le masquage dans le cas de la règle d'interposition ou la visibilité de l'objet dans la règle de taille.

Les données de la neuropsychologie suggèrent néanmoins que les deux classes d'interprétations peuvent rendre compte du traitement des scènes visuelles. Les hypothèses de l'activation d'un schéma global (Biederman, 1981 ; Boyce et al., 1989) ou d'un seul objet critique (Henderson, 1992) sont fondées puisque ces deux types de traitement peuvent être sélectivement perturbés. Farah (1990) distingue deux types de troubles de la perception des scènes visuelles qui se manifestent différemment selon le siège lésionnel. Le traitement global de la scène est

perturbé par une lésion occipito-temporale de l'hémisphère gauche. Le patient peut identifier chaque objet de la scène, mais il est incapable d'en extraire la signification globale. Les patients présentant une lésion pariétale bilatérale focalisent leur attention sur un objet ou une partie d'un objet de la scène et interprètent l'environnement sur la base de ce détail.

CONCLUSION

Dans cette revue, j'ai résumé les données récentes sur les principales étapes de la séquence de traitement des objets. La revue est loin d'être exhaustive puisqu'elle ne prend pas en compte le traitement du mouvement, la modulation attentionnelle du traitement perceptif, la dynamique temporelle des processus cognitifs et les multiples interactions entre les processus. Tout au long de cette description, j'ai essayé de mettre l'accent sur l'énorme quantité de traitements impliqués dans l'identification d'un objet, dont beaucoup sont encore méconnus en raison de l'intérêt récent pour ce domaine de recherche et de la complexité des processus en jeu. Cette complexité a l'inconvénient de rendre difficile l'élaboration d'un modèle théorique ou d'un modèle computationnel qui synthétise et implémente l'ensemble des traitements impliqués dans la perception et la reconnaissance des objets. Elle a aussi un avantage qui est celui de contraindre les chercheurs de différentes disciplines des sciences cognitives à regrouper leurs données pour comprendre les processus sous-tendant la perception des objets.

ABSTRACT

This paper summarizes the present data on visual object perception. Data from neurophysiology, psychophysics, cognitive psychology on normal subjects, neuropsychology and computational models are taken into account. The paper is divided in four main parts. The first part deals with cortical regions involved in object perception. The second

part describes integration and segmentation processes of forms and the mechanisms involved in the computation of contour and surface information. The third part presents models of the organization of stored object representations and data about the role of form and surface information in object recognition. Context effects are presented in the last part which summarizes recent data on the processing of visual scenes.

RÉFÉRENCES

- Antes, J.R., Penland, J.G., & Metzger, R.L. (1981). Processing global information in briefly presented pictures. *Psychological Research*, *43*, 377-292.
- Baylis, G.C., Rolls, E.T., & Leonard, C.M. (1987). Functional subdivisions of temporal lobe neocortex. *Journal of Neuroscience*, *7*, 330-342.
- Biederman, I. (1972). Perceiving real-world scenes. *Science*, *177*, 77-80.
- Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J.R. Pomeranz (Eds.), *Perceptual organization* (pp. 213-254). Hillsdale, NJ: Lawrence Erlbaum.
- Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review*, *94*, 115-147.
- Biederman, I., & Cooper, E.E. (1991). Priming contour-deleted images: evidence for intermediate representations in visual object recognition. *Cognitive Psychology*, *23*, 393-419.
- Biederman, I., & Gerardstein, P.C. (1993). Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1162-1182.
- Biederman, I., Glass, A.L., & Stacy, E.W. (1973). Searching for objects in real-world scenes. *Journal of Experimental Psychology*, *97*, 22-27.
- Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, *20*, 38-64.
- Biederman, I., Mezzanotte, R.J., & Rabinowitz, J.C. (1982). Scene perception: detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*, 143-177.
- Biederman, I., Rabinowitz, J.C., Glass, A.L., & Stacy, E.W. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, *103*, 597-600.
- Boucart, M., Delord, S., & Giersch, A. (1994). The computation of contour information in complex objects. *Perception*, *23*, 399-409.

- Boucart, M., Grainger, J., & Ferrand, L. (1995). Three-dimensional features facilitate object recognition. *Visual Cognition*, 2, 451-478.
- Boucart, M., & Humphreys, G. W. (1992a). The computation of perceptual structure from collinearity and closure: normality and pathology. *Neuropsychologia*, 30, 527-546.
- Boucart, M., & Humphreys, G. W. (1992b). Global shape cannot be attended without object identification. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 785-806.
- Boyce, S.J., & Pollatsek, A. (1992). Identification of objects in scenes: the role of scene background in object naming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 3, 531-543.
- Boyce, S.J., Pollatsek, A., & Rayner, K. (1989). Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 556-566.
- Cavanagh, P. (1991). What's up in top-down processing? In A. Goréa, Y. Frégnac, Z. Kapoula, & J. Findlay (Eds.), *Representations of vision: Trends and tacit assumptions in vision research*. Cambridge: Cambridge University Press.
- De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research*, 52, 317-329.
- De Graef, P., de Troy, A., & d'Ydewalle, G. (1992). Local and global contextual constraints on the identification of objects in scenes. *Canadian Journal of Psychology*, 46, 489-506.
- Desimone, R., & Schein, S.J. (1987). Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of Neurophysiology*, 57, 835-868.
- Desimone, R., & Ungerleider, L.G. (1989). Neural mechanisms of visual processing in monkeys. In F. Bolter & J. Graham (Eds.), *Handbook of neuropsychology* (Vol. 2, pp. 267-299). Amsterdam: Elsevier Science Publishers.
- Dresp, B., & Bonnet, C. (1991). Psychophysical evidence for low level processing of illusory contours and surfaces in the Kanizsa square. *Vision Research*, 31, 1813-1817.
- Eckorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., & Reitboeck, H.J. (1988). Coherent oscillations: a mechanism of feature linking in the visual cortex? *Biological Cybernetics*, 60, 121-130.
- Eckorn, R., & Obermueller, A. (1993). Single neurons are differentially involved in stimulus-specific oscillations in cat visual cortex. *Experimental Brain Research*, 95, 177-182.

- Engel, A.K., König, P., & Singer, W. (1991). Direct physiological evidence for scene segmentation by temporal coding. *Proceedings of the National Academy of Sciences of the USA*, 88, 9136-9140.
- Enns, J.T., & Rensink, R.A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review*, 98, 335-351.
- Farah, J.M. (1990). *Visual agnosia: disorders of object recognition and what they tell us about normal vision*. Cambridge, MA: MIT Press.
- Farah, J.M., McMullen, P.A., & Meyer, M.M. (1991). Can recognition of living things be selectively impaired? *Neuropsychologia*, 29, 185-193.
- Field, D.J., Hayes, A., & Hess, R.F. (1993). Contour integration by the human visual system: evidence for a local "association field". *Vision Research*, 33, 173-193.
- Flores d'Arcais, G.B., & Schreuder, R. (1987). Semantic activation during object naming. *Psychological Research*, 49, 153-159.
- Friedman, A. (1979). Framing pictures: the role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, 108, 316-355.
- Fujita, I., Tanaka, K., Ito, M., & Cheng, K. (1992). Columns of visual features of objects in monkey inferotemporal cortex. *Nature*, 360, 343-346.
- Giersch, A., Boucart, M., Danion, J.M., Vidailhet, P., & Legrand, F. (1995). Effects of lorazepam on perceptual integration of visual forms in healthy volunteers. *Psychopharmacology* (sous presse).
- Gilbert, C.D., & Wiesel, T.N. (1989). Columnar specificity of intrinsic horizontal and cortico-cortical connections in cat visual cortex. *Journal of Neuroscience*, 9, 2432-2442.
- Gray, C.M., König, P., Engel, A.K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338, 334-337.
- Gray, C.M., König, P., Engel, A.K., & Singer, W. (1991). Synchronous neuronal oscillations in cat visual cortex: functional implications. In A. Goréa, Y. Frégnac, Z. Kapoula, & J. Findlay (Eds.), *Representations of vision: trends and tacit assumptions in vision research* (pp. 83-96). Cambridge: Cambridge University Press.
- Gregory, R.L. (1972). Cognitive contours. *Nature*, 238, 51-52.
- Grossberg, S. (1992). Neural facades: visual representations of static and moving form-and-color-and-depth. In G.W. Humphreys (Ed.), *Understanding vision* (pp. 232-271). Oxford: Blackwell.
- Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: boundary completion, illusory figures, and neon colour spreading. *Psychological Review*, 92, 173-211.

- Grossberg, S., & Mingolla, E. (1987). The role of illusory contour in visual segmentation. In S. Petry & G.E. Meyer (Eds.), *The perception of illusory contours*. New York: Springer.
- Hasselmo, M.E., Rolls, E.T., & Baylis, G.C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioral Brain Research*, *32*, 203-218.
- Henderson, J.M. (1992). Object identification in context: the visual processing of natural scenes. *Canadian Journal of Psychology*, *46*, 319-341.
- Heywood, C.A., & Cowey, A. (1987). On the role of cortical area V4 in the discrimination of hue and pattern in macaque monkeys. *The Journal of Neuroscience*, *7*, 2601-2617.
- Hubel, D.H., & Livingstone, M.S. (1987). Segregation of form, color and stereopsis in primate area 18. *The Journal of Neuroscience*, *1*, 3378-3415.
- Hummel, J.E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, *99*, 480-517.
- Humphreys, G.W., Keulers, N., & Donnelly, N. (1994). Parallel visual coding in three dimensions. *Perception*.
- Humphreys, G.W., Lemote, C., & Lloyd-Jones, T.J. (sous presse). An interactive activation-competition model of object naming: effects of object category, priming and simulated lesions on object identification. *Memory*.
- Humphreys, G.W., & Riddoch, M.J. (1987a). The fractionation of visual agnosia. In G.W. Humphreys & M.J. Riddoch (Eds.), *Visual object processing: a cognitive neuropsychological approach* (pp. 281-307). Hillsdale, NJ: Lawrence Erlbaum.
- Humphreys, G.W., & Riddoch, M.J. (1987b). *To see but not to see: a case study of visual agnosia*. London: Lawrence Erlbaum.
- Jolicoeur, P., Gluck, M.A., & Kosslyn, S.M. (1984). Pictures and names: making the connection. *Cognitive Psychology*, *16*, 243-275.
- Jolicoeur, P., Ullman, S., & MacKay, M. (1986). Curve tracing: a possible basic operation in the perception of spatial relations. *Memory and Cognition*, *14*, 129-140.
- Jolicoeur, P., Ullman, S., & MacKay, M. (1991). Visual curve tracing properties. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 997-1022.
- Kovacs, I., & Julesz, B. (1993). A close curve is much more than an incomplete one: effect of closure in figure-ground segregation. *Proceedings of the National Academy of Sciences of the USA*, *90*, 7495-7497.
- Kovacs, I., & Julesz, B. (1994). Perceptual sensitivity maps within globally defined visual shapes. *Nature*, *370*, 644-646.

- Livingstone, M.S., & Hubel, D.H. (1987). Physiological evidence for separate channels for the perception of form, color, movement and depth. *The Journal of Neuroscience*, *7*, 3416-3468.
- Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research*, *32*, 263-273.
- Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. New York: W.H. Freeman and Company.
- Maunsell, J.H.R., & Newsome, W.T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, *10*, 363-401.
- McCormick, P., & Jolicoeur, P. (1992). Capturing visual attention and the curve tracing operation. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 72-89.
- Merigan, W.H., & Maunsell, J.H.R. (1993). How parallel are the primate visual pathways. *Annual Review of Neuroscience*, *16*, 369-402.
- Miyashita, Y. (1993). Inferior temporal cortex: where visual perception meets memory. *Annual Review of Neuroscience*, *16*, 245-263.
- Palmer, S.E. (1975). The effects of contextual scenes on the identification of objects. *Memory and Cognition*, *3*, 519-526.
- Pentland, A. (1986). Perceptual organization and the representation of natural forms. *Artificial Intelligence*, *28*, 293-331.
- Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. 2: Contours bridging gaps. *Journal of Neuroscience*, *9*, 1749-1763.
- Peterhans, E., & von der Heydt, R. (1991). Elements of form perception in monkey prestriate cortex. in A. Goréa, Y. Frégnac, Z. Kapoula, & J. Findlay (Eds.), *Representations of vision: Trends and tacit assumptions in vision research* (pp. 111-124). Cambridge: Cambridge University press.
- Price, C.J., & Humphreys, G.W. (1989). The effects of surface detail on object categorisation and naming. *Quarterly Journal of Experimental Psychology*, *41A*, 797-828.
- Reynolds, R.I. (1985). The role of object-hypothesis in the organization of fragmented figures. *Perception*, *14*, 49-52.
- Riddoch, M.J., & Humphreys, G.W. (1987). Picture naming. In G.W. Humphreys & M.J. Riddoch (Eds.), *Visual object processing: a cognitive neuropsychological approach* (pp. 107-142). Hillsdale, NJ: Lawrence Erlbaum.
- Riddoch, M.J., Humphreys, G.W., Coltheart, M., & Funnell, E. (1988). Semantic systems or system? Neuropsychological evidence re-examined. *Cognitive Neuropsychology*, *5*, 3-25.

- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 305-322.
- Rosch, E., Mervis, C.B., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *7*, 573-605.
- Rumiati, R.I., Humphreys, G.W., Riddoch, M.J., & Bateman, A. (1994). Visual object agnosia without prosopagnosia or alexia: evidence for hierarchical theories of visual recognition. *Visual Cognition*, *1*, 181-225.
- Sanocki, T. (1993). Time course of object identification: evidence for a global-to-local contingency. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 878-898.
- Sartori, G., & Job, R. (1988). The oyster with four legs: a neuropsychological study on the interaction of visual and semantic information. *Cognitive Neuropsychology*, *5*, 105-132.
- Sartori, G., Job, R., & Coltheart, M. (1992). The organization of object knowledge: evidence from neuropsychology. In D.E. Meyer & S. Kornblum (Eds.), *Synergies in experimental psychology, artificial intelligence and cognitive neuroscience. Attention and Performance XIV* (pp. 451-465). Cambridge, MA: MIT Press.
- Schein, S.J., & Desimone, R. (1990). Spectral properties of V4 neurons in the macaque. *The Journal of Neuroscience*, *10*, 3369-3389.
- Shallice, T. (1988). Specialisation within the semantic system. *Cognitive Neuropsychology*, *5*, 133-142.
- Sheridan, J., & Humphreys, G.W. (1993). A verbal-semantic category-specific recognition impairment. *Cognitive Neuropsychology*, *10*, 143-184.
- Tanaka, J., & Bunoski, L. (1993). Is the object recognition system really color-blind? Communication présentée au 34th Annual Meeting of the Psychonomic Society.
- Tanaka, K., Saito, H.A., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the infero-temporal cortex of the macaque monkey. *Journal of Physiology*, *66*, 170-189.
- Ullman, S. (1984). Visual routines. *Cognition*, *18*, 97-159.
- Vitkovitch, M., Humphreys, G.W., & Lloyd-Jones, T.J. (1993). On naming a giraffe a zebra: picture naming errors across different object categories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 243-259.
- von der Heydt, R., & Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex 1. Lines of pattern discontinuities. *Journal of Neuroscience*, *9*, 1731-1748.
- Warrington, E.K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, *107*, 829-854.

- Watt, R.J. (1994). A computational examination of image segmentation and the initial stages of human vision. *Perception*,
- Zeki, S.M. (1990). Functional specialization in the visual cortex: the generation of separate constructs and their multistage integration. In G.M. Edelman, W.E. Gall, & W.M. Cowan (Eds.), *Signal and sense: Local and global order in perceptual maps* (pp. 85-129). New York: John Wiley & Sons.
- Zeki, S., Watson, J.D.G., Lueck, C.J., Friston, K.J., Kennard, C., & Frackowiak, R.S.J. (1991). A direct demonstration of functional specialization in human visual cortex. *The Journal of Neuroscience*, *11*, 641-649.
- Zucker, S.W., & Davis, S. (1988). Points and endpoints: A size/spacing constraint for dot grouping. *Perception*, *17*, 229-247.