

SIMILARITE ET CATEGORISATION

Titre anglais : **Similarity and categorization**

THIBAUT Jean-Pierre^a

Adresse: Université de Liège
Faculté de Psychologie
5, boulevard du Rectorat. Bât 32.
4000 LIEGE
Email: jpthibaut@ulg.ac.be

RESUME

La présente revue de question porte sur les modèles de catégorisation et d'estimation de la similarité. Elle se divise en deux parties. La première traite des modèles fondés sur les notions d'attribut (ou trait) et de dimension. Ces deux types d'unités fondent deux grandes classes d'algorithmes souvent opposés dans la littérature: les modèles en traits et les modèles en dimensions. Dans les représentations dimensionnelles, chaque entité se définit comme un point dans un espace multidimensionnel et la similarité entre deux entités est une fonction de la distance entre ces points. Dans les représentations en traits, une entité est définie en termes de traits et la similarité entre deux entités est une fonction de leurs attributs communs et distinctifs.

La seconde partie porte sur des travaux qui envisagent plus particulièrement comment les sujets déterminent l'espace des traits dans lequel se calcule la similarité. On décrit comment les jugements de similarité sont influencés par le contexte et les théories des sujets portant sur les entités à comparer (ou à catégoriser). On envisage le processus d'alignement par lequel les sujets mettent les entités en correspondance et choisissent les traits qui seront utilisés dans leur comparaison. Enfin, on analyse les travaux qui, récemment, ont distingué catégorisation et jugement de similarité. On propose une conception reposant sur la notion d'espace de traits qui permet de conserver une vue unitaire de ces deux processus tout en expliquant les dissociations obtenues.

Summary :

The present paper is a review of the literature dealing with current models of categorization and similarity estimation. The first part reviews models of concept representations based on notions such as features and dimensions. In the case of dimensional representations, each stimulus is defined as a point in a multidimensional space and the similarity between two stimuli is a function of the distance between the corresponding points in the space. In the case of feature representations, a stimulus is defined in terms of features and the similarity between two stimuli is a function of their common and distinctive features.

In the second part, we analyze the construction of the feature space that subjects refer to when they have to categorize a stimulus or estimate the similarity between two stimuli. We examine the influence of context and of naive theories on the estimation of similarity. The process of alignment is analyzed. We present a unified view of similarity and categorization based on the notion of a feature space while explaining the dissociations obtained between similarity and categorization.

Mots clefs : similarité, catégorisation, concepts, théories.

Key words : similarity, categorization, concepts, theories.

1. INTRODUCTION

Estimer la similarité¹ entre deux ou plusieurs entités est un aspect central de la cognition. La similarité joue un "rôle fondamental dans les théories de la connaissance et du comportement. C'est un principe organisateur par lequel les individus classent les objets, forment des concepts et généralisent" (Tversky, 1977, p.327). Toute tâche de catégorisation ou de reconnaissance d'un objet donné, peut se décrire, en partie, en termes d'une estimation de la similarité entre cet objet et les représentations conceptuelles du sujet. De la même manière, l'interprétation des analogies nécessite également la saisie des aspects communs des entités mises en relation. L'estimation de la similarité repose sur deux facteurs au moins: la fixation d'un espace de traits², ou de dimensions, dans lequel la similarité est calculée, et le calcul de la similarité proprement dit. L'organisation du présent article est fondée sur la description de ces deux étapes. Dans la première partie, on présente les modèles de calcul de la similarité. La seconde envisage le processus de construction de l'espace de traits dans lequel s'effectue le calcul de la similarité ainsi que les variables qui influencent cette construction.

Les conceptions de la similarité peuvent être placées entre deux points de vue, réaliste et subjectiviste. Le premier présuppose que la similarité repose sur les propriétés objectives des entités qui sont comparées. Dans cette perspective, la similarité entre moineaux et mésanges est un donné de la réalité, indépendant de l'observateur: c'est cette similarité objective qui nous fait percevoir ces deux oiseaux comme semblables. Le point de vue subjectiviste postule que la similarité est flexible, susceptible de varier avec les contextes et l'attention accordée à chacun des traits partagés par les entités comparées. Les modèles qui envisagent la similarité uniquement sous la forme d'un calcul opéré sur un ensemble d'unités données *a priori* aux sujets sont proches des conceptions réalistes. Au contraire, les conceptions qui accordent une place au processus de sélection des traits et aux connaissances utilisées par les sujets pour réaliser le jugement se rapprochent des conceptions subjectivistes.

Les estimations de la similarité sont obtenues directement ou indirectement. Elles sont directes dans les jugements de similarité et de différence. Ces jugements sont "sans critère" lorsqu'on demande de coter la similarité (ou la différence) entre deux entités sans spécifier aucun critère de comparaison (e.g., "dans quelle mesure un cerf et une girafe se

ressemblent?"). Ils sont "critériés" lorsqu'on compare deux entités en termes d'un paramètre donné (e.g., "quel est l'animal le plus grand, le cerf ou la girafe?"). Les estimations sont indirectes lorsqu'on les dérivent de résultats obtenus à des épreuves de classement libre dans lesquelles le sujet doit former des catégories en choisissant les critères, ou de résultats obtenus à des tâches d'identification où l'on analyse les stimuli avec lesquels un stimulus est confondu, ou encore lorsqu'on les dérivent de temps de reconnaissance ou de gradients de généralisation des stimuli. Ces indices de similarité ne sont pas toujours corrélés. Ainsi les sujets jugent les chiffres 6 et 9 plus semblables entre eux que 6 et 8, alors que, dans une tâche de reconnaissance, ils confondent plus souvent 6 avec 8 que 6 avec 9. Le jugement de similarité s'explique par l'identité des chiffres 6 et 9, à une rotation près alors que les confusions entre 6 et 8 s'expliquent par le nombre d'attributs perceptifs partagés par ces 2 chiffres. C'est la différence dans les unités utilisées pour les deux tâches qui expliquent cette discordance.

Les descriptions de la similarité reposent principalement sur deux paramètres: (a) la définition et la sélection des caractéristiques des stimuli impliquées dans l'estimation de la similarité. Ces caractéristiques constituent l'espace des descripteurs (dimensions ou traits) dans lequel l'estimation est réalisée (Thibaut & Schyns, 1995); (b) la quantification de la similarité qui peut être modélisée sous forme d'algorithmes. La première partie de la présente contribution porte sur les modèles qui décrivent comment sont combinées les unités (dimensions ou attributs) utilisées par les sujets dans leur estimation de la similarité entre deux entités. Dans ces travaux, le processus de sélection des unités impliquées dans le jugement de similarité reste implicite ainsi que nous le verrons dans les modèles exposés au point 2. Actuellement, cependant, un grand nombre d'auteurs insistent sur la nécessité d'étudier les processus, non triviaux, de sélection des unités: leur thèse est que le calcul de la similarité ne serait qu'un sous-produit des processus de sélection des attributs et des connaissances utilisées pour cette sélection (voir entre autres, Medin, Goldstone, & Gentner, 1993; Murphy & Medin, 1985; Thibaut & Schyns, 1995). Les sujets qui estiment la similarité n'utilisent qu'un sous-ensemble de leurs connaissances sur les entités comparées. Ainsi, un jugement de similarité entre une pomme et une poire, en l'absence de contexte de comparaison spécifié, s'effectuera

par défaut à l'intérieur du champs des fruits, utilisant les propriétés de cette catégorie ou des critères perceptifs immédiats des fruits comparés (couleur, forme). Il est probable que la comparaison de ces fruits ne tiendra pas compte des informations liées au prix de ces deux fruits, ou au fait qu'ils sont utilisés dans la fabrication de certains alcools ou encore d'autres informations qui peuvent devenir pertinentes dans certains contextes (le prix dans un contexte commercial, etc.) (Barsalou, 1987; Medin & Shoben, 1988).

Unités de description

L'estimation de la similarité par un sujet repose sur l'utilisation d'unités de description (ou critères) qui permettent son calcul. Quatre types d'unités sont distinguées dans la littérature: les dimensions, les traits, les parties, et les "touts" (voir Medin & Smith, 1984; Smith & Medin, 1981; Treisman, 1987). Par tout, on entend le stimulus pris dans sa globalité sans décomposition des entités en unités de plus bas niveau. Par dimension, on entend un ensemble de valeurs discrètes ou variant de manière continue comme la longueur, la largeur. Tout stimulus reçoit une et une seule valeur pour chaque dimension caractérisant la catégorie à laquelle il appartient (e.g., un stimulus prend une seule valeur pour la dimension longueur). La valeur zéro est un des niveaux d'une dimension et ne signifie pas que la dimension n'est pas pertinente pour décrire l'objet.

La notion d'attribut, ou trait, prend différentes significations. Dans le sens restreint, on désigne des aspects de la réalité qui prennent deux ou quelques valeurs. Par exemple, le trait *sexe* prend les valeurs *mâle* et *femelle*. D'autres traits prennent plus de deux valeurs comme l'orientation qui peut être verticale, horizontale, ou inclinée. Les valeurs sont mutuellement exclusives puisqu'on ne peut être à la fois mâle et femelle. Dans le sens le plus large, tout aspect d'une entité est un attribut, notamment les parties qui composent les entités qu'elles contribuent à former. Les parties ne sont ni nécessaires et suffisantes, ni mutuellement exclusives. Par exemple, une roue, un moteur, un siège, un pare-brise, sont des parties présentes en même temps dans une même voiture (elles ne s'excluent pas). Par ailleurs, si l'une vient à manquer, l'objet reste une voiture. Dans le cas des parties, la notion de valeur d'attribut n'a guère de sens. Dans la littérature, l'utilisation de l'une ou l'autre acception dépend des stimuli utilisés (voir Barsalou & Hale, 1993).

Formellement, les dimensions peuvent être représentées sous la forme de traits. Les dimensions référant à une intensité comme l'amplitude sonore sont représentées sous la forme d'ensembles de traits emboîtés: par exemple, les sons de grande amplitude contiennent tous les traits des sons plus faibles. Les dimensions qualitatives comme la hauteur d'un son sont représentées par des ensembles "enchaînés": par exemple, un son plus aigu peut avoir des traits en commun avec un son plus grave mais aussi des traits distinctifs propres. Inversement, un grand nombre de traits ne peuvent être décrits sous la forme de dimensions (par exemple, des traits qui réfèrent à une partie, comme "avoir des ailes", "avoir des roues"). Formellement, dans la pratique, dimensions et traits correspondent à des notions théoriques différentes. Les dimensions sont des objets mathématiques dérivés à partir des données des sujets alors que les traits sont des aspects attribués aux entités du réel souvent sur base d'une intuition des chercheurs ou des sujets, c'est-à-dire sans aucune contrainte théorique (voir plus bas).

2. CALCUL DE LA SIMILARITE

Les notions d'attribut et de dimension ont donné naissance à deux classes de modèles de la similarité (et de la catégorisation). Ces modèles sont formulés sous la forme d'algorithmes qui reposent sur l'hypothèse suivante: un objet A est d'autant plus semblable à un objet B qu'il en est proche sur les dimensions qui les décrivent ou qu'il partage de nombreux attributs (ou valeurs d'attributs) avec B. La catégorisation se décrit en termes de similarité: un objet A est classé dans une catégorie X s'il est plus semblable à la représentation dimensionnelle ou en attributs de cette catégorie (représentée sous la forme d'exemplaires ou d'abstraction) qu'aux représentations des catégories contrastives (Smith & Medin, 1981). Les points 2.1. et 2.2. présentent des exemples de ces deux classes de modèles.

2.1. La similarité dans les représentations dimensionnelles

Les analyses multidimensionnelles permettent de décrire un ensemble de stimuli sous la forme d'un ensemble limité de dimensions continues. Une dimension est un axe dans un espace. Par exemple, le champs sémantique des animaux peut se décrire en termes d'axes tels que la férocité, la taille, la proximité avec l'homme, formés chacun d'un ensemble continu de valeurs (e.g., les différentes valeurs que prend la dimension "taille"). Dans l'analyse multidimensionnelle, un stimulus se définit par ses coordonnées dans l'espace des

axes (dimensions) et reçoit une et une seule valeur sur chacune des dimensions qui le définissent (par exemple, un animal donné prend une seule valeur sur la dimension férocité). La proximité entre les objets dans cet espace reflète leur similarité, même si la relation entre distance entre les stimuli dans l'espace et la similarité n'est pas linéaire (voir ci-après; voir également Ashby, 1992; Shepard, 1964; Nosofsky, 1992; Young, 1984). Pour générer ces représentations géométriques, on utilise, notamment, les estimations numériques de la similarité entre stimuli composant des paires (estimations données par un nombre généralement compris entre 1 et 7), ou la matrice des confusions entre paires d'objets (le nombre de fois qu'un stimulus est confondu avec chacun des autres dans une épreuve d'identification). Carroll et Arabie (1980) font la revue des techniques utilisées en psychologie (voir également Kruskal et Wish, 1978; Shiffman, Reynolds, & Young, 1981). Tversky et Gati (1982) soulignent qu'il ne faut pas confondre les aspects descriptifs (décrire un ensemble de données) avec les aspects inférentiels (élaboration d'un modèle psychologique de la similarité et des représentations mentales) de ces techniques.

On utilise ces techniques pour découvrir les dimensions psychologiques censées organiser un ensemble de stimuli. Par exemple, Henley (1969) étudie la catégorie des animaux. Dans un premier temps, un groupe de sujets estime la similarité entre des animaux présentés par paire. On obtient ainsi une estimation de la similarité moyenne entre chaque stimulus et tous les autres stimuli. A partir de la matrice des similarités, des techniques mathématiques permettent l'extraction des dimensions organisant la catégorie des animaux, c'est-à-dire les dimensions qui expliquent une partie significative de la variance des données de similarité. Elles sont ici au nombre de trois. Selon l'auteur, les deux premières dimensions extraites pouvaient être interprétées comme des dimensions référant à la taille et à la férocité alors que la troisième dimension est difficile à interpréter³. Cette troisième dimension illustre une des difficultés posées par les représentations multidimensionnelles. L'analyse des données conduit à l'extraction de dimensions qui n'ont d'autre signification que mathématique. C'est le chercheur qui en donne une interprétation psychologique. La troisième dimension isolée pourrait résulter de la composition de diverses entités conceptuelles qui seraient difficiles à interpréter sous une seule dimension sémantique.

La similarité s'énonce en termes d'une distance entre les paires d'objets dans un espace défini par une métrique dite de Minkowski ou power metric (Reed, 1972; Tversky & Gati, 1982). La distance d_{ij} entre deux objets i et j dans un espace à m dimensions s'exprime par:

$$d_{ij} = \left[\sum_{k=1}^m |x_{ik} - x_{jk}|^r \right]^{\frac{1}{r}}, \text{ où } 1 \leq r \leq \infty \quad (1)$$

où d_{ij} est la distance entre les stimuli i et j , m est un axe orthogonal de l'espace ($m= 1,2,3, \dots, d$), x_{ik} est la projection du stimulus i sur la dimension k . Les valeurs de r généralement envisagées sont 1 et 2 où la valeur 1 désigne une métrique en pâté de maison (city-block) et la valeur 2 une distance euclidienne. La métrique en pâté de maison s'applique lorsque les dimensions qui constituent les stimuli comparés sont séparables (par exemple, la taille et la forme d'un objet). Dans ce cas, la distance entre deux objets est représentée par la somme des différences sur les dimensions prises séparément. La métrique euclidienne s'applique surtout lorsque les dimensions ne sont pas séparables (par exemple, la couleur et la brillance); les estimations des sujets portent sur une distance globale résultant de l'intégration des dimensions qui composent le stimulus (voir Garner, 1974; 1978).

Dans ces modèles, la catégorisation est un cas particulier de jugement de similarité dans lequel l'estimation de la similarité se fait par rapport aux représentations des catégories, qu'elles soient formalisées sous la forme de tendance centrale (prototypes) ou d'exemplaires (Nosofsky, 1992; Reed, 1972). Par exemple, on peut formuler la distance d'un stimulus i par rapport aux prototypes des catégories auxquelles ce stimulus est confronté. La règle de décision est formulée de la manière suivante: un stimulus j appartient à la catégorie 1 si la distance entre j et le prototype de cette catégorie est inférieure à la distance entre j et le prototype de la catégorie 2, distance calculée sur l'ensemble des dimensions qui définissent les catégories. Le prototype d'une catégorie est le point qui correspond à la valeur moyenne prise par chacune des dimensions dans les stimuli de la catégorie. Dans le cas de modèles à exemplaires, pour catégoriser un stimulus j , on calcule la distance entre ce stimulus et tous les stimuli des catégories potentielles.

Cependant, sous la forme ci-dessus, le modèle ne rend pas compte des variations de l'attention accordée à chacune des dimensions qui définissent les stimuli. Les sujets peuvent pondérer chacune des dimensions de manière différente (Attneave, 1950; Shepard, 1964). Par exemple, si les stimuli se définissent par la taille, la couleur, et la forme, un sujet peut réaliser un jugement de similarité en accordant plus d'attention à la couleur, alors qu'un autre sujet accorde plus d'importance à la forme. Pour tenir compte de ces différences dans l'attention, on ajoute un paramètre de pondération w_m qui reflète le poids attentionnel porté par les sujets à la dimension m , où $w_m \geq 0$, $\sum w_m = 1$. La procédure INDSCAL (individual differences scaling) dérive une solution spatiale valable pour le groupe de sujets étudiés ainsi qu'une matrice des poids de chaque dimension pour chaque sujet. Le poids attentionnel attribué à une dimension se marque dans l'espace des dimensions. Des valeurs basses de w_m (attention faible) "rétrécissent" l'espace des représentations le long de la dimension m alors que des valeurs élevées de w_m "étirent" l'espace sur cette dimension.

La distance d_{ij} obtenue dans l'espace dérivé est convertie en une mesure de similarité par une fonction non-linéaire (voir Shepard, 1957; Nosofsky, 1984, 1986). Dans le Generalized Context Model (GCM) de Nosofsky (1986) cette relation est donnée par:

$$\eta_{ij} = \exp(-c \cdot d_{ij}^p) \quad (2)$$

où c est un paramètre de sensibilité et p définit le gradient de similarité (exponentiel pour $p = 1$, et gaussien pour $p = 2$). Le GCM utilise les similarités ainsi obtenues pour rendre compte de performances obtenues dans des tâches de catégorisation en y incluant des paramètres qui représentent l'intervention de variables psychologiques impliquées par la catégorisation. Le modèle complet prend la forme suivante

$$P(R_J / S_i) = (b_J \sum_{j \in C_J} M_j \eta_{ij}) / (\sum_K b_K \sum_{k \in C_K} M_k \eta_{ik}) \quad (3)$$

où η_{ij} représente la similarité entre les exemplaires i et j ; b_j représente le biais de réponse vers la catégorie J , et M_j représente la force avec laquelle l'exemplaire j est stocké en mémoire. Ce dernier paramètre formalise l'intuition selon laquelle la trace d'un exemplaire en mémoire varie en fonction du nombre de présentations des stimuli. Des résultats ont

montré que les performances de classification sont meilleures pour les exemplaires dont la fréquence de présentation est élevée (Nosofsky, 1986).

En résumé, on utilise un ensemble de données de départ (par exemple, une matrice de confusions entre stimuli ou une matrice d'estimations de la similarité) qui est soumis à une analyse multidimensionnelle. Les distances entre stimuli obtenues sont converties en similarités (où les similarités sont des transformées non linéaires des distances) qui sont utilisées pour générer des prédictions portant sur les performances (par exemple, de catégorisation, d'identification) des sujets pour l'ensemble des stimuli.

Propriétés des modèles multidimensionnels

Les approches multidimensionnelles classiques sont sous-tendues par des propriétés métriques qui doivent être respectées. Parmi celles-ci se trouvent: (1) la minimalité selon laquelle la distance entre un point et lui-même est égale à 0; (2) la positivité selon laquelle la distance entre des points distincts est strictement supérieure à 0; (3) la symétrie selon laquelle la distance entre un point a et un point b doit être égale à la distance entre b et a ; (4) l'inégalité triangulaire selon laquelle la somme des distances de i à j et de j à k doit être supérieure ou égale à la distance de i à k , soit $d(i,j) + d(j,k) \geq d(i,k)$; (5) l'additivité des segments selon laquelle pour tout triplet i, j, k placés sur une droite où j se trouve entre i et k , la distance entre i et k est égale à la distance entre i et j plus la distance entre j et k , soit $d(i,j) + d(j,k) = d(i,k)$.

Les jugements des sujets ne se conforment pas toujours à ces propriétés métriques, ce qui diminue la plausibilité psychologique des modèles qui reposent sur elles. Podgorny et Garner (1979) ont montré que certains stimuli sont estimés plus semblables à eux-mêmes que d'autres, résultat qui s'oppose au principe de minimalité. De même, l'hypothèse de symétrie est fréquemment violée, notamment lorsqu'un item peu représentatif est comparé à un item très représentatif. Par exemple, un rouge tirant sur le rose est jugé plus proche d'un rouge pur que l'inverse. De même, on dira "une ellipse est comme un cercle" et non "un cercle est comme une ellipse" ou "la Corée du Nord est comme l'URSS" et non "l'URSS est comme la Corée du Nord". Les jugements de similarité sont plus élevés lorsqu'on compare la Corée du Nord à l'URSS plutôt que l'inverse. L'asymétrie est déterminée par la saillance

des 2 stimuli qui composent la paire. Le stimulus moins saillant (Corée du Nord) est plus proche du prototype (URSS) que l'inverse (Rosch, 1975; Smith & Medin, 1981; Treisman, 1987; Tversky, 1977; Tversky & Gati, 1978, 1982; cf. infra). De la même manière, l'inégalité triangulaire n'est pas respectée (Tversky & Gati, 1982) ainsi que l'exemple suivant le montre: une flamme et la lune (a, b) sont semblables car très lumineuses, la lune et une balle (b,c) sont semblables car rondes, alors qu'une flamme et une balle (a,c) sont très dissemblables.

Les modèles multidimensionnels peuvent rendre compte des asymétries que nous venons de décrire lorsqu'on les envisage dans le cadre de modèles plus larges qui intègrent les processus de traitement impliqués dans le jugement de similarité (voir le Generalized Context Model, GCM, proposé par Nosofsky, 1984, 1986). Dans les cas de violation de l'inégalité triangulaire, on postule que le sujet modifie l'attention qu'il porte à chaque dimension d'un essai à l'autre. Par exemple, une comparaison de la lune et d'une flamme active la dimension "lumière" alors que celle de la lune avec une balle active la dimension "forme". Nosofsky (1991) propose une interprétation des asymétries de similarité basée sur la notion de biais de stimulus. Selon le modèle GCM, ce sont les propriétés des objets individuels qui expliquent les asymétries observées. Le GCM prédit que la force d'une représentation en mémoire d'un stimulus particulier dépend de la fréquence de présentation de ce stimulus. L'activation d'un item à forte représentation en mémoire par un item à faible représentation est plus importante que celle d'un item faible par un item fort, ce qui résulte en asymétries lors des classifications.

Holman (1979) a proposé une série de modèles permettant de rendre compte des asymétries de similarité entre stimuli où la proximité du stimulus i au stimulus j s'énonce:

$$p(i,j) = F[s(i,j) + r(i) + c(j)] \quad (4)$$

où F est une fonction croissante, $s(i,j)$ est une fonction de similarité symétrique, r et c deux fonctions de biais sur les objets individuels. Un des phénomènes que doit pouvoir expliquer un modèle métrique basé sur des biais individuels est celui de la similarité à soi-même (la similarité du stimulus A au stimulus A). Les modèles de distance classiques postulent, en contradiction avec les faits, que la distance d'un stimulus avec lui-même est égale à zéro.

Dans le modèle des biais, des différences dans la similarité à soi-même peuvent apparaître si les biais associés aux items individuels diffèrent. Par exemple, la similarité à soi-même de l'item i sera plus grande que la similarité à soi-même de j si $r(j) + c(j) > r(i) + c(i)$.

En résumé, contrairement à ce qu'affirment Tversky et Gati (1982), des modèles basés sur une distance symétrique adaptés permettent de rendre compte des asymétries dans les jugements de similarité. Cependant, ainsi que le note Nosofsky (1991), la possibilité de modéliser les asymétries de proximité sous la forme évoquée ne suffit pas. Il reste à étudier comment ces asymétries peuvent être décomposées en une part de similarité symétrique et une part de biais individuels.

Représentations multidimensionnelles probabilistes.

Les modèles précédents représentent un objet individuel comme un point unique dans une représentation spatiale définie par un système d'axes. D'autres modèles ont été développés dans lesquels les objets sont représentés par des distributions probabilistes de points dans un espace multidimensionnel. Selon ces modèles, la représentation construite à chaque présentation d'un stimulus peut différer lorsqu'il y a un bruit important dans les systèmes de traitement perceptif ou lorsque les traces en mémoire ne sont pas stables dans le temps. D'où l'utilisation de représentations probabilistes (Ashby & Perrin, 1988; Ashby & Townsend, 1986; Ennis, 1992; MacKay, 1989; Zinnes & MacKay, 1983). Selon la Théorie Générale de la Reconnaissance (GRT) développée par Ashby et Townsend (1986), l'observateur divise l'espace des représentations en zones de réponse définies par des limites de décision (decision boundaries). Toute représentation qui tombe dans une région X donnera une réponse X. Un jugement de similarité entre X et Y est lié à la proportion de la distribution d'une région de réponse X qui recouvre la région de réponse Y. Par exemple, Ashby et Perrin (1988) décrivent une expérience dans laquelle la distance entre les prototypes des distributions X et Y restait constante et le recouvrement des distributions de X et Y variait en fonction des conditions. Les résultats indiquent que la similarité estimée augmente avec la proportion de recouvrement entre X et Y. La catégorisation peut se concevoir dans les mêmes termes. Soit 2 catégories X et Y composées de stimuli variant sur 2 dimensions. Les catégories sont distribuées comme des variables aléatoires normales

bivariées où les membres de la catégorie X prennent des valeurs faibles sur les 2 variables et ceux de la catégorie Y des valeurs élevées. Différents modèles de décision utilisant ces variables peuvent être utilisés. On peut distinguer les modèles de "décisions indépendantes" des modèles à "intégration d'informations". Dans le premier cas, les sujets établissent un critère pour chacune des dimensions prise isolément et combinent les critères obtenus pour donner leur réponse. Dans les modèles à intégration d'informations, les sujets combinent les informations venant des dimensions dans un percept intégré et la décision se fait sur base de cette intégration. Ici également, on distingue les modèles sur base des limites de décision adoptées (Ashby & Gott, 1988; Ashby & Maddox, 1992). Pour tester ces modèles, les auteurs utilisent une technique appelée la "general recognition randomization technic" dans laquelle on ajoute un bruit externe multivarié aux prototypes de chaque catégorie. Les résultats indiquent que les sujets se comportent selon les prédictions du modèle à intégration d'informations. Ashby et Lee (1991) ont également confronté les prédictions de leur modèle à celles dérivées du GCM de Nosofsky. Ils montrent que les modèles sont équivalents dans leurs prédictions de données de catégorisation et d'identification.

Selon leurs auteurs, ces modèles ne sont pas aussi sensibles aux violations des axiomes métriques que les modèles multidimensionnels précédents. Ashby and Perrin (1988) soulignent que, dans leur approche, la similarité n'est pas décrite uniquement en termes d'une mesure de distance mais également en termes des règles de décision adoptées par les sujets. Comme ces règles sont indépendantes du processus de sensation (perception) des stimuli, il semble possible *a priori* d'expliquer les asymétries en termes de ces règles de décision.

Conclusion

En résumé, les approches multidimensionnelles permettent l'extraction d'un espace à plusieurs dimensions dans lequel sont représentés les stimuli. Moyennant un ensemble d'hypothèses supplémentaires, des prédictions sont tirées sur la reconnaissance et la classification des stimuli. Les avantages de cette approche sont liés au mode d'extraction des dimensions: a) certaines dimensions du réel pertinentes pour la catégorisation sont difficiles à décrire en termes verbaux; puisque les dimensions sont extraites directement à partir des

données, les variables linguistiques ne peuvent les contaminer; b) ne sont extraites que les dimensions qui expliquent une partie de la variance des données; c) la représentation des objets dans l'espace obtenu est bien définie, ce qui permet de faire des prédictions quant aux performances dans les tâches étudiées.

Parmi les désavantages se trouvent la difficulté d'interpréter certaines dimensions, notamment lorsque l'on utilise des catégories conceptuellement complexes (voir plus haut, la discussion des données de Henley). On soulignera cependant que la plupart des travaux qui reposent sur ces techniques utilisent des stimuli dont les dimensions sont clairement identifiables (la taille, la couleur, etc.). L'interprétation des dimensions extraites est donc simple. L'essentiel consiste à trouver les valeurs des coefficients qui expliquent au mieux la structure des données en termes des dimensions extraites. Parmi les problèmes, on a également cité ceux liés à la structure métrique des espaces extraits. Même si, formellement, des moyens existent qui permettent de contourner ces difficultés, il reste à donner des facteurs introduits (comme les biais) une interprétation psychologique plausible *a priori* et non *a posteriori*. Enfin, plus généralement, ces approches ne rendent pas compte de la manière dont les dimensions qui composent l'espace sont sélectionnées dans une tâche donnée (Murphy & Spalding, 1995; voir les points 3 et suivants pour une élaboration).

2.2. La similarité et les modèles en traits.

Dans ces modèles, les concepts sont définis sous la forme d'une liste de traits comme "avoir des ailes", "est rouge", "a des roues", etc. La mesure de similarité utilisée repose sur un calcul impliquant les traits communs et les traits distinctifs présents dans les représentations des entités comparées.

Dans un article classique, Tversky (1977) a proposé un modèle de la similarité qui repose sur la notion d'attribut. Il conteste les approches strictement dimensionnelles qui, selon lui, sont applicables aux seuls domaines que l'on peut décrire à l'aide d'un nombre restreint de dimensions. Les catégories de stimuli telles "la personnalité", "les jeux" semblent faire intervenir un "nombre important de traits qualitatifs plutôt qu'un nombre restreint de dimensions quantitatives" (p.328). Chaque objet à comparer est représenté par un ensemble de traits. Soit deux objets a et b, leur similarité est désignée par $s(a,b)$ où

l'échelle s est une échelle ordinale de mesure de la similarité. Dire que $s(a,b) > s(c,d)$ signifie que a est plus semblable à b que c n'est semblable à d .

Le modèle de Tversky (1977) est fondé sur les hypothèses suivantes:

$$a) \quad s(a,b) = F(A \cap B, A - B, B - A) \quad (5)$$

où F est une fonction mesurant l'appariement de deux objets décrits sous la forme d'un ensemble de traits; A est l'ensemble des traits du concept a et B l'ensemble des traits du concept b ; $A \cap B$ désigne les attributs qui sont à la fois dans A et dans B ; $A \cup B$ les attributs qui sont dans A ou dans B ; $A - B$ les attributs qui sont en A mais pas en B ; $B - A$ les attributs qui sont en B mais pas en A .

b) Le modèle fait l'hypothèse que les traits contribuent indépendamment à l'établissement de la similarité. Les traits sont indépendants lorsque $s(a, b) \geq s(a', b')$ ssi $s(c, d) \geq s(c', d')$. Dans la figure 1, ce ne serait pas le cas si le rôle du trait "sourcil" était différent selon que l'on compare $s(a,b)$ à $s(a',b')$ d'une part et $s(c,d)$ à $s(c',d')$ d'autre part. En effet la similarité entre (a,b) et (a',b') ne peut différer de la similarité entre (c,d) et (c',d') que si le trait "sourcil" n'a pas la même importance dans les deux cas (sa contribution à la similarité n'est pas indépendante des situations).

Insérer Figure 1 ici

Si les hypothèses précédentes sont respectées (et d'autres dont nous ne parlerons pas ici, voir Tversky, 1977), alors il existe une transformation S de s telle que:

$$S(a,b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A) \text{ pour tout } \alpha, \beta, \theta \geq 0 \quad (6)$$

La contribution des sous-ensembles de traits (communs et distinctifs) à la similarité est donnée par f qui reflète une saillance à laquelle contribuent, selon Tversky, différentes caractéristiques des stimuli comme l'intensité, la fréquence, la familiarité, le principe de bonne forme, et le contenu informationnel.

Le modèle est compatible avec l'asymétrie de certains jugements de similarité (voir Ortony, Vondruska, Foss, & Jones, 1985, pour une variante du modèle). Dans une tâche d'estimation de la similarité, selon Tversky "on fait naturellement attention au sujet de la

comparaison. Ainsi, on donne plus de poids aux traits de l'objet comparé (a) qu'aux traits du référent (b) auquel on compare (a) ($\alpha > \beta$). Aussi, la similarité est réduite plus par les traits distinctifs de (a) que par les traits distinctifs de (b). Il s'ensuit facilement que chaque fois que $\alpha > \beta$, $s(a,b) > s(b,a)$ ssi $f(B) > f(A)$. L'hypothèse de l'attention (c-à-d, $\alpha > \beta$) implique que la direction de l'asymétrie est déterminée par la saillance relative des stimuli de telle manière que le stimulus le moins saillant est plus semblable au stimulus saillant que l'inverse" (p.333). L'auteur vérifie cette proposition dans des situations où les sujets doivent choisir entre les phrases "la Corée du Nord est semblable à la Chine" et "la Chine est semblable à la Corée du Nord". Ils choisissent généralement la première, car la Chine est une entité plus "saillante"⁴ que la Corée du Nord.

La vérification des propriétés nécessaires à l'application du modèle n'est pas simple. On a dit que le modèle présuppose l'indépendance des attributs. Pour s'assurer de l'indépendance effective des attributs, il faut vérifier que $s(px, qx) > s(p, q)$ où l'adjonction de x à p et q augmente la similarité de p à q , et que $s(p, py) > s(px, py)$ - l'adjonction de x à p et non à py - décroît la similarité. Cette condition n'est pas toujours remplie. Par exemple, un comptable qui fait de l'escalade (px) est plus semblable à un comptable qui joue au basket-ball (py) qu'à un comptable qui n'a pas de passe-temps (p). Bien que les deux passe-temps diffèrent, ils ont également des points communs (notamment le fait d'être des passe-temps) qui contribuent à augmenter la similarité entre px et py (Gati & Tversky, 1984). Introduire des attributs distinctifs peut parfois entraîner l'introduction d'attributs communs et inversement. Les interactions entre traits, évidentes dans l'exemple précédent, peuvent être le résultat de liens conceptuels divers ou varier en fonction du contexte.

Pour détecter les interactions entre attributs, Gati et Tversky (1984) testent la propriété d'échangeabilité (exchangeability), soit $s(px, q) = s(p, qx)$. Cette propriété n'est pas respectée dans l'exemple suivant: un naturaliste partage plus de traits avec un biologiste qu'avec un comptable, si bien que la similarité entre un biologiste (q) et un comptable-naturaliste (px) est plus grande que celle entre un comptable (p) et un biologiste-naturaliste (qx).

Ce résultat montre que le calcul de l'impact de l'adjonction d'attributs dans le cadre du modèle ne peut se faire sans tenir compte des interactions entre les traits composant les stimuli comparés. Le contenu de ces interactions ne peut être déduit de la simple connaissance des traits impliqués car ce sont des entités conceptuellement complexes qui entretiennent des relations variables en fonction du contexte (par exemple, le trait "avoir un passe-temps" peut recouvrir des réalités conceptuelles variables en fonction des autres traits impliqués dans le contexte). En outre, dans la plupart des situations, les traits impliqués dans une comparaison ne sont pas connus puisque, ainsi que nous l'avons signalé plus haut, les traits composant les concepts correspondant ne sont pas connus.

Poids des traits communs et distinctifs

Le modèle de contraste se définissant en termes de traits communs et distinctifs, Gati et Tversky (1984, 1987) ont étudié leur impact respectif dans des jugements de similarité portant sur des paires d'objets composés d'attributs séparables. La procédure expérimentale utilisée repose sur l'ajout d'un attribut donné x (a) dans un seul des deux stimuli à comparer, attribut qui contribue à la dissimilarité entre les stimuli, (b) dans les deux stimuli où il contribue à leur similarité. Si l'on ajoute cet attribut x à bp et bq pour obtenir bpx et bqx , quelle sera la différence entre la similarité des configurations de départ et celle des configurations complétées? Dans les termes du modèle de Tversky, on peut montrer que la contribution de x , $C(x)$, est une fonction $g(x)$ de x . De même, on peut également évaluer le poids d'un attribut ajouté comme attribut distinctif. Soit la paire de stimuli bp et bpy , si on ajoute x au stimulus bp mais pas à bpy , la contribution de x comme attribut distinctif se définit par: $D(x) = f(x)$. La contribution de x comme attribut commun relativement à son impact comme attribut distinctif s'énonce par: $W(x) = C(x)/[C(x) + D(x)] = g(x)/[g(x) + f(x)]$ dont les valeurs se situent entre 0 et 1.

Gati & Tversky(1984), afin d'étudier le poids d'un attribut commun, ajoutent un attribut (e.g., un nuage) à deux stimuli représentant chacun un paysage. Ils demandent ensuite l'estimation de la similarité des stimuli (paysages) avant ou après l'ajout de cet attribut. De même pour la contribution d'un attribut distinctif. On compare les valeurs de W obtenues pour des descriptions où interviennent des traits communs, des attributs substitués

et des attributs ajoutés. On utilise des descriptions verbales (par exemple, des descriptions de personnes) et des stimuli dessinés (par exemple, des visages, des scènes maritimes). Les stimuli verbaux donnent une valeur de W supérieure à .50, valeur qui montre l'importance accordée aux attributs communs par rapport aux attributs distinctifs. Les résultats s'inversent pour les stimuli imagés où W prend une valeur inférieure à .50 (voir également Gati et Tversky, 1987). Cette asymétrie pourrait être liée aux tâches où les stimuli verbaux seraient généralement impliqués dans la catégorisation alors que les stimuli picturaux demanderaient généralement de la discrimination. De la même manière, les stimuli verbaux sont traités sériellement et de manière discrète puisque présentés séquentiellement alors que les stimuli imagés sont appréhendés dans leur totalité. Dans le cas des stimuli imagés, les traits communs disparaîtraient à l'arrière plan au profit des traits distinctifs, expliquant la valeur de W . Cependant, il faut souligner que cette explication ne justifie pas pourquoi le traitement sériel des stimuli verbaux privilégie les traits communs.

Relations, attributs, jugements de similarité et de différence

Medin, Goldstone, et Gentner (1990) montrent que toutes les unités conceptuelles ne contribuent pas de la même manière au calcul de la similarité. Ils étudient le rôle respectif des relations (définies comme une connexion entre plusieurs entités) et des attributs dans les jugements de similarité. *A priori*, formellement, une relation devrait contribuer aux jugements de similarité comme un trait. Chaque relation, comme chaque trait, est comptabilisée comme un composant conceptuel: les sujets peuvent compter les relations que partagent deux entités ou les distinguent. Medin et al. (1990) montrent au contraire que ce sont des entités psychologiquement distinctes qui contribuent séparément à la similarité et la différence. Dans une expérience où l'on présente trois stimuli (A,B,C), A (stimulus de référence) et B partagent des attributs alors que A et C partagent une relation. Par exemple, le stimulus de référence A est composé de trois carrés rouges, le stimulus B de deux carrés rouges et d'un bleu, le stimulus C de trois carrés bleus. Les stimuli A et B partagent deux attributs (deux carrés rouges) alors que les stimuli A et C partagent la relation "les trois carrés qui composent un stimulus sont de même couleur". Le sujet doit choisir, de B et de C celui qu'il estime "plus semblable", ou "plus différent" du stimulus de référence (A). Les

sujets choisissent plus souvent le stimulus qui partage une relation avec A (soit C) lors des jugements de similarité et de différence. Ce résultat ne respecte pas la complémentarité entre jugement de similarité et jugement de différence puisque C est jugé à la fois plus semblable et plus différent de A que B (voir ci-après). Goldstone, Medin et Gentner (1991) font l'hypothèse que ces résultats ne peuvent s'expliquer que si on distingue similarité d'attribut et similarité de relation. Ils suggèrent que le rôle d'un nouvel élément diffère selon qu'il contribue à augmenter (diminuer) le nombre d'attributs ou de relations partagés par des stimuli. Pour le montrer, ils comparent trois modèles définissant le rôle des attributs et des relations dans l'estimation de la similarité. Le premier modèle repose sur la notion d'indépendance définie plus haut. Ce modèle est comparé (a) au modèle MAX selon lequel lorsque deux entités partagent un grand nombre de relations (attributs), le sujet donne plus de poids aux relations (attributs) dans ses jugements de similarité; (b) au modèle MIN selon lequel le sujet donne plus de poids aux attributs partagés par deux entités lorsqu'elles partagent un grand nombre de relations et inversement donne un poids plus grand aux relations lorsque les entités partagent un grand nombre d'attributs. En d'autres termes, le modèle MAX attribue un poids important aux éléments partagés en grand nombre, le modèle MIN, à l'inverse, donne du poids aux informations peu partagées. Ces deux modèles reposent donc sur la distinction entre relations et attributs, puisqu'ils postulent que ces deux types de descripteurs ne sont pas traités de manière identique.

A partir d'expériences reposant sur des tâches de choix forcé, Goldstone et al. (1991) confirment les résultats obtenus par Medin et al. (1990) en montrant que les sujets choisissent plus souvent les similarités relationnelles que les similarités d'attributs et montrent que les réponses confirment plus souvent le modèle MAX que le modèle MIN. En d'autres termes, les résultats confirment le modèle MAX. Ils indiquent que ce sont les informations les plus nombreuses qui contribuent le plus aux jugements de similarité contrairement aux prédictions du modèle MIN qui attribue un poids plus important aux informations les moins nombreuses dans les stimuli.

Si ces résultats contredisent la propriété d'indépendance, ils ne permettent pas de rejeter le modèle de contraste de Tversky (1977). Pour adapter ce dernier modèle, il suffit

d'y ajouter une hypothèse de traitement qui permet d'attribuer un poids à chaque composant en fonction des autres composants qui participent à une situation. Dans le modèle, cela signifierait que $f(A \cap B)$ n'est pas une somme calculée sur les composants partagés par A et B quand plusieurs types de composants sont impliqués. Des fonctions non linéaires dans lesquelles des nouveaux composants partagés prennent un poids plus important si le nombre de composants partagés est déjà important permettent de rendre compte de la fonction f (voir le modèle cubique de Hintzman, 1986, la règle multiplicative de Medin & Schaffer, 1978, etc.).

Similarité et différence

Si un modèle permet la quantification de la similarité entre deux entités, il doit aussi permettre celle de leur différence. Théoriquement, ces deux notions sont complémentaires. Trivialement, deux entités très semblables devraient être peu différentes. Les corrélations entre ces deux mesures, proches de (-1), confirment cette intuition. Cependant Tversky (1977) et Medin, Goldstone, et Gentner (1990) ont montré que la différence n'est pas nécessairement l'inverse de la similarité. Tversky (1977) propose une liste de paires de pays dont certaines sont composées de pays connus (USA-URSS) et d'autres de pays peu connus (Corée-Malaisie). A chaque essai, on propose deux paires de pays, l'une composée de pays connus, l'autre de pays moins connus. Les sujets d'un premier groupe doivent sélectionner la paire dont les pays sont plus semblables entre eux; un second groupe de sujets doit choisir la paire composée des pays les plus différents entre eux. Si les jugements de similarité et de différence sont complémentaires, le pourcentage des jugements "plus semblable" additionné à celui des jugements "plus différent" devrait, pour une paire dominante donnée, être égale à 100%. Or la somme obtenue est égale à 113.5 qui diffère statistiquement de 100. Par exemple, 67 % des sujets du groupe "similarité" jugent que les pays de la paire "Allemagne de l'Ouest-Allemagne de l'Est"⁵ se ressemblent plus que les pays de la paire "Ceylan-Népal", tandis que 70 % des sujets du groupe "différence" estiment que les pays "Allemagne de l'Ouest-Allemagne de l'Est" sont plus différents l'un de l'autre que "Ceylan et Népal". Ces résultats suggèrent que le sujet fait plus attention aux traits communs (distinctifs) dans les jugements de similarité (de différence) entre objets. Confrontés à des pays connus, les sujets

trouvent des caractéristiques qui les rendent semblables ou différents alors qu'ils ne trouvent pas ces caractéristiques pour les pays peu connus. De la même manière, Medin et al. (1990) ont montré qu'un stimulus C partageant une relation avec un stimulus A pouvait en être à la fois plus semblable et plus différent qu'un stimulus B partageant des attributs avec A (voir ci-dessus).

Les explications suggérées ci-dessus mettent en lumière les caractéristiques de beaucoup de modèles basées sur la notion de trait prise de manière générale (qu'ils soient attributs ou relations). En résumé, on postule un ensemble de composants conceptuels susceptibles de guider les estimations des sujets ainsi que des pondérations de ces composants qui varient en fonction des tâches. C'est là que se situent les problèmes liés à ce type de modèle. La notion de trait n'étant pas contrainte, tout aspect du réel peut être utilisé comme descripteur d'une entité. Généralement, ces composants sont convoqués *a posteriori* et ne sont pas explicités. De la même manière, les mécanismes de pondération des traits communs et distinctifs ne reçoivent pas de justification psychologique *a priori*. En outre, l'inégale saillance des traits est une notion mal définie dans le modèle. Tous ces problèmes rendent difficiles les prédictions dans les tâches de catégorisation ou d'identification (Ashby & Perrin, 1988) (voir Schyns, Goldstone, & Thibaut, sous presse; Thibaut, 1991; Thibaut & Schyns, 1995, pour une discussion des contraintes nécessaires à la définition d'un trait). Les avantages et les faiblesses des modèles en traits sont complémentaires: ils payent leur puissance par un manque de contrainte qui rend difficile l'établissement d'un corpus de prédictions précises.

3. SIMILARITÉ, FLEXIBILITÉ, CONTEXTE ET THÉORIES (OU EN QUOI VOS ENFANTS ET VOTRE ARGENT SE RESSEMBLENT-ILS?)

Le poids d'un trait ou d'une dimension peut varier en fonction des contextes, des entités comparées (cfr. infra). Les modèles comportent généralement un paramètre de pondération qui permet de modifier le poids, la saillance des traits ou dimensions dans un jugement de similarité ou une catégorisation. Selon Tversky (1977), les différentes dimensions psychologiques qui contribuent à la saillance d'un trait sont l'intensité, la fréquence, la familiarité, le principe de bonne forme, et le contenu informationnel de ce trait.

L'intensité et le principe de bonne forme indiquent que les caractéristiques perceptives d'un trait en font un aspect du stimulus plus ou moins prégnant pour le sujet (voir Schyns, Goldstone, & Thibaut, sous presse, pour une discussion du rôle des paramètres perceptifs dans la sélection des attributs).

La fréquence d'un attribut peut se diviser en 2 types: la fréquence absolue de l'attribut dans la réalité qui n'est pas un bon indice de l'importance d'un attribut puisque de nombreux aspects récurrents dans un ensemble de stimuli, contribuant à leur fréquence absolue, ne sont pas encodés par le sujet. Plus intéressante est la fréquence d'un attribut telle qu'elle est estimée par les sujets avec tous les biais qui peuvent affecter cette estimation. En fait, la fréquence estimée est proche de la familiarité qui, comme l'a montré Barsalou (1985, 1987), reflète l'expérience du sujet avec un stimulus.

Un autre paramètre est le contenu informationnel d'un stimulus. Puisqu'il n'existe aucune unité absolue en termes de laquelle on puisse le traduire, cette notion ne peut être que relative. Si, *a priori*, un stimulus particulier peut être décrit en termes d'une infinité d'attributs, le contenu informationnel est lui aussi potentiellement infini. C'est le contexte dans lequel le stimulus est présenté qui permet de restreindre le nombre des attributs envisagés aux seuls attributs pertinents pour ce contexte. C'est la solution qui est privilégiée par les auteurs qui lient similarité et utilisation d'une théorie: l'estimation de la similarité entre des entités repose sur l'utilisation de certains traits critères qui doivent être fixés au moment où l'estimation a lieu.

De nombreuses études ont révélé les effets de contexte, suggérant une flexibilité importante des jugements de similarité. Barsalou (1982, 1983) a montré que les sujets estiment très semblables des items *a priori* très différents comme bijoux et enfants lorsqu'ils sont placés dans la catégorie *ad hoc* des "choses à emporter avec soi en cas d'incendie". Roth et Shoben (1983) comparent l'estimation de la similarité de *lait* et *thé* à *café* dans deux contextes. Dans le premier, ils présentent la phrase "pendant qu'elles buvaient leur boisson à la pause de 10 heures, les secrétaires cancanaient", dans le second, la phrase "avant de commencer sa journée, le conducteur de camion pris une boisson et un gâteau au relais

routier". Dans le premier contexte, les sujets estiment que *thé* est plus semblable à *café* que *lait*, alors que dans le second contexte, on obtient le résultat inverse.

Lorsqu'on compare une paire d'objets, les autres items qui appartiennent au même ensemble influencent l'estimation de la similarité de la paire. Par exemple, la similarité entre *orange* et *pomme* est moins élevée si on présente cette paire dans un groupe de stimuli où se trouvent des termes proches comme *prune* que dans un groupe composé de termes plus éloignés comme *haricots* (Sjoberg, 1972; Tversky, 1977).

Le niveau d'expertise influence les jugements de similarité effectués par un sujet. Sjoberg (1972) a comparé des experts de l'habillement et des experts en canidés dans une tâche de jugement de similarité de paires formées de chiens ou d'habits. Les experts jugent les paires de leur domaine d'expertise moins semblables que celles hors de leur domaine.

Différentes tâches de similarité ne donnent pas toujours des résultats corrélés. Les propriétés conceptuelles des stimuli sont plus importantes dans les jugements de similarité que dans d'autres mesures de similarité (confusions des stimuli). Beck (1966) montre qu'un T vertical est jugé plus semblable à un T incliné qu'à un L droit, alors qu'il est plus souvent groupé perceptivement avec un L droit.

Les exemples précédents sont interprétables en termes des traits utilisés pour réaliser la comparaison. Les effets de contexte résultent de la sélection d'un sous-ensemble de traits parmi ceux qui constituent l'espace des traits dans lequel le sujet effectue l'estimation. Le sujet peut utiliser des informations plus ou moins directement accessibles (par exemple, une banane *est jaune* et *arrive par camion au marché*). Dans l'exemple de Barsalou (1983) cité ci-dessus, la similarité hors contexte entre un bijou et un enfant ne sera pas très élevée puisque les descripteurs spontanément évoqués pour chacune des entités ne se recouvrent pas. Dans le contexte de la catégorie des *entités à emporter lors d'un incendie*, le trait *avoir une valeur personnelle* devient très saillant. Les résultats obtenus par Beck (1966) peuvent recevoir la même analyse. Le jugement de similarité entre T droit et T penché et L droit se fait sur une base conceptuelle et l'identité conceptuelle entre les deux T reçoit un poids important. Par contre, dans une tâche de groupement perceptif, ce sont les informations perceptives qui reçoivent un poids important (voir Thibaut & Schyns, 1995).

Théories et similarité

L'analyse des effets de contexte requiert une identification du répertoire des traits connus des sujets. Or un répertoire exhaustif des traits connus utilisables pour décrire un ensemble de concepts ne peut être construit puisqu'il est possible de créer de nouveaux descripteurs (Schyns et al., sous presse) lorsqu'une situation le demande ou de combiner des descripteurs existant. Si n'importe quel prédicat ou combinaison de prédicats peut jouer le rôle de critère de similarité ou de différence, il sera toujours possible de postuler un sous-ensemble de traits qui rend compte d'un effet de contexte particulier.

Confrontés à ce problème théorique, plusieurs auteurs ont mis en cause la valeur heuristique de la notion de similarité qu'ils jugent incomplète tant que n'ont pas été spécifiées les raisons pour lesquelles les "choses" se ressemblent. Par exemple, dans la section précédente, on a expliqué la catégorisation en termes de dimensions ou traits de similarité dont on n'a pas expliqué la sélection. Murphy et Medin (1985) ont développé ces critiques (voir également Goodman, 1972; Medin & Wattenmaker, 1987). Selon eux, "sans explication du pourquoi les choses sont semblables, nous restons face à un problème équivalent: un grand nombre de choses apparaissent semblables simplement parce qu'elles appartiennent à la même catégorie. En termes plus pratiques, les estimations de la similarité que font les gens peuvent être influencées par le fait de savoir que les choses qu'ils sont en train de comparer sont (ou non) dans la même catégorie" (p.291). Sans spécification des critères de similarité, l'affirmation selon laquelle la catégorisation est fondée sur l'appariement de traits est sans contenu car il n'y a pas de réponse unique à la question "dans quelle mesure un stimulus est semblable à un autre ?". L'exemple suivant révèle bien la portée du problème. Pour la comparaison d'un chien et d'un chat, on pourrait utiliser des traits comme *se déplacent à moins de 500 km/h, de 501 km/h; pèsent moins de 1000 kg, de 1001 kg.*

Selon Murphy et Medin (1985), pour réduire l'infinité potentielle des traits, il faut utiliser des informations de haut niveau justifiant les critères qui serviront, ensuite, à l'estimation de la similarité. Ces informations seraient contenues dans les théories implicites, "c'est-à-dire un grand nombre d'explications mentales plutôt qu'un point de vue

scientifique complet et organisé" (p.290). Ces savoirs sur le monde donnent une cohérence aux catégories et justifient les connexions existant entre les traits d'un concept en fournissant un schéma causal explicatif. Ce sont les théories qui donnent la cohérence conceptuelle et non la similarité. Cette notion de théorie comme schéma organisateur est proche de notions comme les modèles cognitifs idéaux (MCI) de Lakoff (1987a, 1987b), les "schémas" de Rumelhart (1975), les "frames" de Minsky (1975), ou les "scripts" de Schank (1982), Schank et Abelson (1977).

Une théorie permet de donner un sens à la présence simultanée de plusieurs traits dans une représentation conceptuelle. Elle est plus qu'un concept ou qu'un ensemble de traits liés ainsi que le suggère l'exemple suivant. Si un quidam, à la fin d'une soirée bien arrosée, se jette tout habillé dans une piscine, le parterre des spectateurs en conclut qu'il est soûl. Ces spectateurs utilisent leur théorie portant sur les effets de l'alcool pour déduire que le comportement résulte d'une consommation excessive de boissons alcoolisées et non un trait conceptuel comme *saute dans une piscine* hypothétiquement présent dans le concept de *soûl*.

Les théories fournissent des critères de similarité plausibles dans un grand nombre de situations. Elles unifient les "catégories *ad hoc*" dont les membres sont, *a priori*, peu semblables entre eux (Barsalou, 1985, 1987, 1989). Par exemple, si on retire le lien qui unifie la catégorie *ad hoc l'ensemble des choses à emporter lorsqu'il y a le feu* dans laquelle se trouvent *les enfants, le chien, le portrait de grand-mère lors de sa dernière chasse aux éléphants, argent et bijoux*, il ne reste qu'un amalgame d'objets sans lien (Barsalou, 1985). Ces connaissances organisées permettent de faire des prédictions sur les événements qui devraient se produire à un moment donné (Bower, Black, & Turner, 1979), sur ce que l'on devrait trouver à un endroit donné (Brewer & Treyns, 1981).

Medin, Wattenmaker, et Hampson (1987) montrent que les regroupements d'objets en catégories varient si on propose ou non des thèmes qui permettent d'unifier les objets que les sujets doivent trier. Selon les auteurs, les thèmes activent des théories utilisées pour regrouper les différents items qu'on leur présente en fonction des connaissances comprises dans la théorie. Par exemple, les auteurs proposent des descriptions verbales d'individus sous la forme de listes de traits. La tâche consiste à répartir une série de stimuli en deux

catégories. Les résultats indiquent que les catégories construites par les sujets reposent sur plusieurs traits lorsqu'on suggère un thème comme "extraverti-introverti", alors qu'elles ne reposent généralement que sur un seul trait lorsque les descriptions sont présentées sans thème unificateur. Ce dernier sert à rassembler les traits qui, en son absence, resteraient sans lien. De la même manière, la structure des catégories que l'on apprend interagit avec les théories qui sont proposées. L'apprentissage sera facilité (entravé) si les propriétés qui composent les items à catégoriser (ne) prennent (pas) du sens dans le cadre des théories activées par le sujet pendant la tâche (Murphy & Allopena, 1994; Murphy & Spalding, 1995; Wattenmaker, Dewey, Murphy, & Medin, 1986; voir également McCauley, 1987; Schank, Collins, & Hunter, 1986).

Heit et Rubinstein (1994) ont montré que les critères de similarité retenus pour l'induction d'une propriété dépendent des théories sous-jacentes. Les sujets induisent plus facilement la présence d'une propriété anatomique P (comportementale) dans un organisme B étant donné la présence de cette propriété dans un organisme A si les organismes A et B sont estimés semblables anatomiquement (comportementalement) que si A et B sont jugés semblables comportementalement (anatomiquement). Les théories des sujets jouent un rôle dans cette restriction des domaines d'induction: avec une théorie donnée (par exemple, celle qui porte sur les caractéristiques anatomiques) il est plus facile de justifier des généralisations dans le domaine auquel la théorie s'applique que dans d'autres domaines. Les sujets n'utilisent pas un stock de propriétés partagées *préexistant* mais recherchent des propriétés qui, contribuant à la similarité, peuvent être justifiées dans un cadre théorique (voir également Holland, Holyoak, Nisbett, & Thagard, 1986).

Les données de la psychologie développementale suggèrent également que les aspects du réel utilisés comme critères de similarité entre objets se modifient avec l'âge. Par exemple, des enfants de 5 ans diront que les nuages et les éponges sont semblables car "ils sont tous les deux ronds et mous" alors que des adultes invoquent une propriété comme "tous les deux sont remplis d'eau qu'ils rejettent"; les critères invoqués par les enfants réfèrent à des propriétés de surface (forme, couleur, etc.), ceux des adultes aux caractéristiques, essentielles, fonctionnelles des objets comparés. Gentner et Toupin (1986) expliquent ces

résultats par les différences dans les structures conceptuelles de l'enfant et de l'adulte. Carey (1985, 1991; Smith, Carey, & Wiser, 1985) montrent que les critères définissant les concepts se modifient parallèlement aux théories sous-jacentes des enfants dans les domaines concernés (voir également Keil, 1989, 1991). Chi, Feltovitch, et Glaser (1981) étudient ces problèmes dans le contexte du développement de l'expertise et montrent que les critères de similarité utilisés pour la réunion de problèmes dans une même catégorie varient en fonction de l'expertise des sujets. Les novices utilisent des caractéristiques superficielles des problèmes comme critères de classification alors que les experts utilisent des critères de classification profonds, c'est-à-dire qui réfèrent à des principes organisateurs du domaine.

On a également montré que l'espace dimensionnel définissant un ensemble de stimuli diffère en fonction des contextes et/ou connaissances impliqués. Melara, Marks, et Lesko (1992) ont montré que la métrique qui permet la meilleure description d'un ensemble de données - en pâte de maisons ou euclidienne (voir point 2) - dépend des instructions qui sont données aux sujets. Ce résultat indique que l'espace de similarité obtenu dépend des connaissances activées dans une tâche particulière alors que, selon leurs utilisateurs, les analyses multidimensionnelles donneraient des représentations qui ne varient pas en fonction du contexte des stimuli analysés.

En résumé, toutes ces données indiquent qu'on ne peut réduire la similarité et la catégorisation à un calcul sur un ensemble des traits communs et/ou distinctifs. La spécification des critères de calcul est aussi importante car elle contraint le processus de calcul des traits.

L'alignement: une condition nécessaire au calcul de la similarité.

Si le calcul de la similarité nécessite la fixation des critères qui seront utilisés pour le calcul, il importe de décrire le processus par lequel les traits utilisés dans chacune des entités comparées sont mis en correspondance, à savoir le processus d'alignement décrit dans le cadre des travaux sur la compréhension des analogies (Gentner, 1989; Markman & Gentner, 1993; Medin, Goldstone, Gentner, 1993). Aligner deux entités, c'est mettre en correspondance les traits communs et distinctifs locaux (par exemple, les entités comparées ont des pattes), les traits globaux (par exemple, les entités comparées sont deux scènes de

ville composées chacune de constructions hautes et en file), ainsi que les relations entre traits. On place ainsi en correspondance terme-à-terme les éléments des deux représentations comparées ainsi que les connexions entre ces éléments.

Dans ce processus d'alignement, les différences entre les entités comparées sont de deux sortes: les différences alignables (c'est-à-dire les différences qui portent sur des aspects présents dans les stimuli comparés mais qui ne prennent pas les mêmes valeurs: par exemple, deux stimuli se définissent par la couleur, mais l'un est rouge, l'autre bleu) et non alignables (par exemple, un dessin comporte un camion alors que l'autre ne contient aucun élément lié aux moyens de transport). Dans la théorie de l'alignement, les éléments qui ne sont pas alignables ne devraient pas intervenir dans le calcul de la similarité, alors que le modèle de Tversky (1977) tient compte de toutes les différences disponibles dans les objets comparés; dans ce modèle, le poids des différences dans le calcul de la similarité est global, qu'elles soient alignables ou non.

Les traits qui sont utilisés pour réaliser l'alignement doivent, dans certains cas, être construits par l'individu, notamment lorsqu'il est confronté à des stimuli nouveaux pour laquelle aucune segmentation n'est disponible. Ceci nécessite que soient créés de nouveaux descripteurs ou différenciés des descripteurs existant, que ce soit pour des situations qui se décrivent en termes perceptifs ou en termes conceptuels. Prenons, pour illustrer le premier cas, le monde du son et de la Haute Fidélité. Les amateurs de Haute Fidélité peuvent, lorsqu'on leur demande de comparer le son que délivre un élément comme par exemple, un amplificateur, le distinguer du son que produit un autre amplificateur. Lorsqu'on interroge de tels experts, ils peuvent évoquer les dimensions qu'ils utilisent pour procéder à cette différenciation. Ils évoquent la largeur et la profondeur de l'image sonore, la balance tonale, la qualité du suivi mélodique, la dynamique, etc. Comme cette comparaison ne peut se faire que séquentiellement, il faut que ces experts mémorisent les caractéristiques sonores du premier amplificateur pour les comparer à celles du second. Ces comparaisons ne peuvent pas être réalisées par des novices qui généralement disent qu'ils ne perçoivent aucune différence entre les appareils. Quelle différence existe-t-il entre les experts et les novices ? Les experts ont réalisé un apprentissage perceptif au terme duquel est désormais disponible

un espace de traits à partir duquel des stimuli, ici sonores, peuvent être comparés ou catégorisés (voir Schyns et al., sous presse, Thibaut & Schyns, 1995). Dans le domaine conceptuel, l'expertise dans un domaine équivaut souvent à l'acquisition de nouveaux descripteurs ou à une utilisation différentes d'anciens descripteurs. Par exemple, Chi et al. (1981) ont montré que les descripteurs utilisés par des experts et des novices pour classer des problèmes ne sont pas identiques. L'organisation conceptuelle qui soutient l'expertise diffère de celle qui fonde les concepts d'un novice.

Selon Medin et al. (1993), pendant le processus d'alignement, les sujets sélectionnent et élaborent les traits utilisés pour le calcul de la similarité. Dans une expérience, les auteurs présentent un stimulus X qu'un premier groupe de sujets doit comparer à un stimulus A alors qu'un autre groupe doit le comparer à un stimulus B. Les auteurs veulent montrer que les traits cités pour décrire X dépendent du stimulus (A ou B) auquel il est comparé. Deux types de stimuli sont proposés, ambigus et non ambigus: pour les ambigus, les stimuli A et B sont construits pour qu'une propriété de A (B) attribuable à X soit incompatible avec les propriétés de B (A); par exemple le stimulus A comprend 3 "doigts" alors que le stimulus B en comprend 4 (le stimulus X est ambigu en termes de nombre de doigts). Dans les non ambigus, X contient deux traits critiques: un des traits se trouve en A et non en B alors qu'un autre se trouve en B et non en A. Les résultats montrent, pour les stimuli ambigus, que X, lorsqu'il est associé à A (B), est décrit avec des propriétés différentes de (et incompatibles avec) celles citées lorsqu'il est associé à B (A). Si les propriétés de X existaient indépendamment des items de comparaison auxquels X est comparé, les sujets les citeraient quel que soit le stimulus auquel X est associé. La mise en correspondance de X avec A (B) conduit à la découverte de traits différents incompatibles avec ceux identifiés lorsque X est mis en correspondance avec B (A). Ceci indique que la recherche des traits repose d'abord sur la définition d'un espace de caractéristiques qui permet cette comparaison (voir Markman & Gentner, 1993, pour une étude des rapports entre alignement et similarité où ils suggèrent qu'il n'y a pas de procédure générale permettant de dire ce qui compte comme une différence ou comme un trait commun; voir également Goldstone, 1994b, pour une modélisation).

Certains résultats obtenus par Tversky (1977) peuvent s'interpréter en termes d'alignement. Les sujets doivent associer un pays de référence (Autriche) à un des trois pays suivants: Suède, Pologne, Hongrie. Les appariements obtenus sont respectivement 49%, 15%, et 36%. Un autre groupe doit choisir parmi Suède, Norvège, Hongrie. Les appariements sont respectivement 14%, 26%, 60%, soit une diminution des classifications Suède et une augmentation des classifications Hongrie. Le critère d'alignement qui permet de classer les pays en deux groupes diffère dans les deux conditions. Pour le premier groupe, le critère est politique, *communiste vs non communiste*, et permet de rassembler Pologne-Hongrie versus Suède-Autriche. Le remplacement de Pologne par Norvège rend cette répartition impossible mais une comparaison des pays révèle un nouveau critère d'alignement, géographique, qui donne naissance à la répartition Suède-Norvège versus Autriche-Hongrie. Ce critère d'alignement maximise la relation entre les paires de pays et entre les pays à l'intérieur de chacune des paires: les quatre pays sont ainsi reliés par un seul thème. D'autres critères auraient été moins puissants. Par exemple, dans le premier cas, le regroupement d'Autriche et Hongrie est possible sur base géographique alors que ce critère est moins convainquant pour Pologne-Suède bien que plausible puisque ces 2 pays sont plus au nord que les 2 autres (l'utilisation de ce critère explique sans doute cette répartition des quatre pays chez 36% des sujets). Aucun critère ne permettait d'aligner Autriche et Pologne vs Hongrie et Suède. Selon l'explication de Tversky, le sujet cherche à appliquer un principe de diagnosticité qui "se réfère à la signification classificatoire des attributs" (p.342). Cependant, le modèle de contraste n'explique pas le processus de changement de critère: il ne motive pas la modification du poids des traits par les caractéristiques des pays comparés.

4. SIMILARITE ET CATEGORISATION: DEUX PROCESSUS A DISSOCIER?

Dissociation: quelques exemples

Certains auteurs (Rips, 1989) proposent, en suite au débat théorique que nous venons d'évoquer et sur base de résultats expérimentaux, de dissocier jugements de similarité et catégorisation et d'en faire des processus distincts. Rips (1989, 1991; voir également Rips & Collins, 1993) décrit des expériences où jugements de similarité et catégorisation ne

convergent pas. Rips (1989) présente deux types de catégories: (a) les catégories dites "fixées", composées de stimuli dont les valeurs prises par certaines dimensions ne varient pas (e.g., la taille pour les pièces de 25 cents); (b) les catégories dites "variables" dont les valeurs prises par certaines dimensions varient (e.g., la taille pour les pizzas). L'auteur présente un item hypothétique dont la taille est située à mi-chemin entre la taille d'une pièce de 25 cents et celle de la plus petite pizza, item que les sujets doivent classer dans l'une des deux catégories et dont ils doivent estimer la similarité avec chacune des deux catégories. Les résultats montrent que les sujets classent plus souvent cet item dans la catégorie "variable" mais le jugent plus semblable à la catégorie "fixée" (Smith & Sloman, 1994 ne reproduisent ces résultats que très partiellement). De la même manière, Fried et Holyoak (1984) montrent que les sujets classent plus fréquemment de nouveaux stimuli dans des catégories dont les stimuli sont très variables plutôt que dans des catégories aux stimuli peu variables, même s'ils sont plus proches physiquement des membres de cette seconde catégorie (voir Keil, 1989, 1991 pour des exemples développementaux). Rips (1989, 1991) interprète ses résultats dans des termes semblables à ceux de Murphy et Medin (1985) et souligne l'insuffisance de la notion de similarité comme fondement de la classification.

Medin et al. (1993) ont montré que la catégorisation d'un item dans une catégorie et non dans une autre dépend, non seulement de la similarité de l'item à classer avec les items de la catégorie où il peut être classé, mais aussi des relations que les items qui composent les catégories entretiennent entre eux. Par exemple, dans une expérience, un premier groupe de sujets a estimé que des Dobermans sont plus semblables à des rats laveurs qu'à des requins. Un second groupe de sujets classe cependant plus souvent les Dobermans dans un ensemble A composé de {ours, lion, requin} que dans un ensemble B composé {ours, lion, raton laveur}. Pris séparément, les Dobermans sont plus proches des membres de B que des membres de A (similarité inférieure avec "requin" comparé à "raton laveur") résultat qui devrait mener à une classification en B. Le résultat s'explique par la propriété émergente de l'ensemble A (la férocité), propriété que les animaux de A partagent avec les Dobermans, alors que les 3 membres de B ne possèdent pas de propriété commune partagée avec les Dobermans. La présence de cette propriété permet un alignement relationnel entre

Doberman et les animaux de A, alignement impossible avec les animaux de B (voir également Gati et Tversky, 1984).

Les données précédentes ainsi que les critiques de la notion de similarité laissent entendre que doit être abandonné le lien nécessaire entre catégorisation et similarité même si ce lien reste, dans la plupart des cas, privilégié. Medin et Ortony (1989; voir également Goldstone, 1994a) suggèrent que les dissociations obtenues par Rips sont des artefacts liés aux consignes utilisées. Selon Goldstone, la présence du mot "semblable" dans la consigne ne garantit pas que les sujets de Rips (1989) aient réellement estimé la similarité entre les entités. Ils peuvent avoir utilisé une interprétation de cette consigne limitée aux aspects perceptifs des objets. Ce raisonnement s'applique aux dissociations obtenues par Carey (1985) entre similarité et induction de propriétés catégorielles. Par exemple, les enfants et les adultes estiment qu'un singe mécanique est plus semblable à un être humain que ne l'est un poisson. Cependant, une propriété biologique des êtres humains sera plus souvent projetée sur les poissons que sur les singes mécaniques.

Similarité: un fondement de la catégorisation?

L'analyse précédente amène Goldstone (1994a) à reconsidérer les relations entre similarité et catégorisation. La première constituerait un fondement (grounding) de la seconde. Les sujets utiliseraient les similarités superficielles pour découvrir les règles de catégorisation plus profondes, par exemple celles qui relèvent des théories expliquant la catégorie. Ainsi, étudier les similarités génétiques entre différents chiens demande que l'on ait d'abord constaté leurs similarités de surface.

Nous pensons, au contraire, que les similarités de surface ne peuvent fonder les similarités catégorielles profondes. Les raisons en sont multiples: a) deux entités possèdent un grand nombre de dimensions contribuant à leur similarité de surface; un grand nombre d'entre elles sont idiosyncrasiques et ne seront intégrées dans aucun réseau de similarités profondes (par exemple, une similarité de couleur ne signale généralement aucune similarité plus profonde, comparez un ours brun et un ballon brun). Choisir les similarités de surface porteuses d'un sens profond nécessite une théorie préalable qui les justifie. C'est donc la théorie qui justifie les similarités de surface "qui ont un sens profond" et non l'inverse; b)

complémentairement, les similarités de surface ne sont pas données *a priori*. Dans le cas d'entités visuelles, elles résultent d'un découpage du stimulus en unités perceptives pertinentes dont il faut également expliquer l'origine. Schyns, Goldstone, et Thibaut (1995), Thibaut (1991; 1995) ont montré que des stimuli visuels "complexes" peuvent recevoir différentes segmentations en fonction des contextes dans lesquels ils sont présentés. Ces unités, différentes selon les contextes, définissent autant de similarités différentes. En d'autres termes, il n'y a pas de découpage perceptif *a priori* qui serve d'input aux jugements de similarité; c) les entités réelles appartiennent généralement à plusieurs niveaux d'inclusion; par exemple, un chien est un "berger allemand", "un chien", "un mammifère", "un animal", "une entité vivante", "une entité physique". Les théories qui fondent chacun de ces niveaux sont différentes. Ainsi les similarités essentielles (théoriques) qui rassemblent les chiens dans la catégorie des "entités vivantes" ne sont pas celles qui les rassemblent dans la catégorie des "chiens" et les similarités de surface connectées à chacun de ces niveaux de similarité profonde sont différentes. Lorsqu'un sujet est confronté à des entités présentant des similarités de surface, il ne peut connaître *a priori* le niveau de similarité essentielle auquel ces similarités de surface peuvent être connectées, sauf s'il a une théorie permettant d'établir cette connexion. Vouloir légitimer la catégorisation par une similarité intrinsèque, c'est ressusciter une similarité objective indépendante du sujet (voir Goodman, 1972).

Cette conception objectiviste d'une similarité présente dans l'environnement a également été défendue par Rosch (1978). Sa notion de "structure corrélacionnelle de l'environnement" signifie que certains attributs sont corrélés dans l'environnement et que ces corrélations sont le fondement des catégorisations ou, en nos termes, fixent les critères de similarité. La paire "ailes-plumes" est plus probable dans l'environnement que la paire "ailes-poils" car les ailes sont plus fréquemment liées aux plumes qu'aux poils. Selon Rosch, les catégories regrouperaient sélectivement des objets porteurs d'attributs corrélés dans l'environnement. Cependant, cette notion se révèle insuffisamment contrainte puisque, souvent, plusieurs corrélations sont possibles et on ne peut décider *a priori* celles qui seront retenues par le sujet (Keil, 1981; Malt & Smith, 1984). Selon Murphy et Medin (1985), les sujets utilisent leurs théories du monde pour justifier les corrélations retenues comme base

de catégorisation: par exemple, le trait "avec des murs" est connecté au trait "avoir un toit" car, dans les théories d'un sujet, ces deux entités assurent une protection contre les intempéries. Ces liens sont déjà en mémoire et sont utilisés pendant l'apprentissage des concepts pour expliquer les cooccurrences de certains traits. Si aucune théorie n'est disponible pour expliquer une corrélation réelle entre deux attributs, cette corrélation est moins utilisée que là où une théorie fournit des explications (Medin, Altom, Edelson, & Freko, 1982; Murphy & Wright, 1984).

Similarités de surface et profondes

Dans la même perspective, Medin et Ortony (1989) tentent de relier les informations directement accessibles aux propriétés les plus profondes. Ils signalent que le calcul de la similarité entre deux entités repose sur l'utilisation de toutes les propriétés pertinentes pour la description d'un objet, des plus accessibles (perceptivement, comme la forme pour un animal) aux plus cachées (comme le code génétique). Ces deux classes de propriétés peuvent contribuer à la similarité dans des directions opposées. Par exemple, les propriétés accessibles -les caractéristiques visibles- d'un ours et d'une baleine sont très dissemblables et conduisent à les classer dans des catégories différentes alors que d'autres connaissances portant sur des caractéristiques cachées (système de reproduction, code génétique, etc.) permettent de rassembler ces deux animaux dans la même catégorie. Les propriétés superficielles peuvent orienter le sujet vers des catégories qui contredisent les informations profondes, essentielles (baleines classées parmi les poissons). Medin et Ortony postulent un *essentialisme psychologique* qui correspondrait à la croyance en l'existence d'une essence qui définit, fonde certaines catégories (que cette essence existe réellement ou non importe peu). Pour ces auteurs, les propriétés profondes justifient les propriétés de surface, les contraignent. Ils proposent d'inclure dans les représentations conceptuelles un "support d'essence" (essence placeholder) qui pourrait contenir trois types d'informations: (a) des convictions à propos des propriétés nécessaires et suffisantes pour appartenir à telle catégorie; (b) une théorie, dans le sens donné à ce terme par Murphy et Medin (1985); (c) la conviction qu'il existe des experts qui connaissent réellement les propriétés qui font qu'une chose est bien ce qu'elle est ou qui essaient de découvrir ces propriétés.

On soulignera que la croyance en l'existence de propriétés profondes qui forment le coeur d'une catégorie (a), ou d'experts qui connaissent ces propriétés (c) ne signifie pas que ces propriétés existent ou que, le cas échéant, les sujets les connaissent. Lorsque les sujets ont une théorie, au sens de Murphy et Medin (1985), elle ne contient que des informations fragmentaires, voire fausses, comparées à celles qui composent les théories scientifiques. Dès lors, les théories scientifiques ne peuvent justifier la structure des concepts puisque la plupart des gens ne les connaissent pas et ne peuvent les utiliser pour orienter leurs catégorisations. *La catégorisation: un cas de calcul de la similarité ?*

Il nous semble possible de donner une vue unifiée où la similarité et la catégorisation reposerait sur des processus identiques opérant sur un ensemble de traits. Dans cette perspective, la catégorisation est un cas particulier d'estimation de la similarité, la différence entre les deux tâches portant sur les critères utilisés pour chacune de ces deux tâches. Soit, par exemple, une pomme réelle, une pomme en plastique et une salade. Ces trois entités sont représentées en mémoire par m , n , p traits respectivement. Ces ensembles de traits ont une intersection. Pour estimer la similarité entre une pomme et une salade d'une part, et entre une pomme et une pomme en plastic d'autre part, de multiples critères sont utilisables. En termes de certains critères, pomme et salade sont plus semblables entre elles (toutes les deux sont comestibles, composées de cellules, etc.) qu'une pomme en plastic et une pomme réelle, alors que selon certains autres critères pomme en plastic et pomme réelle se ressemblent plus (forme, couleur). La similarité se calcule en termes de n'importe quel(s) trait(s) parmi les traits qui composent la représentation d'une entité particulière. En fonction des jugements de similarité effectués et du contexte, les traits utilisés peuvent différer. Par contre, la catégorisation est plus restrictive puisque fondée sur un sous-ensemble de traits fixé appartenant à l'ensemble des traits qui représentent l'entité à catégoriser. Ainsi parmi les m et n traits qui entrent dans la représentation des pommes réelles et des pommes en plastique, certains sont communs aux deux entités et parmi ceux-ci un sous-ensemble est utilisé pour les classer dans la catégorie des pommes. La catégorisation repose sur une *configuration* particulière de critères de similarité. (Ceci n'implique pas que la catégorisation se fasse sur base d'un sous-ensemble de traits nécessaires et suffisants; seulement que les sujets utilisent

des traits du sous-ensemble de traits qui définissent la catégorie, par exemple pour la pomme les traits m_1 à m_i). Lorsqu'une entité peut être classée dans plusieurs catégories (ainsi une pomme réelle peut être classée, selon les contextes, dans la catégorie des pommes, des fruits, des desserts, etc.), chacune des catégories potentielles est définie par une configuration particulière d'un sous-ensemble de traits de la représentation mentale de l'entité à catégoriser. Pour résumer, la catégorisation est prescriptive (le sujet doit se référer à un sous-ensemble de traits) alors que les jugements de similarité dépendent du contexte de leur production et de la sélection (optionnelle) des traits disponibles dans ce contexte.

Une expérience de Thibaut et Dupont (1996) suggère comment les sujets utilisent diverses informations contenues dans les stimuli pour estimer la similarité et catégoriser. Les sujets apprennent deux catégories (A et B) de 9 stimuli inconnus composés d'une partie supérieure à laquelle sont fixés quatre appendices (des pattes). Une majorité d'exemplaires (6 sur 9) de la catégorie A sont composés d'une partie supérieure qui a une forme donnée (par exemple, une sorte de champignon) présente dans cette catégorie uniquement. Six exemplaires de la catégorie B sont composés d'une partie supérieure qui a une autre forme. Par contre, les trois derniers stimuli des deux catégories sont composés d'une forme supérieure présente dans les deux catégories. Ainsi, la partie supérieure ne peut être utilisée comme critère de classification parfait. Le critère d'apprentissage parfait se trouve au niveau de l'arrangement spatial des pattes (2 pattes + 2 pattes pour la première catégorie, 3 pattes plus 1 pour la seconde) permet la classification parfaite des stimuli. Lorsque, dans la phase de test, on présente un item qui possède la partie supérieure présente uniquement dans la catégorie A mais des pattes dont l'arrangement spatial correspond à la catégorie B, les sujets le catégorisent en B mais le jugent plus semblable à la catégorie A.

Cette expérience suggère que les dissociations entre jugement de similarité et catégorisation ne sont pas le résultat de l'utilisation de processus différents, puisqu'on peut produire ces dissociations en manipulant les informations présentes dans les items d'apprentissage. Pendant l'apprentissage des catégories, les sujets construisent un répertoire de n traits (dans l'expérience précédente, les différentes parties supérieures des stimuli, les pattes et leurs arrangements). Les jugements de similarité peuvent se fonder sur n'importe

quelle combinaison de traits encodés pendant l'apprentissage (dans l'exemple ci-dessus, les pattes, leur répartition, la partie supérieure, sa forme) alors que les traits utilisés pour la catégorisation d'un stimulus X dans une catégorie ne comprennent qu'une partie des traits encodés (la répartition des pattes). Dans la plupart des cas, cependant, catégorisation et jugement de similarité sont réalisés sur base des mêmes traits. Les dissociations apparaissent dans des situations où les traits utilisés pour un jugement de similarité sont très saillants dans le contexte de ce jugement mais ne sont la base d'aucune catégorisation (par exemple, "traverse la rue" peut servir de critère de comparaison pour un chien et un piéton mais ne sera pas utilisé pour leur catégorisation) ou, à l'inverse, lorsque les traits utilisés pour la catégorisation n'ont pas un poids important dans le contexte où s'effectue un jugement de similarité. Dans ces cas, un sujet peut estimer dans un contexte donné qu'une entité X est plus semblable à une entité Y qu'à une entité Z tout en la classant en Z.

Ces dissociations sont le tribut à payer à la flexibilité des jugements de similarité qui peuvent reposer sur des traits peu centraux des concepts ou reconstruits à partir des connaissances des sujets (la propriété "traverse la rue" citée ci-dessus pour la comparaison d'un chien et un piéton n'est probablement stockée dans aucun des deux concepts et est reconstruite à partir des propriétés connues pour chacune des deux entités comparées). L'exemple classique de la baleine, mammifère qui ressemble à un poisson, trouve une description évidente dans le présent cadre : selon certains critères la baleine est plus semblable à un poisson (elle nage, etc.) qu'aux mammifères alors qu'on la classe dans cette dernière catégorie sur base d'autres critères.

En résumé, toute entité du réel peut être envisagée à partir d'une multitude de points de vue différents. Les jugements de similarité réalisés par un sujet dépendent des entités comparées et du point de vue adopté sur ces entités. En fonction de celui-ci, deux entités peuvent s'avérer plus ou moins semblables. Par contre, la classification dans une catégorie particulière dépend de la présence dans le référent à catégoriser d'une configuration de propriétés choisies parmi un sous-ensemble des traits connus pour cette entité. Dès lors, les dissociations entre jugements de similarité et catégorisation ou les discordances entre

différentes estimations de la similarité résultent des traits sélectionnés pour réaliser ces tâches.

5. CONCLUSIONS

Les travaux décrits portent sur deux aspects complémentaires de la similarité: les modèles formels présentés dans la première section se concentrent sur une représentation des similarités entre entités, qu'elle soit exprimée sous la forme de traits ou de dimensions. Les travaux sur les processus de sélection des traits visent plutôt à expliquer, contraindre le sous-ensemble de traits utilisés dans les jugements de similarité. Ces deux sources de travaux ne se sont guère fécondées mutuellement. Même si des modèles formels, tels ceux de Nosofsky (1986, 1991) intègrent des paramètres de biais qui permettent de modéliser l'attention portée à des dimensions particulières, la sélection et l'utilisation de ces dimensions n'est pas abordée. De la même manière, les processus de sélection et de mise en correspondance des dimensions des stimuli étudiés dans les travaux décrits dans la seconde partie n'ont pas été connectés aux modèles de la similarité. Une des raisons repose sans doute sur la difficulté de modéliser une similarité si flexible, si dépendante de propriétés et de contextes changeant. L'avenir passe probablement par une meilleure description de la sélection des traits. Une fois celle-ci mieux décrite, on pourra l'intégrer dans les modèles présentés en première partie: les paramètres de biais n'apparaîtront plus, aux dires mêmes de leurs auteurs, comme des constructions *a posteriori* résultant plus du choix des expérimentateurs que des caractéristiques des sujets.

BIBLIOGRAPHIE

- Ashby, F.G. (ed.) (1992). *Multidimensional models of perception and cognition*. Hillsdale, NJ: Erlbaum.
- Ashby, F.G., & Gott, R.E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 598-612.
- Ashby, F.G., & Lee, W.W. (1991). Predicting similarity and categorization. *Journal of Experimental Psychology: General*, *120*, 150-172.
- Ashby, F.G., & Maddox, W.T. (1992). Complex decision rules in categorization: contrasting novice and experienced performance. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 598-612.
- Ashby, F.G., & Perrin, N.A. (1988) Toward a unified theory of similarity and recognition. *Psychological Review*, *95*, 124-150.
- Ashby, F.G., & Townsend, J.T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154-179.
- Attneave, F. (1950). Dimensions of similarity. *American Journal of Psychology*, *63*, 516-556.
- Barsalou, L.W. (1982). Context independent and context dependent information in concepts. *Memory & Cognition*, *10*, 82-93.
- Barsalou, L.W. (1983). Ad hoc categories. *Memory & Cognition*, *11*, 211-227.
- Barsalou, L.W. (1985). Ideals, central tendency and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *11*, 629-649.
- Barsalou, L.W. (1987). The instability of graded structure. In U. Neisser (Ed.), *Concepts and conceptual development* (pp. 101-140). Cambridge: Cambridge University Press .
- Barsalou, L.W. (1989). Intraconcept similarity and its implications for interconcept similarity. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 76-121). Cambridge: Cambridge University Press.

- Barsalou, L.W., & Hale, C.R. (1993). Components of conceptual representation: from feature lists to recursive frames. In I. Van Mechelen, J. Hampton, R. Michalski, & P. Theuns (Eds.) *Categories and concepts: theoretical views and inductive data analysis* (pp. 97-144). London: Academic Press.
- Beck (1966). Effect of orientation and of shape similarity on perceptual grouping. *Perception and Psychophysics, 1*, 300-302.
- Bower, G.H., Black, J., & Turner, T. (1979). Scripts in memory for texts. *Cognitive Psychology, 11*, 177-220.
- Brewer, W.F., & Treyens, J.C. (1981). Role of schemata in memory for place. *Cognitive Psychology, 13*, 207-230.
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.
- Carey, S. (1991). Knowledge acquisition: enrichment or conceptual change? In S. Carey & R. Gelman (Eds.). *The epigenesis of mind* (pp. 257-291). Hillsdale, NJ: Lawrence Erlbaum.
- Carroll, D., & Arabie, P. (1980). Multidimensional scaling. *Annual Review of Psychology, 31*, 607-649.
- Chi, M.T.H., Feltovitch, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.
- Ennis, D.M. (1992). Modelling similarity and identification when there are momentary fluctuations in psychological magnitudes. In F.G. Ashby (ed.) (1992). *Multidimensional models of perception and cognition* (99-133). Hillsdale, NJ: Erlbaum.
- Fried, L.S., & Holyoak, K.J. (1984). Induction of category distribution: a framework for classification learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 10*, 234-257.
- Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Lawrence Erlbaum.
- Garner, W.R. (1978). Aspects of a stimulus: feature, dimensions and configuration. In E.H. Rosch & B.B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Erlbaum.

- Gati, I., & Tversky, A. (1984). Weighting common and distinctive features in perceptual and conceptual judgments. *Cognitive Psychology*, *16*, 341-370.
- Gati, I., & Tversky, A. (1987). Recall of common and distinctive features of verbal and pictorial stimuli. *Memory & Cognition*, *15*, 97-100.
- Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199-241). Cambridge: Cambridge University Press.
- Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy, *Cognitive Science*, *10*, 277-300.
- Goldstone, R.L. (1994a). The role of similarity in categorization: providing a groundwork. *Cognition*, *52*, 125-157
- Goldstone, R.L. (1994b). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1-28.
- Goldstone, R.L., Medin, D.L., & Gentner, D. (1991) relational similarity and the nonindependence of features in similarity judgments. *Cognitive Psychology*, *23*, 222-262.
- Goodman, N. (1972). Seven strictures on similarity. In N. Goodman (Ed.), *Problems and projects* (pp.437-447). New York: Bobbs-Merrill.
- Heit, E., & Rubinstein, J. (1994). Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 411-422.
- Henley, N.M. (1969). A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior*, *8*, 176-184.
- Hintzman, D.L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411-428.
- Holland, D., Holyoak, K., Nisbett, & Thagard, P. (1986). *Induction*. Cambridge: MA: MIT Press.
- Holman, E.W. (1979). Monotonic models for asymmetric proximities. *Journal of Mathematical Psychology*, *20*, 1-15.

- Keil, F.C. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, 88, 197-227.
- Keil, F.C. (1989). *Concepts, kinds and cognitive development*. Cambridge, MA: MIT Press.
- Keil, F.C. (1991). The emergence of theoretical beliefs as constraints on concepts. In S. Carey & R. Gelman (Eds.). *The epigenesis of mind*. Hillsdale, NJ: Lawrence Erlbaum.
- Kruskal, J., & Wish, M. (1978). *Multidimensional scaling*. Londres: Sage Publications.
- Lakoff, G. (1987a). Cognitive models and prototype theory. In U. Neisser (Ed.), *Concepts and conceptual development* (pp. 63-100). Cambridge: Cambridge University Press.
- Lakoff, G. (1987b). *Women, fire and dangerous things: what categories reveal about the mind*. Chicago: Chicago University Press.
- MacKay, D.B. (1989). Probabilistic multidimensional scaling: an anisotropic model for distance judgments. *Journal of Mathematical Psychology*, 33, 187-205.
- Malt, B.C. & Smith, E.E. (1984). Correlated properties in natural categories. *Journal of Verbal Learning and Verbal Behavior*, 23, 250-269.
- Markman, A.B., & Gentner, D. (1993). Splitting the differences: a structural alignment view of similarity. *Journal of Memory and Language*, 32, 517-535.
- McCauley, R.N. (1987). The role of theories in a theory of concepts. In U. Neisser (Ed.), *Concepts and conceptual development* (pp. 288-309). Cambridge: Cambridge University Press.
- Medin, D.L., Altom, M.W., Edelson, S.M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 8, 37-50.
- Medin, D.L., Goldstone, R.L., & Gentner, D. (1990). Similarity involving attributes and relations: judgments of similarity and difference are not inverses. *Psychological Science*, 1, 64-65.
- Medin, D.L., Goldstone, R.L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, 100, 254-278.

- Medin, D.L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179-195). Cambridge: Cambridge University Press.
- Medin, D.L., & Schaffer, M.M. (1978). Context theory of classification learning. *Psychological Review*, 85, 207-238.
- Medin, D.L., & Shoben, E.J. (1988). Context theory of classification learning. *Psychological Review*, 85, 20, 158-190.
- Medin, D.L., & Smith, E.E. (1984). Concepts and concepts formation. *Annual Review of Psychology*, 35, 114-138.
- Medin, D.L., & Wattenmaker, W.D. (1987). Category cohesiveness, theories and cognitive archeology. In U. Neisser (Ed.), *Concepts and conceptual development* (pp. 25-62). Cambridge: Cambridge University Press.
- Medin, D.L., Wattenmaker, W., & Hampson, S. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology*, 19, 242-279.
- Melara, R.D., Marks, L.E., & Lesko, K. (1992). Optional processes in similarity judgments. *Perception & Psychophysics*, 51, 123-133.
- Minsky, M. (1975). A framework for representing knowledge. In P.H. Winston (Ed.), *The psychology of computer vision* (pp. 211-277). New York: McGraw-Hill.
- Murphy, G.L., & Allopenna, P.D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 904-919.
- Murphy, G.L., & Medin, D. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92, 289-316.
- Murphy, G.L., & Spalding, T. (1995). Knowledge, similarity, and concept formation. *Psychologica Belgica*, 35, 127-144.
- Murphy, G.L. & Wright, J.C. (1984). Changes in conceptual structure with expertise : differences between real-world experts and novices. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10, 144-155.

- Nosofsky, R.M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10, 104-114.
- Nosofsky, R.M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, 115, 39-57.
- Nosofsky, R.M. (1991). Stimulus bias, asymmetric similarity, and classification. *Cognitive Psychology*, 23, 94-140.
- Nosofsky, R.M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology*, 43, 25-53.
- Ortony, A., Vondruska, R., Foss, M., & Jones, L. (1985). Salience, similes and the asymmetry of similarity. *Journal of Memory and Language*, 24, 569-594.
- Podgorny, P., & Garner, W.R. (1979). Reaction time as a measure of inter-intraobject visual similarity: letters of the alphabet. *Perception & Psychophysics*, 26, 37-52.
- Reed, S.K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, 3, 382-407.
- Rips, L.J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21-59). Cambridge: Cambridge University Press.
- Rips, L.J. (1991). Similarity and the structure of categories. In D.J. Napoli & J.A. Kegl (Eds.) *Bridges between psychology and linguistics* (pp. 35-53). Hillsdale, N.J.: Erlbaum, 1991.
- Rips, L.J., & Collins, A. (1993). Categories and resemblance. *Journal of Experimental Psychology : General*, 122, 468-486.
- Rosch, E.H. (1975). Cognitive reference points. *Cognitive Psychology*, 7, 532-547. (b)
- Rosch, E.H. (1978). Principles of categorization. In E.H. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum.
- Roth, E.M., & Shoben, E.J. (1983). The effect of context on the structure of categories. *Cognitive Psychology*, 15, 346-378.
- Rumelhart, D.E. (1975). Notes on a schema for stories. In D.G. Bobrow & A.M. Collins (Eds.), *Representation and understanding: studies in cognitive science* (pp. 211-236). New York: Academic Press.

- Schank, R., (1982). *Dynamic memory: A theory of reminding and learning in computers and people*. New York: Cambridge University Press.
- Schank, R., & Abelson, R. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Erlbaum.
- Schank, R., Collins, G., & Hunter, L. (1986). Transcending inductive category formation in learning. *Behavioral and Brain Sciences*, 9, 639-686.
- Schiffman, S., Reynolds, L., & Young, F. (1981). *Introduction to multidimensional scaling*. New York: Academic Press.
- Schyns, P.G., Goldstone, R. & Thibaut, J.-P.(sous presse). The development of conceptual features. *Behavioral and Brain Sciences*.
- Shepard, R. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1, 54-87.
- Sjoberg, L. (1972). A cognitive theory of similarity. *Goteborg Psychological Reports*, 2 (10).
- Smith, C., Carey, S., & Wiser, M. (1985). On differentiation: a case study of the development of the concepts of size, weight and density. *Cognition*, 21, 177-237.
- Smith, E.E. & Medin, D. (1981). *Categories and concepts*. Cambridge: Harvard University Press.
- Smith, E.E., & Sloman, S.A. (1994). Similarity-versus rule-based categorization. *Memory & Cognition*, 22, 377-386.
- Thibaut, J.P. (1991). *Réurrence et variations des attributs dans la formation des concepts*. Unpublished doctoral thesis, University of Liège, Liège.
- Thibaut, J.P. (1995). The development of features in children and adults: the case of visual stimuli. *Proceedings of the seventeenth Annual Meeting of the Cognitive Science Society*. pp. 194-199. Hillsdale: N.J., Lawrence Erlbaum.
- Thibaut, J.P., & Schyns, P.G. (1995). The development of feature spaces for similarity and categorization. *Psychologica Belgica*, 35, 167-185..
- Thibaut, J.P. & Dupont, M. (1996). Conditions for dissociations between similarity and categorization. Manuscrit en préparation.

- Treisman, A. (1987). Properties, parts, and objects. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance* (pp. 35.1-35.69). London: Wiley.
- Tversky, A., & Gati, I. (1978). Studies in similarity. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 79-98). Hillsdale, NJ: Erlbaum.
- Tversky, A., & Gati, I. (1982). Similarity, separability, and the triangle inequality. *Psychological Review*, 89, 123-154.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327-352.
- Watenmaker, W.D., Dewey, G.I., Murphy, T.D. & Medin, D.L. (1986). Linear separability and concept learning: context, relational properties, and concept naturalness. *Cognitive Psychology*, 18, 158-194.
- Young, F.W. (1984). Scaling. *Annual Review of Psychology*, 35, 55-81.
- Zinnes, J.L., & MacKay, D.B. (1983). Probabilistic multidimensional scaling: complete and incomplete data. *Psychometrika*, 48, 27-48.

NOTES

a. L'auteur remercie Martine Poncelet, Aude Oliva, Serge Larochelle et un expert anonyme pour leur relecture attentive du manuscrit.

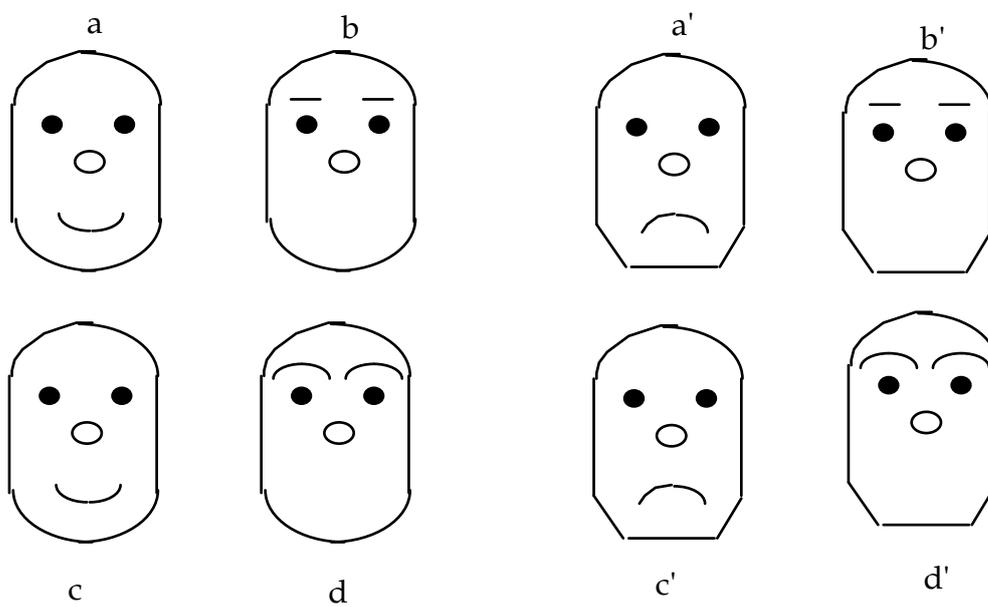
1. Nous préférons le terme similarité à similitude, non pas pour des raisons d'euphonie avec l'anglais similarity, mais parce que similitude réfère à l'identité tandis que similarité désigne une ressemblance n'impliquant pas l'identité. Similitude dans le sens de "posséder de caractéristiques communes" est considéré comme vieilli par le Robert. Le terme semblable n'implique pas cette référence à l'identité.

2. Nous utilisons les mots *traits* et *attributs* comme des synonymes.

3. L'auteur ne considère que les 3 dimensions qui expliquent une part significative de la variance. D'autres dimensions ont pu être extraites mais contribuant peu à l'explication des données. Selon les exemples étudiés, un plus grand nombre de dimensions pourraient émerger.

4. Cette saillance a été estimée indépendamment.

5. L'expérience a été réalisée avant la réunification des deux Allemagnes!



Légendes des figures.

Figure 1. Les visages diffèrent selon 3 attributs. Chacun de ces attributs peut être absent ou prendre une parmi deux valeurs lorsqu'il est présent. Les trois attributs sont le menton, la bouche, et les sourcils. Le menton arrondi est le seul attribut partagé par les paires a, b et c, d, le menton carré par les paires a', b' et c', d'. La bouche souriante est présente en a mais pas en b, en c mais pas en d; la bouche grimaçante est présente en a' mais pas en b', en c' mais pas en d'. Les sourcils droits sont présents en b mais pas en a, en d mais pas en c; les sourcils courbés sont présents en b' mais pas en a', en d' mais pas en c' (adapté d'après Tversky, 1977).