



Categorisation et attribution de proprietes

Edouard Machery

► **To cite this version:**

Edouard Machery. Categorisation et attribution de proprietes. Cahiers Philosophiques de Strasbourg, Université de Strasbourg, 2004, 17, pp.119-147. <ijn_00000508>

HAL Id: ijn_00000508

https://jeannicod.ccsd.cnrs.fr/ijn_00000508

Submitted on 9 Jun 2004

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

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

Catégorisation et Attribution de Propriétés

Edouard Machery

Max-Planck Institute For Human Development (Berlin)
Université de Paris-Sorbonne, Institut Jean-Nicod (Paris)

Le problème

Nous distinguons intuitivement les propriétés que possède un objet et la catégorie à laquelle appartient un objet : par exemple, l'adjectif « rouge » désigne une propriété, tandis que « chien » désigne une catégorie d'objets. Quelle que soit la manière dont cette distinction est théoriquement développée et quelle que soit sa valeur ultime, elle est de nature métaphysique (ou ontologique) : on distingue par là deux types d'entités, les propriétés et les catégories. Ces deux notions sont aussi impliquées dans une distinction de nature *psychologique*, à savoir *la distinction entre l'attribution de propriétés à un objet et la classification d'un objet dans une classe (c'est-à-dire sa catégorisation)*. Il ne s'agit plus de distinguer deux types d'entités, mais deux capacités cognitives. En première approximation, l'attribution de propriétés à un objet (que, par souci de généralité, j'appellerai dorénavant la cible¹) a pour fonction de produire le jugement ou la pensée que la cible possède telle ou telle propriété, tandis que la catégorisation a pour fonction de produire le jugement ou la pensée que la cible appartient à telle ou telle classe. Dans cet article, je me consacre exclusivement à la distinction entre ces deux capacités cognitives.

Comme nous le verrons, *cette distinction est non seulement intuitive, mais elle est aussi postulée, parfois seulement implicitement, par toutes les théories psychologiques contemporaines de la catégorisation*. Toutefois, si ces théories caractérisent précisément les représentations mentales et les processus cognitifs censés réaliser la capacité de catégoriser, *elles ont en revanche peu à dire sur les représentations mentales et les processus qui réalisent la capacité d'attribuer des propriétés* (à ma connaissance). Souvent, elles assument simplement que d'une manière ou d'une autre, avant la catégorisation, les cibles sont identifiées et des propriétés leur sont attribuées. Cet état de fait est à l'origine du problème que je veux soulever et traiter dans cet article : *est-il justifié de postuler une capacité d'attribuer des propriétés à des cibles ?*

Dans quelles conditions ne serait-il pas légitime de postuler une telle capacité ? Postuler une capacité psychologique (en l'occurrence, l'attribution de propriétés à une cible) n'est pas méthodologiquement condamnable si cette capacité est requise par la théorie à des fins explicatives. On pourrait même dire que les théories de la catégorisation offrent des raisons de postuler l'existence d'une capacité cognitive qu'il reste à décrire et à expliquer. De prime abord, cela semble être un progrès théorique. En outre, de telles méthodes sont requises pour l'analyse des capacités cognitives supérieures (ce qui est le cas de la catégorisation), puisque ces capacités dépendent de nombreuses autres capacités cognitives.

En quel sens donc postuler cette capacité serait-il illégitime ? Pour répondre à cette question, il est nécessaire de présenter en quelques mots l'arrière-plan théorique de cet article. Le problème soulevé ici repose sur la méthodologie développée pour les sciences cognitives par

Je souhaite remercier vivement Fr. Longy pour la lecture attentive d'une version antérieure de cet article.

Toute correspondance doit être adressée à edouardmachery@hotmail.com.

¹ On peut en effet catégoriser des propriétés, des événements etc. aussi bien que des objets physiques.

Marr (Marr 1982, 24 sq., voir les discussions de Peacocke 1986 et de Clark 1993, chap. 3). La description d'une compétence cognitive (par exemple, la vision en trois dimensions) se fait en trois étapes. Premièrement, la compétence est divisée en sous-compétences (par exemple, identifier les contours...), articulées les unes aux autres. Ces sous-compétences constituent la compétence analysée (par analogie, la capacité de jouer au tennis est constituée par les capacités de servir, de faire un coup droit etc.)². La compétence analysée (ou les sous-compétences qui la constituent) est décrite en termes fonctionnels : elle transforme une classe d'inputs en une classe d'outputs (par exemple, les stimulations des cellules réceptives oculaires en représentations de l'espace tri-dimensionnel). Deuxièmement, les processus transformant les inputs en outputs sont décrits au moyen d'algorithmes (par analogie, il s'agit de décrire les mouvements par lesquels on fait un coup droit etc.). Si la compétence a été divisée en sous-compétences, il convient de décrire les processus qui réalisent celles-ci. Troisièmement, la réalisation neuronale de ces processus peut être décrite. Chaque niveau décrit la *réalisation* du niveau supérieur en des termes spécifiques.

Le point important pour mon propos est le suivant. La première étape suppose que les compétences cognitives possèdent une *unité fonctionnelle* : intuitivement, elles produisent un type unique d'outputs³. Plus précisément, on dira qu'*une capacité a une unité fonctionnelle si et seulement si la production des outputs qui caractérisent la capacité à partir de ses inputs requiert un seul processus ou plusieurs processus articulés*⁴. Autrement dit, si et seulement si ce n'est pas le cas que la production de certains outputs requiert certains processus tandis que la production d'autres outputs requiert d'autres processus. A ce titre, la reconnaissance des visages et la localisation auditive ne forment pas une capacité fonctionnellement une, puisque les processus qui réalisent la reconnaissance des visages et ceux qui réalisent la localisation auditive ne forment pas une série de processus articulés (voir Lyons 2001 pour une analyse plus détaillée). Au contraire, la reconnaissance visuelle est sans doute une capacité fonctionnellement une, puisqu'il existe une série de processus articulés qui produit la reconnaissance visuelle des objets à partir de l'excitation des cellules rétinales.

La justification de cette contrainte est la suivante. Si une compétence n'est pas fonctionnellement une, sa description au premier niveau distingué par Marr ne peut guider la recherche des processus qui la réalisent (niveau deux). En effet, les outputs requièrent alors différents processus. Une capacité qui n'est pas fonctionnellement une n'a donc pas sa place dans la description cognitive de l'esprit⁵. Assurément, une compétence cognitive peut être, et typiquement est, constituée par plusieurs sous-compétences. Mais cela n'implique en rien que la compétence ne soit pas fonctionnellement une, s'il y a une série de processus articulés qui

² Je parle de *compétence ou capacité complexe* pour une compétence qui est constituée par plusieurs sous-compétences. Autrement, je parle de *compétence ou capacité primitive*.

³ Il y a en outre une contrainte de spécialisation (Lyons 2001) : une capacité (caractérisée par ses inputs et ses outputs) est spécialisée si et seulement si il y a un processus cognitif qui produit uniquement ces outputs à partir de ces inputs. Pour cette raison, reconnaître sa grand-mère ou se souvenir des événements qui se sont passés entre une et deux minutes avant l'instant présent ne sont pas des compétences spécialisées. La justification de cette contrainte est analogue à la justification de la contrainte d'unité fonctionnelle.

⁴ Il y a *deux types d'articulation* (voir Lyons 2001). 1/ Les processus 1,..., n qui réalisent une capacité sont tels que les inputs de 1 sont les inputs caractéristiques de la capacité, les outputs de n sont ses outputs caractéristiques et les inputs de k sont les outputs de k-1. En termes plus intuitifs, les processus forment une chaîne. 2/ Les processus 1,..., n sont tels que les inputs de 1, ..., n-1 sont des sous-ensembles des inputs caractéristiques de la capacité, les outputs de n sont ses outputs caractéristiques, et les outputs de 1, ..., n-1 sont les inputs de n. En termes plus intuitifs, les processus 1 à n-1 sont indépendants les uns des autres, mais sont unifiés en ce qu'ils fournissent les inputs d'un même processus (n).

⁵ Ce qui ne veut pas dire que ces capacités n'existent pas. Il y a autant de capacités que de manière de décrire ce qu'un organisme fait. Par exemple, j'ai la capacité de frapper la lettre « e » de mon ordinateur. Mais de telles capacités n'ont pas de pertinence pour une théorie psychologique.

produisent les outputs de la compétence à partir de ses inputs. C'est le cas de la reconnaissance des visages.

On a donc là *une contrainte sur les capacités qu'il est légitime de postuler : elle doit être fonctionnellement une*⁶. Autrement dit, la capacité postulée ne doit pas être un conglomerat hétérogène de capacités variées aux fonctions diverses. *Dans ce cas en effet, cette capacité n'est pas un posit théorique légitime* : elle n'a pas de rôle à jouer dans la description de l'esprit. L'exemple de la mémoire illustre cette contrainte. La notion de mémoire appartient à la psychologie populaire. Elle est supposée désigner la capacité de posséder des connaissances sur les événements passés. Or, les recherches sur la mémoire suggèrent que cette notion ne désigne pas une fonction homogène de l'esprit, mais recouvre des fonctions psychologiques variées (mémoire épisodique, mémoire déclarative, mémoire à court terme etc.). En fait, plusieurs processus réalisent ce que la mémoire est supposée réaliser : la mémoire des événements singuliers suppose des processus différents de la mémoire des faits, qui est distincte de la modification du comportement en fonction de l'expérience etc. (Tulving 1995). Cela suggère que la notion de mémoire est un posit théorique dont il conviendrait sans doute de se débarrasser au profit des capacités psychologiques qu'elle recouvre.

La contrainte d'unité fonctionnelle suggère donc le problème suivant : ***la capacité d'attribuer des propriétés est-elle une capacité homogène ou hétérogène ?*** Deux réponses opposées sont possibles :

1. ***La thèse de l'homogénéité*** soutient que l'attribution de propriétés est une capacité fonctionnellement une.
2. ***La thèse de l'hétérogénéité*** soutient qu'elle recouvre en fait des capacités variées.

La thèse de l'homogénéité soutient que l'attribution de propriétés est semblable à la capacité de reconnaître les visages : il y a un processus (ou une série de processus articulés) qui peut à lui seul attribuer toutes les propriétés. La thèse de l'hétérogénéité soutient que l'attribution de propriétés est semblable à la mémoire : plusieurs processus non articulés sont nécessaires pour attribuer les propriétés. Comme je l'ai suggéré, ce problème est important. Si la thèse de l'hétérogénéité est correcte, postuler une capacité d'attribuer des propriétés pour rendre compte de la catégorisation est méthodologiquement inapproprié. C'est en fait là la thèse que j'entends soutenir dans cet article : ***la capacité d'attribuer des propriétés doit être éliminée des théories psychologiques dans la mesure où elle n'est pas une capacité homogène***⁷.

Succinctement, ma stratégie argumentative est la suivante. J'examine différents types d'attributions de propriétés. A la lumière de cet examen, je conclus que ces différents types d'attributions requièrent des processus différents – ce qui est incompatible avec la thèse de l'homogénéité.

Dans la première section, j'analyse les théories psychologiques de la catégorisation afin de montrer qu'elles postulent une capacité d'attribuer des propriétés. Dans la deuxième section, j'examine l'hypothèse empiriste selon laquelle l'attribution de propriétés peut être réduite à l'attribution de propriétés sensorielles. Cet examen suggère que l'attribution de propriétés sensorielles requiert des processus *sui generis* qui ne peuvent être étendus aux autres propriétés. Dans la troisième section, je montre que certaines attributions sont en fait des catégorisations, mais que toutes ne peuvent pas l'être. Dans la quatrième et dernière section,

⁶ J'utilise souvent la terminologie suivante : une capacité est *homogène* si elle est fonctionnellement une ; autrement, elle est *hétérogène*.

⁷ Quelle est l'originalité du problème soulevé ici ? La notion de propriété (« feature », « property ») utilisée dans les théories des concepts et de la catégorisation a suscité une vaste littérature critique, expérimentale ou théorique, en psychologie (Murphy et Medin 1985 ; Medin 1989 ; Komatsu 1992 ; Wisniewski et Medin 1994). Néanmoins, ces analyses ne portent pas sur le l'hétérogénéité de la capacité d'attribuer des propriétés.

je présente un modèle schématique de la catégorisation qui rejette l'existence d'une capacité homogène d'attribution.

Première section : les théories contemporaines de la catégorisation postulent toutes une distinction entre deux capacités cognitives, la catégorisation et l'attribution de propriétés

Dans cette section, je n'examine pas une à une les très nombreuses théories psychologiques de la catégorisation. Néanmoins, on peut grouper ces théories en trois classes. Leur examen confirme l'idée que toutes les théories de la catégorisation postulent une capacité d'attribuer des propriétés.

1. Il est possible de distinguer trois types de théories de la catégorisation (Smith et Medin 1981 ; Medin 1989 ; Hampton 1997 ; Murphy 2002). Selon le premier modèle (*catégorisation fondée sur des règles*), la catégorisation repose sur la vérification que la cible satisfait des conditions spécifiées par une **règle**. Selon le deuxième modèle (*catégorisation fondée sur la ressemblance*), la catégorisation repose sur la **ressemblance** de la cible et de la représentation mentale de la catégorie. Selon le dernier modèle (*catégorisation fondée sur l'explication*), la catégorisation repose sur **l'explication** de certaines propriétés de la cible au moyen de la représentation de la catégorie. Cette distinction a été jugée à l'occasion superficielle (Hahn et Chater 1997) : la ressemblance serait impliquée dans ces trois types de théories, qui ne se distingueraient en fait que par le type de ressemblance mis en avant (ressemblance perceptive vs. théorique, par exemple). Néanmoins, elle reste un principe utile pour distinguer différents modèles de la catégorisation.

2. Les théories de la catégorisation fondées sur des règles furent les premiers modèles psychologiques proposés (Hull 1920 ; Bruner et al. 1956), et elles sont naturellement associées avec ce qu'on appelle la conception classique des concepts (Smith et Medin 1981 ; Margolis et Laurence 1999b). En effet, si les représentations des catégories (que nous appellerons parfois, conformément à l'usage en vigueur en psychologie et en philosophie de la psychologie, des concepts) spécifient des conditions séparément nécessaires et conjointement suffisantes pour appartenir à la catégorie représentée, il est naturel que le processus de catégorisation consiste à vérifier que la cible possède chacune des conditions nécessaires spécifiées par le concept. Dans les expériences d'apprentissage de concepts artificiels de Bruner et al., les concepts sont ainsi des fonctions booléennes, le plus souvent des conjonctions, des propriétés qui caractérisent une catégorie d'objets (Bruner et al. 1956). La catégorisation consiste à vérifier que l'objet satisfait bien telle fonction booléenne de propriétés. Elle suppose donc l'attribution de propriétés à la cible. La thèse selon laquelle tous les concepts seraient des représentations des conditions nécessaires et suffisantes d'appartenance a été massivement rejetée dans les années 70⁸. Néanmoins, selon quelques psychologues et selon de nombreux chercheurs en intelligence artificielle, certains concepts seraient conformes à la conception classique (Shanks 1997 ; Pinker et Prince 1999). Naturellement, s'il y a bien de tels concepts, un processus de vérification des conditions spécifiées par le concept reste le modèle le plus naturel de la catégorisation : au lieu d'être un modèle général de la catégorisation, c'est simplement un modèle de la catégorisation impliquant une certaine classe de concepts. Bref, l'explication de la catégorisation fondée sur des règles postule i/ l'existence d'une capacité d'attribuer des propriétés à la cible en sus du

⁸ Medin 1989 et Komatsu 1992 exposent l'histoire récente de la psychologie des concepts et de la catégorisation du point de vue des psychologues. Margolis 1994 offre une approche critique des évolutions récentes. Voir aussi Machery manuscript b.

processus de catégorisation de la cible et suppose ii/ que la catégorisation dépend de la mise en œuvre de la capacité d'attribuer des propriétés.

3. Les théories des prototypes sont les paradigmes des explications de la catégorisation fondées sur la ressemblance (Hampton 2001)⁹. De manière générale, un prototype d'une classe d'objets est une représentation des propriétés les plus fréquentes des membres de cette classe (éventuellement, ces propriétés reçoivent différentes pondérations en fonction de leur importance relative pour la catégorisation, de leur fréquence et/ou de leur saillance parmi les membres de cette classe). Un objet est catégorisé dans la catégorie x (par exemple, dans la catégorie *CHIEN*¹⁰) si sa ressemblance avec le prototype (des chiens) est supérieure à un critère donné (éventuellement sensible au contexte) et excède la ressemblance avec les prototypes des autres catégories. De nombreux modèles ont été proposés pour calculer la ressemblance entre le prototype et la cible (par exemple, les modèles spatiaux de la ressemblance, la règle du contraste de Tversky...¹¹). Néanmoins, ces modèles reposent de manière identique sur une comparaison (« match ») entre les propriétés de la cible et les propriétés représentées dans le prototype¹². En conséquence, selon les théories des prototypes, le processus de catégorisation requiert que la cible (par exemple, un chien perçu visuellement) soit antérieurement représentée comme possédant telle et telle propriété (cet objet dans mon champ visuel a quatre pattes, une queue qui bouge, il a une fourrure, il se meut...).

Les théories qui postulent des exemplaires au lieu de prototypes font aussi appel à la ressemblance¹³. La catégorisation d'une cible dans une classe ne dépend pas de la représentation des propriétés les plus fréquentes de cette classe, mais du souvenir de membres de cette classe qui ont été rencontrés par le passé et dont les propriétés ont été encodées. Il existe différents modèles pour rendre compte de la catégorisation au moyen d'exemplaires (Medin et Shaffer 1978 ; Brooks 1978 ; Nosofsky 1988). Néanmoins, ces différents modèles ont en commun l'idée que les propriétés de la cible sont comparées aux propriétés des exemplaires encodés en mémoire à long terme (MLT). Ces modèles requièrent donc également l'attribution de propriétés à la cible. Le débat avec les modèles qui posent des prototypes porte sur la nature des représentations des catégories (représentation abstraite et générale vs. représentations d'exemplaires singuliers) et non sur le processus de catégorisation¹⁴.

Bref, les modèles de la catégorisation fondés sur la ressemblance postulent i/ l'existence d'une capacité d'attribuer des propriétés à la cible en sus du processus de catégorisation de la cible et supposent ii/ que la catégorisation dépend de la mise en œuvre de la capacité d'attribuer des propriétés.

⁹ Pour une présentation synthétique des théories des prototypes, voir Smith et Medin 1981 ; Hampton 1993 ; Murphy 2002, chap. 3. Ces théories ne s'accordent pas sur la nature exacte du prototype et sur le processus de catégorisation. Les théories des prototypes forment donc une famille de théories.

¹⁰ J'écrirai dorénavant les noms de concepts en majuscules, les noms de propriétés en italiques et les noms de catégories en majuscules italiques.

¹¹ Voir Ahn et Chater 1997.

¹² Sur les difficultés posées par ce processus de « matching », voir Wisniewski et Medin 1994, 268-9.

¹³ Pour une présentation synthétique des théories des exemplaires et pour une comparaison avec les théories des prototypes, voir Medin et Smith 1981 ; Shanks 1997 ; Ross et Makin 1999 ; Murphy 2002, chap. 3. Tout comme les théories des prototypes, les théories des exemplaires sont en désaccord sur la nature des exemplaires et sur le processus de catégorisation.

¹⁴ Notons toutefois que les théories des prototypes et les théories des exemplaires font souvent appel à des algorithmes différents. Les théories des prototypes additionnent en général les propriétés communes à la cible et au prototype (Hampton 1993). En revanche, les théories des exemplaires les multiplient (Medin et Shaffer 1978 ; Ross et Makin 1999). Mais rien ne contraint les théories des exemplaires et des prototypes à utiliser les algorithmes qu'ils utilisent habituellement. Une théorie des prototypes peut parfaitement être combinée avec une règle multiplicative et une théorie des exemplaires avec une règle additive.

4. Le dernier type de modèles de la catégorisation est fondé sur l'explication. En premier lieu, les représentations des catégories sont supposées ressembler davantage à des théories : elles spécifient des propriétés internes (vs. des propriétés perceptives) ; les relations, en particulier causales et fonctionnelles, entre les propriétés sont prises en compte (par exemple, l'information selon laquelle les ailes des oiseaux ont pour fonction de voler est incluse dans le concept OISEAU) ; des informations de nature nomologique sont aussi incluses dans la représentation¹⁵ etc. Ensuite et surtout, la catégorisation ne consiste plus (ou plus seulement) en un processus de comparaison (Murphy et Medin 1985 ; Heit 1997). Au contraire, la représentation de nature théorique fournit des connaissances qui expliquent certaines propriétés de la cible. Le concept qui fournit la meilleure explication des propriétés d'une cible représente la catégorie à laquelle cette cible appartient. Par exemple, le fait qu'un convive saute dans une piscine lors d'une soirée est expliqué par l'effet de l'alcool (Murphy et Medin 1985). Sur la base de cette explication, la cible (le convive) est catégorisée dans la catégorie des personnes ivres¹⁶. A la différence des autres modèles, peu d'algorithmes ont été présentés pour rendre compte de ce type de catégorisation¹⁷ : cela n'est guère surprenant étant donné les difficultés théoriques pour rendre compte de l'inférence à la meilleure explication. De toute évidence, ce processus de catégorisation suppose également l'attribution de propriétés (éventuellement temporaires ou dynamiques) à la cible. Dans l'exemple ci-dessus, le catégorisateur prend pour input le comportement spécifique du convive. Notons que cet exemple montre que la notion de propriété doit être prise en un sens suffisamment large : elle ne doit pas être restreinte aux seules propriétés perceptives (vs. fonctionnelles ou théoriques ou causales...) ni aux propriétés stables (vs. temporaires). Bref, les modèles de la catégorisation fondés sur l'explication postulent également i/ l'existence d'une capacité d'attribuer des propriétés à la cible en sus du processus de catégorisation de la cible et supposent ii/ que la catégorisation dépend de la mise en œuvre de la capacité d'attribuer des propriétés.¹⁸

Puisque la tripartition présentée ci-dessus recouvre toutes les théories contemporaines de la catégorisation, celles-ci postulent toutes i/ l'existence d'une capacité d'attribuer des propriétés à la cible en sus du (ou des) processus de catégorisation de la cible et supposent ii/ que la catégorisation dépend de la mise en œuvre de cette capacité.

Deuxième section : un modèle empiriste de l'attribution de propriétés

Défendre l'homogénéité de l'attribution de propriétés, c'est montrer que sa réalisation ne requiert pas plusieurs processus (non articulés). Dans cette section, je présente, puis critique une hypothèse d'inspiration empiriste conforme à la thèse de l'homogénéité : l'attribution de

¹⁵ Medin et Smith 1985 a introduit cette nouvelle manière de caractériser les concepts. Ce type de théorie est présenté dans Medin 1989 ; Heit 1997 ; Hampton 1997 ; Margolis et Laurence 1999 ; Murphy 2002, chap. 3. Je distingue deux manières de concevoir les concepts comme des théories en introduction de Machery 2002. Il y a diverses manières de rendre précises les idées de base présentées ici.

¹⁶ Wisniewski et Medin 1994 offrent de nombreux exemples de catégorisation qui satisfont ce modèle de la catégorisation (par exemple, p. 260, et surtout p. 269).

¹⁷ Cependant, de nombreux travaux en intelligence artificielle tentent de prendre en compte les effets des connaissances antérieures des sujets dans les tâches cognitives, particulièrement dans la catégorisation (voir Wisniewski et Medin 1994, 272-277). Heit s'efforce ainsi de combiner réseaux de neurones et catégorisation basée sur l'explication (Heit 1997, 2001).

¹⁸ Il est aussi possible que la catégorisation dépende de plusieurs processus cognitifs qui reposent sur des représentations différentes (voir Ross et Makin 1999 ; Machery manuscript a). Il faudrait alors examiner si la catégorisation est vraiment une capacité cognitive une.

propriétés peut être réduite à l'attribution de propriétés sensorielles. L'examen de cette hypothèse offre une première raison de douter de la thèse de l'homogénéité.

1. Avant de décrire ce modèle, il convient de présenter une manière commode de représenter le traitement de l'information par l'esprit au moyen de boîtes et de flèches. Une boîte représente une capacité cognitive. Entre les boîtes sont dessinées des flèches qui représentent les transferts d'information entre les processus qui réalisent les différentes capacités. Une boîte peut contenir plusieurs boîtes qui représentent les sous-compétences qui constituent une capacité donnée. Ceci posé, il est possible de représenter la thèse de l'homogénéité comme suit. Conformément à la thèse de l'homogénéité, ce schéma contient une unique boîte (composée éventuellement d'autres boîtes) pour l'attribution de propriétés. J'appelle cette boîte « attributeur » et « catégorisateur » l'ensemble des processus qui produisent la catégorisation d'une cible¹⁹.

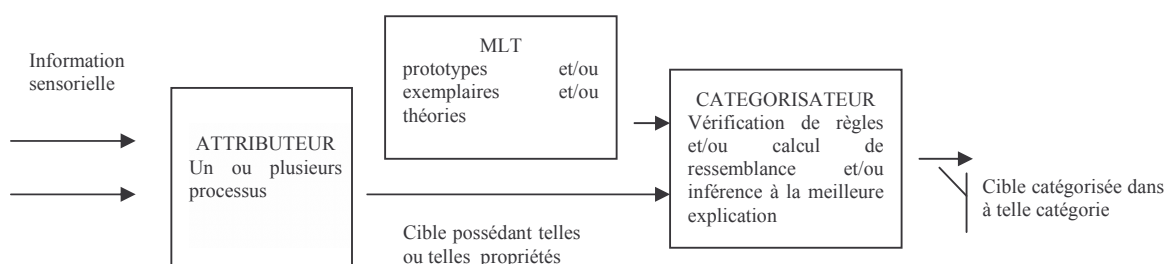


Schéma 1 : la thèse de l'homogénéité²⁰

2. Pourquoi s'intéresser à une conception empiriste de l'attribution de propriétés ? Deux raisons justifient ce choix. Tout d'abord, alors que les modèles contemporains de la catégorisation ne disent presque rien sur l'attribution de propriétés, la conception empiriste de l'attribution de propriétés supporte la thèse de l'homogénéité ; ensuite et surtout, de nombreux psychologues, implicitement ou explicitement, endossent une telle conception. La citation suivante illustre ce point :

« Les développements futurs de la théorie doivent s'écarter des descriptions purement linguistiques des « propriétés » et les remplacer par des valeurs sémantiques qui sont mieux ancrées dans un niveau subsymbolique. (...) *les propriétés elles-mêmes doivent être traduisibles dans des représentations non linguistiques qui ont rapport à l'action et à la perception*, pour que le système représentationnel évite la circularité. Selon Barsalou (1993), l'imagerie visuelle pourrait fournir à la représentation des concepts ce langage représentationnel prélinguistique. » Hampton, 1997, 91²¹

Cette opinion est partagée, plus ou moins explicitement, par de nombreux psychologues (Kosslyn 1978 ; Olseth et Barsalou 1995 ; Solomon et Barsalou 1997 ; Goldstone et Barsalou 1998 ; Barsalou 1999) et, plus rarement, par certains philosophes (Prinz 2002). Enfin, de nombreux modèles en intelligence artificielle endossent la même hypothèse (voir les références dans Wisniewski et Medin 1994, 272-277).

L'hypothèse empiriste est la suivante : *les propriétés sur la base desquelles une cible est catégorisée sont uniquement des propriétés sensorielles ; les propriétés non sensorielles*

¹⁹ Je suppose que la catégorisation est fonctionnellement une.

²⁰ MLT signifie mémoire à long terme.

²¹ Je traduis et je souligne.

mentionnées dans les théories de la catégorisation ne sont qu'une manière de parler des propriétés sensorielles. Par exemple, l'affirmation selon laquelle la catégorisation dans la catégorie *OISEAUX* dépend de l'attribution de la propriété *a des ailes* n'est qu'une manière de parler de l'attribution d'un certain ensemble complexe de propriétés sensorielles. En conséquence, seules des propriétés sensorielles sont attribuées.

L'expression « propriété sensorielle » est ambiguë. Elle désigne les propriétés dont les représentations sont computées par les systèmes perceptifs (*rouge, carré, aigu, salé*), les propriétés dont les représentations sont les outputs des transducteurs²² ou encore toutes les propriétés dont les représentations sont computées au sein des systèmes perceptifs (Prinz 2002)²³. En accord avec la citation de Hampton, j'utilise ici cette expression pour désigner ces trois types de propriétés. L'hypothèse empiriste est donc que l'attribution de propriétés peut être réduite à la détection des propriétés sensorielles. Il n'y a pas de propriétés non sensorielles à attribuer.

3. L'hypothèse empiriste affirme que toutes les propriétés sont sensorielles. Elle réduit donc l'attribution de propriétés à l'attribution de propriétés sensorielles. Cette idée ne soutient la thèse de l'homogénéité que si l'attribution de propriétés sensorielles est elle-même fonctionnellement une. Je présente ici un modèle de l'attribution de propriétés sensorielles selon lequel celle-ci est fonctionnellement une. L'idée de base est la suivante. L'attribution d'une propriété sensorielle à un objet est assimilée à la production de la représentations de cette propriété sensorielle et au liage de cette représentation avec d'autres représentations de propriétés sensorielles. Sur la base de l'output de l'attributeur, les cibles sont catégorisées. Puisque les systèmes perceptifs et le processus de liage sont articulés, l'attribution de propriétés sensorielles est fonctionnellement une. Puisque seules des propriétés sensorielles sont attribuées, l'attribution de propriétés est en général fonctionnellement une. Le schéma suivant résume cette hypothèse.

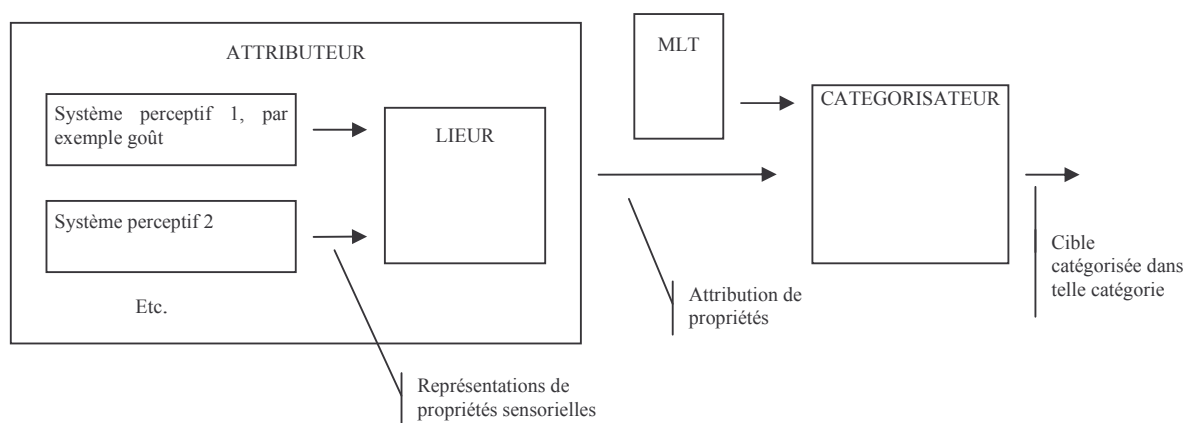


Schéma 2 : un modèle empiriste de l'attribution de propriétés

²² Les transducteurs sont les structures neuronales qui transforment les stimulations sensorielles (par exemple, les stimulations des cônes et des bâtonnets dans l'œil), c'est-à-dire des grandeurs physiques, en représentations (ou les représentations en commandes motrices). Puisqu'ils produisent les représentations qu'utilisent les processus cognitifs, les transducteurs ne sont pas eux-mêmes des processus cognitifs.

²³ Parfois, les propriétés qui peuvent être attribuées sur la base de la perception sont dites sensorielles. Ainsi la propriété *ailé* est-elle à l'occasion qualifiée de propriété sensorielle. On les oppose alors aux propriétés abstraites (par exemple, Wisniewski et Medin 1994, 240).

Si l'attribution de propriétés était réalisée uniquement par les systèmes perceptifs, la thèse de l'homogénéité serait évidemment fautive puisque la plupart des systèmes perceptifs sont (au moins largement) indépendants les uns des autres. Mais l'attribution de propriétés ne peut être réalisée simplement par les transducteurs et les systèmes perceptifs. En effet, la représentation de l'occurrence d'une propriété n'est pas son attribution à une cible : encore faut-il précisément que cette propriété soit attribuée (c'est l'une des objections adressées par Kant à l'empirisme classique dans la déduction transcendantale (B) de la première *Critique de la Raison Pure*). Je fais l'hypothèse que l'attribution des propriétés sensorielles consiste en l'association des représentations de propriétés sensorielles les unes avec les autres. Attribuer des propriétés sensorielles, c'est unifier leurs représentations dans la représentation d'un objet singulier : en sus des mécanismes détectant les propriétés, il faut donc poser des mécanismes les unifiant (c'est ce qu'on appelle le problème du liage, « binding problem »). Les représentations des propriétés sensorielles produites par les systèmes perceptifs sont les inputs communs de ce mécanisme²⁴.

*Ce modèle est inspiré par la littérature psychologique sur la représentation des objets singuliers et des propriétés perceptives*²⁵. Les psychologues travaillant sur ces questions s'accordent en effet sur les points suivants. Les systèmes perceptifs (vision, goût etc.²⁶) sont composés entre autres de systèmes dont la fonction est de détecter l'occurrence de propriétés sensorielles particulières (par exemple, les couleurs, les angles, les contours...) sur la base des informations sensorielles représentées par les transducteurs. Les propriétés sont enregistrées spatialement, c'est-à-dire en fonction de leur position dans l'environnement de l'organisme. Elles sont stockées séparément les unes des autres (Treisman et Gelade 1980 ; Treisman 1992). Treisman a ainsi suggéré que chaque type de propriétés était stocké dans des cartes distinctes de l'environnement (Treisman 1992, 158-159). Tant que seule l'occurrence de ces propriétés est comptée par les systèmes perceptifs, aucun objet n'est représenté. Une fois détectées, les propriétés peuvent être liées par un mécanisme, que j'appelle ici le lieu. Le liage (« binding ») consiste à réunir les propriétés en les associant à la représentation d'un objet singulier. La métaphore du fichier est souvent utilisée pour décrire le résultat de ce processus (Kahneman et al. 1992). La représentation d'un objet singulier (par exemple, de cet ordinateur) est en effet semblable à un fichier qui fait référence à un objet singulier et contient des informations à propos de cet objet.

Je ne décris pas le mécanisme de liage des propriétés, puisqu'il n'y a pas de consensus à ce sujet. Deux questions intensément débattues sont particulièrement pertinentes dans le présent contexte : i/ la représentation d'un objet singulier peut-elle être indépendante du liage de propriétés (Pylyshyn 1989, 2001) ou bien ne représente-t-on un objet que lorsque l'on lie ensemble certaines propriétés ? ; ii/ le liage des propriétés dépend-il de l'espace ou bien aussi (ou alternativement) d'autres facteurs (Pylyshyn 1989, 2001) ? Une réponse naturelle à ces questions est la suivante : l'identification d'un objet singulier ne précède pas le liage de propriétés dans un fichier, mais n'est autre que la création d'un fichier pour lier des propriétés – représenter un objet singulier, c'est lier ensemble des propriétés ; les propriétés sont liées en fonction de leur cooccurrence en une même région spatiale : le lieu rassemble les propriétés qui ont été détectées au même endroit (Treisman 1992, 171). Alternativement, l'identification

²⁴ Cela n'implique pas que toutes les représentations de propriétés comptées par les systèmes sensoriels servent d'inputs à ce mécanisme. Certaines représentations de propriétés sensorielles peuvent ne pas être utilisées dans la catégorisation (c'est peut-être le cas des représentations motrices).

²⁵ Voir, en particulier, Kahneman et al. 1992 ; Treisman 1992 ; Pylyshyn 1989, 2001 ; Peterson 2000 ; Scholl 2001 ; Scholl et al. 2001.

²⁶ Sur la notion de système perceptif, voir Prinz 2002. La nature et le nombre de systèmes perceptifs ainsi que la nature des propriétés sensorielles sont des questions empiriques.

d'un objet singulier peut être (partiellement ou totalement) indépendante des propriétés détectées dans l'espace environnant l'organisme et le liage des propriétés peut dépendre de facteurs non spatiaux en sus de la location spatiale (Pylyshyn 1989, 2001 ; Scholl 2001 ; Scholl et al. 2001).

4. *Ce modèle décrit sans doute correctement l'attribution des propriétés sensorielles.* Attribuer une certaine structure tri-dimensionnelle à un objet, c'est détecter cette propriété et la lier avec d'autres propriétés sensorielles. En effet, comme je l'ai souligné précédemment, ce modèle est conforme aux résultats empiriques et théoriques de la littérature sur les représentations des objets singuliers et de leurs propriétés sensorielles.

Ensuite, il est possible que certaines propriétés mentionnées dans les théories de la catégorisation soient conformes à l'hypothèse empiriste, autrement dit puissent être réduites à des propriétés sensorielles.

En revanche, l'hypothèse empiriste, selon laquelle l'attribution de toutes les propriétés se réduit à la détection et au liage des propriétés sensorielles, ne peut être acceptée. Il n'est pas vrai que toutes les propriétés mentionnées par les théories de la catégorisation ne sont qu'une manière de parler des propriétés sensorielles. En particulier, un certain nombre de propriétés semblent pouvoir être attribuées sur la base d'un nombre indéfini d'autres propriétés, comme par exemple *émotionnellement perturbé* (Wisniewski et Medin 1994, 277). Donc par principe, une réduction à des propriétés sensorielles est impossible. C'est aussi le cas des propriétés fonctionnelles, comme par exemple *éclairé*. Ces propriétés correspondent à un nombre indéfini de propriétés sensorielles, puisque les objets qui éclairent ont un nombre indéfini de formes.

Ainsi, le modèle présenté ci-dessus caractérise correctement l'attribution des propriétés sensorielles, mais non de toutes, dans la mesure où il n'est pas vrai que toutes les propriétés peuvent être réduites aux propriétés sensorielles. En conséquence, en sus du modèle présenté ci-dessus, d'autres processus sont requis pour réaliser l'attribution des propriétés. L'attribution de propriétés n'est donc pas homogène.

En résumé, l'hypothèse empiriste soutient qu'attribuer des propriétés, c'est en général détecter des propriétés sensorielles et les lier. Cela suppose que les propriétés mentionnées dans les théories de la catégorisation ne devraient être que des manières de parler des propriétés sensorielles. Cependant une telle réduction est extrêmement douteuse. Puisque toutes les propriétés ne sont pas sensorielles et puisque l'attribution de propriétés sensorielles semble reposer sur des processus sui generis, l'attribution de propriétés repose sur plusieurs processus.

Troisième section : catégorisation et attribution

L'examen de l'hypothèse empiriste offre une première raison de douter de la thèse de l'homogénéité : l'attribution de propriétés sensorielles semble requérir des processus *sui generis*, dès lors que toutes les propriétés ne sont pas sensorielles. Cela suggère qu'il existe plusieurs processus réalisant l'attribution de propriétés. Dans cette section, j'examine un autre aspect de l'attribution des propriétés.

1. Le point essentiel est le suivant. *Pour nombre de propriétés que mentionnent les psychologues lorsqu'ils illustrent leurs théories de la catégorisation, l'attribution de ces propriétés ne se distingue pas de la catégorisation de la cible dans des catégories correspondant à ces propriétés.* Pour prendre un exemple caricatural (voir 2. pour de véritables exemples), l'attribution à la cible de la propriété *a des ailes* permet sa catégorisation dans la catégorie *OISEAUX*. Mais en quoi l'attribution de cette propriété est-

elle différente de la catégorisation de la cible dans la catégorie *ANIMAUX AILES* ? De toute évidence, les mécanismes qui réalisent la catégorisation (section un) sont parfaitement à même de catégoriser une cible dans la catégorie *ANIMAUX AILES*, s'il y a un concept (prototype, exemplaire ou théorie) de cette catégorie.

2. J'illustre ce problème par trois exemples tirés de la littérature psychologique.

Exemple 1. Considérons d'abord le cas de la catégorisation d'un objet dans la catégorie *CAROTTE*. Smith et al. ont proposé la caractérisation suivante du concept *CAROTTE* (d'après Smith et al. 1988) :

Attributs	Valeurs	
Couleur 0,75	Orange	25
	Rouge	2
Forme 0,5	Longue	20
	Fine	10
Texture 0,3	Croquante	15
	Dure	5
Comment est-elle mangée ? 0,3	En salade	25
	Crue	5

Le concept est représenté au moyen i/ d'attributs (par exemple, couleur), qui représentent des types de propriétés, ii/ de pondérations des attributs en fonction de leur importance, iii/ de valeurs, qui sont les propriétés elles-mêmes (par exemple, orange), et iv/ de votes (par exemple, 25) qui représentent la fréquence d'une propriété donnée parmi les référents du concept.

Le processus de catégorisation est supposé être le suivant. L'attributeur attribue un ensemble de propriétés, par exemple *orange, fin, dur*, à un objet indépendamment identifié. L'output des processus qui constituent l'attributeur sert d'input au catégorisateur. Sur la base du concept, le catégorisateur classe la cible. Maintenant, considérons de plus près les propriétés qui doivent être attribuées à la cible afin qu'elle soit catégorisée. Le problème soulevé est le suivant : l'attribution de la propriété *être mangé en salade* à la cible peut être très naturellement décrite comme une catégorisation de cette même cible dans la catégorie *OBJETS QUI SE MANGENT EN SALADE*. Les processus suggérés de la catégorisation (processus de comparaison, vérification des règles, inférence à la meilleure explication) semblent pouvoir rendre compte de la catégorisation de la cible dans cette dernière catégorie. Supposons par exemple que l'explication de la catégorisation au moyen de prototypes est correcte. Il se peut que nous possédions un prototype des objets qui se mangent en salade sur la base duquel nous classifions la cible dans la catégorie *OBJETS QUI SE MANGENT EN SALADE*. Attribuer la propriété *être mangé en salade* à la cible n'est alors rien d'autre que la catégoriser dans la catégorie, *OBJETS QUI SE MANGENT EN SALADE*, selon le processus normal de catégorisation.

Exemple 2. Un autre exemple illustre ce problème. De manière informelle, Murphy et Medin (1985) illustrent la catégorisation fondée sur l'explication avec l'exemple du convive ivre : sur la base de l'attribution de la propriété dynamique *se jeter dans une piscine tout habillé*, nous catégorisons le convive dans la catégorie *PERSONNES IVRES*. Murphy et Medin supposent donc que le processus par lequel nous tenons quelqu'un pour ivre est un processus de catégorisation. Or *être ivre* peut figurer à titre de propriété dans l'explication de nombreuses catégorisations, par exemple dans la catégorisation d'un individu dans la

catégorie *PERSONNES VIOLENTES*. De nouveau, la catégorisation et l'attribution ne se distinguent pas.

Dans les exemples 1 et 2, l'attribution de la propriété à la cible ne se distingue pas aisément de la catégorisation de la cible dans une autre catégorie. D'autres cas sont problématiques, comme le montre l'exemple 3.

Exemple 3. Wisniewski et Medin demandent aux sujets de catégoriser des dessins d'enfants au sein de différentes catégories (dessins faits par des enfants créatifs vs. non créatifs, par exemple) et d'expliquer les raisons de leur catégorisation (Wisniewski et Medin 1994). L'une des raisons avancées pour catégoriser un dessin dans la catégorie *DESSINS FAITS PAR UN ENFANT URBAIN* (vs. *CAMPAGNARD*) est que la personne dessinée porte un vêtement urbain (Wisniewski et Medin 1994, 262). Autrement dit, la propriété *contient un personnage qui porte un vêtement urbain* est attribuée au dessin. Mais cette attribution d'une propriété se distingue mal de la catégorisation d'une partie du dessin, à savoir du vêtement de l'un des personnages, dans la catégorie *VÊTEMENT URBAIN*. De nouveau, l'attribution de propriétés semble reposer sur les processus de catégorisation. A la différence des deux cas précédents, cependant ce n'est pas la cible elle-même qui est catégorisée (ce n'est pas le dessin qui est catégorisé), mais certaines propriétés de la cible, à savoir la forme de certaines parties de la cible. Cela suppose que certaines propriétés ont déjà été attribuées à la cible (formes, division en parties articulées...). Ces propriétés sont catégorisées (telle forme de telle partie est un vêtement urbain) : c'est là apparemment l'un des processus consistant à attribuer une propriété à la cible (ici le dessin). Sur la base de ces catégorisations, la cible est catégorisée. Bref, dans les deux premiers exemples, l'attribution de propriétés à la cible est identique à sa catégorisation dans des catégories, tandis que dans le troisième exemple, l'attribution de propriété est identique à la catégorisation de certaines propriétés de la cible. Dans les deux cas, mais de manière différente, l'attribution ne semble pas être une capacité cognitive différente du catégorisateur.

3. *S'il est vrai que l'attribution de certaines propriétés à une cible consiste à catégoriser cette cible dans les catégories correspondant à ces propriétés, l'attribution de propriétés ne peut être une capacité homogène.* Et ce pour deux raisons. Tout d'abord, si la catégorisation repose sur l'attribution de propriétés et si l'attribution de toutes les propriétés est un processus de catégorisation, il y a là une régression à l'infini. Donc l'attribution de *certaines* propriétés n'est pas un processus de catégorisation. Donc plusieurs processus réalisent l'attribution des propriétés. Supposons que je catégorise une cible dans la catégorie *OISEAUX* parce qu'elle a des ailes. Par hypothèse, l'attribution de la propriété *est ailée* est elle-même une catégorisation. D'après les théories de la catégorisation, cette catégorisation suppose l'attribution de propriétés à la cible, par exemple une certaine forme. Si l'attribution de cette propriété est elle-même une catégorisation, elle repose sur l'attribution d'autres propriétés. Sous peine de régression à l'infini, l'attribution de certaines propriétés doit être différente de la catégorisation. Deuxièmement, le modèle empiriste de l'attribution de propriétés présenté dans la deuxième section est relativement convaincant pour l'attribution des propriétés sensorielles. Il est probable que leur attribution consiste en la détection et le liage des propriétés sensorielles. Mais les mécanismes de détection et de liage ne sont pas identiques aux mécanismes de la catégorisation.

En résumé, certaines attributions de propriétés semblent être des catégorisations. Deux cas ont été distingués : dans certains cas, attribuer une propriété à une cible, c'est catégoriser la cible dans une catégorie correspondant à la propriété tandis que dans d'autres, c'est catégoriser certaines propriétés de la cible. Sous peine de régression à l'infini, l'identité de

l'attribution de certaines propriétés et de la catégorisation implique que l'attribution d'autres propriétés ne soit pas un processus de catégorisation. Donc l'attribution de propriétés n'est pas fonctionnellement une.

Quatrième section : un modèle de la catégorisation sans attributeur

Le modèle que je souhaite présenter rejette la thèse de l'homogénéité : l'attribution de propriétés à une cible n'est pas une capacité homogène. Au contraire, cette notion recouvre différentes capacités cognitives. Parmi celles-ci, il faut inclure les processus de catégorisation, conformément aux exemples donnés dans la troisième section. En outre, l'attribution de propriétés sensorielles repose sur le modèle empiriste présenté dans la deuxième section.

Dans cette section, je présente d'abord le schéma décrivant cette hypothèse ; dans un second temps, j'en commente les principaux éléments ; finalement, je propose quelques considérations en faveur de ce modèle.

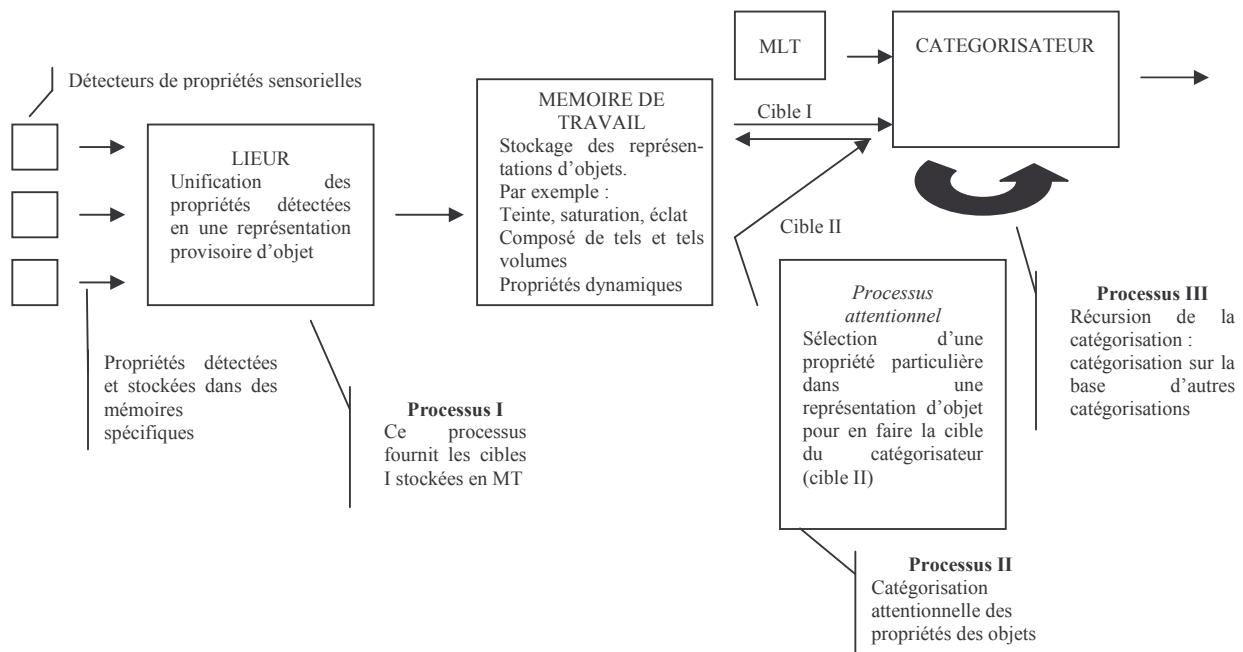


Schéma 3 : la catégorisation sans attributeur

1. Le schéma 3 de la catégorisation propose que l'attribution de propriétés à une cible recouvre *trois* processus différents qui réalisent trois capacités différentes²⁷. La première est le liage des représentations de propriétés sensorielles dans la représentation d'un objet singulier. La deuxième est la sélection d'une propriété dans une représentation d'un objet singulier par des processus attentionnels *top-down*²⁸, sa catégorisation comme étant telle propriété et le stockage de sa catégorie dans cette représentation. Le troisième processus est le processus de

²⁷ Ce schéma est trop simple pour rendre compte de la catégorisation des événements et des situations, puisque celles-ci impliquent plusieurs objets. Il faudrait ajouter un mécanisme, qui groupe les parties de la scène visuelle en unités qui seraient ensuite catégorisées. Ce mécanisme pourrait opérer sur les outputs des lieux. Pour ne pas rendre le schéma encore plus compliqué et spéculatif, j'ai préféré négliger cet aspect important de la catégorisation.

²⁸ Un processus est *top-down* s'il affecte le cours de processus cognitifs situés en amont, c'est-à-dire plus proches des stimulations sensorielles.

catégorisation lui-même : dans certains cas, attribuer une propriété, ce n'est rien d'autre que catégoriser la cible. Examinons de plus près ces différentes capacités.

A/ Le liage des propriétés en une représentation d'objet (processus I). Ce premier processus correspond au modèle décrit dans la deuxième section. Les systèmes perceptifs détectent et représentent les propriétés sensorielles. Les représentations produites par les différents systèmes perceptifs servent d'inputs à un mécanisme de liage, le lieu. Trois points importants doivent être précisés. *Tout d'abord, ces mécanismes ne rendent compte que de l'attribution de certaines propriétés.* Cela est conforme à la thèse de l'hétérogénéité. Ensuite, dans la mesure où *certaines propriétés non sensorielles peuvent peut-être se réduire à des propriétés sensorielles*, - leur attribution peut relever de ces mêmes mécanismes. Autrement dit, l'hypothèse empiriste a un champ d'application assez large. Enfin, *le liage des propriétés dans la représentation d'un objet singulier est fort différent des processus de catégorisation postulés par les théories de la catégorisation.* En effet, le liage ne repose pas sur des propriétés attribuées antérieurement à l'objet, contrairement à tous les processus supposés de catégorisation (section 1) ; au contraire, l'attribution consiste à associer des propriétés les unes avec les autres (sans doute en partie sur la base de leur localisation). En conséquence, même si l'attribution de certaines propriétés est une catégorisation, il n'y a pas de régression à l'infini.

L'output du lieu (processus I) est le premier type d'inputs que reçoit le catégorisateur (cible I). Dans ce cas, la catégorisation porte sur une cible singulière. Elle est catégorisée sur la base des propriétés qui composent le fichier. Eventuellement, des processus *top-down* peuvent moduler l'importance des propriétés qui constituent le fichier ou sélectionner certaines d'entre elles²⁹.

B/ la catégorisation des propriétés de la cible (processus II). Le deuxième type de processus que recouvre l'attribution de propriétés est le suivant : on catégorise une propriété antérieurement attribuée à la cible. En effet, dans certains cas, nous classifions certaines cibles sur la base de la classification préalable de (certaines de) leurs propriétés. Un exemple intuitif suggère ce point : il peut arriver que nous classifiions un objet dans telle espèce de pomme parce que nous lui avons attribué une couleur spécifique. Nous classifions une pomme dans la catégorie *CANADA* parce que nous lui avons attribué la propriété *rouge sanguin*. Ma suggestion est que l'attribution de la propriété *rouge sanguin* correspond à la catégorisation de certaines propriétés stockées dans le fichier mental de cette pomme, par exemple, la possession d'une certaine teinte, d'une certaine saturation et d'un certain éclat. Une propriété, par exemple ce triplet, est prise comme cible et catégorisée comme étant telle propriété, par exemple *rouge sanguin*.

Les mécanismes qui rendent compte de la catégorisation des objets s'appliquent tout naturellement à la catégorisation des propriétés. De fait, les explications de la catégorisation au moyen de prototypes sont nées des études sur la catégorisation des couleurs (Heider 1972³⁰). Je fais donc l'hypothèse que les mécanismes de catégorisation ne sont pas dédiés à la catégorisation des objets physiques, mais à la catégorisation en général. Ils s'appliquent aux propriétés, aux actions, aux événements... Les propriétés fournissent donc le second type d'inputs que reçoit le catégorisateur (cible II).

Le deuxième type de processus pourrait donc être le suivant : attribuer une propriété, c'est i/ sélectionner attentivement des propriétés stockées dans la représentation d'un objet singulier ; ii/ les catégoriser au moyen des processus impliqués dans la catégorisation ; et iii/ stocker la

²⁹ Lorsque l'objet n'est pas perçu, mais décrit verbalement, on peut émettre l'hypothèse qu'une représentation d'objet est produite.

³⁰ Notez que Heider et Rosch sont une seule et même personne.

catégorie de la propriété dans la représentation de l'objet (d'où la flèche qui revient du catégorisateur vers la mémoire de travail dans le schéma 3). Sur la base des propriétés représentées dans le fichier d'objet, y compris la catégorie de la propriété, l'objet est ensuite catégorisé. Le processus II permet donc d'obtenir un fichier d'objet (en d'autres termes, une description d'un objet singulier) différent de celui qui est obtenu à partir des seuls processus sensoriels.

L'attribution de la propriété *rouge sanguin* à une pomme peut être expliquée par ce type de mécanisme. C'est aussi le cas de l'exemple 3 de la troisième section (attribution de la propriété *contient un personnage qui porte un uniforme urbain* au dessin à catégoriser). Ce mécanisme est aussi en mesure d'expliquer la catégorisation des objets sur la base de l'attribution d'actions, de mouvements, etc.. Les propriétés dynamiques sont sélectionnées par l'attention, puis catégorisées dans telle ou telle catégorie de mouvements (d'actions...); finalement, les catégories des propriétés dynamiques sont stockées dans les représentations d'objets (Cavanagh et al. 2001). De manière générale, attribuer des propriétés abstraites (voir Wisniewski et Medin 1994) repose peut-être sur un processus de type II.

Je fais l'hypothèse (spéculative) que la catégorisation des propriétés, et donc le deuxième type d'attribution de propriétés, suppose un processus attentionnel. Néanmoins, le processus proposé pourrait être reformulé indépendamment de cette hypothèse : les propriétés stockées dans les représentations des objets singuliers pourraient être automatiquement sélectionnées par des routines préattentionnelles, puis catégorisées³¹. Le processus serait alors le suivant : i/ des routines préattentionnelles sélectionnent des propriétés dans les représentations des objets singuliers ; ii/ ces propriétés sont catégorisées par le catégorisateur ; iii/ les catégories des propriétés sont stockées dans les représentations d'objets. Néanmoins, certaines données empiriques récentes suggèrent que la catégorisation des mouvements, donc des propriétés dynamiques contenues dans les représentations des objets singuliers, requièrent l'attention (Cavanagh et al. 2001), ce qui est plutôt consistant avec le mécanisme suggéré initialement.

C/ La récursivité de la catégorisation de la cible (processus III). Enfin, le dernier type de processus que recouvre l'attribution de propriétés est la catégorisation de la cible elle-même. Autrement dit, les attributions de propriétés à la cible sont dans certains cas des catégorisations de la cible. Sur la base des catégories dans lesquelles la cible a été classée, la cible est de nouveau catégorisée. Par exemple, l'attribution de la propriété *ivre* à un individu n'est autre que la catégorisation de cet individu dans la catégorie *INDIVIDUS IVRES*. Sur la base de cette catégorisation, je peux catégoriser cet individu dans la catégorie *INDIVIDUS DANGEREUX*.

Je fais donc l'hypothèse que la catégorisation est *réursive* (ce qu'exprime la flèche en gras qui va de l'attributeur à l'attributeur dans le schéma 3). Les cibles peuvent alors être de type I comme de type II. Il est possible de catégoriser un objet singulier, puis de catégoriser cet objet sur la base de cette catégorisation ; il est aussi possible de catégoriser une propriété sur la base de sa catégorisation dans une autre catégorie et de stocker la catégorie finale dans le fichier d'objet.

Les exemples 1 et 2 de la troisième section peuvent être expliqués par ce type de mécanismes. Par exemple, un légume peut être catégorisé comme une carotte parce que la propriété *est mangé en salade* lui est attribué. Cette attribution correspondrait à la catégorisation de la cible dans la catégorie *OBJETS QUI SE MANGENT EN SALADE*. De manière récursive, l'objet est catégorisé dans la catégorie *CAROTTE* sur la base d'une catégorisation antérieure.

³¹ De même, la catégorisation des objets ne requiert l'attention que si le liage est un processus attentionnel. Si le liage est réalisé par des processus préattentionnels, les objets peuvent être catégorisés indépendamment des processus attentionnels. Le rôle de l'attention peut aussi dépendre de la nature des cibles.

2. Pourquoi admettre ce modèle de la catégorisation ? De toute évidence, ce modèle est à la fois schématique et spéculatif. En effet, il ne rend pas compte de toutes les catégorisations (particulièrement de la catégorisation des événements) et ne spécifie pas les différents processus.

Néanmoins, quatre points supportent cette proposition. Premièrement, les éléments qui composent ce modèle sont consistants avec, et même inspirés par, des théories cognitives largement supportées par des données empiriques. C'est en particulier le cas des processus de type I, le liage des propriétés sensorielles. J'ai soutenu en effet que c'est l'une des capacités que recouvre l'attribution des propriétés. Or l'existence de processus liant ces propriétés est conforme aux résultats empiriques et théoriques des théories de la perception. De même, l'hypothèse selon laquelle le catégorisateur consiste en des mécanismes qui ne s'appliquent pas uniquement aux objets, mais vaut pour toutes les catégorisations, est confortée par les résultats de Rosch (Heider 1972).

Deuxièmement, les processus de type II doivent être postulés pour la catégorisation des propriétés et des actions. Par exemple, nous catégorisons les mouvements (en général, les propriétés dynamiques) comme étant telle ou telle action, et telle couleur perçue comme étant telle nuance de mauve. Cela suppose l'identification d'une propriété et sa catégorisation. J'émetts simplement l'hypothèse qu'attribuer des propriétés, c'est dans certains cas catégoriser d'autres propriétés.

Troisièmement, ce modèle fournit une explication satisfaisante de certains caractères de l'attribution de propriétés. En particulier, Wisniewski et Medin ont noté que les sujets interprétaient différemment les dessins d'enfants qu'ils étaient supposés catégoriser (Wisniewski et Medin 1994, 266) : par exemple, ils ne caractérisaient pas de la même manière les vêtements dessinés. Selon le schéma 3, de nombreuses attributions reposent sur les processus de catégorisation (processus de type II et III). Or, la variabilité inter- et intra-individuelle de la catégorisation des entités qui ne satisfont pas clairement un concept (les entités atypiques) est une caractéristique bien connue (McCloskey et Gluksberg 1978 ; Hampton 1998). Puisque les sujets doivent attribuer des propriétés à des dessins d'enfants, la caractéristique rapportée par Wisniewski et Medin est prédite par le modèle suggéré ici : si les dessins sont schématiques (ce qu'ils sont, si l'on se fie à leur reproduction dans Wisniewski et Medin 1994), ils sont des exemples d'entités atypiques. Donc si l'attribution de propriétés à ces dessins repose sur les mécanismes de catégorisation, l'instabilité mentionnée est prévisible.

Finalement, ce modèle n'est pas simplement *ad hoc*, il implique certaines prédictions. En particulier, puisque les mécanismes de la catégorisation jouent un rôle central dans l'attribution des propriétés, les caractéristiques de la catégorisation mises au jour par les théories de la catégorisation devraient être présentes dans un certain nombre d'attributions de propriétés. En particulier, on devrait trouver des degrés de typicalité dans la possession d'une propriété.

Conclusion

Les théories contemporaines de la catégorisation distinguent deux capacités cognitives différentes, la catégorisation et l'attribution de propriétés, puisque la catégorisation repose dans tous les modèles proposés sur les propriétés attribuées aux cibles. Or, en vertu de la méthodologie proposée par Marr pour les sciences cognitives, les capacités doivent être fonctionnellement unes. Cette contrainte s'applique à la capacité postulée d'attribuer des propriétés. J'ai donc examiné les deux hypothèses suivantes :

A/ **La thèse de l'homogénéité** : la capacité d'attribuer des propriétés est fonctionnellement une.

B/ **La thèse de l'hétérogénéité** : la capacité d'attribuer des propriétés est un conglomerat hétéroclite de capacités cognitives fort différentes. Aucune capacité de l'esprit n'a pour fonction d'attribuer des propriétés. Au contraire, des capacités variées ont des fonctions différentes confondues par la notion d'attribution de propriétés.

Je me suis efforcé de montrer que **la thèse de l'homogénéité est fausse** et qu'elle est avantageusement remplacée par un modèle de l'attribution qui ne la suppose pas. Ce modèle caractérise l'attribution de propriétés de la manière suivante :

1/ Attribuer certaines propriétés, c'est détecter des propriétés sensorielles et les lier. Certaines propriétés apparemment non sensorielles peuvent se réduire à des propriétés sensorielles.

2/ Attribuer certaines propriétés, c'est catégoriser d'autres propriétés. Les propriétés abstraites et les actions sont sans doute attribuées de cette manière.

3/ Attribuer certaines propriétés à une cible, c'est catégoriser la cible elle-même.

Bibliographie

- Barsalou, L. W., 1999, Perceptual symbol systems, *Behavioral and Brain Sciences*, 22, 577-609.
- Brooks, L. 1978, Nonanalytic concept formation and memory for instances, in E. Rosch et B. Lloyd, éd., *Cognition and Categorization*, Hillsdale, New Jersey : Erlbaum.
- Bruner, J., Goodnow, J., et Austin, G., 1956, *A Study of Thinking*, New York : John Wiley et Sons.
- Cavanagh, P., Labianca, A., et Thornton, I., 2001, Attention-based visual routines : sprites, *Cognition*, 80, 47-60.
- Clark, A., 1993, *Associative Engines, Connectionism, Concepts, and Representational Change*, Cambridge, Mass.: MIT Press.
- Goldstone, R., et Barsalou, L. W., 1998, Reuniting perception and conception, *Cognition*, 65, 231-262.
- Hahn, U., et Chater, N., 1997, Concepts and Similarity, in K. Lamberts and D. Shanks, éd., 1997.
- Hampton, J. A., 1993, Prototypes models of concept representation, in I. van Mechelen, J. A. Hampton, R. S. Michalski, P. Theuns, éd., *Categories and concepts : theoretical views and inductive data analysis*, London : Academic Press.
- Hampton, J. A., 1997, Psychological Representation of Concepts, in M. A. Conway, éd., *Cognitive Models of Memory*, Cambridge, Mass. : MIT Press.
- Hampton, J. A., 1998, Similarity-Based Categorisation and Fuzziness of Natural Categories, *Cognition*, 65, 137-165.
- Hampton, J. A., 2001, The role of similarity in natural categorization, in U. Hahn and M. Ramscar, éd., *Similarity and Categorization*, Oxford : OUP.
- Heider, E. R., 1972, Universals in color naming and memory, *Journal of Experimental Psychology*, 93, 10-20.
- Heit, E., 1997, Knowledge and Concept Learning, in K. Lamberts and D. Shanks, éd., 1997.
- Heit, E., 2001, Background Knowledge and models of categorization, in U. Hahn and M. Ramscar, éd., *Similarity and Categorization*, Oxford : OUP.
- Hull, C. L., 1920, Quantitative aspects of the evolution of concepts, *Psychological Monographs*, 28, 1.
- Kahneman, D., Treisman, A., et Gibbs, B. J., 1992, The Reviewing of Object Files: Object-Specific Integration of Information, *Cognitive Psychology*, 24, 175-219.
- Kant, E., 2001, *La Critique de la raison pure*, traduction A. Tremesaygues et B. Pacaud, sixième éd., Paris : PUF.

- Komatsu, L. K., 1992, Recent Views of Conceptual Structure, *Psychological Review*, 112, 3, 500-526.
- Kosslyn, S. M., 1978, Imagery and Internal Representation, in E. Rosch and B. B. Lloyd, eds., *Cognition and Categorization*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lamberts, K., et Shanks, D., eds., 1997, *Knowledge, Concepts and Categories*, Cambridge, Mass. : MIT Press.
- Lyons, J. C., 2001, Carving the Mind at its (Not Necessarily Modular) Joints, *The British Journal for the Philosophy of Science*, 52, 277-302.
- Machery, E., 2002, Compositionnalité, Composition Conceptuelle et Combinaison Prototypique, *In Cognito, Revue francophone internationale en Sciences Cognitives*, 22, 5-26.
- Machery, E., manuscript a, Concepts are Not a Natural Kind.
- Machery, E., manuscript b, Concept Theories : Old Ideas in Disguise ?.
- Margolis, E., 1994, A Reassessment of the Shift from the Classical Theory of Concepts to Prototype Theory, *Cognition*, 51, 73-89.
- Margolis, E., et Laurence, St., eds., 1999a, *Concepts, Core Readings*, Cambridge, Mass. : MIT Press.
- Margolis, E., et Laurence, St., 1999b, Concepts and Cognitive Science, in E. Margolis et St. Laurence, 1999a.
- Marr, D., 1982, *Vision*, San Francisco: Freeman.
- McCloskey, M., et Glucksberg, S., 1978, Natural categories: Well-defined or fuzzy sets?, *Memory and Cognition*, 6, 462-472.
- Medin, D. L., 1989, Concepts and Conceptual Structure, *American Psychologist*, 44, 1469-1481.
- Medin, D. L., et Shaffer, M. M., 1978, A context theory of classification learning, *Psychological Review*, 85, 207-238.
- Murphy, G. L., 2002, *The Big Book of Concepts*, Cambridge, Mass: MIT Press.
- Murphy, G. L., et Medin, D. L., 1985, The role of Theories in Conceptual Coherence, *Psychological Review*, 92, 904-919.
- Nosofsky, R. M., 1988, Similarity, frequency and category representations, *Journal of Experimental Psychology : Learning, Memory and Cognition*, 14, 54-65.
- Olseth, K. L., et Barsalou, L. W., 1995, The spontaneous use of perceptual representations during conceptual processing, dans les *Proceedings of the seventeenth annual meeting of the Cognitive Science Society*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Peacocke, Ch., 1986, Explanation in computational psychology: Language, perception, and level 1.5, *Mind and Language*, 1, 2, 173-193.
- Peterson, M. A., 2000, Object Perception, in E. B. Goldstein, éd., *Blackwell Handbook of Perception*, Malden, Mass. : Blackwell Publishers, Inc.
- Pinker, St., et Prince, A., 1999, The Nature of Human Concepts : Evidence from an Unusual Source, in R. Jackendoff, P. Bloom et K. Wynn, eds., *Language, Logic and Concepts*, Cambridge, Mass. : MIT Press.
- Prinz, J. J., 2002, *Furnishing the mind: Concepts and their perceptual basis*, Cambridge, MA: MIT Press.
- Pylyshyn, Z. W., 1989, The role of location indexes in spatial perception : a sketch of the FINST spatial index model, *Cognition*, 32, 65-97.
- Pylyshyn, Z. W., 2001, Visual indexes, preconceptual objects, and situated vision, *Cognition*, 80, 127-158.
- Scholl, B. J., 2001, Objects and attention : the state of art, *Cognition*, 80, 1-46.
- Scholl, B. J., Pylyshyn, Z. W., Feldman, J., 2001, What is a visual object ? Evidence from target merging in multiple object tracking, *Cognition*, 80, 159-177.

- Shanks, D. R., 1997, Representation of categories and concepts in memory, in M. A. Conway, éd., *Cognitive Models of Memory*, Cambridge, Mass. : MIT Press.
- Smith, E. E., et Medin, D. L., 1981, *Categories and Concepts*, Cambridge, Mass. : Harvard UP.
- Solomon, K. O., et Barsalou, L. W., 1997, Productivity and propositional construal as the meshing of embodied representations, *Behavioral and Brain Sciences*, 20, 38-39.
- Treisman, A., 1992, L'attention, les traits et la perception des objets, in D. Andler, éd., *Introduction aux sciences cognitives*, Paris : Gallimard.
- Treisman, A., et Gelade, G., 1980, A feature integration theory of attention, *Cognitive Psychology*, 12, 97-136.
- Tulving, E., 1995, Organization of memory: Quo vadis?, in M. S. Gazzaniga, éd., *The Cognitive Neurosciences*, Cambridge, Mass.: MIT Press.
- Wisniewski, E. J., et Medin, D. L., 1994, On the interaction of theory and data in concept learning, *Cognitive Science*, 18, 221-281.