

Jean-Michel FORTIS

Image mentale et représentation propositionnelle

Introduction

L'existence d'images mentales ne paraît pas pouvoir être mise en doute si l'on se fie aux données de l'introspection. Cependant, l'existence même de l'image mentale n'est pas ce qui est généralement en cause lorsqu'est posé le problème de la réductibilité de l'image à une structure propositionnelle, et cela même chez les anti-imagistes les plus acharnés (Pylyshyn, 1973, p. 2). Il s'agit bien plutôt de savoir si les informations véhiculées par l'image et la manière dont on suppose qu'elles sont extraites de l'image découlent des propriétés quasi perceptives de l'image. Le fait que l'interprétation du contenu d'une image est attribuable à des représentations dont le format n'est pas celui de l'image conduit souvent à poser que les caractères quasi perceptifs de l'image ne sont que des épiphénomènes. De là, on conclut que la ressemblance de l'image à ce dont elle est l'image est secondaire pour identifier l'objet de l'image.

Même chez les imagistes, on trouvera ainsi des positions variées sur ce qu'est l'isomorphisme de l'image. Ceux qui tiennent pour l'épiphénoménalité de l'image pensent qu'il y a bien équivalence informationnelle de l'image et des représentations utilisées par les processus qui l'interprètent ; de là, ils sont naturellement conduits à penser que ces représentations servant à l'interpréter suffisent par elles-mêmes. Les auteurs qui sont de cet avis soutiennent aussi que ces représentations ont un format propositionnel.

Dans une première partie, l'image sera examinée du point de vue des arguments propositionnalistes. Au cours des paragraphes qui

suivent, nous examinerons la notion de ressemblance, en distinguant divers types d'isomorphismes et en explicitant la notion de "medium analogique". Nous verrons quel genre d'isomorphisme diverses expériences nous incitent à attribuer à l'image. Nous critiquerons certains arguments tendant à démontrer qu'il y a nécessairement un niveau propositionnel de représentation sans lequel l'image n'est ni interprétable, ni traductible. Nous nous efforcerons ensuite de montrer que certaines conceptions de l'isomorphisme de l'image et du caractère analogique du medium des images proviennent d'une dévaluation de la notion de ressemblance qui, loin d'être récente, s'inscrit dans une tradition de pensée. Les arguments des anti-imagistes seront enfin confrontés à ceux, empiriques et théoriques, des auteurs qui nient l'épiphénoménalité de l'image.

Dans une deuxième partie, nous présenterons les positions des imagistes, et discuterons de l'intégration d'un composant propositionnel dans leurs modèles.

1. Critique des thèses propositionnalistes

1.1. Sur l'isomorphisme des représentations¹

Quelle sorte d'isomorphisme peut-on attribuer aux représentations imagées sur la base des résultats expérimentaux acquis ? Cette question est au centre de la controverse entre propositionnalistes et imagistes, car une définition restrictive de l'isomorphisme est compatible avec une conception propositionnaliste du format des images mentales.

Avec Shepard (Shepard & Chipman, 1970 ; Shepard, 1975), nous pourrions distinguer plusieurs formes d'isomorphismes pouvant exister entre une représentation (y compris matérielle, comme un tableau) et ce qu'elle représente. L'isomorphisme concret d'ordre 1 correspond à la représentation d'un ensemble de propriétés de l'objet représenté par des propriétés au moins perceptuellement semblables : nous voulons dire que les propriétés physiques *et* leur interaction avec les mécanismes perceptifs sont

¹ Je remercie P. Livet et M. Denis de leurs commentaires, qui m'ont heureusement amené à corriger certains points et à compléter mon information.

cause que ces propriétés physiques sont dites représenter un objet (le rouge peint est ainsi perceptuellement semblable au rouge réel). L'isomorphisme abstrait d'ordre 1 peut se définir comme la représentation de propriétés par des éléments fonctionnellement équivalents de la représentation : des parties de la représentation correspondent (sans leur être perceptuellement identiques au sens précédent) à des parties de l'objet, et, les relations entre ces parties sont préservées par la représentation. En un sens cependant, ces parties de la représentation peuvent être perceptuellement semblables à ce qu'elles représentent : au sens où le rouge perçu est semblable au rouge imaginé. Enfin, l'isomorphisme d'ordre 2 est propre à des représentations dans lesquelles les *différences* entre les propriétés des objets représentés sont corrélées à des différences parallèles dans la représentation¹. Les isomorphismes abstraits d'ordre 1 et d'ordre 2 définissent le rapport d'une représentation à son monde représenté, et non pas la nature de la représentation elle-même ; ils n'impliquent donc aucune préférence pour une réalisation physique particulière de la représentation.

Selon les paradigmes expérimentaux, on peut attribuer à l'image un isomorphisme abstrait d'ordre 1 ou un isomorphisme d'ordre 2. Dans le cas des expériences de Shepard et Chipman (1970), où sont comparées les similarités de forme entre des objets (en l'occurrence les états américains) et les jugements de similarité effectués par des sujets à propos des mêmes objets. De ce type de paradigme, il semble qu'on ne puisse induire qu'un isomorphisme d'ordre 2. En revanche, dans d'autres expériences (par exemple, Kosslyn, Ball & Reiser, 1978), on peut constater que le temps d'accès à une partie de la représentation (des objets représentés) à partir d'un point représenté est corrélé à la distance effective de cette partie à partir du point de départ réel correspondant. De là, on

¹ C'est cet isomorphisme qui caractérise, selon Shepard et Chipman (1970), les représentations dites "imaginées". Leurs expériences mettent en effet en évidence que des jugements portant sur la similitude de formes (en l'occurrence, les frontières d'états américains) sont à peu près identiques, que les formes en question soient perceptivement présentes ou qu'elles soient simplement imaginées par les sujets. De ce paradigme expérimental on ne peut, en effet, induire qu'un isomorphisme d'ordre 2.

est amené à conclure au moins à une préservation des relations métriques de l'objet représenté dans la représentation, donc à un isomorphisme abstrait d'ordre 1. Or, nous avons vu que l'isomorphisme abstrait d'ordre 1 n'entraîne aucune spécification quant au format de la représentation, et que l'interprétation d'une image est distincte de son lien isomorphique à ce qui est le signifiant de cette image ou, pour emprunter un terme de la phénoménologie, à ce qu'elle présente. Mais ces questions sont parfois confondues ; de ce que la ressemblance de l'image au percept qu'elle revivifie ne permet pas de déterminer son interprétation ou son usage cognitif, on conclut parfois à tort que la représentation identifiée à tort à une image doit avoir un format semblable à ce qui sert à l'interpréter. Ce genre de confusion est particulièrement remarquable chez Friedman (1978) lorsque l'effet symbolique de distance obtenu pour des mots abstraits¹ lui suggère que cet effet ne peut avoir aucune base d'imagerie :

"Puisque les effets de distance présents ont été obtenus avec un matériel abstrait dont les représentations étaient presque certainement propositionnelles par leur structure, la question n'est plus de savoir *si* cela [i.e. le recours à une base de données analogique] est possible. En revanche, puisque les effets de distance obtenus reflétaient une information relative à des grandeurs absolues, il est plus opportun de demander comment une base de données abstraite, amodale, et propositionnelle peut produire des différences dans les temps de réaction qui reflètent des informations sur la grandeur absolue." (1978, p. 441).

¹ L'effet symbolique de distance correspond à l'allongement du temps de comparaison de mots dont les référents sont proportionnellement moins éloignés sur la dimension à comparer (le phénomène a été découvert et baptisé par Moyer & Bayer, 1976). Inversement, la comparaison est d'autant plus rapide que les items sont plus éloignés sur cette même dimension. Friedman retrouve ce résultat en utilisant des mots abstraits (tels que "hate", "death", "lose" et "allow", "peace", "joy") et en classant ces mots sur la dimension évaluative bon-mauvais. Elle constate un effet de la distance évaluative séparant les termes d'un couple de mots pris dans une liste ordonnée suivant cette dimension, et un effet de la position ordinale de chaque terme (c'est-à-dire de sa distance à la "bonne" ou à la "mauvaise" extrémité).

Si une représentation (par exemple une droite) sert à symboliser une dimension abstraite, il ne s'ensuit pas que cette représentation ne puisse être une image.

On constate chez Kosslyn une attitude plus rigoureuse à l'égard des propriétés du "tampon visuel". Les propriétés assimilant celle-ci à une perception sont rattachées au "tampon visuel" (*visual buffer*), dont le statut théorique est parfois ambigu¹. Les propriétés découlant du rapport de l'image à son référent (comme le temps de scanning sur une carte mentale, la relation linéaire entre le temps de réponse et l'angle de deux figures dans des tâches de comparaison mentale) sont elles rapportées à un médium fonctionnel, entendant par là que ce médium ne fait que *préserver* des qualités de l'espace perceptif dont il est par nature distinct. Il est vrai qu'on ne peut strictement déduire de ces expériences qu'une similitude fonctionnelle entre "l'espace" de la représentation et l'espace représenté. Dire que le tampon visuel est un médium fonctionnel correspond ainsi à l'engagement théorique minimal sur le format des représentations imagées. Comme on ne peut déduire de la nature spatiale des informations qu'il y a un écran dans le cerveau, Kosslyn conclut que le tampon visuel simule fonctionnellement un espace visuel ?²

¹ Le modèle utilise la métaphore du tube cathodique, qui est censée n'avoir pas de rapport avec la théorie elle-même (cf. Kosslyn, 1981 : 51). Cependant, cette métaphore sert aussi à faire des prédictions enveloppant des propriétés du modèle qui ne devraient être que des détails d'implémentation, et qui pourtant accèdent ainsi à un statut théorique. De fait, lorsque Kosslyn demande si les transformations effectuelles sur une image se font à partir de la matrice de surface ("l'écran") ou depuis les coordonnées polaires servant à coder les images, il suppose que dans la première hypothèse le nombre supérieur de cellules de la matrice pour une grande image entraînera une plus grande quantité de traitement et donc un temps de traitement plus important. (cf. 1980 : 290)

² On trouvera chez Pinker une définition fonctionnelle (c'est-à-dire par la préservation de relations métriques) de l'image qui laisse également ouverte la possibilité que l'image soit une structure si abstraite que ses qualités phénoménales proches de celles d'un percept soient tout à fait inessentiels (1984b).

"Les images se produisent dans un *medium spatial*¹ qui est fonctionnellement équivalent à un espace de coordonnées (peut-être euclidiennes). Cela ne signifie pas qu'il y a littéralement un écran dans le cerveau. Plutôt, la façon dont on accède à un point est telle que les propriétés sur les intervalles d'un espace physique sont préservées dans au moins deux dimensions. (...) Un exemple parfait de cela est un tableau bi-dimensionnel stocké dans les banques de mémoire d'un ordinateur : il n'y a pas de matrice physique dans les banques de mémoire, mais en raison de la manière dont on accède aux cellules, on peut parler avec raison des relations inter-cellules en termes d'adjacence, de distance, et autres propriétés géométriques." (1980, p. 33).

Dans ce passage, Kosslyn sous-entend manifestement que la façon dont l'image représente l'espace peut être très abstraite. Mais la perception n'est pas non plus spatialement étendue quoiqu'elle serve à percevoir l'espace. N'est-elle pour autant qu'une simulation fonctionnelle de l'espace ? En fait, Kosslyn mélange une question portant sur la nature même de l'image (peut-elle être spatiale ?) et une question portant sur la nature de son lien référentiel (l'image représente-t-elle l'espace ?). Il importe, pour définir la nature de l'image, non de la comparer à un contenu, mais à une autre représentation, une précaution dont Denis a bien souligné la nécessité (cf. Denis, 1989, pp. 69-70). Et d'ailleurs, en d'autres

¹ Cette expression de "*medium*" pourrait être prise en mauvaise part comme hypostasiant un système fonctionnel, dont, d'autre part, les propriétés sont phénoménales, et donc qui serait le support (?) des images telles qu'elles apparaissent à la conscience. On serait alors en droit de faire à cette notion les mêmes reproches qu'à toute entité mentale issue d'une approche à la fois phénoménologique et objectiviste des événements mentaux, à savoir qu'elle implique que les propriétés phénoménales des images sont a priori identifiées à des propriétés du traitement de l'image : le *medium* serait alors défini comme un écran mental dont les propriétés observées de l'image reflètent la nature. Cependant, cette expression de "*medium*" doit s'entendre en un sens très large de "structure qui accueille des formes particulières de structures de données. (...) Les *media* peuvent être spécifiés en se référant à leur *formatage* et à leur *accessibilité*. Le *formatage* pose des restrictions sur les types de structures de données que peut recevoir un *medium*. (...) Les caractéristiques d'*accessibilité* dictent aux processus leur mode d'accès aux structures de données dans un *medium*." (Kosslyn, 1981 : 47). Le terme de "*medium*" est donc une manière commode de parler d'un type de structure de données et des transformations opérables sur celles-ci.

endroits, Kosslyn assimile (correctement, à notre avis) l'image à un quasi percept (par exemple, 1980, p. 18). Par la suite, Kosslyn s'est prononcé beaucoup plus directement pour une interprétation neurophysiologique du tampon visuel, et a cherché à l'identifier à un ensemble de structures cérébrales, à savoir les aires suivant les aires visuelles V1 et peut-être V2 qui projettent topographiquement les motifs rétiniens (cf. Kosslyn, Flynn, Amsterdam & Wang, 1990 ; Kosslyn, 1991 ; Kosslyn & Koenig, 1992).

Du fait que le concept d'isomorphisme est plus approprié pour décrire le lien référentiel d'une représentation que la nature même de cette représentation, il ouvre la voie à une caractérisation fonctionnelle de la représentation imagée ; or, définir une représentation par le type d'information qu'elle préserve est compatible avec une réduction propositionnaliste (ou, au niveau du "langage-machine", digitale ; cf. la citation de Kosslyn ci-dessus) de cette représentation. En ce sens, ce concept est insuffisant pour distinguer l'image et la représentation propositionnelle ; aussi, nous nous tournerons de préférence vers la notion "d'analogique", qui nous paraît plus adéquate.

1.2. Image et medium analogique

Il arrive que l'on qualifie l'image de représentation analogique. Cette notion d'analogique recouvre en fait trois acceptions distinctes. En un premier sens, on appelle analogique une représentation dont les dimensions varient selon une loi physique et dont ces variations ont un contenu informationnel (le paradigme est ici le contraste entre machines digitales et machines analogiques). En un deuxième sens, "analogique" veut dire de manière assez vague "qui préserve des propriétés physiques de l'objet à représenter" (ce qui fait du medium fonctionnel au sens de Kosslyn ou au sens de Pinker un medium analogique). Enfin, plus spécifiquement, "analogique" se dit d'une représentation dont les parties sont corrélées à des parties de l'objet représenté de telle sorte que la représentation ressemble quant à son apparence à l'objet représenté¹.

¹ Palmer (1978 : 295-296) a sans doute tort de dire que selon cette dernière conception, si A est au-dessus de B dans le monde extérieur, la représentation de A

Pylyshyn (cf., notamment, 1981) a présenté un argument visant à montrer qu'il n'y a pas de médium pour lequel les dimensions à représenter (l'organisation spatiale : distance, figure, orientation) impliqueraient qu'il obéisse à des contraintes spécifiques de type analogique (au premier sens du terme). Son raisonnement est le suivant : Certains paramètres des représentations sont liés entre eux de manière systématique ; par exemple, on pose généralement que la latence de vérification d'un détail à partir d'un point focalisé sur une carte mentale est une fonction linéaire de la distance (présumée "mentale") séparant ce détail de ce point (cf. Kosslyn, Ball & Reiser, 1978). Or, il y a deux manières de rendre compte de ces liaisons de paramètres. Soit le sujet sait "tacitement" que ces relations entre temps et distance sont valides dans le monde réel et il tente, dans les tâches d'imagerie mentale, "de recréer aussi précisément que possible les événements qui se produiraient si vous étiez réellement en train d'observer un événement" (Pylyshyn, 1981, p. 31) Soit les relations nomo-logiques découvertes proviennent des propriétés physiques du médium utilisé pour produire les images (il s'agit alors d'un médium analogique au sens un). Or, les images sont altérables par des représentations dont on suppose qu'elles ne sont pas réalisées dans ce médium analogique. Les images sont, dit Pylyshyn, *cognitivement pénétrables*. Par conséquent, le médium des images n'obéit pas exclusivement à des relations nomologiques liant ses propriétés physiques. Il s'ensuit que ce médium n'est pas analogique.

En dépit de Pylyshyn, nous pensons que peut exister un médium indépendant du code propositionnel et analogique en un certain sens, différent du sens restrictif que Pylyshyn impose à ce terme. Plus précisément, les propriétés physiques de ce médium ne sont pas exclusivement celles qui permettent d'expliquer les relations observables entre les différents paramètres d'une représentation. Doivent s'ajouter à ces propriétés les caractéristiques

doit être physiquement au-dessus de la représentation de B (et par "physiquement", il semble bien entendre "physiologiquement", comme le contexte l'indique). Palmer confond ici les parties phénoménales de la représentation et l'implémentation de ses éléments.

computationnelles des processus opérant dans ce medium. Si les objets (par exemple, les assemblages de cubes des expériences de Shepard et Metzler, 1971) étaient représentés dans un système de coordonnées centré sur les objets eux-mêmes, les sujets n'auraient pas à réorienter les objets pour les comparer. Si les objets étaient représentés dans un système de coordonnées centré sur la rétine, il suffirait d'amener l'un perpendiculairement à l'autre pour pouvoir les comparer (en comparant les angles faits par une droite issue, par exemple, de la fovéa, et par les axes principaux des deux objets). Or, le rythme et l'ampleur des réorientations opérées par les sujets dans les expériences de Shepard et Metzler suggèrent que les sujets font mentalement tourner les objets jusqu'à les mettre en parallèle¹. Cela suggère qu'un système de coordonnées cartésiennes (centré sur l'observateur, comme Marr, 1982, le suggère) est plutôt utilisé par le medium imaginal. Il paraît donc de toute façon impossible d'expliquer les résultats de Shepard et Metzler uniquement par les propriétés physiques du medium. Or, cela n'empêche pas que ce medium puisse être "analogique", en un sens différent de celui admis par Pylyshyn et que nous exprimerions ainsi : une propriété d'un processus de transformation d'image est analogique lorsqu'elle est une fonction de (par exemple est proportionnelle à) quelque propriété physique du processus opérant la transformation du percept correspondant ; on appelle "propriété physique d'un processus opérant sur des représentations" une propriété dont la valeur dérive de propriétés réelles des objets réels. Notre définition de l'image comme représentation analogique laisse ouverte la possibilité d'une représentation profonde de celle-ci qui ne soit pas de nature "analogique", au sens où ce terme s'oppose à "digital".

¹ Dixon et Just (1978) ont trouvé que les jugements d'identité portant sur une dimension de deux stimuli distincts pouvaient être sujets à l'interférence des différences, pourtant non pertinentes, que les stimuli ont sur d'autres dimensions. Cette interférence ne pouvait provenir de ce que les dimensions n'étaient pas traitables séparément ; en effet, dans une tâche de classification accélérée, elles étaient susceptibles d'être sélectionnées et traitées apparemment sans effet d'interférence. Ainsi, lorsque des sujets ont à juger de l'identité de deux stimuli, ils auraient tendance à égaliser les deux stimuli sur toutes les dimensions non pertinentes pour la tâche, même si celles-ci peuvent être traitées séparément.

Nous reviendrons ci-après sur cette question de la nature des représentations profondes de l'image.

Après cette mise au point conceptuelle, il nous reste à rappeler les critiques qui ont été adressées à l'argument de la connaissance implicite. Pylyshyn pense que les résultats expérimentaux des imagistes (il pense aux expériences de parcours mental, ou *mental scanning*, et de rotation mentale) révèlent surtout : (a) la connaissance tacite qu'ont les sujets des événements réels qu'on leur demande d'imaginer (b) leur capacité à rythmer l'exécution de la tâche de façon que la durée des procédures utilisées corresponde à celle des événements réels. Les sujets interpréteraient la tâche d'imagerie comme signifiant qu'ils doivent rapporter ce qui se serait passé dans des circonstances perceptives similaires (Pylyshyn, 1984, pp. 239s). Pour réfuter cette argumentation, on s'est attaché à montrer que : les sujets ne peuvent pas prédire les résultats expérimentaux en question ou font des prédictions incorrectes — les résultats sont similaires en l'absence de consignes d'imagerie, dans des situations où la nature de la tâche induit probablement les sujets à utiliser l'imagerie — des propriétés de l'image ont avec la perception des similarités et des différences qui sont difficilement prédictibles. Ainsi, en ce qui concerne le premier type de réfutation, Denis et Carfantan (1985) ont montré que les prédictions par les sujets des résultats de diverses expériences (sur le parcours mental, la rotation mentale et la vérification d'attributs sur des objets imagés à différentes tailles) étaient mauvaises. Reed, Hock et Lockhead (1983) ont constaté que des sujets qui pouvaient prédire le rapport entre le temps de parcours et l'allongement du "trajet" mental étaient incapables d'anticiper les effets produits sur le temps de réponse par un changement de la forme du parcours. Denis et Carfantan (1985, p. 57) notent qu'il demeure possible que les sujets ne récupèrent la connaissance tacite requise que lorsqu'ils sont déjà engagés dans la tâche. De fait, Pylyshyn remarque qu'on peut affirmer du "rappel, dans tous les domaines, que les faits pertinents ne peuvent être récupérés sans passer par une séquence particulière d'indices d'accès" (1984, p. 246).

Un deuxième ensemble d'arguments tend à montrer plus directement que les images sont formées à partir d'un codage propositionnel. Ces arguments sont de deux ordres : les uns concernent la génération, la décomposabilité de l'image en parties "interprétées" — d'autres invoquent la nécessité d'une interlangue permettant de traduire un code dans l'autre. Ces thèses seront abordées dans les sections suivantes.

1.3. *Interprétabilité de l'image*

L'image n'est-elle qu'un épiphénomène ? On a tenté de l'établir en montrant que l'image est au mieux redondante et vient en surcroît d'un codage propositionnel. La différence essentielle entre certains imagistes, qui admettent un pré-codage symbolique (tels Kosslyn et Denis), et les propositionnalistes tient alors à ce que ces derniers considèrent que l'image véhicule seulement des informations redondantes ou inutiles alors que les premiers pensent que l'imagerie facilite le traitement (par l'actualisation additionnelle des traits figuratifs chez Denis) ou que des propriétés "émergentes" peuvent apparaître lors de l'inspection de l'image (cf. Kosslyn & Pomerantz, 1977, pp. 64-65).

D'autre part, de ce que certaines figures imaginées ne sont pas réinterprétables comme leurs contreparties perceptives, certains psychologues ont conclu que les images ne sont pas observées et ensuite interprétées mais sont en quelque sorte intrinsèquement signifiantes (Chambers & Reisberg, 1985). Cependant, d'autres auteurs ont trouvé que des formes émergentes pouvaient bien être détectées sur des images que les sujets créaient par superposition, déplacement, addition ou déletion de parties (Finke, Pinker & Farah, 1989). Savoir jusqu'à quel point l'image est une structure observable et réinterprétable semble donc être une question qu'on ne peut trancher *a priori*. Il ressort des études mentionnées que les formes non réinterprétables auraient pour caractéristique de n'être recatégorisables qu'en maintenant le percept *total* "en vue". D'autre part, les formes émergentes, dans les expériences de Finke, Pinker et Farah, n'étaient pas de celles qui imposaient des réinterprétations de leurs parties, par suite de contraintes interprétatives issues de la

figure entière obtenue, contrairement à celles utilisées par Chambers et Reisberg.

De tout cela, il s'ensuit que l'image pourrait être un produit *sui generis* des systèmes de traitement visuel, pourvu de caractéristiques propres.

Une argumentation majeure a été avancée par Pylyshyn (1973, p. 9) pour justifier que l'image est la projection d'une structure propositionnelle : les parties créées lors du processus de génération de l'image n'ont pas à être découvertes par une inspection active de l'image et lorsqu'une image est lacunaire, ce sont des parties ou des attributs perceptifs qui font défaut dans leur intégralité. De plus, une multitude d'images peuvent être informationnellement équivalentes ou exploitées de la même manière par l'appareil cognitif. L'information commune doit être *extraite* de l'image et non lue simplement sur chaque image puisque la multiplicité indéfinie des images correspondant à une signification particulière et le caractère non systématique ou productif de cette multiplicité interdisent que chacune soit liée à une signification particulière. Il s'ensuit, selon Pylyshyn, que l'image est composée de parties et d'attributs déjà interprétés et que l'image ne peut être le support de la signification. En effet, le fait que l'image est engendrée ou rappelée par parties significatives ou attributs univoques prouve qu'elle n'est pas un tableau perceptif, dont la dégradation aurait par exemple pour effet de conserver tous les détails de la scène originale mais avec une précision moindre. Et la synonymie d'images distinctes montre que l'information exploitée ne correspond pas forcément à celle qui est présente dans l'image.

En ce qui concerne le premier argument, il paraît supposer que l'identification d'un objet est obligatoirement une étape propositionnelle. Il est exact et bien établi que les configurations porteuses de signification paraissent se comporter différemment lors de l'effacement d'un percept. Pritchard (1961) rapporte que dans une expérience sur la fragmentation perceptive de mots, les suites significatives ont plus de chance de résister plus longtemps à l'effacement. Paivio (1971, p. 102) note que dans nombre de ces expériences, un étiquetage verbal implicite, concourant par un mécanisme de rétroaction à stabiliser le percept, pourrait rendre

compte de la plus grande stabilité des configurations signifiantes. En effet, selon lui, un codage qu'il appelle "référentiel" (liant l'image à son nom) ou un codage "associatif" (liant des représentations hétérogènes) ont, dans ces expériences, le temps de se développer. On peut aussi donner de ce phénomène une explication invoquant les "assemblées neuronales" postulées par Hebb, pour qui le sens est "l'activité au sein d'un ensemble d'assemblées interconnectées qui, en tant que système de complexité supérieure, peuvent être actives plus longtemps" (Hebb, 1963, p. 13 ; cité dans Paivio, *ibid.*, p. 101). En tout cas, l'explication la plus satisfaisante du phénomène de stabilité perceptive des configurations signifiantes ne doit pas nécessairement faire usage d'un niveau propositionnel de représentation. Il est probable que si Pylyshyn fait l'hypothèse d'un tel niveau, c'est qu'il conçoit l'étiquetage verbal ou l'image comme autant de *codages* possibles pour un même contenu informationnel¹.

Des arguments empiriques ont paru apporter une caution supplémentaire à la thèse que l'image est organisée propositionnellement, en suggérant que la nature analogique de l'image ne contraignait pas systématiquement les processus opérant sur elle. Certaines expériences montrent ainsi que la durée du changement de focalisation d'un point à un autre sur une image ne dépend pas seulement de la distance entre ces points. Ainsi, dans une expérience où les sujets, à qui on demande de diriger leur "regard" mental depuis un point donné, doivent indiquer la position de cibles environnantes, on constate que les cibles les plus proches du point initial et celles qui lui sont directement opposées sont détectées plus rapidement ; en revanche, les cibles placées à une distance intermédiaire sont localisées plus lentement (Hintzman, O'Dell & Arndt, 1981, p. 149). Les auteurs concluent que

"les cartes mentales ne sont pas strictement holistiques, mais consistent en représentations orientées de manière spécifique et — au moins en partie — de propositions reliant des paires d'objets."

¹ Ainsi, il soutient que "les phrases comme les figures doivent être interprétées avant de devenir des contenus conceptuels. La raison en est qu'il y a un nombre indéfini de figures et de phrases qui sont cognitivement équivalentes." (1973 : 7)

Dans des expériences similaires où les sujets doivent, depuis une position imaginaire, donner l'orientation d'une cible par rapport à une autre, Pylyshyn a noté que la corrélation existant habituellement entre la distance et le temps de vérification n'était pas observable (Pylyshyn, 1981, pp. 38-40). Dans ces deux séries d'expériences, les résultats obtenus avec l'instruction d'imagerie sont différents de ceux obtenus dans la condition perceptive, lesquels se conforment à la corrélation attendue. Il est donc impossible de les expliquer en invoquant la saillance perceptive de certaines orientations. Il s'ensuit, dans l'esprit des expérimentateurs, que la saillance d'orientations particulières sur l'image est une propriété non perceptive, donc propositionnelle, de la représentation elle-même, non des processus exploitant cette représentation. On voit que l'argument comporte beaucoup de suppositions, que les auteurs ne cherchent d'ailleurs pas à justifier. Remarquons simplement que la thèse que l'image est une représentation analogique n'implique nullement que tous les processus opérant sur l'image réagissent aux propriétés analogiques de l'image. Les processus de génération d'une image qui suivent des instructions verbales ne peuvent être en ce sens des processus "analogiques". De plus, comme nous le verrons, les propriétés de l'image qui l'assimilent à la perception ne sont pas absolument identiques aux propriétés correspondantes de la perception.

La thèse de Pylyshyn a en outre l'inconvénient d'impliquer que l'identification d'une partie significative d'une scène perceptive est une étape propositionnelle. Selon cette thèse, en effet, le découpage propositionnel de la scène imaginée détermine totalement son découpage en objets visuels. Il se peut que les deux décompositions coïncident jusqu'à un certain point, c'est-à-dire que les termes propositionnels désignent de fait des unités perceptives ; mais cela n'implique pas que la constitution des objets perceptifs et l'étiquetage propositionnel correspondent à un même stade de traitement. Or, il est vraisemblable que l'identification d'un objet passe par une étape purement perceptive et non symbolique. La distinction entre un stade post-sensoriel et précédant la catégorisation elle-même paraît bien établie. Une série d'études neuropsychologiques ont montré que la capacité à apparier

correctement des objets d'un point de vue perceptif était dissociable de la capacité à catégoriser (cf., notamment, Warrington & Taylor, 1978 ; McCarthy & Warrington, 1990). Le stade d'analyse perceptive (ou "apperceptif") a été mis en évidence par la comparaison des performances de sujets souffrant de lésions droites ou gauches à deux types de tâches : dans l'un, les sujets doivent appairer des photographies d'objets pris sous des angles inhabituels avec des vues plus conventionnelles ; dans l'autre, les sujets doivent associer une photographie à un membre d'une paire d'objets semblables d'apparence, mais dont les fonctions diffèrent (par exemple une chaise à accoudoirs est à associer soit à une chaise roulante de malade, qui a des accoudoirs aussi, soit à une chaise longue). On constate que les erreurs des cérébrésés droits sont attribuables à un défaut de reconnaissance perceptive : leurs erreurs sur la tâche d'appariement fonctionnel sont très largement prédictibles d'après leur taux d'erreurs à la tâche d'appariement perceptif. En revanche, les erreurs de cérébrésés gauches (plus précisément, ceux atteints de lésions à l'hémisphère postérieur gauche) se révèlent imputables à peu près exclusivement au défaut d'association fonctionnelle. Warrington et Taylor attribuent la connaissance de la fonction à l'aire "sémantique" de l'hémisphère gauche, la liant essentiellement à la capacité linguistique. Mais il n'est pas du tout certain que l'on soit en présence, chez les cérébrésés gauches, d'une perturbation de la capacité à nommer implicitement les objets. Si Warrington et Taylor trouvent bien une corrélation entre la capacité à nommer un objet et celle à associer des objets par la fonction, cette corrélation est relativement faible. D'autre part, Shallice (1988a : 193) objecte que là où le sujet a le choix entre seulement deux items (ce qui est le cas dans la tâche d'appariement fonctionnel), il est peu probable que ce soit le choix des étiquettes verbales possibles, parmi lesquelles le sujet doit choisir, qui suffise à détériorer les performances. Quoi qu'il en soit de l'origine de ce type de connaissance, et de la nature linguistique de la connaissance fonctionnelle, il demeure que des patients souffrant d'agnosies graves touchant la catégorisation d'objets sont capables de réussir la tâche d'appariement perceptif. Par conséquent, le découpage de la scène perceptive en objets ne semble pas être basé sur une analyse propositionnelle au sens où cette analyse implique que

l'objet soit implicitement étiqueté. On peut certes supposer que la construction du percept est entièrement propositionnelle avant même sa catégorisation, et qu'un ensemble de vues sous différents angles sont codées en mémoire avec leurs seuils de reconnaissance propres (cf. Warrington & James, 1986). Ainsi, on peut montrer qu'un codage propositionnel est assez puissant pour être équivalent à une image transformable de diverses façons. Le problème est plutôt de limiter sa puissance et de montrer pourquoi certains paramètres des représentations imagées sont toujours liés entre eux, et pourquoi certaines propriétés des images sont prédictibles à partir d'une analogie entre image et percept. Ainsi, quand il s'agit de juger de la forme d'un objet, la taille et l'orientation affectent le jugement des sujets. Quand il s'agit de vérifier qu'un objet possède bien une certaine propriété, la taille présumée de l'objet dans l'image influence le temps de vérification (cf. Kosslyn, 1975 et 1976).

Enfin, même si l'image est composée à partir de structures propositionnelles et est interprétée par des processus faisant appel à de telles représentations, il ne s'ensuit pas pour autant que le format de l'image permette seulement de retraduire ce qui a déjà été exprimée sous forme propositionnelle. Nous allons maintenant examiner ce problème de la traductibilité de l'image.

1.4. La traductibilité de l'image

Si l'image est intégralement traductible dans un code propositionnel, il peut sembler qu'elle est redondante. Si, de plus, on n'a aucun moyen direct d'assurer que les sujets ont bien recours à une représentation imagée dans certaines circonstances, alors on peut craindre qu'il soit impossible de déterminer quel est le format de représentation utilisé. Anderson (1978) a développé une argumentation tendant à démontrer qu'il était a priori possible d'exprimer un modèle d'un format dans un modèle d'un autre format et ainsi de rendre indistinguables les deux modèles quant aux discriminations que leurs représentations respectives contiennent. Nous la résumerons, et tenterons de montrer que les hypothèses de départ ne sont pas fondées.

Supposons, dit Anderson, que l'on puisse définir un modèle M des représentations mentales en spécifiant l'ensemble des fonctions

opérant sur elles. Ces fonctions lient les représentations internes (notées I) aux stimuli (S), aux réponses (R) et à d'autres représentations. Soit alors l'ensemble $\mathbf{I} = \{I_1, I_2, \dots, I_i, I_j, \dots, I_n, \dots\}$ spécifiant l'ensemble des représentations internes du modèle. Les fonctions définies sur le modèle seront la fonction Encodage (notée $E / E(S_i) = I_i$), la fonction Transformation (notée $T / T(I_i) = I_j$) et la fonction Décodage (notée $D / D(I_i) = R_i$). On pose les conditions pour qu'il existe un modèle M^* défini par \mathbf{I}^* tel que \mathbf{I}^* préserve les distinctions internes de \mathbf{I} : on dira qu'il existe une bijection f de \mathbf{I} dans \mathbf{I}^* qui possède une inverse calculable telle que $E(S) = f(E^*(S))$. Le fait que f soit une bijection assure que toute distinction au sein de \mathbf{I} est préservée dans \mathbf{I}^* ; la condition que $E(S) = f(E^*(S))$ garantit que E et E^* assignent une représentation correspondante aux mêmes stimuli. Si on définit en outre T^* et D^* comme égales, respectivement, à $f \circ T \circ f^{-1}$, et à $D \circ f^{-1}$, on peut alors déduire que les fonctions définies sur M^* sont exactement parallèles aux fonctions de M . En effet, $E^*(S_i) = f(I_i)$, $T^*(f(I_i)) = f(I_j)$ et $D^*(f(I_i)) = R_i$. On notera que les données comportementales ne permettent pas de savoir quel modèle choisir, puisque les liaisons S - R sont les mêmes dans M et M^* .

Si l'on considère que la fonction f décrite plus haut exprime les relations formelles liant les représentations correspondantes de deux modèles, il semble de prime abord plus difficile de proposer un critère objectif susceptible de les départager. Anderson remarque ainsi (1978, pp. 271s) que les données sur la spécialisation hémisphérique n'autorisent aucune conclusion sur le format des représentations. Il se pourrait en effet que les informations soient localisées selon leur contenu (visuo-spatial dans l'hémisphère droit, par exemple) et non d'après leur format. Il en conclut que des considérations méta-théoriques (faire l'économie d'un format, donner une conception unitaire de l'esprit, au moyen d'un format permettant une modélisation globale, comme dans le système ACT) fournissent les meilleurs critères de choix. Cependant, en l'occurrence, les arguments empiriques ne nous font pas totalement défaut, car nous avons de bonnes raisons de penser que le cortex visuel lui-même est impliqué dans la formation des images (cf.

Farah, 1988). Or, il est peu probable que le traitement visuel lui-même soit propositionnel (la théorie computationnelle la plus développée, celle de Marr (1982), n'est pas une théorie propositionnelle, au sens où les propriétés utilisées seraient des traits exprimables sous forme prédicative ; une théorie telle que celle de Guzmàn (1969) peut être considérée comme propositionnelle, dans la mesure où elle formule sous forme de *règles symboliques* les lois de construction d'objets dont les bords se recoupent ; mais elle n'est pas une théorie de tous les stades d'analyse perceptive, comme celle de Marr).

A l'encontre d'Anderson, nous pensons qu'il n'est pas possible de poser qu'un des modèles préserve les distinctions internes de l'autre sans spécifier les processus effectuels sur les représentations. Il faut, en effet, distinguer les contenus implicites et explicites. Deux modèles peuvent avoir des représentations au contenu informationnel effectivement représenté équivalent¹. Mais ils peuvent différer par les informations exploitables à partir de ces représentations. Ces informations sont implicites particulièrement en raison des processus qui peuvent s'appliquer sur les représentations qui les véhiculent : songeons à la différence entre la représentation procédurale d'un trajet (après la banque, tourner à droite, puis continuer tout droit jusqu'au premier croisement etc.) et sa représentation sous forme de carte mentale. L'information "x habite au point L" peut n'être pas explicitement représentée dans le format procédural. Dans ce cas, les distinctions explicites diffèrent mais non les distinctions implicites. Or, la différence entre ces deux types de distinctions n'est spécifiable qu'en précisant les processus effectuels sur les représentations. Le raisonnement d'Anderson suppose que l'on puisse définir les représentations explicites sans spécifier les processus effectuels sur ces représentations puisqu'il pose que ces représentations constituent un ensemble énumérable indépendamment des fonctions E, T et D. En outre, si Anderson sous-entend que les deux modèles sont implicitement équivalents (seulement, donc, explicitement distincts), alors il doit être possible

¹ Le contenu informationnel effectivement représenté peut être défini comme l'information extractible sans effectuer de processus de transformation sur les structures existantes.

de les distinguer autrement que par les opérations supplémentaires de traduction, dès lors que les opérations effectuant la récupération de la même information sont différentes. Si Anderson sous-entend que les deux modèles sont explicitement équivalents, il faut qu'ils soient implicitement non équivalents. Sinon, les processus définis sur ces deux modèles seront également équivalents et les deux modèles seront identiques. Comme l'intention d'Anderson n'est pas de montrer que deux modèles identiques sont identiques au changement notationnel près (au sens où $(x)(\exists y)Pxy$ est équivalent à $(z)(\exists w)Pzw$), soit il suppose que les deux modèles sont implicitement (seulement) équivalents, et donc distinguables par des processus autres que ceux de traduction et de retraduction ; soit il suppose que les modèles sont explicitement équivalents, et donc là aussi distinguables par leurs processus.

Pylyshyn (1979b) a formulé une objection montrant qu'il est très improbable qu'un modèle préserve les distinctions informationnelles de l'autre. Supposons une théorie imagiste dans laquelle S_1 et S_2 sont encodés par I_{12} , et S_3 par I_3 ; posons, d'autre part, que dans une théorie propositionnelle, S_2 et S_3 sont encodés par P_{23} . Forçons la correspondance entre les I et les P à la bijectivité, de telle sorte que S_2 est encodé par I_{12} et P_{23} tandis que S_3 est encodé par I_3 et P_{23} ; la distinction entre I_{12} et I_3 est donc perdue dans la traduction propositionnelle. On a donc ici un cas de correspondance bijective entre un ensemble I et un ensemble I^* , respectant les conditions d'Anderson, qui pourtant ne préserve pas les distinctions internes.

1.5. L'image et la notion de ressemblance en philosophie

La notion d'image a connu une certaine importance dans la théorie philosophique des représentations. Même chez des auteurs posant la prééminence du concept, des mécanismes abstractifs propres à l'image étaient reconnus (cf. "l'idée-normale" chez Kant, qui est conçue comme une représentation moyenne abstraite des percepts successifs). Chez les atomistes de l'Antiquité, les "images" des choses qui nous informent sur leur être sont des simulacres

matériels qui interagissent physiquement avec nos organes sensoriels. Chez les empiristes (notamment Hume) l'essentiel des processus de connaissance est dévolu à l'image. L'image permettait en effet de rendre compte aisément de plusieurs aspects des représentations : d'une part, elle laisse entrevoir une explication physique de la transmission de l'information par les sens — d'autre part, la ressemblance de l'image à son référent assure que l'encodage ne peut être arbitraire — enfin, l'image est "lue" sans médiation, puisqu'elle ressemble au percept dont elle est issue. Or, malgré l'attrait présenté par une théorie de la connaissance sinon basée sur l'image, du moins prenant celle-ci comme le modèle de toute représentation, le rôle de l'image a été constamment dévalué.

Les tentatives de réduire l'image à un épiphénomène se situent dans une certaine tradition de pensée (peut-être l'ensemble de la philosophie occidentale, marquée par son logocentrisme), qui tend à souligner la nature de signifiant de l'image. Ainsi, la thèse que le sens de l'image lui vient de sa nature descriptive et de sa ressemblance à ce qu'elle figure a été constamment critiquée. Les objections majeures qui ont été opposées à la théorie de la valeur icônique de l'image sont les suivantes : (1) d'une part, si l'image est bien un quasi-percept, elle doit différer du percept par son intensité ou sa vivacité ; (2) d'autre part, l'interprétation d'une image peut lire sur celle-ci des propriétés que l'image ne représente pas exclusivement ou qu'elle ne représente qu'en tant que signe. L'image est donc un quasi-signé, dont la ressemblance plus ou moins arbitraire avec son objet est le produit d'une conscience active ou d'une intention signifiante.

On peut montrer d'abord que l'image n'a pas la détermination et la précision supposées d'un tableau perceptif, qu'elle est lacunaire et que les caractères qui lui manquent sont de ceux qui ne peuvent faire défaut à une représentation perceptive sans contradiction. Selon l'exemple fameux d'Alain, il m'est impossible de compter les colonnes du Panthéon sur l'image que j'en ai. Quel sens cela a-t-il de poser comme quasi-percept une représentation dont le nombre de parties est indéterminée ? C'est là une lacune qui paraît en outre indépendante du plus ou moins de vivacité de l'image. Cet argument se trouve ainsi chez Dennett, dans un passage bien caractéristique de la doxa philosophique en la matière :

"Considérez le tigre et ses rayures. Je peux rêver , imaginer ou voir un tigre zébré, mais est-ce que mon tigre phénoménal doit avoir un nombre particulier de rayures ? Si voir ou imaginer, c'est avoir une image mentale, alors l'image du tigre *doit* — obéissant aux règles des images en général — faire apparaître un nombre de rayures défini, et l'on devrait pouvoir tirer le problème au clair en se posant des questions comme "plus de dix ?" ou "moins de vingt ?". (1969, p. 136)

Dennett en conclut naturellement que l'image est en réalité une description. Plus généralement, l'argument vise à prouver que l'image est lacunaire d'une manière qu'on n'attendrait pas si elle n'était qu'un percept moins intense : dans ce cas, elle pourrait être trouble ou moins définie mais certainement pas *vague*. Précisément, l'attrait du format propositionnel réside essentiellement dans le fait que le *vague* et le *général* sont par lui représentables. De plus, les propriétés de l'image qui devraient être corrélées si elle était bien un quasi-percept sont représentables séparément, de sorte que la certitude que j'ai d'avoir à l'esprit l'image du Panthéon (et que j'attribue à "l'intensité" de l'image) peut être séparée du fait que je suis incapable d'en compter les colonnes. Or, sur une image quasi perceptuelle, intensité et définition devraient être corrélées. Ainsi, une description propositionnelle peut représenter le vague simplement en ne spécifiant pas certains paramètres, et le général, simplement par l'utilisation de symboles linguistiques. Cette façon de justifier l'adoption d'un code propositionnel se retrouve explicitement dans des travaux préoccupés par la représentation des gabarits perceptifs (cf. Humphreys & Bruce, 1989, pp. 62-63).

L'exemple favori des philosophes, dans lequel s'exprime le deuxième type d'objection, et qui parcourt toute l'histoire de la philosophie (d'Aristote à Berkeley, Hume, Locke, Kant et Husserl), est celui de l'image du triangle : mon image du triangle est particulière, mais cette image peut me permettre de raisonner sur le triangle en général ; par conséquent, les particularités de mon image (que le triangle soit équilatéral, rectangle, scalène) sont indifférentes pour l'interprétation générale que j'en fais ; et la décision d'interpréter l'image d'un point de vue général est le fruit d'une intention signifiante qui constitue l'image en tant qu'image de

quelque chose¹. Ainsi, Sartre insiste beaucoup sur le fait que les éléments de l'image sont créés par l'acte même de "quasi observation". En outre, si l'image était totalement déterminée avant d'être "inspectée", des parties d'elle-même demeureraient cachées à la conscience. Quel sens cela a-t-il de poser que nous avons à l'esprit des constituants de représentations auxquels nous n'accédons pourtant jamais ?

Chez certains auteurs (Husserl, Sartre), l'image est ainsi définie par un mode de rapport conscient à l'objet visé, dans lequel l'objet est appréhendé comme non-existant ; cela permet de distinguer l'image de la perception et de l'hallucination, et de répudier le principe de l'assimilation de l'image à une perception affaiblie².

Dans le cas de l'image du triangle, c'est au concept qu'est confiée la fonction d'interpréter l'image, ou comme chez Kant, à un intermédiaire entre le concept et l'image, le schème. Le schème est une règle de formation des images dans l'intuition pure de l'espace (l'espace en tant qu'indépendant des objets dans l'espace, ou l'espace antérieur à la coexistence des objets), c'est-à-dire, dans la pensée de Kant, dans l'intuition pure du temps puisque la possibilité

¹ Ainsi, Husserl note-t-il : "L'expression simpliste d'images internes (par opposition aux objets extérieurs) ne doit pas être tolérée dans la psychologie descriptive (ni *a fortiori* dans la phénoménologie pure). Le tableau n'est une image que pour une conscience constituante d'image, c'est-à-dire qui seule confère à un objet primaire et lui apparaissant dans la perception la "valeur" ou la "signification" d'une image au moyen de son aperception imaginative (fondée dans ce cas sur une perception)." (1962, vol.2, tome 2 : 229)

² Ainsi, Sartre, commentant les expériences du type de celles de Perky (1910) (qui tendaient à susciter des confusions entre des perceptions faibles et des images), écrit : "Ces recherches n'auraient de sens que si l'image était une faible perception. Mais elle se donne *comme image*, toute comparaison d'intensité entre elle et la perception est donc impossible. On ne sait à qui reconnaître la plus grande incompréhension : à l'expérimentateur qui pose de semblables questions ou au sujet qui répond docilement." (1986 : 109). Sartre paraît reprocher au sujet de ne pas s'être élevé au concept d'image correct (c'est-à-dire le sien) ou bien d'avoir une expérience perceptive qu'il prend pour une expérience d'image : il est piquant de voir un phénoménologue accuser un sujet d'avoir un rapport inauthentique à son objet phénoménal. En tout cas, le problème est ramené à une question de définition (que doit-on entendre par "image"?), sans que ces expériences pourtant intéressantes mais mettant en question le dogme de la *conscience* constituante, soient exploitées.

de toute synthèse (y compris spatiale) est la permanence du sujet (dit "transcendantal", c'est-à-dire à l'origine de toute représentation particulière). Cette règle délimite de manière pure (c'est-à-dire sans passer par l'expérience actuelle d'un triangle dans l'espace extérieur) ce qui doit valoir comme triangle. Il en va de même pour les concepts qui sont issus de l'expérience des objets externes :

"Dans le fait, nos concepts sensibles purs n'ont pas pour fondement des images des objets, mais des schèmes. Il n'y a pas d'image d'un triangle qui puisse être jamais adéquate au concept d'un triangle en général. En effet, aucune image n'atteindrait la généralité du concept en vertu de laquelle celui-ci s'applique à tous les triangles, rectangles ou non, mais elle serait toujours restreinte à une seule partie de cette sphère. (...) Le concept de chien signifie une règle d'après laquelle mon imagination peut exprimer en général la figure d'un quadrupède, sans être astreinte à quelque chose de particulier que m'offre l'expérience, ou mieux à quelque image possible que je puisse représenter *in concreto*." (1944, pp. 152-153).

Il n'est alors guère surprenant que Kosslyn, qui a intégré dans sa théorie une composante propositionnelle destinée à décrire les objets imagés, voie dans le schème un précurseur direct de ses "fiches propositionnelles" (Kosslyn, 1980, p. 451). Chez Kant, l'interprétation de l'image d'un concept comme celui de triangle est fixée par l'application des schèmes à la forme pure de l'espace. En dernier ressort, *l'arbitraire* de l'interprétation est contraint par les conditions de possibilité de toute expérience, dont Kant fait la théorie. Comme Kosslyn le remarque (1980, p. 452), si des processus interprétatifs pouvaient faire passer n'importe quelle image pour la représentation de n'importe quoi, alors c'est que des processus non imaginaires constitueraient déjà une représentation non conventionnelle (à la différence des symboles linguistiques) des objets et, dans ce cas, l'utilité de représentations imagées n'existerait plus. Kosslyn, quant à lui, fait l'hypothèse

"qu'une fonction interprétative innée et/ou des genres sélectifs d'expérience déterminent quels aspects de l'image seront pris en compte par le regard intérieur." (ibid.)

Ainsi, l'affirmation que l'image est un signifiant plus ou moins arbitraire dont l'interprétation est contrainte de manière innée peut

apparaître comme une version moderne de la détermination de l'image par le schème.

Nous allons maintenant examiner les arguments empiriques tendant à montrer la parenté de l'image et de la perception, puis nous nous interrogerons sur les rapports entre les aspects perceptifs de l'image et les représentations (linguistiques ?) porteuses de signification.

2. Sur la nature de l'image

2.1. Image et perception

Nous entendons ici par "perception" à la fois un mode de représentation aux propriétés particulières et un type de représentation dont les opérations sont guidées en partie par l'environnement. L'image (visuelle) semble avoir des propriétés idiosyncratiques qui l'apparentent à la vision et partage avec le percept la caractéristique que les processus opérant sur elle sont contraints en partie par la nature de l'objet représenté (l'imagerie est ainsi une sorte d'observation). Notons encore une fois qu'il importe de ne pas confondre la constitution de l'image (la nature de son format, de ses parties, les modules de traitement entrant dans sa génération), et ce qui lui communique son sens. Il s'agit ici d'explorer la constitution de l'image (son *format*) à la lumière d'études empiriques.

Quatre ensembles de preuves sont invoqués pour établir la parenté de l'image et de la perception. Le premier montre que les propriétés analogiques (voir aussi supra) de l'image se retrouvent dans la perception ; le deuxième rassemble les expériences où un effet de l'imagerie sur une activité perceptive est obtenu (ou vice versa). Le troisième comprend les données suggérant que des propriétés de la vision se retrouvent dans l'activité d'imagerie. Le quatrième réunit des données neuropsychologiques et neurophysiologiques mettant en évidence l'importance du module visuel dans l'imagerie et le partage de ressources communes par la perception visuelle et l'imagerie.

2. Sur la nature de l'image

2.1. Image et perception

Nous entendons ici par "perception" à la fois un mode de représentation aux propriétés particulières et un type de représentation dont les opérations sont guidées en partie par l'environnement. L'image (visuelle) semble avoir des propriétés idiosyncratiques qui l'apparentent à la vision et partage avec le percept la caractéristique que les processus opérant sur elle sont contraints en partie par la nature de l'objet représenté (l'imagerie est ainsi une sorte d'observation). Notons encore une fois qu'il importe de ne pas confondre la constitution de l'image (la nature de son format, de ses parties, les modules de traitement entrant dans sa génération), et ce qui lui communique son sens. Il s'agit ici d'explorer la constitution de l'image (son *format*) à la lumière d'études empiriques.

Quatre ensembles de preuves sont invoqués pour établir la parenté de l'image et de la perception. Le premier montre que les propriétés analogiques (voir aussi supra) de l'image se retrouvent dans la perception ; le deuxième rassemble les expériences où un effet de l'imagerie sur une activité perceptive est obtenu (ou vice versa). Le troisième comprend les données suggérant que des propriétés de la vision se retrouvent dans l'activité d'imagerie. Le quatrième réunit des données neuropsychologiques et neurophysiologiques mettant en évidence l'importance du module visuel dans l'imagerie et le partage de ressources communes par la perception visuelle et l'imagerie.

2.1.1. Propriétés analogiques de l'image

Au sens "d'analogique" que nous avons défini plus haut, nous dirons que sont analogiques les propriétés de l'image dont découlent l'effet de distance mentale, l'effet de taille, et l'effet d'écart angulaire.

L'effet de distance correspond à la relation linéaire observée entre la latence de vérification d'un item sur une carte mentale que des sujets ont dû mémoriser, et la distance effective séparant cet item du point de la carte sur lequel on avait prescrit aux sujets de focaliser leur attention (Kosslyn, Ball & Reiser, 1978). La même

relation existe pour des objets présentés en trois dimensions : les temps de parcours mental augmentent linéairement avec les distances tri-dimensionnelles et non pas avec les distances bi-dimensionnelles de la configuration quand on l'observe d'un certain point de vue (Pinker & Kosslyn, 1978). L'effet de taille a été découvert par Kosslyn, à qui nous empruntons le résumé de ses expériences :

"Dans ces expériences, les sujets imageaient une scène contenant un animal-cible (tel qu'une oie) à côté (a) d'un animal-contexte (un éléphant ou une mouche), (b) d'une matrice (de quatre ou seize cellules), ou (c) d'un ensemble de chiffres (deux ou quatre — voir Kosslyn, 1975). Quand les sujets imageaient une cible à côté d'un éléphant à la même échelle, il leur fallait plus de temps pour voir des propriétés de la cible (par exemple, les pattes d'un colley) que lorsque la cible était à côté d'une mouche à la bonne échelle. Ce résultat était inversé quand les éléphants étaient imagés petits et les mouches énormes. De même, il fallait plus de temps pour voir des propriétés d'animaux imagés relativement petits à côté d'une matrice simple ou complexe ou de deux ou quatre chiffres ; les effets dus à la taille comparée étaient les mêmes quels que soient la complexité de la matrice adjacente ou le nombre de chiffres inclus dans l'image." (Kosslyn, 1980 : 225).

L'effet de taille est donc distinct de tout effet dû à la complexité de l'image. Cet effet n'est pas dû non plus à la force d'association liant la propriété à vérifier à l'animal, car on constate que la taille des propriétés influence les temps de vérification, indépendamment de la force d'association, dans une tâche impliquant une activité d'imagerie. Enfin, l'effet de taille paraît attribuable à la résolution du médium plutôt qu'à la plus ou moins grande facilité à localiser le point d'attache de la partie test. En effet, il faut moins de temps à des sujets pour imaginer une partie située dans son contexte que seule (Kosslyn, *ibid.* : 241-246).

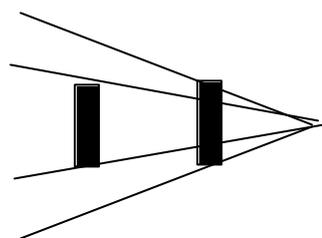
L'effet d'écart angulaire correspond à la relation linéaire existant entre le temps mis pour vérifier que deux figures sont identiques et la différence angulaire entre ces deux figures. Dans les expériences de Shepard et Metzler (1971), qui ont les premiers mis en évidence le phénomène, des sujets étaient confrontés à des paires de figures en trois dimensions, soit identiques, soit non identiques, constituées de cubes accolés les uns aux autres ; quand les deux formes étaient identiques, elles se distinguaient par la plus grande rotation imprimée à l'une par rapport à l'autre, soit dans le plan du dessin, soit en profondeur. Par la suite (Cooper & Shepard, 1973), ces résultats ont été étendus à la comparaison entre un stimulus test et une représentation en mémoire, avec cependant une divergence notable : lorsque des caractères alphanumériques sont utilisés, la relation entre l'écart angulaire et le temps de vérification (en l'occurrence, dire si la figure est à l'endroit ou à l'envers) n'est pas linéaire, bien que le dernier augmente effectivement en fonction du premier. La courbe de cette fonction a une forme d'arc, indiquant que plus l'écart angulaire est

grand et plus le temps de vérification augmente. On constate cependant que la non linéarité est due aux conditions dans lesquelles les sujets ne connaissent pas à l'avance l'identité et l'orientation du stimulus test. S'ils les connaissent au préalable, la relation trouvée est bien linéaire. Kosslyn (1980 : 305-323) pense que dans cette condition les sujets se livrent à un appariement de formes-gabarits, alors que dans les autres conditions où leur fait défaut l'une au moins de ces informations, ils tournent mentalement le stimulus-test jusqu'à une position où le haut du caractère est suffisamment discriminable pour permettre de juger de l'orientation du caractère. Le temps de vérification pour des figures peu déviées de l'orientation droite canonique est donc très court, comparativement à des figures plus éloignées de cette orientation standard. La familiarité des figures à comparer a donc un effet sur les temps de rotation. Nous retrouverons cette importance de la familiarité quand nous discuterons l'importance de la complexité des figures sur le traitement des images.

Ces expériences montrent toutes que l'image mentale doit être parcourue pour livrer des informations qu'elle contient implicitement. Il est cependant difficile de dire jusqu'à quel point l'esprit est capable de lire des propriétés visuelles émergeant d'une image transformée mentalement. Pinker et Finke (1980) ont montré que des sujets étaient capables de décrire avec assez de précision la configuration formée par quatre objets suspendus (et placés à diverses profondeurs) qu'on leur avait demandé de faire tourner de 90°. En outre, des sujets peuvent aussi parcourir mentalement des objets selon une nouvelle perspective, de sorte que leur temps de parcours mental reflète non la distance perçue, mais la distance imaginaire (Pinker, 1980, expérience 3).

2.1.2. Interactions entre images et perceptions

Un deuxième groupe d'expériences tend à montrer des interactions entre une activité imaginative et une activité perceptive. Soit les deux activités coopèrent, soit elles entrent en conflit (donnant lieu à des phénomènes d'interférence). On en conclut alors qu'image et perception partagent les mêmes capacités de traitement. Freyd et Finke (1984) ont trouvé qu'imaginer un gabarit facilitait une tâche de comparaison. Dans leur expérience, certains sujets, qui avaient à juger de la longueur de deux lignes formant une croix, devaient imaginer un carré de longueur inférieur à celles des deux lignes. Quand des discriminations fines de longueur étaient à faire, l'image du carré se révélait faciliter la tâche. Wallace (1984) rapporte que l'illusion de Ponzo se produit également quand les sujets se contentent d'imaginer les droites convergentes.



L'illusion de Ponzo

Plus nombreuses sont les expériences suggérant un partage des ressources visuelles et imaginatives par l'effet d'interférence. Le pionnier de ces travaux est Brooks (1967) : rappelons que dans ses protocoles, les sujets avaient à accomplir une tâche de décision lexicale sur des items présentés auditivement, ou à tracer mentalement une lettre (un "F") à partir du coin inférieur gauche, puis devaient indiquer pour chaque coin rencontré s'il était au bas ou en haut de la figure. On demandait aux sujets de répondre soit en disant "oui" ou "non", soit en faisant un geste de la main, soit en montrant un "y" (pour "yes") ou un "n" présentés visuellement. Ce dernier mode de réponse s'est révélé interférer le plus avec la tâche de tracé mental, tandis que la réponse parlée interférait le plus avec la tâche de décision lexicale. Segal et Fusella (1970) ont découvert que la détection de motifs perceptifs était dégradée quand des sujets devaient se livrer à une activité d'imagerie dans la même modalité. Eddy et Glass (1981) ont montré que la vérification de l'exactitude ou simplement de la cohérence d'informations contenues dans des phrases à valeur d'imagerie élevée était ralentie lorsque ces phrases étaient présentées par écrit. Ils en concluent que la lecture interfère avec l'imagerie. Corballis et McLaren (1982) ont mis en évidence que l'effet consécutif dû à un disque en rotation affectait les performances des sujets dans une tâche par ailleurs similaire à celle de Cooper et Shepard présentée ci-dessus. Lorsque l'effet consécutif était dans une direction congruente, la rotation mentale du caractère était accélérée alors qu'elle était ralentie quand l'effet consécutif était en sens inverse.

2.1.3. Propriétés de la vision partagées par l'imagerie

Un troisième ensemble de résultats expérimentaux suggère que des propriétés de la modalité visuelle se retrouvent dans l'imagerie. Finke et Kosslyn (1980) sont parvenus à estimer la forme et la taille du champ de résolution de l'imagerie : les sujets devaient focaliser en imagination une paire de points variant avec les expériences par leur écartement, déplacer leur "regard" de leur point de focalisation initial et dire jusqu'à quelle distance les points leur apparaissaient séparés. Le champ de résolution obtenu est absolument similaire à celui de la condition perceptive analogue, où les sujets déplacent leur regard sur un champ réel où sont réellement présents les points à discriminer (ces résultats ont été confirmés et étendus ; cf.

Finke & Kurtzman, 1981b). Quoique l'intensité d'un stimulus ne paraisse pas bien préservée dans une image, on a pu montrer (Crovitz, Rosof & Shiffman, 1971) qu'il y avait dans l'imagerie un phénomène analogue au point de fusion critique (fréquence à laquelle une lumière clignotante apparaît continue), sauf que la fréquence au point de fusion perceptive est beaucoup plus élevée. Une preuve plus directe d'une équivalence de l'imagerie et de la perception a été établie par Finke (1979 et 1980). L'idée était de montrer que des erreurs de coordination visuo-motrice suivant une adaptation à une phase de vision déformée pouvaient se produire dans des conditions faisant intervenir l'imagerie. On a montré d'abord que lorsque des sujets doivent pointer du doigt une cible vue à travers un prisme déformant, ils surmontent leurs erreurs initiales, ou du moins deviennent de plus en plus précis. Quand ces mêmes sujets reviennent ensuite à des conditions perceptives normales, ils commettent des erreurs de désignation dans la direction opposée aux erreurs initiales de la condition en vision déformée : si, par exemple, leur doigt pointait trop à droite de la cible dans la condition en vision déformée, il pointe ensuite trop à gauche lors du retour à la vision normale. Si on a renseigné les sujets sur la direction et l'amplitude de leurs erreurs lors de la phase d'adaptation à la vision déformée, les erreurs de coordination de la main adaptée sont transférées à la main non adaptée. Si, lors de la phase d'adaptation, on demande aux sujets de simplement imaginer leurs erreurs ("imagery feedback"), au lieu de les constater de visu ("perceptual feedback"), ils commettront lors de la phase de réadaptation le même genre de fautes de visée que dans la condition perceptive ; en outre, l'effet de transfert à la main non adaptée s'observera aussi.

2.1.4. Travaux neurophysiologiques et neuropsychologiques sur l'image

Une dernière gamme de données expérimentales étayant la similitude de l'imagerie et de la perception visuelle nous vient de travaux de neuropsychologie et de neurophysiologie. Nous reprendrons ici les résultats les plus univoques, exposés en particulier dans les revues de Farah (1988), Paivio (1986, ch.12) et Denis (1989, ch.III-7).

Chez des sujets qui devaient visualiser mentalement une promenade dans un quartier, Roland et Friberg (1985) ont trouvé une augmentation du flux sanguin dans les aires postérieures du cerveau comprenant le lobe occipital (le cortex visuel) et les aires postéro-supérieures du lobe pariétal et inféro-temporales du lobe temporal ; or, certains auteurs considèrent que ces dernières aires correspondent respectivement à la localisation et à l'identification d'un stimulus visuel (cf. Ungerleider & Mishkin, 1982). Utilisant la technique de l'électroencéphalographie, Davidson et Schwartz (1977) ont mis en évidence une atténuation maximale du rythme α dans les aires du cortex visuel lors d'une tâche où les sujets devaient

imaginer des flashes lumineux ; de plus, des conditions de contrôle ont permis de montrer que cette atténuation n'est pas due purement à une augmentation de la charge attentionnelle. La participation des aires visuelles à l'accomplissement d'une tâche d'imagerie a aussi été démontrée par la mesure des potentiels évoqués. Farah, Peronnet, Weisberg et Perrin (1988) ont constaté une augmentation de positivité marquée dans les aires occipitales dans une condition où des sujets avaient à imaginer les référents de mots présentés par écrit. Cette augmentation est spécifique à la tâche d'imagerie après les premières 450 msec de traitement, et différencie celle-ci d'une condition de contrôle où les sujets se contentent de lire les mêmes mots.

La présence de déficits corrélés de l'imagerie et de la vision est également attestée. Ainsi DeRenzi et Spinnler (1967) ont trouvé que des patients incapables d'exécuter certains tests visuels (assortir des couleurs entre elles ; repérer une figure colorée sur un fond d'une autre couleur) se révélaient également incapables de répondre à des questions impliquant la mémoire visuelle (par exemple : "de quelle couleur est une mandarine ?). Dans l'analyse d'un cas désormais fameux, Beauvois et Saillant (1985) ont trouvé des déficiences parallèles de la vision et de l'imagerie des couleurs chez un patient incapable de réussir des tâches "croisant" la modalité visuelle et la modalité verbale : alors que les performances étaient normales sur des tâches impliquant seulement des associations visuelles entre couleurs ou verbales entre noms de couleurs, elles se dégradèrent nettement lorsqu'il s'agissait de répondre verbalement à des questions semblant requérir la médiation de l'imagerie ("quelle est la couleur d'un cornichon ?"), de nommer une couleur donnée ou de montrer une couleur énoncée verbalement. Notons que ce cas enveloppe non seulement un partage des ressources de la vision des couleurs et de l'imagerie correspondante, mais aussi des relations intermodales avec les noms de couleur communes à l'imagerie et à la vision. Bisiach et Luzzati (1978) ont pu établir que des cérébrolésés droits affectés d'une négligence de leur champ visuel gauche omettaient systématiquement de mentionner, dans des descriptions de mémoire d'un lieu familier (la Piazza del Duomo à Milan), les objets situés à gauche de leur champ imaginaire, quel que soit le point de vue adopté (en l'occurrence, selon que les sujets regardaient vers la cathédrale ou depuis celle-ci). Bisiach, Luzzati et Perani (1979), avec le même type de patients et des sujets normaux, ont construit en outre une série de tests révélant que les capacités d'imagerie des sujets souffrant d'héminégligence étaient fortement empêchées : deux dessins figurant des nuages simplifiés sont passés l'un après l'autre derrière une fente qui n'en rend visible qu'une section à la fois ; il s'agit alors de déterminer si les deux formes sont identiques ou non, tâche qui est supposée exiger le maintien en mémoire et l'addition des images mentales de parties successives ; les expérimentateurs constatent que les sujets cérébrolésés jugent difficilement de la similitude des deux dessins successifs. Enfin, Levine, Warach et Farah (1985) ont pu constater que des déficits sélectifs touchant la localisation ou l'identification d'objets visuels se retrouvaient quand ces mêmes

objets devaient être imaginés. Ainsi, un patient se révèle incapable de distinguer les formes d'objets, d'animaux familiers et des visages *ou* de les décrire de mémoire, alors qu'il peut décrire avec précision les situations relatives de lieux sur une carte. Inversement, un patient qui a des difficultés à situer les objets s'avère incapable de localiser de mémoire des endroits de son voisinage, de décrire la position de villes sur une carte ou de montrer du doigt un objet proche si on lui bande les yeux. En revanche, la description d'objets, d'animaux et de visages lui est aisée.

2.1.5. Conclusion

Ces quatre groupes d'expériences concourent à prouver que l'image est bien un phénomène visuel, en ce qu'elle revivifie au moins partiellement un percept donné. Or, ce que l'image présentifie est distinct de son interprétation. En amont, l'interprétation peut être guidée par une consigne verbale, ou par une génération intentionnelle de l'image ; en aval elle est déterminée par les processus d'interprétation de propriétés perceptives émergeant de la "lecture" de l'image. En tout cas, l'image se distingue de la perception par sa liaison en amont à un composant verbal, et, en aval, par son usage cognitif, qui peut lui aussi impliquer ce composant verbal : entre le processus de génération de l'image et le stade d'inspection de celle-ci¹, l'image aurait essentiellement pour fonction de fournir des informations figuratives non disponibles automatiquement. Certaines expériences montrent que le recours à l'imagerie a peu de chances de se produire quand les informations à exploiter, même si elles sont figuratives, ont fait l'objet d'un

¹ Farah (1984) donne de bonnes raisons de croire que les stades de la génération et de l'inspection de l'image sont effectivement distincts. Ainsi, à l'appui de l'existence d'un déficit sélectif de la génération, on constate que certains patients (des cérébrolésés postérieurs gauches) peuvent reconnaître ou décrire des objets présentés visuellement mais ne peuvent pas les décrire de mémoire, les dessiner à l'énoncé de leur nom ou détecter par l'introspection toute activité d'imagerie (ibid. : 256-259). La mise en évidence d'un déficit spécifique de l'inspection (ou lecture) de l'image est plus délicate puisqu'un tel déficit implique à l'évidence une dégradation très étendue. Néanmoins, si l'on fait appel à l'explication la plus parcimonieuse, la détérioration du processus d'inspection suffirait à rendre compte des patients qui ne peuvent ni décrire ni dessiner des objets évoqués de mémoire ou présentés visuellement (ibid. : 264).

surapprentissage (cf. Kosslyn, 1980, pp. 351s ; Denis, 1989, pp. 118s).

A notre avis, l'image est bien une représentation analogique, c'est-à-dire une représentation dont les propriétés font qu'un processus opérant sur l'image est analogique. En tant que représentation analogique, l'image est essentiellement liée au percept, et dorénavant nous entendrons par "représentation analogique" toute représentation dont les propriétés physiques dérivent de celles d'un percept.

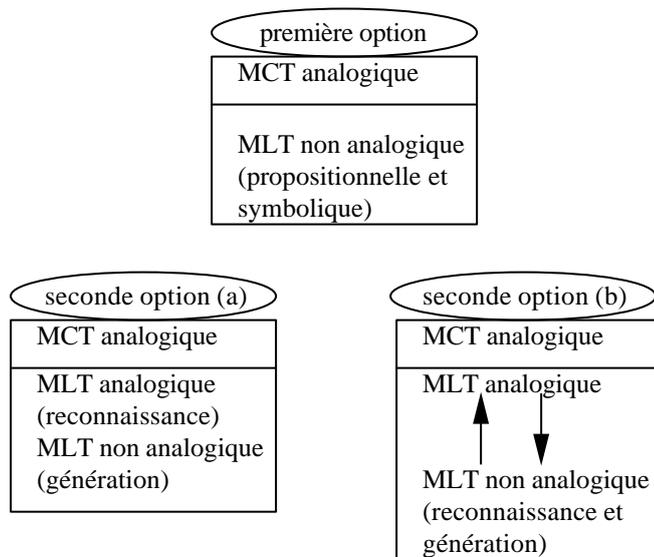
Dans les paragraphes qui suivent, nous examinerons les rapports entre les représentations linguistiques et l'image, en nous attachant au mode de composition visuo-linguistique des parties de l'image, particulièrement au stade de la génération de celle-ci. Mais auparavant, il nous faut examiner un dernier point touchant un éventuel codage propositionnel de l'image : même si les représentations imagées à court terme et actualisées dans une tâche spécifique sont analogiques (non propositionnelles), il demeure possible qu'elles soient issues de représentations profondes codées dans un format amodal, ou bien qu'une mémoire visuelle (modale, analogique) à long terme coexiste avec une mémoire amodale à long terme.

2.2. Représentations propositionnelles profondes et mémoire visuelle

Les propriétés assimilant image et percept pourraient caractériser l'image en tant que représentation de travail, alors que les représentations à long terme coïncideraient au moins en partie avec des descriptions propositionnelles, et les représentations profondes décrivant l'apparence des objets seraient non analogiques¹ ; comme alternative, des propriétés modales analogiques pourraient aussi caractériser la représentation visuelle à long terme, sans exclure que cette mémoire visuelle puisse coexister avec des représentations

¹ Les représentations non analogiques sont soit les représentations propositionnelles (du type symbole linguistique), soit les représentations symboliques (par exemple coordonnées polaires, transformations de Fourier) utilisées dans un modèle pour décrire les apparences d'un objet ou de ses parties.

amodales. Du reste, il n'est pas toujours possible de savoir laquelle de ces deux options est préférée par tel ou tel auteur : si Denis ne se prononce pas sur la présence d'une mémoire analogique à long terme, il paraît restreindre l'activité d'imagerie à l'actualisation de traits non spécifiques (en effet, ces derniers sont censés appartenir à un système sémantique comprenant aussi les traits non figuratifs reflétant la connaissance du monde qu'a le sujet ; cf. 1989, p. 136)¹ En tout cas, l'argument essentiel des imagistes intégrant un composant propositionnel dans leurs théories est que le traitement des propriétés figuratives d'un concept peut *aussi* s'effectuer sans le secours de l'imagerie. De là, ils supposent qu'un même code amodal sous-tend les représentations visuelles et linguistiques des propriétés figuratives. Les différentes options théoriques peuvent être figurées par le diagramme suivant :



¹ C'est, cependant, seulement dans son ouvrage précédent que Denis (1979) identifie nettement ce système sémantique commun à une mémoire propositionnelle : "On peut admettre en parlant ainsi des activités d'imagerie, que l'on parle de processus spécifiques opérant sur des structures de représentation par elles-mêmes amodales et descriptibles en termes propositionnels." Dans l'ouvrage de 1989, il semble moins catégorique mais retient tout de même le qualificatif de "propositionnel" pour caractériser la nature des composants du système unique (cf. p.47).

Plaçons-nous dans l'hypothèse de deux mémoires telle qu'elle est prônée par la seconde option théorique. Soit la mémoire analogique est utilisée pour la reconnaissance visuelle (d'objets externes ou imaginés) et la mémoire propositionnelle pour la génération d'images, soit les deux mémoires sont utilisées conjointement dans la reconnaissance visuelle et dans la génération d'images. La première version de l'option "mixte" n'est absolument pas orthodoxe. Dans une perspective neuropsychologique, elle ne peut s'appuyer que sur des sujets chez qui on observe une dissociation entre reconnaissance et génération d'images. D'autre part, la présence de traits figuratifs amodaux devrait, chez ces mêmes sujets, être corroborée par un déficit conjoint de l'accès verbal et de l'accès "visuel" (c'est-à-dire non verbal, comme lorsque le patient doit dessiner un objet de mémoire) à une même représentation figurative. Ce type de dissociations s'observe effectivement chez le patient HJA, étudié par Humphreys et Riddoch (1987 ; Ellis & Young, 1988, ch.2). HJA peut copier des dessins qu'il ne reconnaît que très difficilement ou pas du tout mais s'avère capable de dessiner correctement de mémoire les objets qu'il ne parvient pas à identifier normalement ; en outre, il peut définir et *décrire verbalement* ces mêmes objets. Cependant, si cette théorie est économique pour expliquer les performances de HJA, elle est au contraire plus coûteuse lorsqu'il s'agit de rendre compte des déficits corrélés de la reconnaissance et de la mémoire visuelle que nous avons énumérés plus haut. De plus, Farah (1984) a montré qu'il était possible de rendre compte d'un large ensemble de déficits de l'imagerie en faisant l'hypothèse d'une mémoire à long terme commune pour la génération et l'inspection des images visuelles. Enfin, Pinker a suggéré (1984) que, puisque les images, comme les représentations à long terme d'objets perceptifs, semblaient être centrées sur l'observateur, il était plus économique que les processus de reconnaissance et ceux transformant les images visuelles partageassent les mêmes représentations.

Si, maintenant, nous adoptons l'optique de l'utilisation conjointe des deux mémoires dans la génération de l'image, dans l'ignorance relative où nous sommes de la décomposition précise des processus de génération, nous risquons d'être rejetés à la première option, celle posant des descriptions non analogiques à long terme

et des représentations analogiques à court terme. Nous serions en effet incapables de distinguer l'option "représentations analogiques à long terme + représentations non analogiques à long terme + processus de génération" et l'option "représentations non analogiques à long terme + processus de traduction du non analogique à l'analogique + représentations analogiques à court terme + processus de génération". Dans notre esprit, la théorie de Kosslyn correspond à la première option ; la théorie de Marr pourrait correspondre à la deuxième option, si elle était augmentée de processus analogiques : elle inclut en effet un niveau où les formes des objets sont représentées en tant que formes et non en tant que coordonnées symboliques. D'autres auteurs sont clairement favorables à des représentations analogiques à long terme (par exemple qui préservent les distances dans l'espace à trois dimensions entre des objets vus de l'observateur ; cf. Pinker, 1984b : 53).

La seule manière envisageable d'évaluer les mérites respectifs des options 1 et 2b serait de mettre en évidence une dissociation des processus non analogiques et des processus analogiques intervenant dans la génération de l'image. D'après Kosslyn (1987), le codage des relations topologiques et de l'attachement des parties d'un objet imagé (que nous désignerons dans la suite par le nom de "système mérologique") serait par excellence l'œuvre de représentations propositionnelles. En effet, les relations topologiques et d'attachement des parties ne sont codables que par des symboles suffisamment généraux pour être insensibles à la variété extrême des orientations de ces parties dans la réalité¹. Comme ces relations catégorielles sont proches des représentations linguistiques, elles doivent impliquer des régions voisines du cerveau et donc se trouver plutôt représentées dans l'hémisphère gauche. Comme, par ailleurs, les apparences des parties des objets sont codées par des listes de coordonnées polaires, il s'ensuit qu'il n'y a pas place dans cette théorie pour des représentations non analogiques à long terme.

¹ Dans ses articles récents, ainsi que dans l'ouvrage de 1992 écrit en collaboration avec Koenig, Kosslyn passe sous silence l'éventualité d'un codage propositionnel des relations topologiques des parties.

A l'appui de la thèse des représentations mérologiques de nature propositionnelle, on peut constater une tendance des déficits spécifiques de la génération, voire aussi de la reconnaissance, d'objets imagés à se situer dans l'hémisphère gauche, alors que les déficits de l'orientation spatiale seraient plutôt localisés dans l'hémisphère droit (voir le panorama de Farah, 1984) ; or, certaines perturbations de l'imagerie paraissent effectivement ressortir à un défaut d'assemblage des parties et à une désorganisation de l'accès à ces parties. Ainsi, le patient Ch. T. de Deleval, De Mol et Noterman (1983), qui souffre d'une lésion occipitale gauche, déclare-t-il :

"Si on me demande d'imaginer la tête d'une vache, je sais qu'elle a des oreilles et des cornes, mais je ne peux revoir leurs places respectives. De même je ne peux déterminer le nombre de doigts d'une patte de grenouille alors que je manipule tous les jours cet animal au laboratoire. Si je veux me représenter un objet aussi commun qu'une clef, alors que je sais parfaitement à quoi elle sert, je suis incapable d'en avoir une vue mentale complète." (Deleval, De Mol & Noterman, 1983, p. 71).

A l'appui d'une dissociation du système mérologique et du système de coordonnées métriques, Kosslyn (1987) s'autorise des résultats de deux expériences montrant un traitement différent d'une même tâche par l'un ou l'autre hémisphère. Dans la première expérience (loc. cit., p. 624), il s'agit, pour un sujet dont le corps calleux a été sectionné (par conséquent, dont les deux hémisphères peuvent être considérés comme dissociés), de juger si une lettre capitale se compose seulement de traits droits ou possède des traits courbes. On ne fournit au sujet visuellement que la minuscule correspondante, de sorte qu'il doit accéder à une image mentale de la majuscule pour pouvoir juger. L'hypothèse est que le sujet forme une image visuelle de la lettre majuscule en joignant les parties qui la constituent. Si l'hémisphère gauche est bien responsable de l'attachement des parties d'un objet complexe, on attend que ses performances soient en l'occurrence supérieures à celles de l'hémisphère droit. On constate que tel est le cas. Plusieurs situations de contrôle étaient employées pour éliminer la possibilité qu'un autre système de traitement soit responsable des résultats

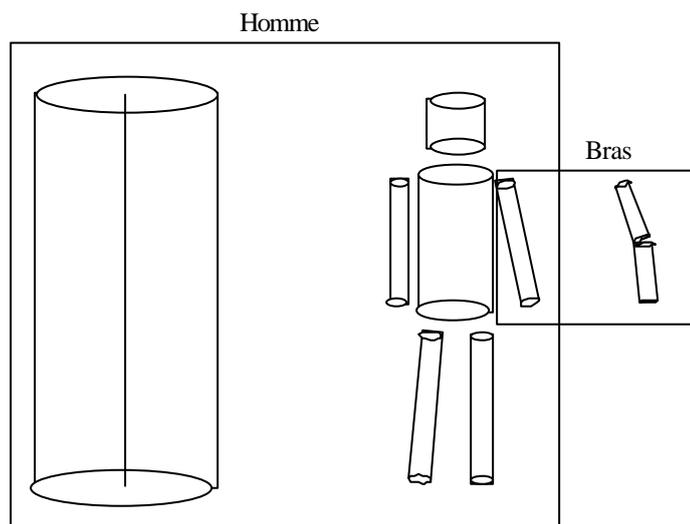
différents obtenus par les deux hémisphères : — on latéralise la présentation visuelle d'une majuscule et on pose au sujet la même question que précédemment ; on constate que les résultats sont presque parfaits pour les *deux* hémisphères ; le sujet peut donc encoder, inspecter, évaluer et répondre — on latéralise la présentation de minuscules et on demande au sujet d'apparier celles-ci avec des majuscules sur un tableau voisin ; on constate que le sujet convertit correctement la minuscule en la majuscule correspondante (stade d'accès) — on latéralise des mots de trois lettres, puis, deux secondes plus tard, on donne au sujet la position de la lettre à classer, toujours suivant le même critère. Les deux hémisphères obtiennent de bons résultats et peuvent donc tous deux retenir et inspecter la représentation d'une lettre capitale. Par élimination, on conclut que seul le stade de la génération de l'image était à incriminer dans le traitement différentiel de la tâche par l'un ou l'autre hémisphère.

Dans la seconde expérience (ibid. : 631 ; voir aussi Kosslyn et al., 1989 et Hellige & Michimata, 1989), il s'agit de comparer les performances respectives des deux hémisphères dans deux types de tâches : dans l'une le sujet doit juger si un point est contenu ou non dans une figure irrégulière ; dans l'autre, le sujet doit déterminer si le point est proche ou loin de cette même figure. Kosslyn présume que la première tâche exige le recours au système mérologique-topologique, tandis que la seconde implique l'utilisation du système de coordonnées. Par conséquent, on s'attend à ce que les relations d'inclusion/exclusion soient mieux traitées par l'hémisphère gauche, alors que les rapports de proximité/éloignement seraient préférentiellement traités par l'hémisphère droit. On constate que les résultats confirment ces attentes. D'autres expériences basées sur les mêmes principes ont depuis lors corroboré ces résultats (cf. Hellige & Michimata, 1989 ; Kosslyn, Koenig, Barrett, Backer Cave, Tang & Gabrieli, 1989).

Il est notable que ces expériences, aussi bien que le cas neuropathologique cité, ne permettent pas d'induire quoi que ce soit sur le format supposé des représentations mérologiques. L'attribution d'un format propositionnel à de telles représentations est seulement basée sur l'hypothèse (1) que des relations très générales liant des parties font nécessairement appel à un code

symbolique de type propositionnel —(2) que la proximité du propositionnel et du linguistique entraîne le partage par ces deux codes de capacités communes —(3) que ces capacités communes rapprochent les représentations mérologiques des modules linguistiques du cerveau, dont on sait qu'ils sont généralement localisés à gauche (ce que Kosslyn appelle "l'effet boule de neige").

Un problème essentiel est donc de savoir sous quelle forme sont stockées les informations visuelles exploitées par les processus de génération de l'image. Il est souvent admis qu'une représentation permettant d'imaginer ou de reconnaître un objet sous des orientations diverses doit à la fois être centrée sur l'objet (plutôt que relative à un point de vue) et préserver les propriétés tri-dimensionnelles de cet objet (lesquelles sont effectivement utilisées par le sujet pour changer de point de vue sur un objet : cf. Pinker, 1980 et Pinker & Finke, 1980). Selon Marr (1982), un modèle-3D centré sur un objet est formé par un système hiérarchique de cônes dont les axes sont spécifiés à chaque niveau par leur point d'attache au cône de niveau supérieur. Les cônes de premier niveau sont issus des axes principaux d'élongation de l'objet, les positions des cônes de niveau immédiatement inférieur sont placés par rapport à ces axes et ainsi de suite, jusqu'au degré de détail souhaité. Le diagramme ci-dessous illustre trois niveaux de spécification pour la forme de l'être humain :



D'après Marr, 1982, p. 306.

Si une telle représentation peut être un moyen de reconnaître ou d'engendrer des formes complexes, il n'est nullement évident qu'elle soit l'unique recours. Utilisant leur technique des vues inhabituelles présentées à des cérébrolésés droits et à des individus normaux, Warrington et James (1986) ont constaté que le taux de reconnaissance des deux groupes ne dépendaient pas du raccourcissement éventuel de l'axe d'élongation de l'objet. Leur thèse est alors que la reconnaissance se base sur des faisceaux de traits perceptifs (des contours visuels) codés par des représentations centrées sur l'objet, et sur la représentation des positions relatives de ces traits. La position de ces faisceaux de traits sur la rétine permet d'estimer l'orientation de l'objet. Les cérébrolésés se distingueraient des sujets normaux en ce qu'un plus grand nombre de traits leur serait nécessaire pour identifier un objet, ce qui expliquerait que leurs performances ne se différencient des sujets normaux que quantitativement (c'est-à-dire seulement par une élévation des seuils de reconnaissance des mêmes objets) :

"Dans le contexte de la reconnaissance d'objet, nous suggérerions que des traits peuvent être conceptualisés comme des faisceaux relativement *uniques* de contours visuels (par exemple courbure, angle) ou comme des ensembles de faisceaux de ce genre assortis de leurs positions relatives, indépendamment de leur orientation absolue dans l'espace. Ainsi, des traits distinctifs peuvent spécifier l'identité structurelle d'un objet, et ce de manière directe, et en même temps, leurs positions rétinienne peuvent être utilisées pour spécifier leurs orientations dans un espace centré sur l'objet." (1986, p. 365).

Chez Warrington et James, contrairement à Marr, la catégorisation des formes de parties est en puissance effectuable par un autre système que le calcul des relations métriques fines entre ces formes. Il semble bien qu'il faille effectivement admettre plusieurs routes possibles de la reconnaissance d'objet ; Humphreys et Riddoch ont trouvé une dissociation, chez des cérébrolésés droits et le patient HJA, entre une condition où des photographies d'objets doivent être appariées en dépit du

raccourcissement d'un axe, et une condition où l'appariement doit s'effectuer malgré la faible saillance d'un trait distinctif ; de là, ils ont conclu à l'existence d'une double route qu'emprunterait la reconnaissance visuelle (cf. Humphreys & Riddoch, 1984 et 1985). Chez les sujets normaux, la route la plus rapide dans des circonstances particulières l'emporterait (Humphreys & Bruce, 1989, pp. 78-82).

Ces thèses sont compatibles avec l'idée qu'il y a bien au moins deux systèmes distincts traitant les traits, interprétés comme des parties saillantes, et l'orientation, et que ces systèmes peuvent fonctionner de manière dissociée ; elles sont aussi compatibles avec l'hypothèse, avancée par Biederman, que la reconnaissance d'un objet visuel dépend de l'identification de parties élémentaires et que cette identification est peu sensible aux variations métriques, à l'orientation et à la dégradation intra-partie¹ (1987, p. 118). Chez Biederman et chez Warrington, donc, l'identification d'un objet au stade apperceptif ne nécessite pas un codage propositionnel, en dépit de la relative indépendance du point de vue qui doit être propre à la représentation élaborée lors de la reconnaissance. En revanche, le format dans lequel sont couchées ces relations d'attachement est, chez Kosslyn, propositionnel, parce que ces parties entrent dans des relations d'attachement qui sont générales (i.e. insensibles aux variations possibles).

Les rapports entre le composant linguistique et l'image seront examinés dans la section suivante, et nous reviendrons sur ce problème d'une imagerie abstraite, c'est-à-dire éloignée d'une représentation définie centrée sur l'observateur.

2.3. Composition et complexité de l'image

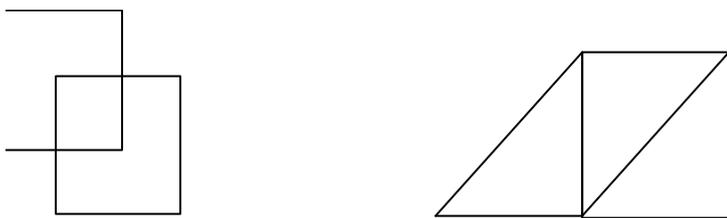
Le problème de la composition d'une image complexe en parties plus simples est étroitement lié à celui des relations entre le langage

¹ Biederman arrive à cette conclusion après avoir constaté que le brouillage de l'identification des parties élémentaires (l'effacement de leurs sommets) dégrade nettement plus les performances de sujets que la destruction d'un contour visuel au sein d'une même partie ; enfin, il semble qu'à nombre de composants égal, le point de vue n'ait pas d'influence sur les performances (ibid. : 135 et 141).

et l'imagerie. En effet, il importe de savoir si la composition de l'image obéit, au moins partiellement, à une conceptualisation linguistique et si le composant linguistique intervient au stade de la génération ou de l'inspection de l'image. Or, sur les effets dus à la composition par parties de l'image et le stade auquel ces effets se manifestent, les auteurs qui ont abordé la question s'opposent. Alors que Kosslyn pense que la complexité de l'image fait varier son temps de génération, Paivio (notamment 1986, p. 130) est d'avis que ce paramètre n'a d'influence qu'à l'étape de la réponse. La controverse sur ce dernier point pourrait se résumer ainsi : l'image est-elle engendrée de manière globale et simultanée (comme une "entité holistique", pour reprendre les termes de Paivio, *ibid.*) ou bien est-elle produite par un rattachement de composants effectué en partie séquentiellement ? Si l'on se fie aux travaux expérimentaux pertinents, force est de constater que l'opinion de Kosslyn est mieux soutenue. Son hypothèse d'un effet de la complexité au stade de la génération est testable, selon lui, de la manière suivante : si une image est disponible en totalité et simultanément au stade de la génération, engendrer une image ne signifie rien de plus que lire une scène mentale déjà prête. Par conséquent, les effets dus à la taille et à la complexité de l'image au stade de la génération devraient être similaires à ceux observés au stade de l'inspection. Or, on constate que la taille de l'image n'affecte pas son temps de génération et que l'augmentation de ce temps est principalement dépendante de la quantité de détails que les sujets y ont placés. Si l'image était globalement disponible au stade de la génération, niveau de détail et taille interagiraient en sorte que la complexité d'une image allongerait d'autant plus les temps de génération que celle-ci est produite à une taille subjective plus grande (cf. Kosslyn, 1980, pp. 97 et 107-111). Dans la seule expérience pertinente que cite Paivio pour étayer sa thèse (Hoffmann et al., 1983), la complexité de l'image est manipulée seulement indirectement par le nombre de traits figuratifs attribués aux concepts ; les sujets ne sont donc pas invités à contrôler le niveau de complexité de leur image elle-même. Les trois critères que Paivio considère essentiels pour établir que l'image est une structure "holistique" (la mémoire des relations spatiales, l'absence de contraintes séquentielles pour la récupération de l'information, les

effets d'intégration et de décomposition d'une information intégrée ; cf. Paivio, 1986, p. 165) concernent non le stade de la génération de l'image mais le stade de l'inspection ; au stade de l'inspection, Paivio, n'exige pas que l'image soit lue simultanément. Il note d'ailleurs, à notre avis correctement, que la lecture de l'image est parallèle non au sens où les informations y sont traitées simultanément mais au sens où l'ordre de traitement "n'est pas séquentiellement contraint par la structure de représentation elle-même" (ibid., p. 60). De même, lorsqu'on dit que l'image est une structure parallèle, il faut entendre qu'elle est engendable à partir de n'importe quelle partie. Nous nous pencherons maintenant sur l'influence des composants linguistiques et perceptifs dans la composition de l'image.

L'étude de l'influence du composant linguistique sur la composition de l'image a été entreprise surtout par Kosslyn. Il s'agit pour lui de montrer comment différentes descriptions verbales d'une image affectent le temps de génération de celle-ci. Dans une de ces expériences, les sujets lisent une des descriptions possibles de formes en mosaïque, comme ci-après :



D'après Kosslyn, 1980 : 100.

Les descriptions désignent un nombre plus ou moins grand de parties, selon que les figures sont décomposées en carreaux adjacents ou en formes transparentes superposables. On constate qu'en général, les sujets prennent d'autant plus de temps à imaginer les figures que la description comporte plus d'éléments. Lorsqu'un motif perceptif est donné *avant* la description verbale que l'on prescrit aux sujets de suivre dans la tâche subséquente d'imagerie, la décomposition verbale du motif influence malgré tout le temps de génération de l'image de ce motif (Kosslyn, ibid. : 107). Dans une

autre expérience, des motifs composés d'une même lettre (X ou Y), de X et Y arrangés en lignes et en colonnes divisées ou non en deux groupes (définis par un espace passant au milieu du motif) doivent être mentalement modifiés d'après des instructions verbales. Lorsque ces prescriptions font intervenir un plus grand nombre d'éléments, les latences d'imagerie sont allongées. De plus, "le temps additionnel requis pour chaque ligne ou colonne décrites s'ajoutait au (était indépendant du) temps supplémentaire dû au nombre d'unités d'ordre supérieur [i.e. les groupes partageant éventuellement les motifs] mentionnées dans la description." (ibid.).

Ce dernier résultat suggère clairement que la génération de l'image procède par unités organisant hiérarchiquement les motifs.

Un problème se pose, cependant, quant à l'interprétation de ces résultats. Supposons que l'on appelle "visuo-linguistiques" les termes désignant des "bonnes formes" visuelles. Il se pourrait alors que l'on ait mis en évidence non pas tant l'influence de la description verbale *per se* sur la génération d'une image, que l'influence mixte des termes visuo-linguistiques. Dans ce cas, des résultats similaires à ceux de ces expériences ne pourraient être obtenus que si les termes utilisés réfèrent à de "bonnes formes" visuelles ; ainsi, on peut considérer que les termes de "ligne", "colonne" et "groupe", dans les expériences ci-dessus, exploitent les lois de bonne continuité et de proximité de la *Gestalttheorie*. L'importance de cette segmentation visuo-linguistique serait d'autant plus grande que les images d'objets classifiables en catégories hiérarchiques se différencieraient par leurs parties. Deux genres de données nous inclinent à une telle conclusion : d'une part, le fait que "former l'image d'un objet correspondant au niveau de base d'une hiérarchie est apparemment une activité cognitive privilégiée", parce que la formation d'une image prend le moins de temps pour des mots correspondant à des exemplaires du niveau de base (Denis, 1989 : 148 ; Hoffmann, Denis & Ziessler, 1983). D'autre part, on a pu constater que, dans des expériences qui estimaient la prégnance relative des parties par rapport aux autres attributs d'un objet, les parties constituaient la majorité des attributs spontanément énumérés pour des objets du niveau de base, différenciaient davantage les membres au niveau de base d'une même catégorie,

tout en étant corrélées au même degré que les autres attributs avec la typicalité (les autres attributs n'étant donc pas de meilleurs prédicteurs de la prototypicalité des objets qui les possèdent), tandis que les objets du niveau subordonné tendaient à partager les mêmes parties mais à se différencier par les autres attributs (Tversky & Hemenway, 1984 : études 1, 2 et 3). Comme les expériences menées par Tversky et Hemenway utilisaient un matériel verbal pour décrire les parties aussi bien que les autres attributs, on ne peut en déduire qu'un effet majeur de la segmentation visuo-linguistique sur la composition et la différenciation des objets appartenant au niveau de base.

Il nous semble, toutefois, que les expériences où deux descriptions d'une *même* forme entraînent des latences d'imagerie inégales mettent au moins en évidence une influence du linguistique sur le choix du découpage visuel d'une forme donnée. D'autre part, il est possible de contraindre des sujets à un découpage d'une image assez arbitraire et d'observer, malgré le peu de naturel des parties ainsi distinguées, le même type d'effets que précédemment (cf. Kosslyn, *ibid.* : 102). De plus, ces effets ne sont pas corrélés aux relations structurelles faisant dépendre la génération d'une partie de la génération préalable d'une autre, montrant ainsi la plasticité de l'imagerie (*ibid.* : 179).

Il est clair que l'imagerie est affectée par les qualités de forme et qu'elle doit comprendre des principes de la *Gestalttheorie*. Aux stades de l'inspection et de la transformation, on sait que des sujets reconnaissent plus facilement une partie d'une figure mémorisée quand elle est une bonne forme (Reed, 1974) ou identifient plus aisément une forme synthétisée à partir de "bonnes" sous-parties (Palmer, 1977) ; on sait aussi que le temps mis pour effectuer la rotation d'une partie, dont on doit décider de l'appartenance à une figure, dépend de la qualité (au sens de la *Gestalttheorie*) de cette partie (Pylyshyn, 1979a). De tout cela, on peut raisonnablement conclure que l'imagerie est affectée par le nombre de sous-unités impliquées dans une description verbale, que l'effet du nombre d'unités n'est pas dû au recoupement linguistique de la "bonne" analyse visuelle (même si cette correspondance visuo-linguistique

favorise le traitement), enfin, que la synthèse et la reconnaissance de sous-unités dépend de leurs qualités de forme.

Plus généralement, il y a plusieurs manières de concevoir le rapport du linguistique et de l'analyse visuelle, selon que l'on pose une correspondance bijective, ou un recoupement inexact entre les formes visuelles et leurs termes descriptifs. D'une part, il paraît clair que la correspondance est parfois inexistante, puisqu'il y a des configurations d'objets imagés qui ne sont pas exactement descriptibles par le langage : songeons à des objets formant une constellation complexe, par exemple. D'autre part, là où existe effectivement une correspondance, les termes linguistiques étant généraux, puisque le mot "rectangle" désigne un nombre infini de formes dont les rapports entre les côtés sont différents, il semble également que la correspondance ne puisse être bijective. En réalité, il est assez arbitraire de distinguer le cas où la correspondance n'existe pas et le cas où elle n'est pas bijective : ainsi, la relation "à côté de" subsume-t-elle les relations "à gauche de" et "à droite de", ou bien n'est-elle pas imageable ? Dans l'optique d'une théorie propositionnelle du sens, une relation spatiale générale est représentable par une image annotée par une structure propositionnelle indiquant ce pour quoi l'image doit être prise, son "interprétation"¹. La relation pourrait aussi être représentée par une imagerie abstraite du type de celle conceptualisant, d'après Langacker ou Lakoff, les scènes visuelles de façon à les conformer aux pouvoirs représentationnels d'une langue (Langacker, 1987 :

¹ Ainsi, Jackendoff (1987a et 1987b) suppose que l'interprétation d'une image, qu'il assimile au modèle-3D de Marr, est fournie par son annotation propositionnelle ; simultanément, il énonce le principe que "si le modèle-3D doit être le composant du système visuel qui est le plus directement responsable de notre capacité à parler de ce que nous voyons, il doit être suffisamment riche en pouvoir d'expression pour fournir toutes les distinctions visuelles que nous pouvons exprimer par le langage." (1987a : 105) Cette position paraît incohérente : si l'image a par elle-même un pouvoir d'expression susceptible de recouvrir toutes les distinctions linguistiques-spatiales, on ne voit plus pourquoi il faudrait systématiquement lui adjoindre une annotation propositionnelle.

133s ; Langacker & Casad, 1985 ; Lakoff, 1987)¹. Enfin, les termes descriptifs non imageables pourraient être inéliminables, c'est-à-dire constitueraient un mode de conceptualisation sans contrepartie dans la modalité visuelle. Ce problème n'est soluble que par une étude approfondie des relations entre la langue, la vision et l'imagerie. Nous avons vu dans la section précédente qu'il ne se limitait d'ailleurs pas au stade de l'interprétation réciproque d'un code par l'autre, mais s'étendait à la question du rôle du non analogique dans la génération de l'image. Néanmoins, nous pouvons indiquer brièvement quelques difficultés que rencontrera la première option théorique.

Tout d'abord, la liaison de l'image au composant propositionnel peut s'effectuer à plusieurs niveaux : chez Kosslyn, la composition de l'image est exprimée par des symboles propositionnels désignant les parties. Chez Denis, les traits figuratifs supposés codables propositionnellement sont également d'un niveau élevé (par exemple "a quatre pattes", "couvert de poils", "possède une couleur caractéristique" etc. ; cf. Denis, 1984). Mais chez d'autres auteurs (par exemple Miller & Johnson-Laird, 1976), le codage propositionnel recoupe les éléments de l'analyse perceptive à un stade *antérieur* à celui où la scène est globalement perçue par le sujet ; ainsi, un mouvement complexe est composé perceptivement de mouvements élémentaires, chaque mouvement élémentaire est identifié par un terme propositionnel, et la formule du mouvement complexe en fonction de ces mouvements élémentaires définit les conditions où un terme de la langue peut désigner ce mouvement (ibid. : 81-84). Alors que chez Denis et Kosslyn le propositionnel est exprimable directement par des mots de la langue, chez Miller et Johnson-Laird il *conditionne* et recoupe l'application du langage. De

¹ L'idée d'une imagerie abstraite est-elle foncièrement anti-intuitive? Les recherches sur l'imagerie étudient essentiellement des images précises, définies et souvent surprises. A l'opposé, certaines images ont un caractère d'indistinction, qui transparait par exemple dans le compte rendu suivant : "Prolétariat : j'avais une image étrange, une étendue plate et noire, et, au-dessous, une mer roulant obscurément, un flot indéterminé, quelque chose comme une masse sombre et épaisse roulant de lourdes vagues. Que signifiait la masse? L'extension dans le monde entier, quelque chose comme un dynamisme latent." (in Flach, 1925 ; cité par Sartre, 1986 : 192-3)

trois choses l'une : soit l'expression des traits figuratifs en langage propositionnel se laisse guider par la langue, et ainsi "long et mince" font deux traits et non pas un ; soit les traits propositionnels figuratifs découlent d'une analyse des traits effectivement utilisés dans le traitement visuel ; soit les traits propositionnels appartiennent à un code amodal à part dans lequel sont retraduites les distinctions visuelles et linguistiques. La thèse implicite de Denis et Kosslyn est qu'à un certain niveau, traits perceptifs et traits propositionnels se recouvrent. Si les traits propositionnels découlent de l'analyse perceptive, ils sont probablement distincts des traits linguistiques, et le code propositionnel n'est pas structuré comme la langue. Par exemple, dans un langage propositionnel dont le vocabulaire est dérivé de l'analyse visuelle, la relation "à l'intérieur de" correspondrait au terme marqué dans une langue, parce que la relation converse est plus facilement détectable parmi des distracteurs, et "saute aux yeux" plus aisément (cf. Treisman & Gormican, 1988 : exp. 12). Or, dans nos langues, il n'existe aucune raison de supposer que la relation "hors de" est privilégiée. Bien entendu, on peut alors poser que le code propositionnel recoupe à la fois les traits perceptivement pertinents et les traits linguistiques et s'en remettre à la troisième solution. Faute de préciser le niveau auquel le codage propositionnel intervient, la solution de facilité consiste à penser que le code propositionnel exprime toutes les distinctions présentes dans les modalités ; telle est l'hypothèse du langage de la pensée comme sens commun (*sensus communis*), chez Fodor (1975), ou des structures conceptuelles chez Jackendoff (1983).

Conclusion

Au terme de ces réflexions, la question du code amodal propositionnel ne peut nullement être considérée comme résolue, même dans le cadre du thème présent. Au stade de la génération de l'image, l'évaluation des facteurs modifiant la complexité et la composition de l'image a conduit, nous l'avons vu, à regarder celle-ci comme un phénomène à la fois quasi perceptif et influençable par le langage. Par là a été réintroduite la question du code amodal.

Notons que la notion même d'amodalité demeure mal définie : les niveaux de traitement où interviendrait le code amodal sont à l'évidence hétérogènes. Il est, en effet, foncièrement différent de décomposer les parties des représentations visuelles propositionnellement, comme, par exemple, chez Palmer (1975), ou de faire de la proposition la description profonde du sens de l'énoncé, comme chez Fodor (1975) ou Kintsch (1974 et 1988). Si le code amodal est constitué de propositions et de leurs termes, est-il structuré comme un langage ou bien n'est-il qu'un catalogue de traits attribués à des objets ? Doit-on s'aider des prédicats linguistiques pour identifier ces traits, ou bien aussi des traits présents dans l'analyse perceptive ? Il est possible qu'un code propositionnel obéisse à cette double contrainte, soit influencé à la fois par la forme de surface de phrases décrivant une scène perceptive, et par des propriétés intrinsèques de l'espace visuel de l'homme (cf. par exemple, Clark, Carpenter & Just : 1973). Il est pourtant difficile de déterminer si cette double contrainte est effectivement acceptable, sauf à admettre une certaine transparence de la langue à l'égard des contraintes cognitives, dont le code propositionnel pourrait hériter. Certains s'autorisent cette audace (en particulier Jackendoff, 1983).

Il faut garder à l'esprit que des représentations amodales pourraient être issues de représentations modales et en être abstraites, jusqu'à avoir des propriétés absolument anti-intuitives déterminées par le médium de représentation lui-même (nous songeons ici à une représentation connexionniste)¹. L'image elle-même peut être considérée comme étant à un degré d'abstraction plus bas, éliminant seulement des caractéristiques du percept qui appartiennent à un bas niveau de traitement (cf. Hebb, 1968 et Finke, 1980 et 1985).

Quant au thème des liens de l'image et du propositionnel, il nous semble que la discussion devrait désormais s'orienter sur les points

¹ Il est implicitement compris, dans la philosophie de la psychologie cognitive, que les représentations qui sont porteuses de signification sont liées entre elles d'abord en vertu de leur contenu. La possibilité d'un déterminisme cérébral de la forme des représentations qui conditionnerait certaines propriétés de leur contenu est proche du paradigme connexionniste.

suivants : — une étude plus avant des facteurs influençant la complexité cognitive de l'image — la possibilité de représentations amodales abstraites des représentations modales, enfin, et plus largement, — une évaluation globale des théories du code amodal, en particulier parce que les interactions entre le langage et les représentations modales non verbales sont souvent supposées s'effectuer par un code amodal, lequel est même parfois identifié au support ultime du sens.

Jean-Michel FORTIS
84, rue Damrémont
75018 Paris

Bibliographie

- ANDERSON J.R. (1978) Arguments concerning representations for mental imagery, *Psychological Review*, 85, pp. 249-277.
- BEAUVOIS M.F. & SAILLANT B. (1985) Optic aphasia for colours and colour agnosia : a distinction between visual and visuo-verbal impairments in the processing of colours, *Cognitive Neuropsychology*, 2(1), pp. 1-48.
- BISIACH E. & LUZZATI C. (1978) Unilateral neglect of representational space, *Cortex*, 14, pp. 129-133.
- BISIACH E., LUZZATI C. & PERANI D. (1979) Unilateral neglect, representational schema and consciousness, *Brain*, 102, pp. 609-618.
- BROOKS L.R. (1967) The suppression of visualization by reading, *Quarterly Journal of Experimental Psychology*, 19, pp. 289-299.
- CHAMBERS D. & REISBERG D. (1985) Can mental images be ambiguous?, *Journal of Experimental Psychology : Human Perception and Performance*, 11, pp. 317-328.
- COOPER L.A. & SHEPARD R.N. (1973) Chronometric studies of the rotation of mental images, in W.G Chase, éd., *Visual information processing*, New York, Academic Press.
- CORBALLIS M.C. & McLAREN R. (1982) Interaction between perceived and imagined rotation, *Journal of Experimental Psychology : Human Perception and Performance*, 8, pp. 215-224.
- CROVITZ H.F., ROSOF D. & SHIFFMAN H. (1971) Timing oscillation in human visual imagery, *Psychonomic Science*, 24, 87-88.
- DAVIDSON R.J. & SCHWARTZ G.E. (1977) Brain mechanisms subserving self-generated imagery : electrophysiological specificity and patterning, *Psychophysiology*, 14, pp. 598-601.

- DELEVAL J., DE MOL J. & NOTERMAN J. (1983) La perte des images souvenirs, *Acta Neurologica Belgica*, 83, pp. 61-79.
- DENIS M. (1979) *Les images mentales*, Paris, Presses Universitaires de France.
- DENIS M. (1984) Propriétés figuratives et non figuratives dans l'analyse de concepts, *L'Année Psychologique*, 84, pp. 327-345.
- DENIS M. (1989) *Image et cognition*, Paris, Presses Universitaires de France.
- DENIS M. & CARFANTAN M. (1985) People's knowledge about images, *Cognition*, 20, pp. 49-60.
- DENNETT D.C. (1969) *Content and consciousness*, New York, Humanities Press.
- DERENZI E. & SPINLER H. (1967) Impaired performance on color tasks in patients with hemispheric lesions, *Cortex*, 3, pp. 194-217.
- DIXON P. & JUST M.A. (1978) Normalization of irrelevant dimensions in stimulus comparisons, *Journal of Experimental Psychology : Human Perception and Performance*, 4, pp. 34-46.
- EDDY J.K. & GLASS A.L. (1981) Reading and listening to high and low imagery sentences, *Journal of Verbal Learning and Verbal Behavior*, 20, 333-345.
- ELLIS A.W. & YOUNG A.W. (1988) *Human cognitive neuropsychology*, Londres & Hillsdale, Lawrence Erlbaum Associates.
- FARAH M.J. (1984) The neurological basis of mental imagery: a componential analysis, *Cognition*, 18, pp. 245-272.
- FARAH M.J. (1985) Psychophysical evidence for a shared representational medium for mental images and percepts, *Journal of Experimental Psychology : General*, 114, pp. 91-103.
- FARAH M.J. (1988) Is visual imagery really visual? Overlooked evidence from neuropsychology, *Psychological Review*, 95, pp. 307-317.
- FARAH M.J., PERONNET F., WEISBERG L.L. & PERRIN F. (1988) Brain activity underlying mental imagery: an ERP study, *manuscrit*.
- FINKE R.A. (1979) The functional equivalence of mental images and errors of movement, *Cognitive Psychology*, 11, pp. 235-264.
- FINKE R.A. (1980) Levels of equivalence in imagery and perception, *Psychological Review*, 87, pp. 113-132.
- FINKE R.A. (1985) Theories relating mental imagery to perception, *Psychological Bulletin*, 98, pp. 236-259.
- FINKE R.A. & KURTZMAN H.S. (1981a) Area and contrast effects upon perceptual and imagery acuity, *Journal of Experimental Psychology : Human Perception and Performance*, 7, pp. 825-832.
- FINKE R.A. & KURTZMAN H.S. (1981b) Mapping the visual field in mental imagery, *Journal of Experimental Psychology : General*, 110, pp. 501-517.
- FINKE R.A., PINKER S. & FARAH M.J. (1989) Reinterpreting visual patterns in mental imagery, *Cognitive Science*, 13, pp. 51-78.

- FINKE R.A. & SCHMIDT M.J. (1977) Orientation-specific color aftereffects following imagination, *Journal of Experimental Psychology: Human Perception and Performance*, 3, pp. 599-606.
- FINKE R.A. & SCHMIDT M.J. (1978) The quantitative measure of pattern representation in images using orientation-specific color aftereffects, *Perception and Psychophysics*, 23, pp. 515-520.
- FLACH A. (1925) Über symbolische Schemata im produktiven Denkprozess, *Archiv für die ges. Psychologie*, LII, pp. 369-440.
- FODOR J.A. (1975) *The language of thought*, Cambridge, MA., Harvard University Press.
- FREYD J.J. & FINKE R.A. (1984) Facilitation of length discrimination using real and imagined context frames, *American Journal of Psychology*, 97, pp. 323-341.
- FRIEDMAN A. (1978) Memorial comparisons without the "mind's eye", *Journal of Verbal Learning and Verbal Behavior*, 17, pp. 427-444.
- GUZMAN A. (1969) Decomposition of a visual scene into three-dimensional bodies, in A. Grasselli, éd., *Automatic interpretation and classification of images*, New York, Academic Press.
- HEBB D.O. (1963) The semiautonomous process: its nature and nurture, *American Psychologist*, 18, pp. 16-27.
- HEBB D.O. (1968) Concerning imagery, *Psychological Review*, 75, pp. 466-477.
- HELLIGE J.B. & MICHIMATA CH. (1989) Categorization versus distance: hemispheric specialization for processing spatial information, *Memory and Cognition*, 17(6), pp. 770-776.
- HINTZMAN D.L., O'DELL C.S. & ARNDT D.R. (1981) Orientation in cognitive maps, *Cognitive Psychology*, 13, pp. 149-206.
- HOFFMANN J., DENIS M. & ZIESSLER M. (1983) Figurative features and the construction of visual images, *Psychological Research*, 45, pp. 39-54.
- HUMPHREYS G.W. & BRUCE V. (1989) *Visual cognition: computational, experimental and neuropsychological perspectives*, Londres & Hillsdale, Lawrence Erlbaum Associates.
- HUMPHREYS G.W. & RIDDOCH M.J. (1984) Routes to object constancy: implications from neurological impairments of object constancy, *Quarterly Journal of Experimental Psychology*, 36A, pp. 385-415.
- HUMPHREYS G.W. & RIDDOCH M.J. (1985) Authors' corrections to "Routes to object constancy", *Quarterly Journal of Experimental Psychology*, 37A, pp. 493-495.
- HUMPHREYS G.W. & RIDDOCH M.J. (1987) *To see but not to see: a case study of visual agnosia*, Londres & Hillsdale, Lawrence Erlbaum Associates.
- HUSSERL E. (1962-1969) *Recherches Logiques*, Paris, Presses Universitaires de France. [trad. fr. de *Logische Untersuchungen* (1913) Halle a.d.S., Max Niemeyer]

- JACKENDOFF R. (1983) *Semantics and cognition*, Cambridge, MA, M.I.T. Press.
- JACKENDOFF R. (1987a) On beyond zebra : the relation of linguistic and visual information, *Cognition*, 26, pp. 89-114.
- JACKENDOFF R. (1987b) *Consciousness and the computational mind*, Cambridge, MA., M.I.T. Press.
- KANT E. (1944) *Critique de la raison pure*, Paris, Presses Universitaires de France. [tr. fr. par J. Tremesaygues et B. Pacaud de *Kritik der reinen Vernunft* (1781-1787)]
- KERST S.M. & HOWARD J.H. (1978) Memory psychophysics for visual area and length, *Memory and Cognition*, 6, pp. 327-335.
- KINTSCH W. (1974) *The representation of meaning in memory*, Hillsdale, MA, Lawrence Erlbaum.
- KINTSCH W. (1988) The role of knowledge in discourse comprehension : a construction-integration model, *Psychological Review*, 95, pp. 163-182.
- KOSSLYN S.M. (1975) Information representation in visual images, *Cognitive Psychology*, 7, pp. 341-370.
- KOSSLYN S.M. (1976) Can imagery be distinguished from other forms of internal representations? Evidence from studies of information retrieval times, *Memory and Cognition*, 4, pp. 291-297.
- KOSSLYN S.M. (1980) *Image and mind*, Cambridge, MA, Harvard University Press.
- KOSSLYN S.M. (1981) The medium and the message in mental imagery : a theory, *Psychological Review*, 88, pp. 46-66.
- KOSSLYN S.M. (1987) Seeing and imagining in the cerebral hemispheres : a computational approach, *Psychological Review*, 94, pp. 148-175.
- KOSSLYN S.M. (1991) A cognitive neuroscience of visual cognition : further developments, in R.H. Logie & M. Denis, éd.
- KOSSLYN S.M., BALL T.M. & REISER B.J. (1978) Visual images preserve metric spatial information : evidence from studies of image scanning, *Journal of Experimental Psychology : Human Perception and Performance*, 4, pp. 47-60.
- KOSSLYN S.M., FLYNN R.A., AMSTERDAM J.B. & WANG G. (1990) Components of high level vision : a cognitive neuroscience analysis and accounts of neurological syndromes, *Cognition*, 34, pp. 203-277.
- KOSSLYN S.M. & KOENIG O. (1992) *Wet mind : the new cognitive neuroscience*, New York, The Free Press.

- KOSSLYN S.M., KOENIG O., BARRETT A., BACKER CAVE C., TANG J. & GABRIELI J.D.E. (1989) Evidence for two types of spatial representations : hemispheric specialization for categorical and coordinate relations, *Journal of Experimental Psychology : Human Perception and Performance*, 15(4), pp. 723-735.
- KOSSLYN S.M. & POMERANTZ J.R. (1977) Imagery, propositions, and the form of internal representations, *Cognitive Psychology*, 9, pp. 52-76.
- LAKOFF G. (1987) *Women, fire and dangerous things : what categories reveal about the mind*, Chicago, IL, University of Chicago Press.
- LANGACKER R.W. (1987) *Foundations of cognitive grammar, vol.1 : theoretical prerequisites*, Stanford, Stanford University Press.
- LANGACKER R.W. & E.H. CASAD (1985) 'Inside' and 'outside' in Cora grammar, *International Journal of American Linguistics*, 51, pp. 247-281. [repris dans R.W. Langacker (1991)]
- LANGACKER R.W. (1991) *Concept, image and symbol : the cognitive basis of grammar*, Berlin & New York, Mouton de Gruyter.
- LOGIE R.H. & DENIS M., éd. (1991) *Mental images in human cognition*, Amsterdam, Elsevier Science Publishers.
- MARR D. (1982) *Vision*, New York, Freeman.
- MCCARTHY R.A. & WARRINGTON E.K. (1990) *Cognitive neuropsychology : a clinical introduction*, San Diego, CA., Academic Press.
- MILLER G. & JOHNSON-LAIRD P. (1976) *Language and perception*, Cambridge, Harvard University Press.
- MOHLER C.W. & WURTZ R.H. (1977) Role of striate cortex and superior colliculus in visual guidance of saccadic eye movements in monkeys, *Journal of Neuropsychology*, 40, pp. 74-94.
- MOYER R.S. & R.H. BAYER (1976) Mental comparison and the symbolic distance effect, *Cognitive Psychology*, 8, pp. 228-246.
- NORMAN D.A., RUMELHART D.E. et al. (1975) *Explorations in cognition*, San Francisco, Freeman.
- PAIVIO A. (1971) *Imagery and verbal processes*, New York, Rinehart & Winston.
- PAIVIO A. (1986) *Mental representations : a dual coding approach*, New York, Oxford University Press.
- PALMER S.E. (1975) Visual perception and world knowledge : notes on a model of sensory-cognitive interaction, in D.A. Norman, D.E. Rumelhart et al.
- PALMER S.E. (1977) Hierarchical structure in perceptual representation, *Cognitive Psychology*, 9, pp. 441-474.
- PALMER S.E. (1978) Fundamental aspects of cognitive representations, in E. Rosch & B.B. Lloyd, éd., *Cognition and categorization*, Hillsdale, NJ, Lawrence Erlbaum Associates.

- PASIK P. & PASIK T. (1982) Visual functions in monkey after total removal of visual cerebral cortex, in W.D. Neff, *Contributions to sensory physiology*, vol. 7, New York, Academic Press.
- PERKY C.W. (1910) An experimental study of imagination, *American Journal of Psychology*, 21, pp. 422-452.
- PINKER S. (1980) Mental imagery and the third dimension, *Journal of Experimental Psychology : General*, 109(3), pp. 354-371.
- PINKER S. (1984) Visual cognition : an introduction, *Cognition*, 18, pp. 1-63.
- PINKER S. & FINKE R.A. (1980) Emergent two-dimensional patterns in images rotated in depth, *Journal of Experimental Psychology : Human Perception and Performance*, 6(2), pp. 244-264.
- PINKER S. & KOSSLYN S.M. (1978) The representation and manipulation of three-dimensional space in mental images, *Journal of Mental Imagery*, 2, pp. 69-84.
- PRITCHARD R.M. (1961) Stabilized images on the retina, *Scientific American*, 204, pp. 72-78.
- PYLYSHYN Z.W. (1973) What the mind's eye tells the mind's brain : A critique of mental imagery, *Psychological Bulletin*, 80, pp. 1-24.
- PYLYSHYN Z.W. (1979a) The rate of "mental rotation" of images : a test of a holistic analogue hypothesis, *Memory and Cognition*, 7, pp. 19-28.
- PYLYSHYN Z.W. (1979b) Validating computational models : a critique of Anderson's indeterminacy of representation claim, *Psychological Review*, 86, pp. 383-394.
- PYLYSHYN Z.W. (1981) The imagery debate : analogue media versus tacit knowledge, *Psychological Review*, 88, pp. 16-45.
- PYLYSHYN Z.W. (1984) *Computation and cognition*, Cambridge, MA., M.I.T. Press.
- REED S.K. (1974) Structural descriptions and the limitations of visual images, *Memory and Cognition*, 2, pp. 329-336.
- REED S.K., HOCK H.S. & LOCKHEAD G.R. (1983) Tacit knowledge and the effect of pattern configuration on mental scanning, *Memory and Cognition*, 11, pp. 137-143.
- SARTRE J.-P. (1986) [1940] *L'imaginaire*, Paris, Gallimard.
- SEGAL S.J. & Fusella V. (1970) Influences of imaged pictures and sounds on detection of visual and auditory signals, *Journal of Experimental Psychology*, 83, pp. 458-464.
- SHALLICE T. (1988) *From neuropsychology to mental structure*, Cambridge, Cambridge University Press.
- SHEPARD R.N. (1975) Form, formation and transformation of internal representations, in R.L. Solso, éd., *Information processing and cognition : the Loyola Symposium*, Hillsdale, NJ, Lawrence Erlbaum Associates.

- SHEPARD R.N. & CHIPMAN S. (1970) Second order isomorphism of internal representations : shapes of states, *Cognitive Psychology*, 1, 1-17.
- SHEPARD R.N. & METZLER J. (1971) Mental rotation of three-dimensional objects, *Science*, 171, pp. 801-703.
- TREISMAN A. (1986) Properties, parts and objects, in K. Boff, L. Kaufman & J. Thomas, éd., *Handbook of perception and human performance*, chap. 35, pp. 1-70.
- UNGERLEIDER L.G. & MISHKIN M. (1982) Two cortical visual systems, in D.J. Ingle, M.A. Goodale & R.J.W. Mansfield, éd., *Analysis of visual behavior*, pp. 549-586, Cambridge, MA, M.I.T. Press.
- WALLACE B. (1984) Apparent equivalence between perception and imagery in the production of various visual illusions, *Memory and Cognition*, 12, pp. 156-162.
- WARRINGTON E.K. & JAMES M. (1986) Visual object recognition in patients with right hemisphere lesions : axes or features?, *Perception*, 15, pp. 355-366.
- WARRINGTON E.K. & TAYLOR A.M. (1973) The contribution of the right parietal lobe to object recognition, *Cortex*, 9, pp. 152-164.
- WARRINGTON E.K. & TAYLOR A.M. (1978) Two categorical stages of object recognition, *Perception*, 7, pp. 695-705.
- WEISKRANTZ L., WARRINGTON E.K., SANDERS M.D. & MARSHALL J. (1974) Visual field recovery from scotoma in the hemianopic field following a restricted occipital ablation, *Brain*, 97, pp. 709-728.