

DE L'HYPOTHESE DE SAPIR-WHORF AU PROTOTYPE : SOURCES ET GENESE DE LA THEORIE D'ELEANOR ROSCH*

Jean-Michel Fortis

Introduction

La théorie du prototype a pour origine les travaux d'Eleanor Rosch dans les années 70. Elle doit à son milieu d'émergence, l'université de Harvard, sa thématique de départ, la catégorisation, et plus précisément la catégorisation des couleurs. Nous verrons que la dette va au-delà de cette thématique, et que des aspects importants de la théorie ont été anticipés dans les travaux de Bruner, Lenneberg, et Brown (tous à Harvard).

La théorie de Rosch naît au sein d'un débat théorique précis, celui entourant l'hypothèse dite de Sapir-Whorf. Elle prolonge plusieurs tentatives visant à tester empiriquement cette hypothèse au moyen des méthodes de la psychologie. Sa discipline d'origine est donc en premier lieu la psychologie, et ses références intellectuelles viennent d'abord de celle-ci.

Pour comprendre la portée de la théorie, il est important de dissocier deux aspects: (1) la prise de position de Rosch dans le débat sur l'hypothèse de Sapir-Whorf, et le rôle de cette prise de position dans la théorie du prototype; (2) l'outil théorique que constitue cette théorie, qui peut être décrit et évalué indépendamment de la prise de position de Rosch sur l'hypothèse de Sapir-Whorf. Dans la suite, j'examinerai ces deux aspects tour à tour. La suite de cet article sera donc divisée en deux grandes sections : la première (parties A et B) sera consacrée au débat sur l'hypothèse de Sapir-Whorf et à l'intervention de Rosch sur la question. Il s'agira essentiellement des recherches sur la catégorisation des couleurs. La seconde (parties C et D) décrira l'outillage théorique qui a été progressivement élaboré à partir de cette prise de position, et a été enrichi pour tenir compte d'autres catégories. Ce sera aussi l'occasion de voir ce que Rosch doit à ses prédécesseurs dans l'élaboration de la notion de prototype et d'autres éléments de la théorie.

Avant d'en venir à la prise de position de Rosch sur l'hypothèse de Sapir-Whorf, je présenterai d'abord le contexte historique qui lui sert de cadre. Ce contexte est double: il s'agit en premier lieu, comme je l'ai mentionné, des recherches psychologiques sur l'hypothèse de Sapir-Whorf, en particulier des études portant sur la relation entre le lexique des couleurs et la reconnaissance ou la mémorisation des couleurs. Il s'agit en second lieu de la publication de l'étude de Berlin et Kay (1969) sur les *basic color terms*, qui exerça une forte influence sur la position personnelle de Rosch. Ce contexte historique permet de comprendre dans quelles circonstances a été élaborée la notion de prototype.

A. L'hypothèse Sapir-Whorf et la catégorisation des couleurs

La théorie du prototype s'inscrit initialement dans une série de recherches sur la catégorisation des couleurs, recherches dont le cadre méthodologique a été posé, au cours des années 50, dans les travaux d'Eric Lenneberg, Roger Brown et John Roberts. Ces travaux répondent à une préoccupation centrale : mettre à l'épreuve l'hypothèse de Sapir-Whorf sur les rapports de la langue et de la pensée en formulant cette hypothèse de façon à la rendre testable par les méthodes de la psychologie.

A.1. Whorf et le principe de relativité

* Je remercie Danièle Dubois pour la discussion que j'ai eue avec elle et sa lecture attentive de cet article. Merci aussi aux relecteurs, qui m'ont permis de compléter les références bibliographiques.

Il n'entre pas dans mon propos de présenter Whorf, ni de retracer en détail sa filiation ou ses antécédents. Cette tâche a été accomplie ailleurs.¹ On peut dire que sa perspective relativiste se situe dans un mouvement intellectuel ayant des racines historiques profondes, où s'exprime à la fois une certaine défiance à l'égard des pièges et illusions du langage, et, parfois, une confiance dans la capacité humaine à réformer son usage du langage. Cette ambivalence se trouve chez Sapir (Joseph 2002), et elle a présidé à la critique du langage menée à l'époque du *linguistic turn* (au début du 20^{ème} siècle).

La situation historique de Whorf ayant été brossée à très grands traits, voyons comment Whorf formule son "nouveau principe de relativité" (1956 : 214)² :

"le système linguistique d'arrière-plan (en d'autres termes, la grammaire) de chaque langue n'est pas seulement un instrument de reproduction servant à vocaliser des idées mais il est lui-même formateur d'idées, le programme et le guide de l'activité mentale de l'individu, de son analyse des impressions, de sa synthèse de ses ressources mentales. La formulation d'idées n'est pas un processus indépendant, strictement rationnel au sens ancien du terme, mais fait partie d'une grammaire particulière et diffère d'une grammaire à une autre dans des proportions plus ou moins grandes. Nous découpons la nature selon les tracés que notre langue a dessinés sur elle. Nous ne découvrons pas les catégories et les types que nous isolons à partir du monde des phénomènes parce qu'elles se trouveraient sous le nez de tout observateur ; au contraire, le monde se présente à nous sous la forme d'un flux kaléidoscopique d'impressions qui doit être organisé par nos esprits — c'est-à-dire en grande partie par nos systèmes linguistiques mentaux (...) Nous sommes ainsi confrontés à un nouveau principe de relativité, selon lequel les mêmes preuves physiques ne conduisent pas tous les observateurs à la même image de l'univers, à moins que leurs ressources linguistiques ne soient similaires ou puissent être ajustées de quelque manière" (Whorf 1956 [1940] : 212-3).

Ce texte semble attribuer au "système linguistique" la fonction que Kant avait dévolue aux catégories et aux formes de l'intuition (espace et temps) : celle de fournir des cadres d'appréhension des "phénomènes" et d'organiser le divers de l'expérience ("le flux kaléidoscopique des impressions"³). Certains l'ont interprété dans ce sens radical (Black 1959). De fait, Whorf s'intéressait aux conceptions fondamentales de l'univers véhiculées par la langue et aux idées les plus générales les sous-tendant (temps, espace et matière). Mais il pensait aussi que des effets cognitifs et comportementaux beaucoup plus spécifiques sont enveloppés dans les systèmes linguistiques. Il pouvait s'agir, plus simplement, de l'influence que peut avoir sur le comportement un mot qui prête à confusion et qui invite à catégoriser une situation dans les termes d'une autre (voir ci-après le cas de *empty*). Surtout, l'interprétation radicale a contre elle le fait qu'à plusieurs reprises Whorf postule une couche perceptive universelle où opèrent les principes d'organisation de la Gestalttheorie (voir sur ce point Lee 1996).

Sans qu'il soit possible de nous attarder ici sur l'exégèse de Whorf, nous retiendrons l'idée que la langue organise notre vision du monde (peut-être seulement au-delà de cette couche perceptive), sélectionne des aspects de l'expérience et sert à catégoriser cette expérience.⁴

¹ Pour une présentation de Whorf, voir Lee (1996). Koerner place Whorf dans une tradition qui remonte à Locke et Leibniz, et dont Humboldt, Boas et Sapir sont d'importants jalons. Joseph (2002) s'intéresse davantage aux sources possibles plus immédiates, y compris hétérodoxes (Korzybski et la théosophie). Lucy (1992) décrit en détail l'évolution des recherches en matière de relativité linguistique depuis Whorf inclus.

² Evidemment par analogie avec la théorie de la relativité d'Einstein. L'analogie se trouvait déjà chez Sapir (1949 [1924] : 159).

³ La formule évoque peut-être William James lorsqu'il décrit les impressions du bébé "as one great blooming, buzzing confusion" (James 1950 [1890] : 488).

⁴ Cette involution linguistique de Kant n'est d'ailleurs pas sans rappeler le néokantisme. Comme Cassirer (1910 [1977]), Whorf était préoccupé par la remise à jour qu'imposait la nouvelle science à nos cadres de pensée. La

Cette idée générale et vague, renommée, peut-être à la suite de Hoiijer et de Carroll, hypothèse de Sapir-Whorf⁵ est la version largement monétarisée de l'hypothèse qui a suscité les recherches dont il sera traité ici.

A.2. Whorf lu par Lenneberg

Lorsque Lenneberg s'empare de Whorf, il transforme le *principe* de relativité en une *hypothèse* de travail. Sa propre formulation de l'hypothèse est sur le mode interrogatif. C'est une question qu'il entend soumettre à l'expérimentation.

Voici comment il formule l'hypothèse : "La structure d'une langue donnée affecte-t-elle les pensées (ou la puissance de penser [thought potential]), la mémoire, la perception, la capacité d'apprentissage de ceux qui parlent cette langue ?" (Lenneberg 1953 : 463).

Ce qui intéresse Lenneberg n'est donc pas de savoir comment nos contenus de pensée sont modelés par notre langue ou conditionnent une vision du monde, mais de déterminer si la langue facilite ou inhibe le traitement de ces contenus par des facultés cognitives. En prélude à la présentation de sa propre stratégie, Lenneberg va d'abord expliquer pourquoi il a dû se séparer de la démarche de Whorf. C'est surtout au procédé de traduction, abondamment employé par Whorf, qu'il s'attaque.

Pour souligner l'hétérogénéité des conceptions du monde contenues dans des langues amérindiennes, Whorf avait coutume de gloser littéralement des morphèmes pour en faire ressortir l'étrangeté. Pour Lenneberg, l'étrangeté des langues amérindiennes est une conséquence de la manière qu'a Whorf de gloser les énoncés en traduisant chaque morphème par son sens général. On obtiendrait des paraphrases tout aussi étranges, observe-t-il, si on substituait aux morphèmes de nos langues leurs gloses générales. Par exemple, quand Whorf donne d'un énoncé en langue apache correspondant à *c'est une source qui goutte* la glose littérale *en tant qu'eau ou source, de la blancheur se déplace vers le bas*, il ne tient compte que des sens généraux des différents morphèmes (Whorf 1956 [1941b] : 241). Or, objecte Lenneberg, la paraphrase générale et littérale de *it is a dripping spring* serait tout aussi étrange. En outre, poursuit Lenneberg, les morphèmes ne sont ici que des béquilles à usage mnémotechnique servant à construire, dans le contexte de la situation, un ensemble d'associations plus riches que leurs gloses générales. Réduire la signification d'un énoncé à une addition de gloses générales laisse échapper la signification qui est construite dans le contexte d'une situation et qui est justement la contrepartie cognitive qui intéresse Whorf (Lenneberg 1953 : 466). Ce sont donc plutôt les associations en rapport avec la situation qui comptent. Or, celles-ci renvoient à des données de la culture et de l'environnement physique qui sont souvent incomparables d'une société à l'autre.

De même, reprenant l'exemple fameux de Whorf sur *empty*⁶, Lenneberg impute la faute non à l'équivocité du mot, mais à l'ignorance du sujet qui précipite l'accident. C'est selon lui le contexte cognitif de l'individu, qui est cause de l'accident, et non le mot, qui est plutôt un phénomène stylistique, laissé à la liberté du locuteur.

description de grammaires hétérodoxes a donc une portée épistémologique et propédeutique, celle de nous inviter à embrasser des visions du monde alternatives, et d'accomplir ainsi le décentrement que la science moderne exige.

⁵ Selon Koerner (2002 : 39), Hoiijer (1954) pourrait être le premier à avoir parlé d'hypothèse de Sapir-Whorf, mais ce serait plutôt Carroll qui, dans sa préface à son anthologie de Whorf, aurait popularisé l'appellation (Carroll 1956 : 27). Il faut noter que Lenneberg (1953) renvoie à Whorf, non à Sapir.

⁶ Whorf, qui travaillait pour une société d'assurances, avait observé dans des constats d'accidents que des explosions semblaient provenir de ce que des bidons contenant en réalité un résidu de gaz avaient été malencontreusement considérés comme *vides*. Il avait attribué ces accidents à l'équivocité du mot *vide*, qui renvoie à la fois au fait qu'un contenu a été vidé (mais peut laisser des résidus) et à un sens plus abstrait, synonyme d'*inerte, sans effet* (Whorf 1956 [1941a]).

Il est clair que pour Lenneberg on ne peut s'attaquer à la question de la relativité linguistique à partir d'énoncés où on ne sait pas dans quelles proportions le sens des morphèmes et le contexte culturel, environnemental ou individuel contribuent à construire la signification globale. Ce que Lenneberg demande, c'est un domaine référentiel qui jouerait le rôle d'un pur donné indépendant de ces contextes, et dont le découpage par la langue ne laisserait pas de latitudes au locuteur (contrairement à *empty*). Il cherche donc à isoler un tel domaine, comparable de langue à langue, et ces aspects de l'énoncé qui relèvent de qu'il appelle la *codification*, c'est-à-dire des distinctions sémantiques que la langue *contraint* un locuteur à communiquer (Lenneberg 1953 : 467-8). Dans cette perspective, la découverte d'un authentique effet whorfien consisterait dans la mise en évidence d'une relation entre la codification et un comportement qui serait lui-même l'indice d'une différence de traitement : "Si l'investigateur s'intéresse, comme moi, à la cognition, il cherchera à déterminer les relations qui existent entre la codification et le genre de comportement qui caractérise la mémoire, la reconnaissance, la résolution de problème, la formation de concepts et la perception, avec l'espoir de montrer que certaines spécificités de ces processus ne peuvent être expliquées que par les spécificités de la codification" (Lenneberg 1953 : 468).

La procédure serait donc la suivante : on sélectionne un domaine, on détermine quelles sont les réponses linguistiques contraintes (ou les plus contraintes) à des aspects de ce domaine, et on observe si ces réponses ont un effet sur le comportement. Pour des psychologues comme Lenneberg et son collègue Brown, ce dernier point est crucial : comme le notera Brown (1976 : 128-9) chez Whorf, les preuves des effets de la langue sur la cognition non-linguistique faisaient cruellement défaut et l'argumentation semblait replier la langue sur elle-même. La mise en évidence d'un effet cognitif était susceptible de rompre cette circularité.

Le domaine que choisit Lenneberg est le spectre des couleurs, dont il présume qu'elles peuvent être *discriminées* de la même manière par tous les êtres humains, quoique les *groupements* (*groupings*) de couleurs varient de langue à langue (Lenneberg 1953 : 468-9). Le terme de groupement renvoie très probablement à la notion psychologique de *catégorie*, c'est-à-dire à une classe de stimuli considérés comme équivalents du point de vue de certains de leurs attributs. Catégoriser, c'est alors déterminer ces attributs et ainsi l'extension d'une classe. Chez Bruner par exemple (Bruner et al. 1956), une tâche de catégorisation consistait à soumettre à des sujets des cartes portant un motif variant sur plusieurs dimensions (par ex. forme et couleur) et à leur demander de trouver le "concept", c'est à dire le principe de classement déterminé à l'avance par l'expérimentateur (par exemple 'carré et rouge'). Le problème est qu'en assimilant la couleur à un "concept" et la catégorie correspondante à une classe de stimuli, Lenneberg suppose que nommer des couleurs revient à classer les stimuli selon un principe de classement déjà connu, c'est-à-dire en fonction de leurs teinte, saturation et brillance. Or, il ne va nullement de soi que la couleur soit vraiment une dimension pertinente pour certaines langues, ni qu'elle se réduise aux dimensions de teinte, saturation et brillance.⁷

Dans le cadre de l'expérience sur les couleurs, les réponses cherchées sont les noms données aux couleurs, et un des comportements testés peut être la reconnaissance. Il s'agira ainsi de savoir si les couleurs nommées facilitent la reconnaissance, plus exactement si les couleurs les mieux codables sont les plus faciles à reconnaître ultérieurement à une première présentation. Selon Lenneberg, le dispositif n'a pas besoin d'une comparaison inter-langues : une langue suffit à mettre en évidence un effet whorfien. En somme, une même langue contient en elle-

⁷ Ce point, soulevé déjà par Boas, a été souligné par plusieurs auteurs. Saunders (1992) donne un aperçu de ces objections et des références. Il faut noter ainsi que dans certaines cultures les couleurs sont ancrées sur des objets privilégiés (par ex. le bétail) ou ne sont pas séparables d'autres dimensions, par exemple, dans un système ancré sur la végétation, le sec, le frais, le mûr, le pourri. Les pastilles du nuancier Munsell constituent alors un matériau artificiel qui préjuge de ce qu'est la couleur.

même, avec des noms de couleurs plus ou moins codables, la diversité qu'on chercherait inutilement dans plusieurs langues (Brown 1976 : 129). Dans cette perspective, la comparaison interlinguistique ajoute de l'eau au moulin relativiste mais n'est plus centrale pour la démonstration. En outre, le problème de la traduction est éliminé, parce que le protocole "objectifie les sens intuitionnés des formes en prenant soin de les rapporter à des stimuli de l'environnement. Ainsi, il est possible (du moins dans certains cas) de spécifier le sens par référence aux propriétés physiques des stimuli" (Lenneberg 1953 : 471). On voit que pour la rendre testable, l'hypothèse de Sapir-Whorf a changé de nature : le "sens" est ancré sur les propriétés physiques du référent et il ne s'agit pas de comprendre ce que ce sens doit au fonctionnement grammatical des noms ou adjectifs de couleur, ni si la couleur est une dimension pertinente de la culture considérée. En fait, et comme le fera Rosch plus tard, Lenneberg semblait assimiler purement et simplement *sens* et *référence*. Signe de cette candeur, Brown (1976 : 125) nous apprend incidemment que c'est Chomsky qui les invitera à distinguer *meaning* et *reference*.

A.3. Mise en place du protocole : l'épreuve "intra-culturelle"

Le protocole de Brown et Lenneberg (1954) est la mise en œuvre des principes méthodologiques développés par Lenneberg dans l'article que nous venons de présenter. Le but de leur expérience est d'évaluer l'effet d'une variable, la codabilité d'une couleur, sur une autre, la reconnaissance de cette couleur. L'expérience est "intra-culturelle" au sens de Lenneberg, c'est-à-dire concerne une seule langue, l'anglais, et des sujets américains. Alors que jusqu'ici le matériel traditionnel pour étudier la dénomination des couleurs consistait en papiers ou brins de laine colorés, Brown et Lenneberg emploient pour la première fois le nuancier de Munsell, où les couleurs sont étalonnées rigoureusement en fonction de trois dimensions (teinte, brillance et saturation).⁸ Les couleurs sélectionnées pour la tâche sont les meilleurs exemples, déterminés par un jury de sujets, d'une série de teintes toutes à saturation maximale (rouge, orange, jaune, vert, bleu, violet [purple], rose, marron [brown]). La codabilité d'une couleur est définie par un éventail de mesures : la longueur du

⁸ Voir Saunders (1992) pour une mise au point historique sur les recherches antérieures à Lenneberg et leurs protocoles. Saunders (1992 : 14) remarque qu'en choisissant le nuancier Munsell Lenneberg ne fait que traiter l'un des modes d'existence de la couleur, celui de la couleur de *surface* (et non, par exemple, la lumière colorée ou la couleur d'un milieu "dématérialisé"). La pastille de couleur du nuancier a un statut très particulier : elle entre dans un système différentiel dont la prise de connaissance par le sujet lui suggère de rejeter certaines dimensions comme non pertinentes (le brillant, la texture ou la matière) ; elle joue aussi le rôle d'étalon de *la* couleur, c'est-à-dire de la couleur dont les manifestations spécifiques dans tel ou tel contexte seraient des incarnations (Dubois & Cance 2009). En tant qu'elle doit être prise comme neutralisant certaines dimensions perceptives, comme un élément d'un système et comme étalon, elle fonctionne comme *signe* d'une couleur (voir aussi Latour 1993 sur ce point). Comme le notent Dubois et Cance, c'est un entour culturel qui investit une pastille de son statut de signe et lui fait jouer le rôle qu'elle est amenée à avoir dans l'épreuve psychologique : "La pastille de couleur comme représentation adéquate d'une couleur doit alors être conçue non comme une représentation naturelle de la couleur (ou des couleurs) mais comme un produit d'activités cognitives non seulement psychologiques, individuelles, et comme la (re)matérialisation d'une abstraction historiquement produite par une culture, qui l'investit d'un statut ontologique, soit de représentation d'une réalité du monde, soit d'une catégorie de l'entendement ou d'une vérité attestée parce que fondée dans la science physique" (Dubois & Cance 2009 : 80). Que les *termes* de couleur retenus dans ces expériences soient considérés comme donnant accès à *la* couleur désignée ne fait qu'ajouter au caractère très spécifique du contexte expérimental, dès lors qu'en dehors de ce contexte les façons de désigner les couleurs varient en fonction de leur mode de donation matériel (surface, lumière, milieu), du dispositif de présentation (image bidimensionnelle ou objet "réel" par ex.), du contexte discursif (*rouge Pompéi* si on parle de peintures plutôt que de rouge à lèvres), et, pour les couleurs de surface, selon qu'elles sont vues comme pigment, comme entité colorée (*sable*), comme coloration (*blond*), ou sont pourvues de valeurs symboliques (*rouge Marilyn* pour un rouge à lèvres ; cf. Dubois & Cance 2009).

nom (plus le nom est court, plus la couleur est censée être codable⁹) ; le temps mis pour la dénommer ; le degré d'accord inter- et intra-subjectif sur le nom approprié. La tâche de reconnaissance sera d'identifier, parmi 120 couleurs, 4 couleurs présentées antérieurement, certaines faisant partie de ces meilleurs exemples, d'autres noms.

Les résultats confirment le lien supposé entre la codabilité et le taux de réussite à la tâche de reconnaissance. Mais que disent-ils de l'hypothèse de Sapir-Whorf ? Comme le note Lucy (1992 : 141), celle-ci semble bien avoir été dissoute sur le chemin. Si l'on était resté fidèle à Whorf, l'hypothèse aurait dû porter sur la possibilité qu'une langue soit *constitutive* des catégories perceptives, mais non sur le fait que des mots puissent faciliter le traitement de couleurs "objectives". A cet égard, l'effet le plus fort était à chercher dans les jugements de typicalité préalables à l'expérience et qui avaient permis de sélectionner les meilleurs exemples de couleur. Pour un adepte du principe de relativité, et à une époque qui ne songeait pas encore à des couleurs focales universelles, la saillance manifeste de ces couleurs, reflétée dans l'uniformité des jugements de typicalité, aurait pu passer pour un effet particulièrement puissant des discriminations imposées par les noms.

Enfin, faute d'une comparaison avec une autre langue que l'anglais, on ne peut savoir si les noms de couleur sont la cause d'une meilleure reconnaissance ou sont seulement l'expression d'une meilleure discriminabilité. C'est précisément la lacune qu'entend combler l'étude dont nous allons parler maintenant.

A.4. L'épreuve "inter-culturelle"

L'étude de Lenneberg et Roberts (1956) est un pas décisif en direction des théories universalistes de Berlin et Kay, et de Rosch. Non qu'elle ait été écrite dans un esprit universaliste, bien au contraire, mais elle introduit un protocole et des notions, et obtient des résultats qui préparent aux recherches universalistes.

Les sujets sont des anglophones et des Zunis (ethnie du Nouveau Mexique, leur langue est non affiliée) à qui on propose, selon un protocole innovant, d'indiquer directement les contours de leurs catégories de couleurs sur un large échantillon du nuancier de Munsell.¹⁰ Ils doivent aussi indiquer sur le nuancier la ou les pastille(s) de couleur qui leur semblent les plus typiques de la catégorie. Comme la langue des Zunis ne distingue pas le jaune et l'orange, Lenneberg et Roberts font l'hypothèse qu'ils ne les reconnaîtront pas non plus dans la tâche de mémorisation. Cette attente est pleinement satisfaite mais leurs résultats font apparaître un phénomène qui n'avait pas été anticipé : certaines de couleurs les plus typiques sont identiques de langue à langue, même si les frontières des catégories diffèrent. Quoique cette observation présage des résultats qu'obtiendront Berlin et Kay (1969) et anticipe clairement leur notion de couleur focale, Lenneberg et Roberts n'en tirent aucune conclusion qui irait à rebours de leur perspective relativiste.

Deux aspects essentiels des travaux futurs sont désormais en place : le protocole et les couleurs focales. Il faut y ajouter un troisième élément : la typicalité. Dès sa thèse (dont un résumé se trouve chez Lenneberg 1957), Lenneberg, qui anticipait très nettement Rosch, invoquait pour la description des catégories de couleur une conception non classique : non pas la formation d'ensembles de couleurs aux frontières nettes, mais de catégories aux limites

⁹ L'inspiration vient ici de Zipf (1935), qui mettait en relation la longueur d'un mot avec sa fréquence. Or, Lenneberg pense que la fréquence d'un mot est un bon index de la disponibilité de la catégorie à laquelle il renvoie. Autrement dit, la catégorisation d'une couleur, entendue ici comme sa dénomination, est facilitée si la fréquence du nom est forte.

¹⁰ L'échantillon contenait 320 pastilles (40 teintes x 8 niveaux de brillance) à saturation maximale (Berlin & Kay 1969 utiliseront le même échantillon, augmenté de 9 teintes achromatiques; cf. Saunders 1992 pour une discussion). Au moyen d'un marqueur, les sujets devaient dessiner le contour sur une feuille d'acétate placée sur le nuancier.

chevauchantes et des exemplaires plus ou moins typiques. Il étendait même, comme le fera Rosch, la notion de degrés de typicalité aux concepts en général : “la meilleure caractérisation des concepts, écrivait-il, est d’en faire des zones de typicalité croissante et décroissante sur un continuum de stimulation” (Lenneberg 1957 : 2, cité par Brown 1976 : 135).

On le voit, les recherches de Lenneberg, Brown et Roberts sont comme la matrice dont découleront la théorie de Berlin et Kay et, par leur intermédiaire, la forme initiale de la théorie du prototype. Mais des tensions au sein des tenants du relativisme vont les conduire à infléchir leur position et préparer le terrain à un renversement de perspective.

A.5. Le retournement universaliste

Les acquis de Brown et Lenneberg ne tardèrent pas à être disputés. Peu de temps après, Burnham et Clark (1955) publièrent une étude qui contredisait expressément la thèse d’une corrélation positive entre codabilité et reconnaissance, en utilisant cette fois des couleurs moyennement saturées. Lenneberg lui-même, pour des raisons qui tenaient surtout à ses recherches sur les rapports entre maturation cognitive et le développement du langage chez l’enfant, s’éloignait de son relativisme initial pour adopter une position beaucoup plus biologisante, qu’il qualifiait de “néokantienne” (Lenneberg 1962).

Certains chercheurs (Lantz et Steffle 1964, Steffle et al. 1966¹¹) tentèrent de surmonter la contradiction entre les résultats de Brown et Lenneberg (1954) et Burnham et Clark (1955) en présentant un nouvel index, différent de la codabilité et plus proche de l’acte de communication. Cet index, la précision communicationnelle (*communication accuracy*) reposait sur l’efficacité avec laquelle la description d’une couleur par un locuteur permettait à un autre sujet d’identifier ladite couleur dans un nuancier. Ils trouvèrent que le nouvel index prédisait à la fois les résultats de Brown et Lenneberg et ceux de Burnham et Clark. Malgré ce résultat, cette approche n’emporta pas la décision à long terme. Les réflexions de Brown (1976 : 145) permettent de comprendre pourquoi : en se focalisant sur l’acte de communication, les recherches de Lantz et Steffle portaient non sur un système de catégorisation mais sur l’efficacité individuelle à communiquer. Autrement dit, l’index de précision communicationnelle était trop psychologique, et pas assez linguistique.

Mais le retournement universaliste est surtout le signe que le contexte épistémologique était en train de changer. L’expansion de la grammaire générative (en particulier à partir d’*Aspects*) créait un climat favorable à une perspective universaliste.¹² La psychologie cognitive, qui tissait des liens avec l’intelligence artificielle et un modèle computationnel de l’esprit, s’enhardissait dans le champ de la “pensée” et de la sémantique et tendait à promouvoir une conception individualiste, potentiellement physicaliste, des “concepts”.¹³ La croissance des

¹¹ Toujours la filière de Harvard : Lantz et Steffle étaient alors étudiants à cette université (Brown 1976).

¹² En outre, la grammaire générative était perçue par les linguistes comme réintroduisant la sémantique dans la grammaire (ce dont témoigne l’avènement de la sémantique générative, perçue au début comme prolongeant les travaux de Chomsky). Dans ce cadre, la recherche de dénominations universelles des noms de couleurs pouvait apparaître comme un pan de la quête des universaux substantiels.

¹³ En psychologie, un ensemble de facteurs conspirent ainsi à la naturalisation de la pensée. Plusieurs livres dont le titre fait référence à la pensée paraissent dans les années 50, et marquent un renouveau d’intérêt pour des processus mentaux qui vont au-delà de la perception. Bruner, à Harvard, a joué un rôle important dans le lien entre recherches sur la perception, théorie de la catégorisation et de la résolution de problèmes (Bruner, Goodnow & Austin 1956). Le modèle computationnel de l’esprit facilite évidemment cette résurgence (voir en psychologie Neisser 1967). Certains travaux sur la mémoire et la catégorisation emploient des formalismes qu’on retrouve en simulation artificielle (Collins & Quillian 1969). Des études sur la mémoire commencent à parler de regroupements d’items mémorisés en fonction de rapports sémantiques (Tulving 1968). Cette convergence participe à l’impression générale d’un “tournant cognitif”, associé chez ses acteurs à un retour aux processus mentaux “supérieurs”. Le thème de la catégorisation, intermédiaire entre perception et pensée, sert en quelque sorte de transition.

études ethnoscientifiques, favorisée par les analyses componentielles des systèmes de parenté, dont la rigueur impressionnait les anthropologues, contribuait à orienter l'anthropologie vers une problématique référentialiste. Enfin, toujours en anthropologie, commençait à réémerger une vision évolutionniste et quasi organiciste de la culture comme prolongeant le processus d'adaptation à l'environnement et d'exploitation de ses ressources (Sahlins & Service 1960). Ainsi se constituait le terreau qui nourrirait le modèle universaliste de Berlin et Kay (1969), modèle auquel Rosch empruntera la notion de couleur focale.

A.6. Berlin et Kay (1969)

L'étude de Berlin et Kay est une enquête comparative sur la catégorisation des couleurs portant sur 98 langues. Pour 20 langues, les données sont de première main (mais souvent recueillies auprès d'un seul informateur, de surcroît connaissant l'anglais) ; les matériaux des 78 autres langues sont glanés dans des sources littéraires diverses, parfois anciennes. La première phase est la collecte de noms de couleur qui, filtrés et épurés selon une série de critères, sont réduits à une liste de *basic color terms*, c'est-à-dire de termes jugés comme particulièrement saillants dans la langue considérée.¹⁴ Le protocole expérimental reprend celui de Lenneberg et Roberts : les informateurs, auprès desquels on avait collecté des noms de couleur doivent indiquer les contours de la catégorie et le(s) meilleur(s) exemple(s) de la catégorie.

Comme Lenneberg et Roberts à une échelle plus réduite, les auteurs observent une certaine invariance des couleurs typiques (ou focales, *foci*) correspondant aux *basic color terms* mais des variations aux frontières des catégories. Présumées universelles, les couleurs focales suggèrent l'existence d'un déterminisme biologique (Berlin et Kay 1969 : 109). D'autre part, la typologie des inventaires de noms de couleur, augmentée des informations collectées dans les sources livresques, fait apparaître plusieurs systèmes récurrents que Berlin et Kay interprètent comme autant de stades dans une évolution qui mène d'un système dual (réduit à blanc / noir) à un système pleinement différencié de 11 couleurs dénommables par des termes de base.¹⁵ Cette évolution refléterait un développement économique et technologique, les systèmes les moins différenciés étant propres aux peuples "primitifs" (sic) de ce point de vue, tandis que les systèmes plus complexes (dont, évidemment, l'anglais) seraient caractéristiques des "nations les plus civilisées du monde" (Berlin 1970 : 14).

D'une méthodologie discutable (voir le choix des informateurs), parfois infidèle à ses propres règles (dans la sélection des termes de base), pleine d'erreurs de toutes sortes (Saunders 1992), accusée de redécrire les données disponibles pour les conformer à l'hypothèse de départ (Hickerson 1971), l'étude de Berlin et Kay n'en a pas moins reçu un accueil globalement favorable (Saunders 1992). Sa valeur a été amplifiée par le succès de la théorie du prototype et sa réduction à un modèle neurophysiologique (Kay & McDaniel 1978), deux développements postérieurs qui ont contribué à la légitimer (voir par exemple Lakoff 1987).

¹⁴ Les critères sont les suivants : un terme de base doit être monolexémique, ne doit pas être un sous-ordonné (comme *écarlate* : 'sorte de rouge'), ne doit pas être d'application restreinte à un domaine (comme *blond*), doit être psychologiquement saillant (facilement élicité). En réalité, ces critères n'ont pas été appliqués de manière stricte (voir Saunders 1992 pour un examen critique). Enfin, le mot *terme* signale que Berlin & Kay ne s'intéressent pas à la catégorie lexicale des désignations qu'ils obtiennent.

¹⁵ L'évolution postulée était la suivante (Berlin & Kay 1969 : 4) :



Elle a depuis connu de multiples modifications sans que ses fondements aient été remis en cause par ses auteurs (Kay 1975, Kay & McDaniel 1978, Kay et al. 1997).

Depuis Lenneberg jusqu'à Berlin et Kay, nous sommes passés du relativisme à l'universalisme. Après avoir récapitulé les étapes de ce cheminement, Brown note justement : "Il y a une fascinante ironie dans cette recherche en ce qu'elle a débuté dans un esprit fortement imprégné de relativisme et de déterminisme linguistique, et en est venue maintenant à défendre une position d'universalisme culturel et d'insignifiance du linguistique" (Brown 1976 : 152).

Pour notre propos immédiat, l'important est que la théorie universaliste de Berlin et Kay a donné l'impulsion aux recherches de Rosch. C'est vers elle que nous allons maintenant nous tourner.

B.1. Les débuts de Rosch

En guise de préliminaire, il n'est pas sans intérêt de retracer brièvement la carrière d'Eleanor Rosch. Après des études de philosophie qu'elle conclut par un mémoire sur Wittgenstein, elle s'oriente vers la psychologie clinique, qu'elle part étudier à Harvard (Rosch 1999b). Elle abandonne rapidement la psychologie clinique pour entreprendre une thèse, sous la direction de Roger Brown¹⁶, qu'elle soutient en 1969 (Rosch Heider 1969). Son travail porte à la fois sur les aspects sociaux et cognitifs du développement linguistique. Il met en rapport les capacités communicatives d'enfants avec leur manière d'aborder certaines tâches, et cette variable est corrélée à son tour à des différences de classe sociale. Dans un article qui fait suite à sa thèse, Rosch entend montrer que les enfants de classe défavorisée traitent moins bien certaines tâches en raison d'une précipitation excessive à répondre, d'un mauvais *tempo* (Heider 1971a). Elle suggère que le meilleur remède à cette impulsivité réside dans une prise de conscience de la stratégie adaptée à la tâche. Les notions de tempo et d'impulsivité, l'angle de recherche et les protocoles expérimentaux choisis par Rosch reflètent l'influence de Jerome Kagan, alors à Harvard, et devenu depuis l'un des psychologues développementalistes les plus célèbres (cf. Kagan 1964, 1966).

L'année de sa thèse, Rosch accompagne son mari, l'anthropologue Karl Heider¹⁷, sur son terrain de recherche, un peuple de l'ouest de la Nouvelle Guinée, les Dani. Elle décide de mettre à profit son séjour en étudiant ses hôtes (Rosch 1999). Son intérêt se porte sur les interactions mère-enfant et surtout sur la catégorisation des couleurs et des formes. Ce dernier thème, une fois de plus, était très "harvardien". Comme nous l'avons vu, la catégorisation était une des préoccupations centrales de Brown et Lenneberg (tous deux à Harvard, rappelons-le), mais aussi de Jerome Bruner, professeur à Harvard depuis 1952, qui avait écrit sur la question de la catégorisation un ouvrage marquant (Bruner et al. 1956). Nous aurons l'occasion de revenir sur l'influence de Bruner.

La rencontre des Dani va permettre à Rosch de prolonger ses recherches sur la catégorisation des couleurs chez les enfants, recherches qu'elle venait d'entamer. L'impulsion provenait de l'ouvrage récemment publié de Berlin et Kay (1969), qui semblait ouvrir la voie à une contestation du "nominalisme" que tendaient à promouvoir les études de Brown, Lenneberg, Lantz et Steffle. Ces auteurs sont les interlocuteurs que dans ses articles Rosch convoque, jouant Berlin et Kay contre les "nominalistes".

¹⁶ Le même Roger Brown qui collabora avec Lenneberg sur la reconnaissance des couleurs en fonction de leur codabilité (Brown et Lenneberg 1954).

¹⁷ Fils du psychologue autrichien Fritz Heider (1896-1988). Fritz Heider connaissait Egon Brunswik, autrichien d'origine et psychologue comme lui, dont le fonctionnalisme probabiliste allait exercer une influence sur Rosch, comme nous le verrons (section D.6).

Un premier article de Rosch sur la question de la catégorisation des couleurs avait montré que les enfants choisissaient préférentiellement de désigner les couleurs focales de Berlin et Kay lorsqu'on leur demandait de montrer la couleur correspondant à un nom fourni par l'expérimentateur, ou simplement de choisir une couleur parmi d'autres. Il leur était aussi plus facile de retrouver une couleur donnée parmi un choix d'échantillons possibles si la couleur testée était focale (Rosch Heider 1971b). Si Rosch considérait que ces résultats venaient bien conforter la position de Berlin et Kay, c'est-à-dire celle de couleurs focales universelles, ils n'apparaissent pourtant pas concluants. Dans un tel dispositif, et de l'aveu même de Rosch, la part du déterminisme cognitif et celle d'un biais linguistique possible étaient difficiles à séparer, puisque les enfants avaient pu être influencés par les catégories de l'anglais.

A cet égard, les Dani arrivaient à point nommé. Leur langue semblait réduire la dénomination des couleurs à deux termes, associés surtout à la brillance (l'un pour les couleurs "chaudes", l'autre pour les couleurs "froides"), et paraissait dépourvue de noms pour les couleurs focales de Berlin et Kay. Vu cette pauvreté du vocabulaire, confiné à ce contraste binaire entre couleurs "chaudes" et couleurs "froides", Rosch raisonnait que la mise en évidence d'un meilleur traitement des couleurs focales ne pouvait être que d'origine cognitive. Ainsi, les Dani permettraient d'exclure l'hypothèse d'un biais linguistique.

Sur ce dernier point, il faut observer d'emblée que Rosch a eu tendance à exagérer dans la présentation de ses expériences la pauvreté du lexique des couleurs des Dani, peut-être afin de rendre ses résultats d'autant plus éclatants. Il lui arrive en effet d'affirmer que le lexique des couleurs est chez les Dani essentiellement réduit à deux termes, citant à l'appui de cette affirmation l'étude ethnographique de son mari, Karl Heider. Cependant, cette affirmation n'est confirmée ni par la lecture des passages que Karl Heider consacre aux couleurs dans son livre (Heider 1970 : 49, 175, 289), ni par les remarques incidentes que fait Rosch elle-même dans une étude méthodologique menée parallèlement à ses expériences (Rosch 1972c). Recueillant les dénominations données par les Dani aux échantillons des couleurs saturées, Rosch fait les observations suivantes : "Les exemples de *mili* [le terme pour les couleurs "froides"] étaient situés avec constance parmi les verts et les bleus les plus sombres. Toutefois, il est apparu que *mola* [le terme pour les couleurs "chaudes"] avait deux points focaux ; le plus communément, un rouge sombre, et sinon un rose pâle. (...) Le second résultat inattendu était qu'en plus des termes de base *mili* et *mola* environ la moitié des informateurs employaient de façon régulière des noms de couleur à peu près équivalents aux termes anglais "rouge", "jaune" et "bleu". (...) Vingt informateurs (50 %) employaient un terme pour "rouge"." (Rosch 1972b : 451).

Rosch tentera d'annuler cette modeste mais embarrassante libéralité lexicale en choisissant dans certaines expériences ceux des sujets qui n'emploient que *mili* et *mola*, ce qui n'exclut pas une connaissance passive des termes (Rosch 1972a : 16 ; Rosch 1973a : 332).¹⁸ On note aussi que *mola* a deux couleurs focales (rouge sombre, rose pâle).¹⁹ Quant à la variabilité des réponses, elle est attribuée à un système en évolution, où les noms de couleur sont en train de recomposer leur découpage du spectre (Rosch 1972b).

Enfin, il convient de noter, si le lecteur ne l'a déjà fait, qu'à la paire *noir / blanc* de Berlin et Kay Rosch a substitué l'opposition *couleur froide / couleur chaude*, sans pour autant questionner ce changement ni se demander si la couleur est vraiment une dimension pertinente dans la langue des Dani. En réalité, de la nature linguistique des termes de couleur, de leur fonctionnement dans la langue, Rosch ne nous apprend rien (et son époux guère plus).

¹⁸ Saunders (1995) laisse entendre que l'emploi par Rosch d'une variété appauvrie de la langue dani (le *Police Dani*, usitée dans les relations entre Hollandais et Dani) a pu contribuer à la simplification des réponses, donc à une tendance à réduire les choix lexicaux à l'alternative *mili / mola*.

¹⁹ Ce fait conduira Berlin et Kay à modifier leur théorie initiale et à introduire la notion de *catégorie composite* (Berlin et al. 1997).

Le rapport entre leur emploi en situation réelle et en situation expérimentale n'est pas davantage abordé.

B.2. Recherches sur les Dani

Les recherches sur les Dani se développent dans deux directions, toutes deux destinées, dans la perspective de Rosch, à mettre à l'épreuve l'hypothèse de Sapir-Whorf. D'une part et dans un premier temps, il s'agit de vérifier si leur mémoire des couleurs est différente de celle des anglophones, ce qui, si tel était le cas, plaiderait pour une influence de la langue (Rosch Heider & Olivier 1972).

D'autre part, et dans un second temps, leur langue étant dépourvue de noms "stables" pour les couleurs focales, ils sont, à l'égard des catégorisations possibles de l'espace des couleurs, des sujets que Rosch considère comme vierges. On peut donc leur enseigner différentes partitions possibles de cet espace. Ces partitions seront considérées comme "naturelles", si elles sont construites autour des couleurs focales de Berlin et Kay (1969), ou "non naturelles" sinon (par ex. in Rosch Heider 1973a : 332). On examinera alors si les Dani ont une préférence pour les partitions naturelles.

B.3. Influence des catégories de couleur sur la reconnaissance et la mémoire

En ce qui concerne le premier type d'interrogation, qui est de savoir si les sujets sont influencés par leur langue lorsqu'ils doivent se remémorer une couleur, le protocole de Rosch est le suivant : elle soumet à ses sujets américains et dani deux tâches de dénomination, consistant à donner les noms des couleurs de deux échantillons de tailles différentes, et une tâche de mémorisation, qui demande aux sujets de désigner dans un échantillon la couleur qu'ils ont vue trente secondes plus tôt. Sur la base des résultats obtenus, elle construit deux matrices, l'une où la case (i,j) indique le nombre de fois où la couleur i a suscité comme réponse la couleur j dans la tâche de mémorisation ; l'autre où la case (i,j) indique le nombre de sujets qui ont donné le même nom aux deux couleurs. Un traitement de ces données permet d'en élaborer une représentation tridimensionnelle et de visualiser le degré de similarité des réponses fournies par les sujets dans ces deux types de tâches.

Rosch raisonne que si l'hypothèse "whorfienne" était valide, les deux configurations seraient similaires pour une même langue, et différentes d'une langue à l'autre. Or, constate-t-elle, un rapprochement entre les configurations obtenues aux tâches de mémorisation par les deux cultures fait apparaître plus de ressemblance que toute autre comparaison (Rosch & Olivier 1972 : 348). La conclusion est qu'il y a bien un effet proprement cognitif.

Incidemment, on remarque tout de même des effets interprétables dans un sens "whorfien" : les deux configurations les plus proches (selon cette fois une autre mesure, numérique, de proximité) sont celles des réponses "mémorisation" et des réponses "dénomination" chez les Dani. Et la plus faible corrélation entre dénomination et mémorisation chez les Américains, délicate à interpréter, est laissée inexplicite. Quoi qu'il en soit, les résultats que Rosch retient, c'est la similitude des réponses à la tâche de mémorisation, et le fait que ces réponses paraissent refléter les similarités visuelles des couleurs. Sur ce dernier point, Rosch parle d'*isomorphisme* des représentations mnémoniques, reprenant le terme que Shepard et Chipman venaient d'employer dans une étude sur les images mentales (Shepard et Chipman 1970).

Dans une autre vague d'expériences, Rosch (1972a) va marier la dimension comparative et interlinguistique de l'approche de Berlin et Kay au type de procédure expérimentale de Brown et Lenneberg (1954). Elle reprend à Brown et Lenneberg leur concept de codabilité, mais,

point important, elle la définit essentiellement par la longueur du nom donné à une couleur par les sujets, et par le délai de dénomination (le temps mis à trouver le nom). Elle écarte donc l'une des dimensions retenues par Brown et Lenneberg, celle d'accord intersubjectif. Or, l'accord intersubjectif était, pour Brown et Lenneberg, le facteur essentiel. Elle rejette de même le facteur de *communication accuracy*, mis en avant par Lantz et Steffire.²⁰ S'estompent alors les facteurs relevant de l'aspect communicatif du langage, moins assimilables dans une théorie pour laquelle la causalité va du physiologique au linguistique. Pour Rosch, il s'agit de fermer la possibilité que les sujets s'aident du langage pour accomplir la tâche de mémorisation en exploitant la précision et la stabilité des noms correspondant aux couleurs les plus codables. Cette possibilité est fermée par le fait de choisir un large échantillon de langues de familles différentes : si les mêmes couleurs sont toujours les plus facilement codables quelle que soit la langue, ce ne peut être une coïncidence. Il doit donc s'agir d'un biais cognitif, ce dont on peut s'assurer en voyant si les couleurs les plus facilement codables sont aussi les plus faciles à traiter dans une tâche non linguistique.

La première tâche est de mesurer la codabilité pour un échantillon de langues (indo-européennes, sino-tibétaines, austronésiennes, afro-asiatiques). Rosch constate qu'elle est maximale pour les couleurs focales quelle que soit la langue : les couleurs focales sont celles qui ont la dénomination la plus courte et sont nommées le plus rapidement.

La seconde tâche est de chercher si les couleurs focales sont traitées plus facilement dans une tâche présumée être non linguistique. Pour mettre en évidence cette portée cognitive, Rosch conserve le dispositif expérimental de Brown et Lenneberg en recourant à une tâche de reconnaissance au cours de laquelle le sujet doit retrouver dans une mosaïque de 160 pastilles la couleur qui lui a été présentée. Elle a soin de choisir, cependant, un intervalle relativement long (30 sec.) entre la présentation de la couleur cible et celle de la "mosaïque". Brown et Lenneberg avaient en effet montré que cet intervalle était favorable à l'obtention d'une corrélation élevée entre codabilité et reconnaissance, ce qui est précisément le résultat que Rosch souhaite établir.

Les résultats sont conformes aux espérances : les couleurs focales sont plus facilement reconnues. Certes, les couleurs focales sont aussi celles qui ont le plus haut degré de codabilité, et cette corrélation avait justement conduit Brown et Lenneberg à une conclusion relativiste. Cette conclusion, cependant, ne peut tenir, dès lors que les couleurs les mieux codables sont stables de langue à langue. Bref, l'expérience de Rosch reproduit en miniature l'évolution qui a mené de Brown et Lenneberg (1954) à Berlin et Kay (1969).

B.4. Des couleurs aux formes : la notion de "catégorie naturelle"

Les nouvelles expériences que nous allons aborder se distinguent des précédentes sur deux points. D'une part, elles testent directement l'acquisition des catégories de couleur par des sujets présumés "vierges", les Dani. D'autre part, elles élargissent l'investigation à un nouveau champ, celui des formes. C'est ce nouveau champ qui motive apparemment l'adoption du terme de *prototype*, que Rosch emploie initialement pour les formes ou lorsqu'elle parle à la fois des formes et des couleurs. Mais voyons d'abord le cas des couleurs. L'idée de Rosch est d'enseigner aux Dani des catégories, c'est-à-dire de leur apprendre à associer des noms (choisis dans le lexique des noms de clans²¹) à des groupes de couleurs, et

²⁰ Une décision injustifiée, selon Lucy (1996). La précision communicationnelle (*communication accuracy*) était en effet prédictive des taux de reconnaissance pour des échantillons de saturations différentes. Ainsi, note Lucy (1996 : 182), "Heider [= Rosch] ignore la plus puissante mesure de codabilité [= la précision communicationnelle] et ne démontra pas que sa propre explication des résultats antérieurs marcherait aussi pour d'autres échantillons."

²¹ Les Dani ayant rejeté des noms dépourvus de sens, Rosch a dû faire usage des noms de clans.

aussi à des formes. En ce qui concerne les couleurs, les catégories “naturelles” sont construites autour des couleurs focales (par ex., en simplifiant, bleu-vert, bleu, bleu tendant vers le violet), les catégories “non naturelles” autour de couleurs “internominales” (par exemple jaune, jaune-vert, vert) ou comme ayant des couleurs focales seulement à leur périphérie (par exemple, rouge, rouge-violet, violet, où le rouge est focal). Notons aussi que les couleurs d’une même catégorie peuvent varier en brillance, et non pas seulement en teinte, comme dans les exemples qui viennent d’être mentionnés (la brillance est une dimension présumée importante pour les termes de couleur dani).

En ce qui concerne les formes, Rosch choisit comme *prototypes* présumés trois “bonnes formes”, c’est à dire trois figures géométriques régulières (un cercle, un carré et un triangle équilatéral) qui vont servir de schémas générateurs. Elle construit ensuite plusieurs ensembles. Dans un premier ensemble, chacune de ces figures est le point de départ de diverses transformations qui en modifient le contour : telle ligne est interrompue, une ligne droite est transformée en courbe ou un arc de cercle est linéarisé, des côtés sont allongés etc. Les figures issues par transformation des bonnes formes sont ensuite employées à leur tour comme patrons générateurs de nouvelles figures, les bonnes formes étant intégrées à ces nouvelles séries, mais sans avoir cette fois le rôle de patron générateur. Rosch constitue ainsi deux grands types d’ensembles : le premier ensemble, où les bonnes formes servent à engendrer une série de figures ; et les autres ensembles, où ce sont des figures “dégénérées” qui sont les points de départ de leurs séries. Dans ces derniers ensembles, Rosch considère que les bonnes formes ont le statut d’éléments périphériques. Comme pour les couleurs, les sujets devront apprendre à associer un nom de clan à une catégorie.

Qu’il s’agisse des formes ou des couleurs, les résultats obtenus par Rosch corroborent son hypothèse : les catégories “naturelles” sont apprises plus facilement et donnent lieu à moins d’erreurs que les autres catégories, et au sein de chaque catégorie (naturelle ou non), le prototype est le membre dont le nom est appris le plus facilement, et, du moins pour les formes, le membre choisi comme le plus typique.²²

Il est vraisemblable que le terme de *prototype* a été introduit parce que la seconde expérience, celle portant sur les formes, illustre bien un emploi de ce terme en psychologie, celui du prototype considéré en tant que *schéma* générateur d’une série d’items artificiels (comme chez Attneave 1957 ; voir la section suivante). Mais, ce qui est une nouveauté, le terme subsume maintenant la notion de bonne forme (issue de la Gestalttheorie) et celle de couleur focale. Il permet d’élaborer la notion plus générale de catégorie “naturelle” : une catégorie naturelle est une classe dont le prototype est une bonne forme ou une couleur focale.

Simultanément, l’hypothèse d’une origine neurophysiologique de cette structure catégorielle commence à recevoir un certain soutien empirique. Dans l’article en question ici, Rosch peut s’appuyer sur les résultats présentés par McDaniel à Harvard en 1972, où McDaniel montrait que la réponse neuronale à des couleurs atteignait un pic pour certaines couleurs focales (jaune, bleu et vert).²³ A cet égard, il est à noter que le début de convergence entre la neurophysiologie et la psychologie des couleurs vient opportunément : il est contemporain des

²² Les Dani se refusèrent à choisir le membre typique dans le cas des couleurs. Rosch ne fournit guère d’indications sur les réactions des Dani durant les tests. Saunders (1995) cite une étude de Sorenson (1976), menée également en Nouvelle Guinée, et qui montre qu’un interrogatoire expérimental peut être une situation déstabilisante pour des sujets appartenant à une culture où cette situation ne va nullement de soi.

²³ Chez De Valois & Jacobs (1968), qui prolongeaient la théorie proposée au 19^{ème} siècle par Hering, quatre types de cellules neuronales déterminaient la perception des couleurs et étaient centrées, dans certaines conditions d’activation, sur la perception du bleu, du jaune, du rouge et du vert (la brillance était codée par deux autres types de cellules). Cependant, leurs expériences étaient menées avec des lumières monochromatiques, non avec des couleurs de surface (aux longueurs d’onde complexes). Ni les modèles de De Valois et Jacobs ni celui de McDaniel ne rendent compte des autres termes de base (rose, orange, marron, violet). Rosch elle-même reconnaît qu’il y a là un problème (1977 : 18).

premiers doutes, certes modérés, qu'exprime Rosch sur la validité des recherches de Berlin et Kay et qui font suite au compte rendu dévastateur qu'en avait donné Hickerson (1971).²⁴ Cette convergence vient donc faire, en quelque sorte, contrepoids aux objections que l'on pouvait opposer à Berlin et Kay et engage un processus de légitimation réciproque des travaux respectifs de Berlin et Kay, de Rosch et de McDaniel (Saunders 1992, 1995, 2000). C'est à partir de cette conception universaliste et naturalisante des catégories perceptives que Rosch va développer l'outillage théorique qui constitue sa théorie du prototype. Tournons-nous d'abord vers la notion et le terme de prototype.

C.1. Le prototype avant Rosch

Rosch n'est pas la première à employer le terme de prototype pour désigner l'item qui, au sein d'un ensemble, a un statut distinctif. Pour mieux apprécier la singularité de sa position, il est important de comprendre l'infléchissement qu'elle a fait subir à la notion dont elle a hérité. Il faut donc revenir à ses antécédents historiques.

A l'origine, la notion de prototype a d'étroites affinités avec celle de *schéma*. Comme l'articulation entre les deux notions est clairement présentée chez Attneave (1957), c'est avec lui que nous entamerons notre périple historique, commençant pour ainsi dire, à mi-parcours.

Le dessein d'Attneave est de déterminer si l'apprentissage ou la reconnaissance de configurations peuvent être facilités par la familiarisation préalable avec une structure dont ces configurations sont des variations. Cette structure ou forme génératrice de variations est ce qu'Attneave appelle un "schéma". Pour illustrer son propos, il fait la comparaison suivante, qu'il emprunte à Hebb (1949) :

"Hebb souligne l'importance que revêt l'acquisition du schéma d'une classe d'objets qu'il faut différencier les uns des autres : ainsi, les Chinois "se ressemblent tous" aux yeux d'un observateur occidental qui n'a pas vu beaucoup de Chinois. Conçu de cette manière, le schéma consiste, au moins en partie, en une représentation de la tendance centrale ou de la forme commune à la classe d'objets en question. Si l'observateur dispose d'un étalon subjectif du visage humain obtenu à partir du "moyennage" de visages américains [sic], il se peut qu'il apprenne un nouveau visage en fonction de la manière et du degré selon lesquels il dévie de ce schéma. (...) Cependant, si on l'immerge tout à coup dans une population chinoise, son étalon ne sera plus central, et les nouveaux visages en dévieront plus ou moins dans la même direction."

Le schéma auquel Attneave fait appel ici a plusieurs caractéristiques : c'est une forme acquise, construite par extraction d'une tendance centrale, correspondant à une sorte de valeur moyenne des instances auxquelles le sujet a été exposé. Le schéma est donc une structure fluente, qui se modifie avec l'expérience. Il fonctionne comme un étalon de comparaison, qui facilite la discrimination et la reconnaissance de formes qui en sont des variations. Enfin, et inversement, il contrarie le traitement de formes qui en dévient uniformément, ce qui implique que les sujets intègrent une mesure de la variabilité des instances en même temps qu'ils abstraient le schéma sous-jacent.

Pour démontrer la validité de cette notion de schéma, Attneave propose à des sujets une tâche d'association formes – mots. Les sujets seront séparés en deux groupes : les uns seront exposés à des schémas, c'est-à-dire à des structures ou formes dont les items expérimentaux seront des variations, tandis que les autres sujets ne passeront pas par cette étape préalable. Attneave soumet ensuite ses sujets à une tâche consistant à appairer des structures et des formes à des noms (telle structure s'appelle *Sam*, telle autre *Joe*...).

Les schémas sont des tableaux de lettres et des polygones :

²⁴ Rosch (1973a : 331) : "Bien que les données de Berlin et Kay ne constituent pas un soutien sans équivoque de leurs thèses (cf. Hickerson 1971), leurs thèses méritent qu'on s'y arrête et réclament plus de recherche."

	S		K	Y	G
F		Y	P	Z	K
P			O	F	
R		I		I	
	D				W



FIGURE 1 : types de schémas utilisés par Attneave (1957)

Les variations sont obtenues par le changement d'une lettre sur les tableaux, et par déformation pour les polygônes. Peut-être parce que le schéma est générateur d'une série de variantes, Attneave convient de l'appeler aussi *prototype*.

Les résultats de l'expérience corroborent l'hypothèse de départ, ce qui implique qu'un modèle d'apprentissage fondé uniquement sur des associations entre items est invalide. Attneave en conclut que l'apprentissage procède par l'abstraction d'un prototype qui correspond à la *tendance centrale* d'une classe. Ce prototype est en même temps un *schéma*, dans la mesure où les variations sont définies par rapport à lui.

A la suite d'Attneave, un certain nombre d'études reprendront l'idée du prototype générateur d'une série. Certaines expériences, bien connues en psychologie et citées par Rosch, montreront ainsi que l'exposition à la série engendrée suppose l'abstraction, au cours de l'apprentissage d'une série, du prototype qui a servi à la générer (Posner et Keele 1968). Le matériel utilisé sera typiquement constitué de stimuli visuels, souvent "analogues" (par ex. des nuages de points chez Posner et Keele 1968 ; pour une revue, cf. Cauzinille-Marmèche et al. 1990).

C.2. La notion de schéma

D'où Attneave tire-t-il sa notion de schéma ? La réponse n'est guère mystérieuse, elle est même l'une des filiations les mieux reconnues dans le sérail des sciences cognitives : Bartlett l'a introduite pour rendre compte de phénomènes ayant trait à la réélaboration de représentations mnémoniques. Le terme même de *schéma* est emprunté par Bartlett au neurologue Henry Head, chez qui il renvoyait au schéma corporel. Voyons maintenant comment ce passage de relais s'est effectué.

Dans un article célèbre écrit en collaboration avec Gordon Holmes (1911), Head distinguait deux types de schéma, le *schéma postural* et le *schéma superficiel*. Il décrivait le schéma postural comme un étalon "produit par les postures et les mouvements antérieurs (...) par rapport auquel tous les changements de posture suivants sont évalués avant qu'ils ne deviennent conscients. (...) Des modifications continues nous permettent de construire un modèle postural de nous-mêmes qui change constamment. Chaque nouvelle posture ou mouvement est enregistré sur ce schéma plastique, et l'activité du cortex met chaque nouvel ensemble de sensations évoqué par la posture modifiée en rapport avec ce schéma" (Head & Holmes 1911 : 186). Quant au schéma superficiel, il correspondait à un modèle de la surface du corps permettant de localiser une stimulation sur cette surface.

Ce que Bartlett retient du schéma de Head est l'idée qu'une structure passée agit sur la reconnaissance du présent, l'assimilant en quelque sorte, tout en se déformant constamment sous son influence. Il préfère d'ailleurs au terme de *schéma* ceux d'*organized setting*, qui, selon lui, évoquent davantage la dynamisme inhérente à la notion qu'il cherche à circonscrire.²⁵

²⁵ "Schéma renvoie à une organisation active des réactions, ou des expériences passées, organisation qu'on doit toujours supposer à l'œuvre dans toute réponse adaptée de l'organisme. Ainsi, lorsqu'un comportement est

Une des préoccupations majeures de Bartlett est de rendre compte des déformations observées dans la reproduction des stimuli. Il observe par exemple que la forme de gauche ci-dessous est parfois reproduite comme celle de droite, qui évoque le cadre d'un tableau.



FIG. 2 : un stimulus (à g.) et sa reproduction (à dr.) in Bartlett (1932 : 20)

Ces déformations témoignent à ses yeux d'une reconstruction active, par laquelle un stimulus est assimilé à une forme familière qui, lui préexistant, est aussi appelée *preformed scheme*. Il faut noter aussi qu'assimilation et déformation sont des processus généraux affectant tout matériau mémoriel, et ne sont pas donc pas restreintes aux modalités perceptives. Bartlett consacre ainsi une étude approfondie à la reproduction par des sujets anglais d'un conte amérindien, montrant comment des sujets "rectifient" un matériau non familier conformément à leur héritage propre, culturel et personnel.

Le soupçon pourrait s'immiscer que Bartlett a une vision déterministe de ces processus assimilateurs. En réalité, Bartlett consacre beaucoup d'efforts à contrecarrer une telle vision. Malheureusement sa théorie est sur ce point passablement vague, voire obscure. Bartlett évoque le jeu des instincts, intérêts, idéaux et, d'une façon plus cryptique, d'*attitudes*, lesquelles semblent renvoyer à des modes d'orientation par rapport à des tâches, et sont accompagnées d'un composant affectif. La notion, qui vient vraisemblablement de l'école de Würzburg, emprunte peut-être au concept d'*Einstellung* l'idée d'une direction imprimée à la réponse du sujet par son appréhension de la tâche, appréhension confusément sentie mais non explicite quant aux contenus pensés (Larsen & Bertsen 2000).

Il est une autre raison pour laquelle le schéma ne dit pas tout du processus de reproduction. La reproduction, observe-t-il contient souvent un détail frappant qui n'a pas de lien narratif, inférentiel, conventionnel... avec le reste de la structure reproduite et se situe ainsi hors schéma. Reflétant les intérêts personnels du sujet, ce détail manifeste aussi la faculté qu'a ce sujet de manipuler un schéma en en détachant un élément hors séquence.

Comme souvent, parmi les idées de Bartlett, ce seront les plus simples et les plus faciles à opérationnaliser qui se transmettront. Ainsi, Woodworth (1938), parlera de la mémorisation de formes comme d'un processus de construction d'un schéma avec *correction* associée, pour séparer ce qui est préformé (le schéma, par exemple une bonne forme) et l'élément déviant (par exemple une encoche dans cette forme). Opérationnalisé sous la forme d'une tendance centrale, et généralement avec un matériel visuel, ce schéma deviendra une structure en fonction de laquelle les déviations sont reconnues, et que les déviations modifient à leur tour constamment. Ce sont ces conceptions du schéma comme tendance centrale et de la déviation comme correction apportée à un schéma dont Hebb et Attneave se saisiront. Entre temps, la préoccupation initiale de Bartlett pour les processus de reconstruction et sa tentative de concilier déterminisme et élaboration active seront passées au second plan.

ordonné et régulier, une réponse n'est rendue possible que par sa relation à des réponses similaires qui ont été organisées de manière sérielle, même si ces réponses opèrent en tant que masse unitaire et non l'une après l'autre" (1932 : 201). Il est possible, et même probable, que l'image des représentations agissant en masse soit une réminiscence de la masse aperceptive de Herbart.

C.3. L'extension de la théorie aux "catégories sémantiques"

L'étape suivante de l'évolution de Rosch, esquissée dès 1971 (Rosch 1971c) est cruciale : elle se propose désormais d'étendre son approche au-delà des noms de couleurs, un projet qui est facilité par sa vision du lexique comme nomenclature. Pour accomplir cette généralisation, Rosch a besoin d'une notion de prototype susceptible d'englober des classes "naturelles", comme les couleurs et les formes, et des classes qu'on pourrait penser comme culturellement déterminées.

La notion qui va prévaloir repose sur une double opposition : celle entre concept artificiel et catégorie "réelle" ; celle entre attributs discrets et continus. L'appartenance à un concept artificiel de la psychologie, quand il est défini par des attributs discrets (par ex. 'rouge et carré'), est affaire de tout ou rien : soit un item en est une instance, soit il n'en est pas une instance. Dit autrement, selon ce qu'il est convenu d'appeler la *théorie classique*, les attributs sont des conditions nécessaires et suffisantes de l'appartenance à la catégorie.²⁶ Or, remarque Rosch, les catégories "naturelles" étudiées jusqu'à maintenant ont montré la fausseté d'une telle conception. Quant au concept artificiel constitué d'attributs continus, s'il mène bien à une notion de prototype comme *tendance centrale*, il n'a pas de typicalité "naturelle", dans la mesure où la tendance centrale dépend des exemplaires construits par l'expérimentateur (Rosch 1973b : 112-3).²⁷ Rosch oppose ainsi des catégories dites "réelles" et des catégories artificielles, et caractérise cette opposition par deux propriétés, la typicalité et la naturalité. Les catégories "réelles" sont donc des classes de référents plus ou moins typiques de leur catégorie, et, au moins dans le cas des couleurs et des formes, cette typicalité doit avoir un fondement naturel, c'est-à-dire cognitif (en un sens qui reste à préciser). Rosch ajoute que les catégories "réelles" sont identifiables par le fait qu'elles reçoivent un *nom* dans la langue. Enfin, pour ajouter à la confusion, Rosch identifie purement et simplement le *sens noyau* (*core meaning*) d'une catégorie (c'est-à-dire du *nom* de la catégorie) aux *référents* les plus typiques (1973b : 112).

Dans l'immédiat, c'est le premier aspect, l'existence de degrés de typicalité, que Rosch va s'attacher à démontrer. La naturalité de l'organisation des catégories autres que les couleurs et les formes reste pour le moment une question en suspens.

Pour procéder à son expérience, Rosch dispose d'un outil mis au point par Battig et Montague (1969) : des listes de lexèmes (essentiellement des noms) fournies par des sujets (des étudiants) en réponse aux noms de 56 catégories, et donnant pour les lexèmes produits leur fréquence et leur rang de production dans les réponses. De ces listes Rosch extrait huit catégories (les fruits, les sciences, les sports, les oiseaux, les véhicules, les maladies, les légumes, les crimes), qu'elle appelle *catégories sémantiques*, et choisit pour chaque catégorie six instances de fréquences différentes. Par exemple, la catégorie des fruits comporte la pomme, la prune, l'ananas, la fraise, la figue et l'olive.

Le protocole de Rosch consiste à obtenir de sujets (des étudiants en psychologie) une évaluation de la typicalité de chaque instance, sur une échelle de 1 à 7. Les instructions données sont particulièrement détaillées et constituent en réalité une explication de la notion roschienne de catégorie et du but de l'expérience.²⁸ Le fait que les sujets aient été des

²⁶ Dans son article sur les points de référence cognitifs, Rosch parle explicitement de degrés d'appartenance à une catégorie (1975a : 544), peut-être sous l'influence de la notion d'ensemble flou, prôné par Lakoff. Elle établit ainsi un lien entre typicalité et degré d'appartenance. Dans l'évolution de la théorie, ce lien sera de moins en moins perçu comme systématique (pour une discussion, voir Lakoff 1987).

²⁷ La tendance centrale désigne l'ensemble des mesures qui donnent le centre d'un ensemble de données (par ex. la moyenne arithmétique, la médiane ou le mode).

²⁸ Voici ces instructions (1973b : 131-2) : "This study has to do with what we have in mind when we use words which refer to categories. Let's take the word "red" as an example. Close your eyes and imagine a true red. Now imagine an orangish red... imagine a purple red. Although you might still name the orange-red or the purple-red with the term "red", they are not as good examples of red (as clear cases of what red refers to) as the clear "true"

étudiants en psychologie, dont on peut penser qu'ils avaient déjà une notion de ce que devait être une catégorie dans le protocole, ne pouvait guère qu'accentuer cette *mise en phase* de l'expérimentateur et des sujets.

Pour une instance donnée, les résultats font apparaître une corrélation élevée entre la typicalité et la fréquence obtenue par Battig et Montague (1969).²⁹ Par exemple, la pomme est jugée être un fruit plus typique que l'olive, ou l'aigle un oiseau plus typique que l'autruche. Cette corrélation entre typicalité et fréquence signifie-t-elle que la typicalité se réduit à la fréquence ? Rosch concède, cependant, que la fréquence peut jouer un rôle en élevant la saillance. Mais, objectera-t-on, est-ce la fréquence qui conditionne la saillance des couleurs focales et des bonnes formes ?

Pour le moment, donc, nous n'avons pas d'explication de la typicalité qui vaille pour les catégories "naturelles" et les autres. Seul est clairement affirmé le refus d'identifier typicalité et fréquence. Refuser cette identification, c'est mettre sa foi dans le fait que la typicalité fait sens pour les sujets, que les instructions, par leur précision, ont déterminé les sujets à fonctionner en mode "catégorie roschienne", et non en mode "fréquence", et donc que l'argument de la mise en phase, loin d'être une objection à l'expérience, la conforte plutôt. Tel semble donc être le but premier de l'expérience que nous avons examinée ici : simplement montrer que la typicalité fait sens pour les sujets.

Enfin, cette mise en phase des sujets et de l'expérimentateur est encouragée par le choix des instances et du nom de la catégorie. La composition des listes de mots fournit à l'item à classer un contexte qui est loin d'être neutre, bien que Rosch semble considérer ses taxonomies comme des classifications "naturelles".³⁰ Le soupçon que la constitution des listes reflète l'hypothèse à tester et oriente les résultats est partagé par Rastier, qui dit à ce propos :

"Une liste de mots ne constitue pas un paradigme, mais un syntagme parataxique, et ses membres sont naturellement en interaction contextuelle. Les traits sémantiques sélectionnés dans chacun des items dépendent de la composition de la liste soumise aux sujets d'expérience ou produite par eux. Loin d'être indépendants des frontières de la catégorie, les exemplaires prototypiques en dépendent étroitement : par exemple, dans une classe composée de noms de meubles, 'téléphone' sera évidemment jugé fort peu typique; mais il en irait tout autrement dans une classe où il voisinerait avec 'courrier', 'télex', 'télécopie'. Le degré de typicalité, et les traits saillants qui le déterminent, sont donc en fait *prédéterminés* par l'expérimentateur. (...) Bref, une bonne part des différences de degré de typicalité est due tout

red. In short, some reds are redder than others. The same is true for other kinds of categories. Think of dogs. You all have some notion of what a "real dog", a "doggy dog" is. To me a retriever or a German shepherd is a very doggy dog while a Pekinese is a less doggy dog. Notice that this kind of judgment has nothing to do with how well you like the thing ; you can like a purple red better than a true red but still recognize that the color you like is not a true red. You may prefer to own a Pekinese without thinking that it is the breed that best represents what people mean by dogginess. On this form you are asked to judge how good an example of a category various instances of the category are. The first category is "fruit". On the left side of the page are six different kinds of fruit ; the first is "apple". To the right of apple are seven blanks ; the blank closest to apple is to be checked if an apple is a good example of your idea or image of a fruit. The blank to the extreme right is to be checked if apple fits very poorly with your idea or image of a fruit. The other blanks represent the range in between a very good and very poor fit (...) Mark one and only blank for "apple" etc. Don't worry about why you feel something is or isn't a good example of the category (and don't worry whether it's just you or people in general who feel that way) — just mark it the way you see it."

²⁹ Voici un échantillon des catégories de Battig et Montague : unités de temps et d'espace, éléments chimiques, combustibles, métaux, fleurs, insectes, poissons, serpents, oiseaux, quadrupèdes, parents, parties du corps, meubles, outils de menuiserie, titres militaires, habitations, pays, états américains, villes, universités, assaisonnements, titres militaires, danses, musiques, crimes, parties du discours etc.

³⁰ Du moins pour le niveau de base, qui serait un niveau correspondant au "contexte normal" (Rosch 1978 : 43). Nous reviendrons sur la notion de "niveau de base".

simplement à la disparate des catégories soumises à l'expérience, ou recueillies par des expériences ne reflétant rien des conditions attestées de communication" (Rastier 1991 : 193).

C.4. Le prototype comme gestalt et point de référence

Initialement, la notion de prototype, même étendue aux catégories "sémantiques", est pensée en termes de gradience et de saillance, et les catégories sont décrites de façon holistique : les attributs "naturels" varient sur des dimensions continues, une catégorie contient des membres plus ou moins saillants, de degrés de typicalité inégaux, ou, dit autrement, la catégorie envisagée globalement a une "structure interne" et des contours flous. Les membres des catégories "sémantiques" ne sont pas encore des complexes d'attributs. Cette appréhension de la catégorie est, dirait-on, gestaltiste : une catégorie ou une zone catégorielle est comme un fond aux bords flous sur lequel se détache un prototype.

Cette image gestaltiste de la catégorie est sous-jacente à la comparaison qu'établit Rosch entre le prototype et les notions de *point de référence* (*reference point*) et *point d'ancrage* (*anchoring point* ; Rosch 1975a). Les notions de point de référence et de point d'ancrage semblent avoir été suggérées à Rosch par des passages d'un article de Wertheimer (1938) où celui-ci, soucieux de replacer quantité ou numérosité dans leurs contextes "naturels" et non scientifiques, pointait l'asymétrie de certains jugements de comparaison dans de tels contextes : 101 est proche de 100, ou le "suit", une douzaine peut valoir comme quantité repère etc. De même, une droite orientée à 85° est presque verticale, mais on n'évalue pas la verticalité d'une droite par rapport à la diagonale orientée à 85°. ³¹

Pour opérationnaliser ces notions de point de référence et point d'ancrage, Rosch reprend comme champ d'investigation les exemples de Wertheimer (lignes et nombres) et leur ajoute sa catégorie de prédilection, les couleurs. L'idée est de montrer que les couleurs focales, les lignes à certaines orientations (horizontales et verticales notamment) et les nombres distinctifs (comme les multiples de dix) sont des point de référence, et qu'ils peuvent être identifiés comme tels parce qu'ils donnent lieu au type d'asymétrie noté par Wertheimer. Mais comment mettre en évidence cette asymétrie ? Par quels tests ?

La solution de Rosch repose sur deux idées, qui vont donner lieu à deux tests : si le prototype a tendance à servir d'étalon de comparaison, dans un jugement linguistique du type 'x est quasiment y', y, qui représente l'étalon de comparaison, devra plutôt correspondre au prototype et x devra plutôt correspondre à l'entité à comparer. Il s'agit là d'un premier test, de nature linguistique. De plus, Rosch fait l'hypothèse que lorsqu'il sert d'étalon de comparaison, le prototype joue le rôle d'un point d'attraction : une entité proche du prototype sera jugée plus proche de lui que le prototype ne sera jugé proche de cette entité. On peut alors envisager une tâche où le sujet devra placer l'un par rapport à l'autre le prototype et une instance non prototypique. Si le sujet place l'instance non prototypique près du prototype quand il sert de point de référence fixe, mais les éloigne l'un de l'autre dans le cas inverse (instance fixe, prototype mobile), on conclura que le prototype a bien la fonction de point de référence. On voit que le premier test exploite un contexte linguistique ('x est presque y') et que le second fait appel à une représentation spatiale de la similarité.

C'est Lakoff qui a inspiré à Rosch son premier test, celui de nature linguistique. Lakoff venait en effet de publier une étude sur les degrés d'appartenance d'un item à une catégorie, qu'il proposait d'aborder au moyen du formalisme des ensembles flous de Zadeh (Lakoff 1973). Il s'appuyait lui-même sur une présentation où Rosch avait exposé des résultats préliminaires

³¹ L'exemple de la droite verticale n'est dans la référence citée par Rosch (Wertheimer 1938). Il pourrait provenir de Wertheimer (1912).

sur les degrés de typicalité de membres de catégories “sémantiques”.³² Dans l’étude de Lakoff, les contextes du type ‘x est quasiment y’, ou ‘x est un vrai y’ étaient analysés comme autant de manifestations linguistiques de l’existence de degrés d’appartenance catégorielle. Lakoff appelait *enclosures (hedges)*³³ les expressions modulant l’appartenance catégorielle, c’est-à-dire permettant d’introduire de la gradience (du “flou”) ou, au contraire, de la diminuer. Il n’y avait d’ailleurs nulle contradiction dans le fait de dire qu’une enclosure *manifestait* le fait que la catégorie était floue, et permettait d’introduire du flou ou d’en retirer. Dans l’optique de Lakoff, une catégorie était floue de façon inhérente, et la fonction d’une enclosure était de *situer* l’appartenance d’un item à cette catégorie sur une échelle. Par exemple, Lakoff considérait que l’enclosure *sort of* dans *he is sort of tall* (‘il est plutôt grand’) indiquait que les critères d’application de *tall* étaient “relâchés”, c’est-à-dire s’appliquaient à des tailles inférieures à celles requises normalement par *tall*. Pour les noms, Lakoff opposait par exemple l’enclosure *strictly speaking* (*strictly speaking, a whale is a mammal*), appropriée pour des instances ayant un haut degré d’appartenance à une catégorie, à l’enclosure *regular*, applicable à des instances ayant un degré d’appartenance nulle à une catégorie, comme par exemple dans *Harry is a regular fish* (*Harry n’appartient pas à la catégorie poisson*). D’un autre côté, Lakoff assignait à ces enclosures la capacité d’isoler, de “catalyser” des *types* d’attributs différents (définitoires, propres à la catégorie pour *strictly speaking*, ou bien “connotatifs”, métaphoriques, non critères, pour *regular*).³⁴

Etant donné la perspective de Rosch dans son expérience, les enclosures sont envisagées non comme des catalyseurs de différents types d’attributs mais comme des révélateurs de la structure interne d’une catégorie. Mais, alors que la structure interne est analysée en termes de degrés d’appartenance chez Lakoff, Rosch l’envisage comme un paysage où l’on se repère à partir des points saillants que sont les prototypes. Il s’agit justement de savoir si les prototypes remplissent bien cette fonction de repère.

Le dispositif est simple : on présente au sujet deux couleurs, dont l’une est focale, et on observe si le sujet met la couleur focale en position de x ou de y dans des contextes du type ‘x est presque y’ (‘x is almost y’).³⁵ Il importe de noter que ce sont les *stimuli* eux-mêmes que le sujet doit insérer dans les contextes linguistiques, et non des expressions référant à ces stimuli, en plaçant *physiquement* ces stimuli sur des cartes portant les contextes. Autrement dit, le dispositif suppose soit que les sujets peuvent étirer le sens des enclosures jusqu’à leur attribuer la fonction de marquer les points saillants de la représentation mentale des couleurs (elle-même isomorphe à leur perception) ; soit que ces enclosures, indépendamment du dispositif, fonctionnent de toute façon comme des indicateurs de saillance relative. La première possibilité n’est pas envisagée par Rosch parce qu’elle ne se soucie pas d’analyser la signification des enclosures ou les distorsions que son dispositif pourrait faire subir à cette signification. Et son point de vue représentationnaliste est une fois de plus confirmé : si les

³² Voir Rosch (1971c). Rosch avait présenté les jugements de typicalité recueillis auprès de sujets à qui avaient été soumises des listes d’items relevant des catégories “sémantiques” suivantes : jouets, oiseaux, fruits, maladies, métaux, crimes, sports, véhicules, parties du corps. Rappelons qu’à cette époque, Lakoff, en partie sous l’influence de Ross, s’intéressait aux phénomènes de gradience non seulement dans le lexique proprement dit mais dans la grammaire. Sa théorie de la grammaire floue (*fuzzy grammar*) faisait dépendre l’acceptabilité plus ou moins grande d’une construction d’une hiérarchie de contraintes syntaxiques et lexicales (Lakoff 1973b). L’étude de Lakoff est révélatrice d’une perspective encore logiciste, que signale par ailleurs le recours à Zadeh : le degré de typicalité d’un item est interprété en termes de degré de *vérité* du jugement d’appartenance de cet item à une catégorie.

³³ Le terme d’*enclosure* a été proposé à titre de traduction française de *hedge* par Kleiber et Riegel (1978) (merci à l’un de mes relecteurs de m’avoir signalé cette référence).

³⁴ Ma présentation simplifie l’analyse de Lakoff. En réalité, Lakoff distingue quatre types d’attributs et non deux.

³⁵ Les contextes en question sont : “__ is essentially __” ; “__ is basically __” ; “__ is roughly __” ; “__ is almost __” ; “__ is sort of __” ; “Loosely speaking __ is __.”

noms de couleur tiennent lieu des percepts de couleur, pourquoi des représentations physiques de couleur ne pourraient-elles remplacer les noms qui en tiennent lieu ?

Dans la majorité des cas, les résultats confirment les hypothèses de départ. Toutefois, cette confirmation n'est pas totale. Il arrive en effet que le stimulus servant de point de référence soit la couleur de plus grande saturation, et c'est le plus grand nombre, non le prototype, qui figure dans la position de y dans 'x is almost y'. Ceci suggère que certaines enclosures portent sur certains types d'attributs, qu'elles servent en quelque sorte de catalyseurs. *Almost* paraît ainsi cibler la dimension quantitative ou intensive des instances. Or, pour constituer sa liste d'enclosures, Rosch avait éliminé celles pour lesquelles le dispositif, au yeux de sujets préalablement testés, ne faisait pas sens. Il n'est donc pas interdit de penser que ces enclosures constituent des contextes spécifiquement adaptés aux tâches choisies et masquent, mais imparfaitement, la capacité qu'ont les enclosures de cibler certains types d'attributs.

En outre, les résultats de la tâche linguistique et de la tâche spatiale ne concordent pas : dans la tâche spatiale, les couleurs prototypes de teinte différente ne fonctionnent pas comme points de référence à saturation et brillance égales. Quelle que soit la raison qu'on puisse invoquer, il semble que la proximité au prototype et le choix même du prototype soient fonction de la tâche.

En résumé, sa conception gestaltiste de la catégorie conduit Rosch à négliger les différents types d'attributs que Lakoff avait distingués au moyen d'enclosures, pour ne garder qu'une conception holistique de la catégorie, selon laquelle les instances sont reconnues par rapport à des prototypes. Cette conception est conditionnée à la fois par le champ d'investigation privilégié par Rosch (des catégories à valeurs continues) et par ses préjugés représentationnalistes, qui l'incitent à subordonner l'analyse d'une catégorie à son organisation perceptive et cognitive. Pourtant, les résultats de l'expérience auraient dû suggérer que les enclosures faisaient plus qu'indiquer la proximité d'une instance au prototype, et pouvaient apparaître comme des catalyseurs de certains types d'attributs (comme par exemple la grandeur d'un nombre pour *almost*). Ces résultats auraient pu suggérer aussi que les effets de prototypie étaient modulés par la nature des tâches (linguistique ou spatiale). Ce n'est pas, toutefois, l'interprétation retenue par Rosch.

C.5. La représentation mentale des catégories

Une autre manière d'envisager le prototype comme point de référence est de se placer non au niveau d'un autre exemplaire qui lui serait comparé, mais au niveau de la catégorie qui les subsume tous les deux. Sous cet angle, la catégorie est proche du prototype, ou pour le dire comme Rosch, l'information contenue dans la catégorie a beaucoup en commun avec celle représentée dans le prototype. Cette façon d'envisager la question met en rapport deux entités théoriques dont l'une, le prototype, est censée mieux connue que l'autre, la catégorie, dans la mesure où la représentation cognitive de la catégorie, son "code mental" comme dit aussi Rosch, n'est pas claire.

Comme toujours, la validation expérimentale porte d'abord sur le domaine des couleurs (Rosch 1975d). La méthode choisie est la technique d'amorçage : on fournit au sujet le nom de la catégorie (par ex. "rouge") en guise d'amorce (ou aucune amorce dans la condition contrôle) et on lui demande ensuite de juger si deux stimuli (par ex. deux rouges identiques, deux rouges différents ou un rouge et un bleu, tous plus ou moins "typiques") sont identiques, ou appartiennent à la même catégorie. Si la représentation de la catégorie est proche du prototype, Rosch raisonne que l'amorce facilitera le jugement du sujet quand les stimuli sont identiques et typiques. Cette attente est confirmée et le protocole est purement et simplement étendu aux catégories "sémantiques" du type *oiseau* (Rosch 1975b), avec toutefois une addition notable : les exemplaires sont présentés soit sous forme verbale (par ex. "hirondelle")

soit sous forme visuelle (un dessin d'hirondelle). L'objet de cette addition est d'éclairer la nature du format de représentation des catégories, et de répondre aussi aux préoccupations de l'époque, qui voit l'explosion des recherches sur l'image mentale et leur rapport avec le verbal (Fortis 1994). Sans que nous puissions ici détailler les arcanes du raisonnement, on peut néanmoins dire que la conclusion de Rosch oscille entre deux interprétations des résultats : (1) la représentation cognitive d'une catégorie comme oiseau est "amodale", c'est-à-dire ni imagée ni verbale, mais est plus proche d'une représentation imagée, et le prototype en tient lieu (Rosch 1975b) ; et s'agissant du "niveau de base", comme nous le verrons bientôt, le prototype est décrit comme une image moyenne des membres de la catégorie ; ou bien (2) la seule conclusion nette est que la catégorisation est fonction de la distance au prototype, tandis que la nature du format mental et de la représentation du prototype demeurent obscures. Selon cette seconde interprétation, on peut seulement affirmer que la catégorisation est un traitement modulé par des degrés de typicalité. C'est vers cette conclusion a minima que Rosch s'oriente par la suite (Rosch 1977 et 1978).³⁶

D'autres dispositifs permettent d'évaluer la distance d'un exemplaire au prototype, ou, pour s'en tenir à l'interprétation (2) son degré de typicalité. On observera par exemple que le temps de vérification d'un énoncé comme *une pomme est un fruit* est plus court pour les instances prototypiques (comme *pomme*) et plus long pour les instances moins typiques (comme *pastèque* ; Rosch 1973b).

Nous sommes ici face à une évolution : nous avons vu que la notion de prototype découlait directement de son application à un champ d'investigation, les couleurs, où le prototype était la représentation d'un stimulus. Ainsi, quand Rosch s'occupait des couleurs, des formes, des droites du plan ou des nombres, le prototype était un exemplaire de sa catégorie, c'est-à-dire une couleur focale, une bonne forme, une droite verticale ou horizontale, ou encore un nombre rond. La notion de schéma comme tendance centrale, si elle correspondait à une autre notion de prototype, demeurerait dans le cadre d'une conception réifiante. Mais cette conception s'étend difficilement aux catégories "sémantiques", dans la mesure où ces catégories ne varient pas sur quelques dimensions bien déterminées. Finalement, la notion de prototype devient hétérogène.³⁷

D. L'articulation des catégories en attributs : ressemblance de famille et catégories de base

Nous arrivons maintenant à une étape essentielle du développement de la théorie, marquée par cinq éléments nouveaux : (1) la reconnaissance de l'hétérogénéité des catégories, que Rosch va d'ailleurs s'attacher à réduire ; (2) l'introduction du concept de ressemblance de famille pour rendre compte de la structure interne des catégories ; (3) l'entrée massive des attributs,

³⁶ "Le fait même de parler d'un prototype n'est qu'une fiction grammaticale ; ce à quoi on réfère, ce sont des jugements de prototypicalité. Il n'y a que dans les catégories artificielles qu'existe par définition un prototype" (Rosch 1978 : 40). Voir aussi : "Les prototypes ne constituent pas une théorie de la représentation. (...) Quoique nous ayons suggéré ailleurs qu'il serait raisonnable, étant donné les principes de base de la catégorisation, que les catégories soient représentées par des prototypes qui sont le plus possible représentatifs des instances de la catégorie et le moins possible représentatifs d'items en dehors de la catégorie (Rosch et Mervis 1975, Rosch 1977), cette affirmation reste une formule non précisée tant qu'elle ne fait pas partie d'une théorie spécifique de la représentation. (...) Les prototypes peuvent être représentés par des systèmes à format propositionnel ou imagé" (Rosch 1978 : 40). Rosch ne sera pas toujours fidèle à ses propres avertissements et continuera par la suite à donner des définitions réifiantes du prototype (cf. Gabora, Rosch & Aerts 2008).

³⁷ Cette diversité va concerner non seulement le prototype roschien mais les diverses théories psychologiques qui vont proposer des modèles de la catégorisation compatibles avec les phénomènes de typicalité, y compris pour un même type de catégories. Il y a donc en psychologie plusieurs types de prototypes et de catégories, et pour un type de catégorie plusieurs modèles possibles. Voir les revues de Cordier et Dubois (1981) et Smith et Medin (1981), et la conclusion de cet article.

qui viennent “discrétiser” et structurer la représentation des membres d’une catégorie ; (4) l’importation des notions d’indice et d’information dans la théorie, issues pour l’essentiel du fonctionnalisme probabiliste d’Egon Brunswik, et de la théorie informationnelle de Garner ; (5) la distinction de plusieurs niveaux de catégorisation. Je vais maintenant discuter ces points à la suite.

D.1. L’hétérogénéité des catégories

La conception gestaltiste de la structure interne des catégories tend à niveler les différences entre les différentes sortes de catégories que Rosch a étudiées jusqu’ici : les couleurs, les formes, les catégories “sémantiques”, les lignes et les nombres. Du point de vue de Rosch, ce nivellement pouvait se justifier par la volonté d’opposer sa vision des catégories et de leur structure interne à la conception “classique”, où l’appartenance à une catégorie était affaire de tout ou rien.

Une fois acquise l’existence d’une “structure interne”, il fallait bien reconnaître que les catégories étudiées étaient hétérogènes. Certaines étaient naturelles ou “réelles”, d’autres artificielles. Les prototypes des catégories naturelles étaient censés avoir une origine neurophysiologique. Ceux des catégories artificielles étaient des schémas générateurs conçus pour construire une série de variantes. D’autres catégories, comme les catégories “sémantiques” et les nombres semblaient plutôt être des créations culturelles.

Pour clarifier les choses, Rosch propose de distinguer trois espèces de prototypes : ceux qui sont saillants avant la formation de la catégorie correspondante, et qui ont “probablement” une base physiologique (couleurs, formes, expressions des émotions³⁸) ; ceux qui correspondent à une “tendance centrale” (cf. les expériences avec des catégories artificielles type nuages de points) et qui sont construits préalablement à leur catégorie ; enfin, ceux qui ne précèdent pas la formation de la catégorie et “doivent être formés à partir des items de la catégorie et selon des principes d’apprentissage et de traitement de l’information” (Rosch & Mervis 1975 : 574 ; nous verrons plus loin ce qu’il faut entendre en l’occurrence par *traitement de l’information*).

Ces distinctions mettent en relief un problème qui est demeuré en suspens : le principe de formation des prototypes de catégories “réelles” mais non “naturelles”, c’est-à-dire en particulier des prototypes de catégories dites “sémantiques” peut difficilement être d’ordre neurophysiologique. C’est précisément ce que Rosch et Mervis se proposent d’éclaircir dans cette étude, et nous verrons plus loin de quelle manière elles ont pensé résoudre ce problème.

D.2. La ressemblance de famille

Lorsque Rosch aborde les catégories “sémantiques”, un problème auquel elle ne s’est pas encore confrontée se pose. A un niveau relativement abstrait, le principe qui confère à une catégorie son unité n’est pas toujours clair : il ne suffit pas, dans l’optique de Rosch, de dire que certaines catégories, comme celle des fruits, ont une définition rigoureuse. Son problème est de savoir ce qui nous détermine à juger, dans des conditions naturelles et profanes où la tomate n’est pas (ou est peu) un fruit, que telle entité est un fruit. Encore ce problème serait-il résolu pour les fruits qu’il se poserait avec plus d’acuité pour la catégorie des sports. Qu’est-

³⁸ Rosch et Mervis font allusion ici aux travaux d’Ekman (1971), qui mettaient en évidence une classe d’expressions faciales d’émotions universellement reconnues (joie, tristesse, dégoût, surprise, crainte, et, mais avec des restrictions, colère). Rosch avait fait la connaissance d’Ekman lors d’un passage par Berkeley sur le chemin de la Nouvelle Guinée (K. Heider 2007). Ekman avait persuadé Rosch et son mari Karl Heider de soumettre ses tests sur les expressions faciales d’émotions aux Dani, malgré le fait que, nous apprend Ekman, son hypothèse universaliste ait suscité quelque scepticisme chez les deux chercheurs (Ekman 1971 : 276). Les résultats conforteront l’hypothèse universaliste, et seront interprétés par Rosch comme concordant avec ses propres travaux sur les couleurs et les formes (par ex. dans Rosch 1975c).

ce qui nous permet de catégoriser une activité comme sport ? Qu'ont en commun les activités que nous appelons *sport* ?

C'est chez Wittgenstein que Rosch trouve une voie d'approche qui permet de dépasser cette dernière question et de lui en substituer une autre. S'inspirant de Wittgenstein (qu'elle a étudiée lors de son cursus en philosophie), Rosch considère en effet que la catégorisation n'exige pas que toutes les instances d'une catégorie "sémantique" partagent un attribut commun. Il suffit qu'elles soient liées entre elles par une *ressemblance de famille*.³⁹

Voyons maintenant comment Wittgenstein introduit la *ressemblance de famille* dans le passage classique que Rosch a sans doute à l'esprit :⁴⁰

“§ 66 — Considérons par exemple les processus que nous nommons les « jeux ». J'entends les jeux de dames et d'échecs, de cartes, de balle, les compétitions sportives. Qu'est-ce qui leur est commun à tous ? — Ne dites pas : Il faut que quelque chose leur soit commun, autrement ils ne se nommeraient pas « jeux » — mais voyez d'abord si quelque chose leur est commun. — Car si vous le considérez, vous ne verrez sans doute pas ce qui leur serait commun à tous, mais vous verrez des analogies, des affinités, et vous en verrez toute une série. Comme je l'ai dit, ne pensez pas, mais voyez ! Voyez, par exemple, les jeux sur damier avec leurs multiples affinités. Puis passez aux jeux de cartes : ici, vous trouverez beaucoup de correspondances avec la classe précédente, beaucoup de traits communs disparaissent, tandis que d'autres apparaissent. Si dès lors nous passons aux jeux de balle, il reste encore quelque chose de commun, mais beaucoup se perd. — Tous ces jeux sont-ils « divertissants » ? Comparez les échecs et la marelle. Ou bien y a-t-il en tous une façon de gagner et de perdre, ou une compétition des joueurs ? Songez aux patiences. Dans les jeux de balle, on gagne et on perd ; mais quand un enfant lance la balle contre le mur et la rattrape, ce caractère se perd. Voyez quel rôle jouent l'adresse et la chance. Et combien différentes l'adresse aux échecs et l'adresse au tennis. Songez maintenant aux jeux de rondes : ici il y a l'élément du divertissement, mais combien d'autres caractéristiques ont disparu ! Et ainsi nous pouvons parcourir beaucoup d'autres groupes de jeux ; voir surgir et disparaître des analogies [Ähnlichkeiten].

Et tel sera le résultat de cette considération : nous voyons un réseau complexe d'analogies qui s'entrecroisent et s'enveloppent les unes les autres. Analogies d'ensemble comme de détail.

§ 67 — Je ne puis caractériser mieux ces analogies que par le mot « ressemblance de famille » [Familienähnlichkeiten] ; car c'est de la sorte que s'entrecroisent et s'enveloppent les unes les autres les différentes ressemblances qui existent entre les différents membres d'une famille ; la taille, les traits du visage, la couleur des yeux, la démarche, le tempérament etc. — Et je disais : les « jeux » constituent une famille.”⁴¹

³⁹ L'idée d'une chaîne sémantique liant les acceptions d'un mot non nécessairement reconnu comme polysémique n'est pas nouvelle, mais Rosch semble bien la considérer comme telle, sans doute parce qu'elle ignore l'histoire de la sémantique et que le modèle type de catégorie à attributs discrets qui prévalait alors en psychologie était celui de Bruner. Historiquement, cette idée de chaîne sémantique a d'abord été appréhendée dans une perspective diachronique, dont l'objet était de décrire l'évolution du sens d'un mot (souvent sous l'effet de tropes). Elle est donc consubstantielle aux débuts de la sémantique moderne au 19^{ème} siècle (Nerlich 1992).

Dans leur commentaire des *Investigations Philosophiques*, Baker et Hacker (1980 : 207-8), en philosophes, citent au nombre des précurseurs de Wittgenstein Dugald Stewart et John Stuart Mill (dont le père, James Mill, avait été le disciple de Stewart). L'idée de dérivation sémantique est associée chez Dugald Stewart à une vision empiriste qui accorde la primauté aux acceptions référant aux “objets des sens” (1816 : 262s). Elle est introduite dans une discussion où se trouvent contestées les définitions générales et abstraites. Mais il s'agit en réalité d'une autre perspective que celle de Wittgenstein, puisque celui-ci s'intéresse à des instances qui auraient été sans doute considérées par Stewart comme tombant sous un même sens.

⁴⁰ Rosch et Mervis se contentent de renvoyer aux *Investigations Philosophiques*, sans donner de référence précise. Comme les sections 66 à 71 sont le locus classicus, et que Rosch s'appuie ailleurs sur ces textes (Rosch 1987), il est vraisemblable qu'ils sont bien le fonds dans lequel elle puise. Un autre passage important, et qui anticipe sur celui des *Investigations Philosophiques*, se trouve dans la *Grammaire Philosophique* (I.vi, § 73).

⁴¹ Il n'est pas impossible que l'expression de *ressemblance de famille* soit un écho de Brentano : faisant allusion à la question classique évoquée par Aristote dans la *Métaphysique* Γ, 2, qui est de décrire les rapports entre les sens de *sain* dans *médecine saine, homme sain, urine saine* (analogie *pros hen*, par référence à une chose, ici la santé), Brentano (1816 : 60) parle de *noms de famille* (*Familiennamen*), soulignant l'impossibilité d'une

L'originalité de Wittgenstein est qu'il s'intéresse aux emplois qui semblent subsumés par un sens qu'on serait tenté de considérer comme univoque. Il n'est pas certain qu'une telle univocité soit introuvable et je ne crois pas que Wittgenstein entende montrer qu'une définition générale du jeu est impossible. Il semble plutôt qu'il considère cette définition comme inutile, ou comme, justement, un des jeux de langage que nous jouons et que nous érigerions à tort en modèle. A cet égard, le texte sur le jeu peut être envisagé comme une mise en abîme, du fait que, comme on sait, Wittgenstein dépeint le langage sous les traits d'un ensemble de jeux. Ainsi, une façon d'interpréter le texte, d'ailleurs corroborée par d'autres passages, serait d'y voir un avertissement lancé aux tentatives de réduire l'usage du langage à des jeux archétypaux, comme le jeu de la définition ou encore le jeu de la dénomination.⁴²

Quoi qu'il en soit des motivations de Wittgenstein, il offre à Rosch le moyen de penser la possibilité de catégoriser des instances sur la base de leurs attributs en l'absence d'une définition générale. On remarquera toutefois que dans le réseau de ressemblances décrit par Wittgenstein, on ne trouve pas de prototype.

Il s'agit maintenant de comprendre comment un prototype peut émerger d'une ressemblance de famille. En outre, comme de proche en proche les ressemblances qui lient un jeu à un autre peuvent être de plus en plus lâches, voire inexistantes, il importera aussi de comprendre comment la catégorie peut conserver une certaine cohésion. C'est vers la solution que Rosch apporte à ce double problème, celui de la typicalité et de la cohésion, que nous nous tournons maintenant.

D.3. La formation des catégories : une vision pragmatiste

Commençons par le second problème, celui de la cohésion des catégories. Il s'agit du problème crucial, puisqu'il met en question l'existence même des catégories. En effet, le cas limite du manque de cohésion, c'est la dissolution même de la catégorie.

Rosch est ainsi conduite à s'interroger sur les raisons pour lesquelles nous constituons des catégories. Les réponses qu'elle va formuler croisent de multiples influences qu'on peut ramener à trois approches principales : le pragmatisme, le fonctionnalisme probabiliste et la théorie informationnelle de la catégorisation.

Du pragmatisme, Rosch hérite l'idée que nos catégories existent parce qu'elles sont utiles. Elles sont utiles parce que catégoriser c'est, à partir d'une connaissance limitée, produire des inférences sur l'objet catégorisé et donc prédire les conséquences de notre comportement vis-à-vis de cet objet :

“une catégorie a le maximum d'utilité quand le fait de connaître la catégorie à laquelle une chose appartient permet à l'organisme de connaître autant d'attributs de cette chose que possible. L'utilité de la segmentation d'un même domaine décroît à mesure que diminue le nombre de propriétés prédictibles à partir de la connaissance de la catégorie” (Rosch 1975c : 197).

Le fondement pragmatiste de cette vision de la catégorisation a été explicitement revendiqué par Bruner. En témoigne le passage suivant :

“L'assignation d'un objet à une catégorie a des conséquences heureuses en ce qui concerne le comportement ultérieur qui est dirigé vers l'objet perçu ; voici un objet qui a l'apparence d'une pomme et, de fait, il tient le docteur à distance si on en mange un par jour.”⁴³

définition unique (comme Wittgenstein) mais rapportant les divers emplois de *sain* à un même sens primaire, la santé.

⁴² Sur la variété des jeux de langage, voir le §23. Sur la dénomination comme jeu, voir le §126.

⁴³ Allusion au dicton “an apple a day keeps the doctor away.”

Il faut dire que les philosophes, et notamment le pragmatiste C. S. Peirce ont passé plus de temps à défendre cette conception que les psychologues n'en ont mis à les écouter sérieusement. Le sens [meaning] d'une proposition, comme Peirce l'a remarqué dans sa célèbre étude sur la théorie pragmatique de la signification⁴⁴, est l'ensemble des jugements hypothétiques qu'on peut formuler sur les attributs ou conséquences associés à cette proposition. (...) Le sens d'une chose est ainsi l'assignation d'un objet à un réseau d'inférences hypothétiques portant sur autres propriétés observables, ses effets etc. (...) En apprenant à percevoir, nous apprenons les relations qui existent entre les objets et les événements que nous rencontrons, nous apprenons les bonnes catégories et les bons systèmes de catégories, nous apprenons à prédire et à vérifier quelle chose va avec laquelle" (Bruner 1957 : 126).

Une catégorie utile est donc une catégorie qui autorise de nombreuses inférences, et en particulier une catégorie qui est constituée d'attributs prédictibles les uns des autres. Rosch ajoute que cette utilité croît avec le nombre d'attributs associés par la catégorie.

Le même principe permettait à Bruner d'expliquer la typicalité : de même qu'une catégorie se forme par la conjonction d'attributs cooccurrents, de même les membres typiques sont ceux qui agrègent le plus d'attributs de la catégorie, ou dont les attributs dévient le moins de ceux enregistrés dans la catégorie :

"Un oiseau a des ailes, un jabot, des plumes et des pattes caractéristiques. Mais l'ensemble complet des traits n'est pas nécessaire pour identifier correctement une créature comme oiseau. Si elle a des ailes et des plumes, le jabot et les pattes sont hautement prédictibles. En codant ou en catégorisant l'environnement, nous nous attendons à ce que tous ces traits soient coprésents. C'est cette conception unitaire qui a la propriété configurationnelle ou la qualité figurale [Gestalt property] d'"aviarité" [birdness]. De fait, une fois qu'une configuration a été établie et que l'objet a été identifié en termes d'attributs configurationnels, l'organisme percevant tendra à "rectifier" ou "normaliser" tout attribut qui dévie de ce qui était attendu. (...) Quand cette conception est suffisamment bien établie, elle acquiert la capacité de fonctionner comme un attribut discriminable et qui a apparemment sa forme propre d'irréductibilité. On peut classer les choses selon le degré auquel elles s'approchent de l'aviarité, d'une manière très similaire à ce qu'on fait lorsqu'on classe des longueurs, alors que la longueur est un attribut qu'on peut présumer moins complexe" (Bruner et al., 1956 : 47).

Bien que ce thème des degrés de typicalité ait été peu exploré par Bruner, qui s'intéressait davantage aux stratégies mises en œuvre par les sujets pour identifier des catégories "classiques", il est clairement présent chez lui. On note aussi que la typicalité d'une instance est pensée en termes de proximité à une forme globale, et comme initialement chez Rosch, sa vision de la catégorie "réelle" est gestaltiste.

En conclusion, nous sommes en possession d'un premier principe de formation et de cohésion des catégories : le caractère interprédictible des attributs.

D.4. La ressemblance de famille à l'épreuve expérimentale

Il s'agit maintenant de comprendre comment de la prototypicalité émerge d'une ressemblance de famille. Rosch et Mervis posent trois hypothèses : (1) les membres d'une catégorie ont une ressemblance de famille, c'est-à-dire ont en commun certains attributs mais pas tous ; (2) les membres qui partagent le plus d'attributs avec d'autres sont les plus typiques ; (3) les membres les plus typiques sont ceux qui ont le moins d'attributs en commun avec des membres d'autres catégories.

⁴⁴ Peirce (1878).

Leur protocole est simple : elles demandent à des sujets de lister des attributs qu'ils associent aux membres de diverses catégories. Leurs instructions engagent les sujets à se focaliser sur des attributs objectifs (à lister des parties d'objets ou des activités concrètes accomplies avec eux). Elles comptent ensuite les attributs que chaque membre a en commun avec le reste de la catégorie.⁴⁵ Elle vérifient ensuite si les membres qui ont le plus d'attributs en commun avec le reste de la catégorie sont ceux qui avaient été jugés les plus typiques dans une expérience antérieure (Rosch 1975b). Elles constatent enfin que les résultats confirment les thèses (1) et (2).

La thèse (3) est nettement plus difficile à mettre en évidence. Il faut en effet des catégories contrastives pour juger si un attribut est distinctif d'une catégorie et se trouve peu dans d'autres. Mais quelles sont les catégories qui contrastent avec *chaise* ? Rosch et Mervis (1975) n'ont obtenu de leurs sujets que peu de réponses à ce genre de questions, ou des réponses qu'elles ont jugé "déraisonnables" ou trop "créatives" (ibid. : 584). Leur solution de rechange déplace complètement le problème : elles substituent en effet aux catégories "sémantiques" des catégories artificielles, en l'occurrence des suites de lettres (par ex. PHMQB ou XPHMQ) où une lettre vaut un attribut et pour lesquelles les ressemblances sont donc *perceptives*. Elles ont alors beau jeu de montrer que les sujets apprennent plus facilement des catégories dont les membres ressemblent moins à ceux d'autres catégories. Mais qu'en est-il pour des catégories où les attributs sont en nombre indéfini, les membres susceptibles d'appartenir à plusieurs catégories et les relations de ressemblance contextuellement déterminées ? Quelle est la catégorie de *sardine* : les poissons, les aliments en boîte, les créatures comprimées dans un endroit exigu, les objets susceptibles d'obstruer le port de Marseille ? Si *sardine* change de catégorie selon le contexte, les catégories contrastives aussi. Certes, Rosch obtient bien des listes d'attributs, mais, encore une fois, on ne peut s'empêcher d'observer que l'expérience et ses instructions, et le choix de la catégorie coiffant les membres, sont seulement un contexte parmi d'autres.⁴⁶

Le caractère acontextuel de la théorie est de fait un problème majeur. Les tentatives que fera Rosch pour valider ses normes de typicalité par d'autres procédés que les jugements de sujets et les listes d'attributs ne remédient pas au problème. L'un de ces procédés consiste à demander à des sujets de produire un contexte où figure le nom d'une catégorie, puis, ayant substitué à ce nom celui d'une instance plus ou moins typique, à faire évaluer aux sujets l'acceptabilité de la substitution. A titre d'exemple, le contexte pouvait être "une vingtaine d'*oiseaux* sont souvent perchés sur les fils téléphoniques devant ma fenêtre et le matin ils gazouillent" et parmi les substitutions se trouvaient *hirondelles* (typique) ou *dindes* (moins typique). Entendue comme une mesure du degré auquel un exemplaire représente une catégorie, la substituabilité semble plutôt être un doublon des normes de typicalité : les exemplaires typiques sont ceux qu'on trouve dans des contextes stéréotypés, et ces contextes stéréotypés sont ceux qui identifient les exemplaires typiques. Comme le souligne Rastier, "*le canari et le poisson* évoquera plutôt le cyprin doré prototypique, mais non *le cormoran et le poisson*", et pour le citadin moyen le premier contexte est probablement plus stéréotypé (1991 : 196).

⁴⁵ Les catégories étaient les suivantes :

Furniture (chair, sofa, table, lamp, footstool, piano, cushion, radio, stove, picture...).

Vehicle (car, truck, bus, bicycle, trolley car, tractor, cart, wheelchair, tank, raft, sled, horse, blimp 'ballon dirigeable', wheelbarrow, elevator...).

Fruit (orange, apple, banana, blueberry, lemon, honeydew 'variété de melon', date, coconut, tomato, olive...).

Weapon (gun, knife, sword, bow and arrow, tank, teargas, whip, icepick, fists, rocket, poison, scissors, words, foot, screwdriver...).

Vegetable (peas, carrots, string beans, lettuce, beets, tomato, Lima beans, onion, potato, yam, pumpkin, rice...).

Clothing (pants, shirt, dress, skirt, bathing suit, shoes, tie, mittens, hat, apron, purse, wristwatch, necklace...).

⁴⁶ Voir Rastier (1991) et Barsalou (1982, 1983) pour des critiques allant dans ce sens.

D.5. Indice et validité d'indice : le fonctionnalisme probabiliste

Rosch est maintenant en possession des principes qui permettent d'entrevoir une solution au double problème qui s'était posé (celui de la typicalité et de la cohésion dans une ressemblance de famille) : pris ensemble, ces principes impliquent qu'une catégorie se forme à partir d'attributs interprédictibles, et que la typicalité d'une instance se mesure au nombre d'attributs qu'elle partage avec le reste de la catégorie. Les jeux sont des jeux parce qu'ils ont des attributs interprédictibles, et les jeux les plus typiques sont ceux qui, parmi les jeux, ont le plus d'attributs en commun avec la catégorie, c'est-à-dire avec les autres jeux.

A cette dernière caractéristiques des membres les plus typiques, Rosch ajoute une dimension contrastive : les prototypes seront aussi les membres les plus différents des membres de catégories contrastives. En termes de ressemblance de famille, l'hypothèse se résume ainsi :

“Les membres d'une catégorie en viennent à être considérés comme prototypiques de la catégorie dans son entier dans la mesure où ils ont une ressemblance de famille (ont des attributs qui se chevauchent) avec d'autres membres de la catégorie. Inversement, les items considérés comme les plus prototypiques d'une catégorie seront ceux qui ont le moins de ressemblance de famille avec d'autres catégories et le plus faible degré d'appartenance à ces autres catégories” (Rosch & Mervis 1975 : 575).

L'idée de Rosch est de retraduire la ressemblance de famille en employant la notion de force prédictive d'un attribut. Autrement dit, plus un attribut est spécifique d'une catégorie et plus il est prédictif de l'appartenance d'un item à une catégorie ; et une instance est d'autant plus typique qu'elle a plus d'attributs de ce genre. Pour décrire et mesurer cette prototypicalité, Rosch va maintenant importer le vocabulaire et un mode de calcul issus du fonctionnalisme probabiliste d'Egon Brunswik. Nous reviendrons ci-après sur le fonctionnalisme probabiliste et ses origines. Pour nos besoins immédiats, une présentation cursive suffira.

Selon le fonctionnalisme probabiliste, un attribut proximal ou “palpable” (Beach 1964) est un indice (*cue*) permettant d'inférer une propriété de l'objet concerné ou sa catégorie. Un indice est associé de façon probabiliste à une propriété ou à une catégorie. Dans le cas de l'identification d'une catégorie, la tâche de l'organisme est de déterminer les probabilités que les indices rencontrés soient associés à telle catégorie plutôt qu'à telle autre. Ces probabilités d'indices tiennent compte des catégories contrastives : dans le mode de calcul proposé par Beach, la somme de ces probabilités propres aux indices d'un objet donné relativement à une classe donnée est divisée par la somme des probabilités correspondantes pour toutes les catégories où ces indices sont pertinents. Un indice a donc d'autant plus de chances de signaler l'appartenance à une catégorie donnée que sa probabilité de signaler d'autres catégories est faible. C'est le résultat de ce type de calcul que Beach appelle la “quantité relative de preuve” (*relative amount of evidence*), et que Rosch va dénommer *validité d'indice* (*cue validity*), revenant à la terminologie de Brunswik.

Ainsi, la validité d'indice d'un indice x prédicteur d'une catégorie donnée y croît avec la fréquence d'association de cet indice avec la catégorie y et décroît avec la fréquence d'association de cet indice avec d'autres catégories (Rosch et al. 1976 : 384). La validité d'indice d'un membre est la somme des validités d'indice de ses attributs (c'est-à-dire de ceux qui sont pertinents pour une catégorie donnée). Les prototypes sont les membres qui ont une validité d'indice maximale.

D.6. Origine de la notion de *cue validity*

Si l'on se fie aux références données par Rosch (Rosch & Mervis 1975 ; Rosch et al. 1976) la notion de *cue validity* semble avoir été relayée par Reed (1972) et Beach (1964), qui la tient

lui-même d'Egon Brunswik (1903-1955). Rosch elle-même reconnaît l'héritage de Brunswik mais en déplore le contenu encore inchoatif et programmatique (Rosch 1975c : 198). C'est ici l'occasion de présenter Brunswik.

Né en Autriche, disciple de Karl Bühler, dont il fut l'assistant au *Psychologisches Institut* de Vienne, installé plus tard aux Etats-Unis, Brunswik semble avoir exercé une influence à retardement sur la psychologie américaine (Brunswik, Hammond & Stewart 2001). Il y a eu peu de psychologues brunswikiens, et son magistère paraît n'avoir eu qu'un rôle intellectuel assez diffus (Leary 1989).

Les idées de Brunswik pénètrent aux Etats-Unis en partie grâce à Edward Tolman, un des grands noms de la psychologie américaine. Brunswik avait fait la connaissance de Tolman lors de la visite de ce dernier au laboratoire de Bühler, visite qui faisait suite à plusieurs séjours de Tolman en Allemagne, notamment chez l'un des fondateurs de la Gestalttheorie allemande, Kurt Koffka. Peu après leur rencontre, Brunswik et Tolman collaborèrent à un article (Tolman et Brunswik 1935). Tolman invita ensuite Brunswik en 1935 à Berkeley où ce dernier finit par s'installer définitivement en 1937 (Hammond & Stewart 2001 ; Leary 1989). L'installation de Brunswik aux Etats-Unis accompagnait un mouvement de fond, une véritable *translatio studiorum*, durant laquelle de nombreux savants germanophones, et en particulier les tenants de la Gestalttheorie traversèrent l'Atlantique.⁴⁷

Deux grandes thèses générales caractérisent la théorie de Brunswik : la perception procède par inférence à partir d'indices, et ces inférences sont de nature probabiliste. La nature inférentielle de la perception est une idée léguée par Helmholtz. Quant à la thèse que la connaissance est de nature probabiliste, elle semble avoir été inspirée par Reichenbach, avec qui Brunswik était entré en contact par le fait de leur commune appartenance au Cercle de Vienne (Brunswik 1943 ; Leary 1989).

La psychologie de Brunswik parle d'*objets* distaux, à une époque où les stimuli proximaux sont privilégiés par le behaviorisme, parce qu'on pense que la connaissance des chaînes stimulus-réponse permet de comprendre le mécanisme du comportement et d'envisager une réduction de la psychologie à une psychophysiologie (Brunswik 1955 ; Heider 1939). De Bühler, Brunswik dit retenir l'idée que la perception et le comportement reposent sur une relation sémiotique, par laquelle un représentant ou *indice*, stimulus proximal mais aussi objet distal reconnu comme moyen en vue d'une fin, tient lieu d'un représenté (respectivement, l'objet perçu ou la fin poursuivie ; Tolman et Brunswik 1935).⁴⁸ Un organisme, explique-t-il, repère des couplages causaux entre des indices, par exemple, la disparité des images rétinienne, et une propriété du monde, par exemple, la distance d'un objet. Ces couplages sont seulement probabilistes : le représentant ne mène pas à coup sûr au représenté, le moyen ne permet pas d'atteindre infailliblement la fin. Dans le cas de la vision du relief, la disparité rétinienne est prise en défaut par le stéréoscope, puisqu'elle ne fonctionne plus comme indice d'un relief réel (Tolman & Brunswik 1935). Ces variations dans la "texture causale" de l'environnement peuvent être exprimées au moyen du concept de *validité d'indice* : on dira que la validité de l'indice est plus ou moins élevée selon que l'existence de la propriété du monde est plus ou moins assurée lorsque l'organisme prend en compte cet indice. La proximité, au sens gestaliste de principe de groupement et de constitution d'une figure, est un de ces indices dont la validité peut être mise en évidence: à partir de photographies, Brunswik montre que les lignes (approximativement) parallèles les plus rapprochées appartiennent aux

⁴⁷ Koffka émigre en 1927, Lewin et Wertheimer en 1933, Köhler en 1935 (Mandler 2007).

⁴⁸ Brunswik renvoie à l'évocation par Bühler de la notion traditionnelle de signe, défini comme *id quod stat pro aliquo*. Dans sa *Théorie du Langage*, il salue le tournant sémiotique pris par certains travaux de psychologie, notamment ceux de Tolman et de Uexküll (l'inventeur du concept d'*Umwelt*, c'est-à-dire d'environnement qui, pour un organisme, est signifiant ; Bühler 2009 [1934] : 107-8).

mêmes objets, et non à des objets différents (Brunswik & Kamiya 1953). Comme la perception est de façon ultime un mode d'adaptation de l'organisme à son milieu (ou "ecology"), Brunswik parle aussi de la *validité écologique* d'un indice.

Il semble que ce soit Tolman qui ait amené Brunswik à intégrer dans son modèle la relation moyens-fin et l'objet comme terme du comportement (voir Leary 1989 à ce sujet ; Tolman 1933). Un objet qui se donne comme moyen, diront-ils de concert, joint à ses facettes perceptives (ses *discriminanda*) des aspects "manipulatoires" (des *manipulanda*, par exemple sa "masticabilité") et des aspects utilitaires, en tant qu'il est appréhendé comme moyen (ce sont ses *utilitanda*, par exemple satisfaire la faim). Les *manipulanda* et *utilitanda* devaient, dans l'esprit de Tolman, rendre compte du fait que l'environnement était structuré comme un *champ de moyens-fins*, concept que Tolman disait hériter du psychologue gestaltiste Kurt Lewin (Tolman 1933). En raison de cette relation fonctionnelle qui lie la perception à son objet et le comportement à son objectif, et en vertu de la nature probabiliste de cette relation, Brunswik baptise sa théorie *fonctionnalisme probabiliste* (Brunswik 1955, 1956).

Notons que la notion de validité écologique n'est pas sans affinités avec la vision pragmatiste de la catégorisation. Cette convergence peut expliquer l'attrait qu'a pu avoir la pensée de Brunswik en milieu américain. Bruner reconnaît d'ailleurs en plusieurs endroits l'influence de Brunswik (1957 ; Bruner, Goodnow & Austin 1956).

Pour notre propos, certains aspects des idées de Brunswik revêtent une importance particulière. D'une part, la théorie de Brunswik ne pose pas de relation biunivoque entre indices et propriétés, autrement dit, un même indice peut mener à des propriétés différentes et une même propriété peut être associée à des indices différents (comme c'est le cas pour la perception de la distance). En termes roschiens, les attributs sont distribuables sur plusieurs catégories. D'autre part, l'indice n'est pas un critère permettant de *définir* la propriété. Brunswik a en effet une vision très large de l'indice : la taille, constate-t-il, est utilisée par des sujets comme un indice permettant d'apprécier l'intelligence d'un individu ; or, elle n'est évidemment pas une caractéristique définitoire de l'intelligence (Brunswik 1943). Bruner sera sur ce point le continuateur de Brunswik, quand il distinguera parmi les attributs des attributs *critériaux*, qui permettent à un sujet de décider de l'appartenance d'un item à une catégorie, et des attributs *définitoires*, qui sont du ressort des disciplines compétentes (Bruner, Goodnow & Austin 1956). De même, chez Rosch, un attribut qui n'est pas définitoire peut tout de même intervenir dans la catégorisation d'un item. C'est ce que montre le passage suivant, où Rosch distingue les attributs critériaux (ici, en fait, les attributs définitoires de Bruner !) de la *signification* d'une catégorie, pensée comme *représentation* des référents : "Les sous-catégories qui sont identiques du point de vue des attributs critériaux [= définitoires] peuvent différer dans le degré auquel elles représentent le sens fondamental [*core meaning*] de la catégorie — par exemple, la distinction entre sauvage et domestiqué est sans pertinence pour une définition rigoureuse d'*oiseau* puisque les oiseaux comme les non-oiseaux peuvent être sauvages ou domestiqués ; cependant, les oiseaux sauvages sont considérés comme plus centraux que les oiseaux domestiqués" (Rosch, 1973b : 141-2).

En outre, chez Brunswik, la validité de l'indice est corrélée à sa véridicité, puisqu'un organisme a peu de chances de survie s'il utilise des indices trompeurs. Cette résonance réaliste de l'indice a sa contrepartie chez Rosch dans l'idée que les catégories, au moins à un certain niveau d'abstraction, expriment des "structures corrélationnelles" de l'environnement. Enfin, nous verrons que Rosch associe attributs perceptifs et comportements moteurs d'une manière qui semble faire écho aux *discriminanda* et *manipulanda* de Tolman et Brunswik.

D.7. De l'indice à l'information : la structure corrélacionnelle

A l'époque où Rosch élaborait sa théorie, nous avons vu que la notion de prototype était étroitement liée à celle de schéma, et que cette dernière était associée à celle de tendance centrale et disjointe de celle de validité d'indice. La validité d'indice entraine certes dans le calcul probabiliste de l'appartenance d'un item à une catégorie, mais ne servait pas à singulariser un membre de la série. Ainsi Reed (1972) *opposait*-il la catégorisation par calcul de la distance d'un item au prototype, envisagé comme tendance centrale d'une catégorie, et la catégorisation par calcul de la validité d'indice. Au contraire, Rosch *identifie* ce qu'elle appelle le prototype à l'instance ayant la validité d'indice la plus élevée.

On peut se demander maintenant si un prototype est distinctif parce que ses attributs sont peu partagés par d'autres catégories, ou bien parce qu'il correspond à la configuration la plus riche d'attributs corrélés, ou les deux à la fois. Ces propriétés pourraient en effet être disjointes : un prototype pourrait être constitué de beaucoup d'attributs partagés par des membres de la catégorie mais dont certains seraient peu distinctifs de la catégorie. Rosch n'est pas très claire sur ce point, ce qui pourrait indiquer que pour elle les deux propriétés sont liées. Si un attribut est très distinctif d'une catégorie, il permet de prédire avec une grande probabilité l'appartenance à cette catégorie, et de cette appartenance on peut inférer la présence d'autres attributs. La distinctivité d'un attribut mènerait ainsi à des corrélations entre attributs.

Pour décrire ces corrélations, Rosch va employer des outils venus de la théorie de l'information. Elle dira ainsi que des attributs sont *redondants* lorsque de l'un on peut prédire la présence des autres. Affirmer qu'au sein d'une catégorie les prototypes ont la validité d'indice la plus élevée se traduira en termes informationnels en disant que les prototypes maximisent la redondance de l'information. Pour comprendre ce nouveau langage, un détour par la théorie qui a inspiré Rosch se révèle nécessaire.

D.8. Les noces de l'indice et de l'information : vers une explication génétique

La redondance reflète quelque chose de réel dans l'environnement, que Rosch, à la suite de Garner (1974), appelle la *structure corrélationnelle*. On peut dire aussi que cette structure est porteuse d'une information, en un sens que nous serons amenés à préciser ici.

Vu l'importance de cette notion de *structure corrélationnelle* dans l'économie générale de la théorie, il faudra nous attarder quelque peu sur ce que Garner entend par ces termes, ce qui nous permettra de comprendre quel bénéfice Rosch a pensé en tirer. Je présenterai donc d'abord l'approche de Garner, avant d'en venir à la manière dont Rosch l'a interprété.

Garner n'est pas le premier à appliquer des concepts de la théorie de l'information à la psychologie. Peu de temps après son introduction dans la théorie de l'information, le concept de redondance avait été rapproché par Brunswik de celui d'attributs vicariants (*vicarious cues*) : de même, disait-il, que la redondance permet une transmission plus sûre de l'information, de même le fait de disposer de plusieurs types d'indices pour une même propriété devait assurer une meilleure stabilité des inférences (Brunswik 1955 : 212). A la même époque, Attneave (1954), en partie inspiré par les travaux de James Gibson, avait suggéré d'envisager les régularités dans l'environnement comme des formes de redondance exploitables par l'organisme et avait proposé des principes d'analyse perceptive exploitant ces redondances.

Les travaux de Garner (1962, 1974) cependant, poussent nettement plus loin l'application systématique et formelle des outils théoriques issus de la théorie de l'information. Un exemple permettra d'apprécier l'esprit de ces travaux.

Soit l'ensemble de figures suivant (figure 3). Il est composé de 16 stimuli variant sur quatre dimensions, et chaque dimension a deux valeurs possibles (la forme est circulaire ou carrée, l'ouverture est droite ou gauche, la ligne est verticale ou horizontale, elle est droite ou brisée).

Cet ensemble épuise les combinaisons possibles. Il constitue, selon les termes de Garner, un *ensemble total*.

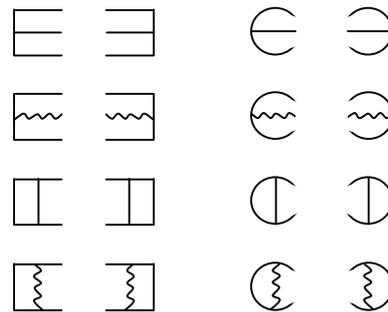


FIG. 3 : Un ensemble total de stimuli à quatre dimensions de deux valeurs chacune (Garner 1974 : 6-7)

Supposons que nous choissions parmi cet ensemble les colonnes 1 et 4, de façon à obtenir un sous-ensemble de huit stimuli (figure 4). Dans ce sous-ensemble, les dimensions et, dans chaque dimension, les valeurs de cette dimension sont équiprobables :

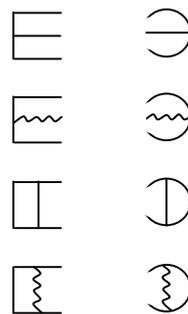


FIG. 4 : Sous-ensemble redondant de l'ensemble total précédent (Garner 1974 : 7)

On remarque que deux valeurs sont corrélées : si la forme du stimulus est carrée, son ouverture est à droite, si la forme est circulaire, l'ouverture est à gauche. On peut exprimer la même chose en disant qu'en supprimant l'une ou l'autre de ces dimensions (c'est-à-dire la forme ou le côté de l'ouverture), par exemple en ouvrant tous les stimuli du même côté, on obtiendrait tout de même huit stimuli différents. Autrement dit, la *structure corrélacionnelle* de ce sous-ensemble entraîne que l'une de ses dimensions de variation est *redondante*. Cette structure corrélacionnelle est une propriété du stimulus, elle est donc *réelle*. De même que James Gibson (1979), Garner passe de l'idée que l'environnement est structuré et que cette structure est de l'information à l'affirmation d'une forme de réalisme.

Conformément à son approche qui part d'un "ensemble total", Garner insiste à plusieurs reprises sur le fait que construire une catégorie, c'est inférer un ensemble plus vaste de variations possibles voire l'ensemble total des variantes, c'est-à-dire des catégories contrastives possibles. Lorsqu'il analyse la classification spontanée de stimuli, il ne manque pas non plus de souligner que le contraste entre catégories semble guider la catégorisation, de sorte que la catégorisation paraît suivre un principe qu'il formule ainsi : "le sujet classe les stimuli de façon à maximiser les différences perçues entre les classes tout en maximisant les ressemblances perçues à l'intérieur des classes" (Garner 1974 : 98). La convergence avec Rosch est frappante et il est possible que Rosch ait trouvé chez Garner de quoi expliquer la

cohésion des catégories et leur séparation les unes d'avec les autres, malgré le caractère lâche de la structure en ressemblance de famille.

Aux yeux de Rosch, la théorie de Garner possède un attrait supplémentaire (Rosch 1975a : 545). Elle explique en termes rigoureux (informationnels) l'asymétrie existant entre le prototype et une autre instance, asymétrie qu'elle avait décrite au moyen des concepts de point de référence et point d'ancrage. Ainsi, sur la figure ci-après, si la flèche signifie "est similaire à", les bonnes formes seront à droite, ce qui est une manière de dire que la relation de similarité est orientée. Il s'agit là d'une asymétrie qui correspond exactement à ce que voulait dire Rosch quand elle faisait jouer au prototype le rôle d'un point de référence.

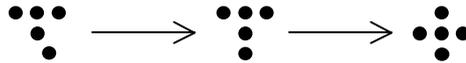


FIG. 5 : un exemple de similarité asymétrique (Garner 1974 : 19).

En outre, la notion de bonne forme reçoit chez Garner une définition rigoureuse. On notera que la forme en croix de la figure 5 est invariante en cas de rotation par rapport à un axe vertical ou horizontal, de sorte que les bonnes formes peuvent être caractérisées comme des structures qui sont invariantes sous un certain nombre de transformations.

Malgré ces concordances entre les deux auteurs, des malentendus subsistent. Il est important de garder à l'esprit que, pour Gardner, la redondance n'est pas synonyme d'avantage cognitif. Tout dépend de la situation considérée, dont les paramètres sont multiples : nature de la redondance (simple ou complexe), du type de tâche (rappel ou discrimination), des dimensions elles-mêmes, du fait de choisir un matériau signifiant ou non signifiant... Or, pour Rosch, la redondance (ou la validité d'indice) est un principe d'efficacité : concentrée sur les prototypes, elle introduit de la discontinuité entre catégories, et ces prototypes simplifient le travail d'inférences à partir d'indices en fonctionnant comme étalons de comparaison.

Si l'on chausse les lunettes de Gardner, dire, comme le fait Rosch, que la redondance est un principe d'efficacité cognitive revient à privilégier un type de situation : celle où le sujet constitue lui-même une classe à partir d'items dont la structure corrélationnelle est simple. Il est donc clair que, pour Rosch, Garner a surtout l'intérêt de suggérer un principe d'explication de la *formation* des catégories : les items qui vont ensemble sont réunis par une structure corrélationnelle. Quant au rôle de la structure corrélationnelle dans le traitement de catégories déjà constituées, c'est chez Garner un point complexe, auquel Rosch ne se confronte pas.

Paradoxalement, Rosch n'hésite pas à s'appuyer sur Garner, qui emploie presque exclusivement des catégories artificielles, pour opposer le monde réel à l'univers des catégories factices telles qu'on les construit en psychologie : "l'opinion selon laquelle les catégories sont à l'origine arbitraires ne serait raisonnable que si le monde était entièrement non structuré, c'est-à-dire si le monde formait un ensemble de stimuli dans lequel tous les attributs possibles du stimulus étaient combinés de manière équiprobable avec tous les autres attributs possibles" (1976 : 242). Du point de vue de Garner, cette manière de s'exprimer est absurde. D'une part, le matériel expérimental de Garner consiste en grande partie en stimuli "arbitraires" aux attributs équiprobables. Mais surtout, dès lors qu'un sous-ensemble d'un "ensemble total" est considéré, de la structure est introduite dans ce sous-ensemble. En fait, ce que Rosch semble surtout retenir de Garner tient à son réalisme : la structure corrélationnelle et la redondance sont des propriétés de l'environnement qui se reflètent dans la structure des catégories.

D.9. Les niveaux de catégorisation

Les catégories “sémantiques” ont introduit un nouveau problème dans la théorie du prototype : les noms qui identifient les catégories et leurs membres sont à différents niveaux d’abstraction, *quadrupède* étant par exemple plus *abstrait* que *chien*. Le problème est alors de comprendre pourquoi ou à quelles fins nous avons besoin de catégoriser les mêmes référents à différents niveaux d’abstraction.

Dans l’environnement intellectuel de Rosch, deux types d’approche ont joué un rôle important sur la manière dont elle appréhende ce problème : de façon certaine, l’étude des taxonomies ethnobiologiques, principalement sous l’égide de Brent Berlin ; et, avec moins de certitude, les réflexions de Roger Brown sur le choix du niveau d’abstraction dans les interactions verbales. Commençons par la notion de *genre* chez Berlin.

Berlin part de la constatation que les classifications d’espèces vivantes, dans diverses cultures, font apparaître trois niveaux saillants : un niveau supérieur, appelé “forme de vie” (*life form*) et articulé en un nombre limité de termes (par ex. *arbre, oiseau, mammifère, poisson*); un niveau inférieur, celui des genres populaires (*folk genera*) comptant de nombreux membres, souvent plusieurs centaines, qui tendent à être désignés par des termes monolexémiques ou des composés non-productifs, et dont la majorité sont monotypiques (ont peu ou pas de subdivisions, par ex. *pin, bar*) ; et un niveau “spécifique”, dont les dénominations sont souvent binomiales (*pin noir, pin parasol, bar blanc, bar rayé*). D’autres niveaux sont plus marginaux : le niveau le plus abstrait (le *règne* in Berlin 1978) coiffant celui des formes de vie est, selon Berlin et al. (1973: 215), rarement nommé (par ex. *plante, animal*). De même, il est peu fréquent que le niveau intermédiaire entre la forme et le genre (par ex. *arbres à feuilles persistantes*) soit exprimé. Enfin, le niveau le plus particularisant est celui des variétés (par ex. *pin noir de Turquie*).⁴⁹

Il est important de comprendre que cette distinction en niveaux ne coïncide pas exactement avec une hiérarchie taxonomique : un taxon qui n’a au-dessus de lui que le niveau du règne ne correspond pas nécessairement à une forme de vie. Il peut être *conceptuellement* générique mais non affilié à une forme de vie déterminée et placé directement sous le taxon le plus élevé (c’est-à-dire sous *plante* ou *animal*). Les classifications populaires comprennent en effet des membres qui sont, en ce sens, aberrants. Cette définition *conceptuelle* du genre est essentielle et prime sur une définition par rang taxonomique. Pour considérer qu’un taxon est un genre, le fait que ce taxon ait peu de subdivisions et corresponde à un faisceau de différences morphologiques, comportementales, d’habitat etc. est discriminant.

Selon Berlin, le niveau générique serait privilégié : nous avons vu qu’il inclut un nombre plus grand de membres, qui sont généralement monotypiques, et qui tendent à être désignés par des termes inanalysables ou peu productifs, à l’instar des *basic color terms*, tandis que les dénominations de niveau spécifique sont souvent binomiales et analysables en genre et différence (Berlin et al.: 217). Les recherches personnelles de Berlin sur la terminologie botanique du tzelal montreraient aussi que les termes génériques sont les plus facilement élicités et les mieux mémorisés. Ils seraient enfin les premiers acquis (Berlin et al. 1973 : 221).

Le genre populaire tient ces prérogatives du fait qu’il correspondrait à des discontinuités perceptives du monde biologique (Berlin 1978 : 11), et ce sont ces discontinuités qui sont exprimées dans sa définition conceptuelle. Que l’origine perceptive de ces discontinuités soit perceptive, c’est ce que vient souligner encore la comparaison du genre avec une gestalt. A son tour, le genre-gestalt est rapproché du prototype roschien (Berlin 1978 : 17), ce qui semble impliquer qu’un genre populaire est en lui-même un prototype.

Berlin ne dit pas que cette explication par les discontinuités perceptives peut être généralisée à tous les niveaux. Parmi les taxa inférieurs aux genres il en est qui ont un certain relief et ont

⁴⁹ J’ai adapté ici les exemples de Berlin, les classifications du français étant parfois différentes.

typiquement une nomenclature simple (comme les genres). Ils le doivent non au fait qu'ils correspondent à des discontinuités perceptives mais à leur importance culturelle, par exemple en tant que cultigènes (Berlin 1978). On entrevoit ici la possibilité que le rôle culturel d'une espèce ou d'un objet puisse contrebalancer l'avantage perceptif censé être propre au niveau de catégorisation privilégié. Cette possibilité met en question l'origine perceptive du niveau privilégié.

D.10. L'éphémère "niveau générique"

Initialement, Rosch appelle "niveau générique" ce qui deviendra ensuite le "niveau de base" (Rosch 1975c). Cette dénomination est un renvoi explicite au genre populaire de Berlin, et entend entériner au plan cognitif les prérogatives que ce niveau avait reçues des travaux anthropologiques de Berlin. Pourtant, son niveau générique ne correspond pas à celui de Berlin. Rosch place en effet la discontinuité majeure au niveau des formes de vie de Berlin, par exemple entre *animal* et *oiseau*, et son niveau générique correspond donc à *oiseau* et non à *rouge-gorge* ou *corbeau*. Pourquoi cette divergence? Rosch choisit de traiter la catégorie *oiseau* comme la catégorie *chien*, c'est-à-dire pense que la relation *chien-doberman* est de même nature que la relation *oiseau-poule*. Or, ainsi que le note justement Kleiber (1994: 246-7), il se pourrait bien qu'une poule soit d'abord identifiée comme *poule* avant de l'être comme *oiseau*, ce dont ferait foi le fait que la reprise anaphorique de *poule* par *oiseau* est difficile dans

? *Paul a une poule dans sa basse-cour. Pierre a également un oiseau mais un dindon.*

Alors qu'une poule est d'abord catégorisée comme *poule*, un doberman est d'abord identifiable comme une sorte spéciale de *chien*. Sa catégorisation comme *doberman* semble demander un contexte où sa distinctivité par rapport à d'autres races de chien est focalisée (Kleiber 1994). Ainsi, le fait que Rosch et Berlin ne s'accordent pas sur le niveau privilégié (le genre) reflète un problème bien réel: il est à craindre en effet que le niveau privilégié de catégorisation d'un item ne soit pas toujours le même d'un item à une autre. Rosch pense le niveau générique à partir d'une catégorie comme celle des chiens, alors que Berlin prend plutôt comme modèle la relation *oiseau-poule*. Que la poule soit un cas marginal d'oiseau n'interagit pas, chez Rosch, avec le rang taxinomique où elle place le niveau de base.

Malgré cette divergence entre Rosch et Berlin, l'idée fondamentale est la même: la classification privilégie un niveau où le découpage opéré est d'origine perceptive et, en ce sens, "naturel". Ce découpage perceptif vaut non seulement pour les êtres vivants mais pour toutes les catégories qui peuvent être représentées par une image visuelle, par exemple pour les artefacts.

Le niveau générique est principalement appréhendé en termes de configuration d'attributs donnant lieu à une forme globale "moyenne" qui en est le prototype, ce qui rappelle irrésistiblement la qualité figurale d'aviarité dont parlaient Bruner et al. (cf. supra):

"Le niveau générique est celui où le gain en corrélation morphologique [*shape correlation*] par rapport au niveau immédiatement supérieur (plus abstrait) est le plus grand. Par exemple, les formes de "chaises" sont probablement beaucoup mieux corrélées que les formes de tous les objets classifiables comme "meubles"; cependant, le gain en corrélation est relativement faible quand on descend au niveau plus concret suivant, aux fauteuils, aux rocking chairs, aux chaises de cuisine etc. (...) Il se pourrait aussi que l'objet de niveau générique soit le niveau le plus abstrait de classification qui peut être codé cognitivement de façon économique, car il se pourrait que ce soit la forme moyenne qui serve de prototype. (...) Ainsi, il se pourrait que l'objet de niveau générique soit le point de rencontre des deux principes généraux de classification qui ont été discutés ici [c-à-d recherche de l'abstraction et reflet des corrélations d'attributs]. Les corrélations d'attributs séparables (tels que "a des ailes", "a un bec", "a deux

pattes” etc.) produiraient des corrélations des attributs intégraux qui définissent une forme.⁵⁰ Si des formes sont suffisamment corrélées, une forme moyenne est une représentation d’une classe d’objets qui fait sens [*meaningful*] et pourrait être codée cognitivement comme prototype” (Rosch 1975c : 201).

Le prototype tel qu’il est défini ici tient du schéma envisagé comme tendance centrale. Or, nous l’avons vu, le prototype doit aussi être saillant ou distinctif. Est-il alors une image moyenne ou une forme très distinctive? Y a-t-il une contradiction dans la théorie? Apparemment, la contradiction se résout au niveau générique : lorsque les instances d’une catégorie ont des formes bien corrélées et sont distinctives, la distinctivité moyenne suffit à assurer celle de la catégorie. Le prototype n’a pas à être particulièrement distinctif, puisque le reste de la catégorie l’est.⁵¹

D.11. Roger Brown et le “niveau moyen d’abstraction”

Sur le sujet des niveaux de catégorisation, il est difficile de cerner l’influence que Brown a pu avoir sur Rosch. Brown a été le professeur de Rosch, et à un moment, celle-ci se pose un problème que Brown s’était posé, mais en des termes différents, quelque vingt années auparavant (Brown 1958b).

Le propos de Brown (1958b) est double : il s’agit de comprendre pourquoi nous désignons une même chose de plusieurs noms, en particulier à différents niveaux de catégorisation, et de quelle manière se négocient les dénominations entre adultes et enfants, dont les intérêts pratiques et les capacités sont différentes.

La solution de Brown est utilitariste. Le nom le plus répandu est au niveau de discrimination qui est approprié à nos besoins pratiques :

“La cuillère est rarement appelée autrement que *cuillère*, quoiqu’elle soit aussi une pièce d’*argenterie*, un *artefact*, et une *certaine cuillère de restaurant mal nettoyée*. Au sein de la société, la préférence pour le mot *cuillère* correspond à la pratique sociale consistant à traiter toutes les cuillères comme équivalentes, mais différentes des couteaux et des fourchettes. Il n’y a pas de noms propres pour les cuillères parce que leur individualité a rarement un sens. Il en va de même pour les ananas, la petite monnaie, les portes, et les taxis. Le nom le plus répandu pour chacun d’eux les catégorise de la façon que la société a besoin de les catégoriser à des fins non-linguistiques. Le nom le plus répandu est au niveau d’utilité ordinaire” (Brown 1958b : 16).

A l’instar de Bruner, Brown fait résider cette utilité dans le pouvoir prédictif de la catégorisation : connaître la catégorie d’une entité, c’est inférer un ensemble d’attributs possédés par cette entité ou diriger l’attention vers des attributs pertinents dans une situation (1958a : 216s). Dire d’une carafe que c’est *du verre*, c’est attirer l’attention vers sa fragilité. Ainsi, en étant catégorisé, un objet se donne comme un support d’actions caractéristiques. Les actions associées à cet objet ont un rôle distinctif, sur lequel Brown insistera ensuite, dans un ouvrage postérieur : une entité catégorisée comme *chat* se donne comme objet à caresser, mais doucement, de peur des griffes. Cette caresse est si distinctive qu’elle n’est ni plus ni moins spécifique que la catégorie *chat* (Brown 1965 : 319).

Cependant, le monde utilitaire de l’adulte n’est pas celui de l’enfant. L’enfant n’a guère besoin, par exemple, de connaître exactement la valeur exacte d’une somme d’argent, et, dans

⁵⁰ Rosch emprunte les notions de dimensions séparables et intégrales à Garner. Taille et couleur sont des dimensions séparables, saturation et brillance ne le sont pas. Les sujets peuvent ignorer une dimension séparable quand ils en utilisent une autre comme base de classification (par ex. quand ils mettent à part tous les objets verts) mais combinent les deux dimensions quand elles sont intégrales. D’autres critères de séparabilité / intégralité figurent chez Garner (1974, chap. 5, 6 et 7).

⁵¹ Ce qui rejoint Berlin qui, nous l’avons vu, considérait le niveau générique comme prototypique.

le cadre d'un échange entre adulte et enfant, nommer cette somme *money* suffira. Si le niveau de catégorisation peut être ainsi le fruit d'une négociation dans laquelle l'adulte s'adapte à l'enfant, il n'en demeure pas moins qu'en général l'adulte impose ses catégorisations à l'enfant. Mais quel est le niveau que tendrait à privilégier l'enfant? Brown pense initialement que le développement cognitif va dans le sens de la différenciation : l'enfant aurait d'abord une catégorie indifférenciée *quadrupède*, mais les catégories utilitaires de l'adulte le contraindraient à descendre à un niveau plus différencié (*chien* ou *chat*). En somme, le développement du vocabulaire semble procéder à partir d'un *niveau moyen d'abstraction* (*middle level*, par ex. *cuillère*, *chien*) qui correspond le plus souvent aux catégories utilitaires de l'adulte, bien que le sens naturel du développement cognitif privilégierait des catégories initiales plus abstraites. Il invoque aussi les actions distinctives d'une catégorie (comme pour la catégorie *chat*), lesquelles définiraient le niveau initial d'acquisition chez l'enfant (Brown 1965).

A l'explication utilitariste et praxique de Brown Rosch substitue les principes qui président à la formation du niveau générique, c'est-à-dire la recherche des catégories les plus inclusives susceptibles de refléter les discontinuités des objets de l'environnement. Du fait que ces principes s'appliquent directement aux catégories perceptives, qu'ils donnent lieu à des images "génériques", et sont dans cette mesure indépendants du langage, Rosch considère qu'ils peuvent expliquer pourquoi les catégories de ce niveau générique peuvent être chez l'enfant un point d'entrée pour les catégories linguistiques. Elle s'attachera donc à montrer que les catégories de ce niveau sont les premières catégories formées, et les instances de ces catégories les premières nommées (Rosch 1975b, 1977). Autrement dit, l'acquisition des catégories est plus déterminée par les capacités cognitives de l'enfant et la structure de l'environnement que par les ajustements et négociations linguistiques entre l'enfant et les adultes, ou l'univers praxique dans lequel l'enfant est immergé, quoique ce dernier joue aussi un rôle pour Rosch, comme nous allons le voir.

D.12. Le "niveau de base"

C'est à l'étude du niveau générique, désormais rebaptisé *niveau de base*, que Rosch consacre un de ses articles les plus longs (Rosch et al. 1976). Le sujet permet en effet d'articuler l'appareil théorique applicable aux catégories "sémantiques" : structure corrélationnelle, validité d'indice, discontinuité morphologique.

Nous avons vu que Rosch caractérisait son niveau générique en termes de forme visuelle, mais aussi en termes de gain informationnel et de corrélations d'attributs. Deux propriétés se dégagent : le niveau générique avait une structure corrélationnelle riche, et il était bien distinctif (*chaise* contrastait plus avec *table* qu'au niveau inférieur *fauteuil* avec *rocking chair*). Nous savons désormais que ces propriétés peuvent être exprimées en termes de validité d'indice : au niveau de base la validité d'indice est maximale. Cette retraduction dans le vocabulaire du fonctionnalisme probabiliste coïncide avec la prise en considération de nouveaux attributs qui sont les interactions motrices associées à une catégorie (par ex. s'asseoir pour une chaise; Rosch et al. 1976; Rosch 1976). Ces interactions, qui font écho aux *actions distinctives* de Brown, aux *manipulanda* de Tolman (voir supra) et plus lointainement à l'*Affordung* de Lewin, participent à la saisie unitaire des objets au niveau de base, puisque Rosch montre que des séquences motrices caractéristiques sont partagées par les objets de ce niveau.

Pour vérifier l'hypothèse que les catégories de niveau de base ont une validité d'indice supérieure à celle des autres niveaux, Rosch emploie sa procédure habituelle: après avoir "épuré" une liste de catégories (nous y reviendrons), elle demande à des sujets de lister les attributs des catégories restantes, en comparant cette fois trois niveaux (surordonné, de base et

subordonné). Pour une catégorie surordonnée comme *outil* (tool), elle obtiendra trois attributs ('fabrique des choses', 'répare des choses', 'métal') mais sept pour une catégorie de niveau de base comme *scie* ('manche' [*handle*], 'dents', 'lame', 'coupe'...) et aucun attribut supplémentaire pour une catégorie subordonnée comme *scie à métaux* (Rosch et al. 1976).⁵² Enfin, pour certaines catégories, la définition taxonomique du niveau de base ne coïncide pas avec sa définition conceptuelle. Ainsi, *oiseau*, qui serait du point de vue taxonomique une catégorie de rang surordonné est conceptuellement une catégorie de niveau de base : le nombre d'attributs donné est typique d'une catégorie de base et les catégories subordonnées contiennent peu d'attributs supplémentaires. On voit ici que Rosch dissocie, comme Berlin, rang taxonomique et rang conceptuel.

A la suite de Berlin aussi, Rosch plaçait au niveau générique une discontinuité morphologique majeure dans la taxonomie. Les membres de catégories supérieures étaient censés avoir des formes visuelles beaucoup plus hétérogènes, et les membres de catégories inférieures des formes peu distinctes de ceux du niveau générique. En d'autres termes, ce niveau devrait constituer "le niveau le plus global auquel il soit possible de former une image mentale qui soit isomorphe à un membre moyen de la classe, ainsi, le niveau le plus abstrait auquel il soit possible d'avoir une image relativement concrète" (Rosch 1976 : 246). Cette intuition est vérifiée en mesurant le taux de chevauchement de tracés figurant les contours de membres d'une même catégorie. Rosch trouve que les membres d'une catégorie au niveau de base se chevauchent plus que ceux d'une catégorie surordonnée, mais que le gain de chevauchement entre le niveau de base et le niveau subordonné est faible. Enfin, le tracé moyen permet d'identifier facilement une catégorie de base, mais non une catégorie surordonnée (Rosch et al. 1976).

Comme l'expérience exigeait que les membres de catégories de base et surordonnées soient encore identifiables à partir d'un tracé, des catégories ont disparu (les fruits par exemple)⁵³. Les catégories restantes se caractérisent justement par des tracés bien distincts, présentés selon une orientation canonique dominante pour une même catégorie (les meubles de face, les véhicules et les animaux de côté).

Cette sélection des catégories pose un problème général. Dans l'expérience où Rosch demandait aux sujets de lister les attributs pour des catégories de différents niveaux, certaines catégories avaient été éliminées : légumes, jouets, armes, serpents. Les catégories restantes étaient : instruments de musique, fruits, outils, vêtements, meubles, véhicules, arbres, poissons, oiseaux. Il est vraisemblable que certaines catégories, comme les légumes, les jouets et les armes, avaient été éliminées parce qu'elles croisaient d'autres paradigmes (Rosch et al.

⁵² Les attributs donnés pour *scie* caractériseraient tout autant *couteau*. En revanche, beaucoup des attributs donnés pour *oiseau* sont plus distinctifs : 'becs', 'des plumes', 'ailes', 'pond des œufs', 'nidifie', 'vole', 'gazouille'. Mais les attributs de *scie* sont-ils vraiment moins distinctifs ? Dans la perspective référentialiste de Rosch, ces attributs devraient référer aux attributs réels des scies par l'intermédiaire de leur image mentale. Or, on ne coupe avec une scie comme avec un couteau, le manche (*handle*) d'une scie est différent de celui d'un couteau etc. De plus, si les manches de scie et les manches de couteau ont quelque chose en commun, où situera-t-on le niveau auquel ces manches se distinguent d'autres manches (par ex. *manche de pioche*, ou en anglais *door handle* ?). Bref, dans le calcul de la validité d'indice, on devrait (1) faire intervenir non seulement l'attribut verbal mais l'image de l'attribut réel qu'il dénote (sans quoi *scie* = *couteau*) ; (2) décider jusqu'à quel point un mot désignant un attribut nomme quelque chose de commun aux référents désignés par ce mot. Cette dernière décision est particulièrement complexe (elle devrait faire intervenir par exemple la gestuelle pour le cas des manches). On voit aussi que le français et l'anglais n'auront pas la même validité d'indice pour *scie*, puisqu'en français *manche* ne s'applique pas aux poignées de porte (contrairement à *handle*).

⁵³ La catégorie des jouets avait disparu depuis longtemps, fort heureusement, car il se peut que pour un citadin un canard en plastique (ou son dessin) ressemble plus au vrai canard que le canard réel. Les catégories restantes étaient les vêtements (avec un niveau de base : pantalon, chemise, chaussure, chaussette), les véhicules (voiture, camion, avion, motocyclette), les animaux (chat, chien, poisson, papillon), les meubles (chaise, table, canapé, lit) (Rosch et al. 1976 : 399).

1976 : 387-8) : un légume est une sorte de plante ou une sorte de nourriture, un ours peut être un jouet ou un animal, un couteau est un ustensile ou une arme etc.

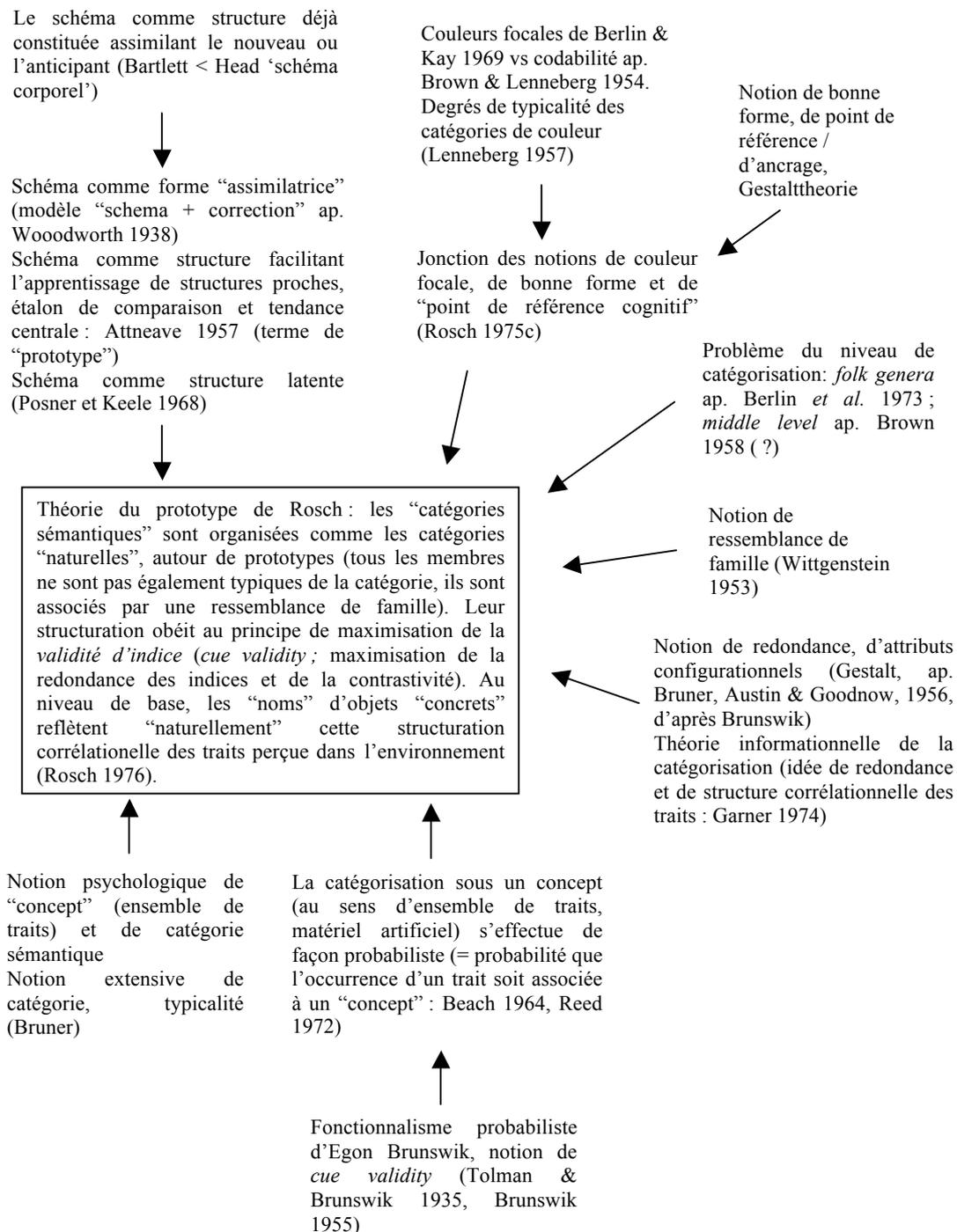
Cette épuration successive fait soupçonner qu'on recherche, avec le niveau de base, une correspondance aussi parfaite que possible entre une forme, une appartenance catégorielle, et un ensemble d'attributs. Autrement dit, toutes les catégories sont pensées sur le modèle de la taxonomie, alors que dans une langue, on a plutôt affaire à des paradigmes, c'est-à-dire à des lexèmes susceptibles d'être membres de plusieurs ensembles contrastifs (pour une critique en ce sens, voir Rastier 1991 : 187).

Pour Rosch, le point de vue taxonomique domine parce qu'il est gagé sur le monde réel, et notamment sur la discontinuité morphologique des *entités* au niveau de base. Retraduite dans le vocabulaire informationnel qui nous est maintenant familier, cette discontinuité morphologique signale une autre rupture, celle entre structures corrélationnelles "réelles" et groupements arbitraires : "nous pensons que le niveau de classification de base, le niveau premier auquel des "découpages" sont faits dans l'environnement, est un compromis entre ces deux niveaux; c'est le niveau le plus général et le plus global pour lequel des catégories sont encore capables de délimiter les structures corrélationnelles du monde réel" (Rosch 1976 : 243).⁵⁴

⁵⁴ La formation du niveau de base et celle du prototype obéissent donc au même principe, sauf que tout prototype n'est pas forcément fondé sur une corrélation "réelle" : "les prototypes de la catégorie semblent être simplement les membres de la catégorie qui reflètent le plus la structure corrélacionnelle de la catégorie dans son ensemble" Rosch (1976 : 248).

D.13. La constitution de la théorie du prototype (résumé)

Notre parcours des concepts de la théorie et la description de leur articulation sont maintenant achevés. Afin d'en avoir une vue synthétique, j'ai récapitulé sur le diagramme ci-dessous les principales influences qui se sont exercées sur Rosch.



Conclusion

La théorie du prototype est une articulation originale de concepts, d'outils descriptifs et de protocoles qui, avant que Rosch ne s'en saisisse, étaient disponibles et flottaient dans l'atmosphère de Harvard, même si ce sont Berlin et Kay qui ont donné la dernière chiquenaude. Il est possible que la libéralité dont Rosch a fait preuve dans ses emprunts n'ait pas peu contribué au succès de sa théorie. Nous avons vu en effet que ses recherches sur les couleurs pouvaient s'appuyer sur Berlin et Kay et bénéficier de leur légitimation progressive. L'extension de la théorie aux catégories "sémantiques", qui s'opéra d'abord dans une perspective holistique (voir la notion de *structure interne* de la catégorie), força l'impression d'une théorie cumulative, de plus en plus englobante.

La théorie de la formation des catégories exploita des idées venues du fonctionnalisme probabiliste auxquelles leur retraduction dans le vocabulaire de la théorie informationnelle (issue de Garner) donnait un tour rigoureux. La greffe prit d'autant mieux que les deux théories tendaient vers une forme de réalisme. Quant à la notion de niveau de base, elle convergea avec des idées défendues en anthropologie par Berlin. Le tout s'inscrivait harmonieusement dans un mouvement anti-relativiste ou anti-nominaliste dont les prémices se faisaient nettement sentir dans les années soixante. La théorie tendait ainsi vers une naturalisation généralisée des catégories d'objets concrets (Dubois et Resche-Rigon 1995). Cette naturalisation, d'abord appliquée aux couleurs et aux formes, s'appuyait pour les catégories "sémantiques" sur des processus censés être universels, comme la maximisation de la validité d'indice. Ainsi, au moins au niveau de base, la théorie semblait découvrir un lien causal médiatisant la structure de l'environnement et les catégories humaines.

Autre avantage, la notion de point de référence a permis à Rosch (1983) de récupérer celle de *représentativité*, mise en avant chez Tversky et Kahneman (1974, 1982) pour rendre compte de certaines anomalies dans les estimations de probabilité fournies par leurs sujets. Tversky et Kahneman observaient en effet que la représentativité, par exemple de stéréotypes sociaux, semblait conditionner la réponse à des estimations de probabilité, bien qu'elle fût logiquement non pertinente. L'articulation précise entre la représentativité et les principes roschiens de formation des catégories n'est pas très claire (Rosch 1983 : 82), mais les deux approches ont en commun de se fonder sur des cas saillants, non sur une appréhension "logique" du raisonnement et des catégories.

La théorie du prototype a eu un immense impact en psychologie et en sémantique cognitives.⁵⁵ Elle n'a pas tardé, toutefois, à être contestée sur le front de la psychologie. Comme elle semblait réduire les attributs, définitoires ou non, à des indices, certains ont voulu rétablir la distinction entre traits définitoires et non définitoires. Cette distinction pouvait s'autoriser d'études montrant que des phénomènes de typicalité étaient observables dans des catégories non floues (comme les nombres pairs par ex. ; Armstrong et al. 1983). Dans d'autres modèles, le nivellement des attributs définitoires et non définitoires ne valait qu'au stade d'identification d'un objet, et lors d'une seconde passe, si nécessaire, les attributs définitoires entraient en jeu pour déterminer l'appartenance catégorielle de l'objet (Smith, Shoben & Rips 1974). Des modèles de traitement alternatifs faisaient de la typicalité non une propriété des représentations mais un effet qui pouvait être expliqué sans passer par la notion réifiante de prototype. Par exemple, dans l'approche dite de l'*exemplaire*, la représentation d'un concept correspondait aux exemplaires partageant un nombre critique de propriétés avec d'autres exemplaires (Smith et Medin 1981). Le lien fait par Rosch entre typicalité et ressemblance de famille fut également contesté, à partir de l'idée que des objets partageant peu de propriétés "objectives" pouvaient néanmoins constituer une catégorie parce qu'ils

⁵⁵ Dubois et Resche-Rigon (1995) recensent dans une base de données américaines, entre 191 et 1995, 129 références contenant le mot-clé prototype, dont 88 renvoient à la notion de "catégorie naturelle" comme variable indépendante.

avaient en commun une action ou un but (par ex. ‘objets à vendre en brocante’ ou ‘choses susceptibles de tomber sur la tête’; Barsalou 1983). Enfin, le rôle du contexte était dans une certaine mesure réévalué. Des recherches montrèrent que la structure d’une catégorie pouvait être remodelée par le contexte. L’existence de structures catégorielles invariantes ou “normales” était ainsi mise en doute (Roth & Shoben 1983).

Avec la métaphore et les cadres (*frames*), le prototype devint une part essentielle de l’outillage théorique de la sémantique cognitive. A une date assez précoce, la théorie du prototype fut connue de plusieurs linguistes qui allaient jouer un rôle important dans l’avènement de la linguistique cognitive. Nous avons vu que Lakoff mentionna les recherches de Rosch dès 1973, dans son article sur les enclosures. L’impact de Rosch paraît ensuite avoir été amplifié par une présentation qu’elle fit à un *Linguistic Institute* de Berkeley en 1975. Fillmore en prit acte (Fillmore 1975), et la démonstration de Rosch poussa Lakoff à se détourner de la sémantique formelle (Ruiz de Mendoza Ibañez 1997), à une époque où il cherchait un modèle théorique. Les premiers travaux de linguistique prenant pour cadre d’analyse la notion de prototype suivirent quelques années plus tard (Coleman & Kay 1981, Brugman 1988 [1981], Bybee & Slobin 1982). Cette adaptation de la théorie du prototype à la linguistique se fit essentiellement dans le cadre de la sémantique lexicale, et particulièrement dans le but de traiter la polysémie (comme c’était le cas chez Brugman pour la préposition-particule *over*), en se fondant sur l’idée que les acceptions sont à un lexème ce qu’étaient pour Rosch les instances à la catégorie (voir Kleiber 1990 sur cette évolution). Dans ces travaux, le sens central est souvent concret ou spatial, ce qui renoue avec la tradition empiriste, qui a toujours accordé un statut privilégié aux significations ayant un rapport avec l’expérience sensible. Il faut noter aussi que cette adaptation n’a pas tout retenu de la théorie du prototype. La notion de validité d’indice, par exemple, n’a pas été importée en linguistique. Elle impliquait de prendre en compte les catégories contrastives, or la sémantique cognitive est souvent sémasiologique et part donc d’un mot. Toutefois, un certain nombre de travaux ont conduit à une vision plus subtile de certains aspects de la théorie du prototype, questionnant par exemple la possibilité de distinguer nettement différents niveaux de catégorisation et la vision taxonomique du lexique (Geeraerts 2010). Conséquence de cette absorption du prototype dans la linguistique cognitive, alors qu’initialement les liens entre la psychologie de la catégorisation et la linguistique étaient assez étroits, ils sont aujourd’hui plus distendus. En contrepartie, l’intégration de la théorie du prototype à la linguistique a permis de renouer avec des thèmes traditionnels de la sémantique lexicale (Geeraerts 1988, 1993), comme l’existence de frontières sémantiques floues, ou le changement de sens à partir d’acceptions “concrètes” censées être prototypiques.

Références

- Armstrong S.L., Gleitman L.R. & Gleitman H. (1983) What some concepts might not be, *Cognition* 13 : 263-308.
- Ash, Mitchell G. & Woodward, William R. (1989) *Psychology in twentieth century thought and society*, Cambridge, Cambridge University Press.
- Attneave, Fred (1954) Some informational aspects of visual perception, *Psychological Review* 61(3) : 183-193.
- Attneave, Fred (1957) Transfer of experience with a class-schema to identification-learning of patterns and shapes, *Journal of Experimental Psychology* 54(2) : 81-88.
- Baker, G.P. & Hacker, P.M.S. (1980) *Wittgenstein, understanding and meaning. Volume 1 of an analytical commentary of the philosophical investigations. Part 1.* Malden, Mass., Blackwell Publishing Ltd.
- Barsalou L.W. (1983) Ad hoc categories, *Memory and Cognition* 11(3) : 211-227.
- Beach, L. R. (1964) Cue probabilism and inference behavior, *Psychological Monographs: General and Applied* 78(5): 1-20.
- Berlin, Brent (1964) *The descriptive semantics of Tzeltal numeral classifiers*. Ph. D., Stanford University.
- Berlin, Brent (1970) A universalist-evolutionary approach in ethnographic semantics. In Fischer, A. (ed.) *Current directions in anthropology, Bulletins of the American Anthropological Association*, vol. 3, n° 3 , part 2, 3-18.
- Berlin, Brent (1978) Ethnobiological classification. In Rosch, Eleanor et Lloyd, Barbara (ed.), *Cognition and categorization*, Hillsdale, NJ, Lawrence Erlbaum Associates, 9-26.
- Berlin, Brent & Dennis E. Breedlove & Peter H. Raven (1973) General principles of classification and nomenclature in folk biology, *American Anthropologist*, 75 : 214-242.
- Berlin B. & Kay P. (1969) *Basic color terms : their universality and evolution*, Berkeley, University of California Press.
- Black, Max (1959) Linguistic Relativity: The Views of Benjamin Lee Whorf, *Philosophical Review* 68 (2) : 228-238.
- Brentano, Franz (1816) *Von der mannigfachen Bedeutung des Seienden nach Aristoteles*. Freiburg im Breisgau, Herdersche Verlagshandlung.
- Brown, Roger (1958a) *Words and things*, New York, The Free Press of Glencoe.
- Brown, Roger (1958b) How shall a thing be called?, *Psychological Review* 65(1) : 14-21.
- Brown, Roger (1965) *Social Psychology*. Collier Macmillan.
- Brown, Roger (1976) Reference. In memorial tribute to Eric Lenneberg, *Cognition* 4 : 125-153.
- Brown, Roger & Lenneberg, Eric (1954) A study in language and cognition, *Journal of Abnormal and Social Psychology*, 49 : 454-462.
- Brugman, Claudia (1988) *The story of over: Polysemy, semantics and the structure of the lexicon*. New York: Garland Press [reprise de (1981) *The story of over*. Berkeley, CA: UC-Berkeley master's thesis].
- Bruner, Jerome (1957) On perceptual readiness, *Psychological Review* 64(2) : 123-152.
- Bruner, Jerome, Goodnow, Jacqueline S. & Austin, George A. (1956) *A study of thinking*, New York, Wiley.
- Brunswik, Egon, Hammond, Kenneth R. & Stewart, Thomas R., eds (2001) *The essential Brunswik: beginnings, explications, applications*, Oxford, Oxford University Press.
- Brunswik, Egon (1943) Organismic achievement and environmental probability, *Psychological Review* 50(3) : 255-72.
- Brunswik, Egon (1955) Representative design and probabilistic theory in a functional psychology, *Psychological Review* 62(3) : 193-217.

- Brunswik, Egon (1956) *Perception and the Representative Design of Psychological Experiments*, 2d ed., rev. & enl., Berkeley, University of California Press.
- Brunswik, Egon & Kamiya, Joe (1953) Ecological cue-validity of “proximity” and of other Gestalt factors, *American Journal of Psychology* 66 : 20-32.
- Bybee Joan L. & Slobin Dan I. (1982) Rules and schemas in the development and use of the English past tense, *Language* 58(2): 265-289.
- Cauzinille-Marmèche Evelyne, Dubois Danièle & Mathieu Jacques (1990) Catégories et processus de catégorisation. In Netchine-Grynberg, Gaby (dir.), *Développement et fonctionnement cognitifs chez l'enfant*, Paris, Presses Universitaires de France, 93-119.
- Coleman, Linda & Kay, Paul (1981) Prototype semantics: The English word lie, *Language* 57(1) : 26-44.
- Collins, Allan M. & Quillian, Ross (1969) Retrieval time from semantic memory, *Journal of Verbal Learning and Verbal Behavior* 8(2) : 240–248.
- Cordier Françoise & Dubois Danièle (1981) Typicalité et représentation cognitive, *Cahiers de Psychologie Cognitive* 1 : 299-333.
- De Valois, R.L. & Jacobs, G.H. (1968) Primate color vision, *Science* 162 : 533-40.
- Dubois, Danièle & Cance, Caroline (2009) Mettre un terme aux couleurs de base. In Dubois, Danièle (dir.), *Le sentir et le dire*, Paris, L’Harmattan: 75-104.
- Dubois, Danièle & Resche-Rigon, Philippe (1995) De la “naturalité” des catégories sémantiques : des catégories “d’objets naturels” aux catégories lexicales, *Intellectica* 20 : 217-45.
- Ekman, Paul (1971) Universals and cultural differences in facial expressions of emotion. In J. K. Cole (Ed.), *Nebraska symposium on motivation*. Lincoln NE: University of Nebraska Press.
- Formigari, Lia (1988) *Language and Experience in 17th Century British Philosophy*, Amsterdam / Philadelphia, John Benjamins.
- Fortis, Jean-Michel (1994) Image mentale et représentation propositionnelle, *Intellectica* 19 : 253-305.
- Gabora, Liane, Rosch, Eleanor, & Aerts, Diederik (2008) Toward an ecological theory of concepts, *Ecological Psychology* 20(1), 84-116.
- Garner, Wendell R. (1962) *Uncertainty and structure as psychological concepts*, New York, Wiley.
- Garner, Wendell R. (1974) *The processing of information and structure*, New York, Wiley.
- Geraerts, Dirk (2010) *Theories of lexical semantics*, Oxford, Oxford University Press.
- Gibson, James J. (1979) *The ecological approach to visual perception*, Boston: Houghton Mifflin.
- Hammond, Kenneth R. & Stewart, Thomas R. (2001) Introduction. In Brunswik, Egon, Hammond, Kenneth R. & Stewart, Thomas R., (eds) (2001), 3-11.
- Head, Henry & Holmes, Gordon (1911) Sensory disturbances from cerebral lesions, *Brain* 34 : 102-254.
- Hebb, Donald (2002) [1949] *The organization of behavior*, Mahwah, NJ, Lawrence Erlbaum Associates.
- Heider, Fritz (1939) Environmental determinants in psychological theories, *Psychological Review* 46(4) : 383-410.
- Heider, Karl G. (1970) *The Dugum Dani: A Papuan Culture in the Highlands of West New Guinea*, Aldine Publishing.
- Hickerson, Nancy P. (1971) Review of *Basic Color Terms: Their Universality and Evolution* by Brent Berlin and Paul Kay, *International Journal of American Linguistics*, Vol. 37, No. 4, pp. 257-270.

- Hoijer, Harry (1954) The Sapir-Whorf hypothesis. In Hoijer, H. (ed.) *Language in culture : Proceedings of a conference on the interrelations of language and other aspects of culture*, Chicago, Chicago University Press, 92-105.
- James, William (1950 [1890]): *The Principles of Psychology*, Vol. 1. New York: Dover.
- Joseph, John, (2002) The sources of the Sapir-Whorf hypothesis. In Joseph, John, *From Whitney to Chomsky. Essays in the History of American Linguistics*, Amsterdam / Philadelphia, John Benjamins, 71-105.
- Kagan, Jerome (1965) Impulsive and reflective children. Significance of conceptual tempo. In J. Krumboltz (ed.), *Learning and the educational process*, Chicago, Rand McNally : 133-161.
- Kagan, Jerome (1966) Reflection-impulsivity: The generality and dynamics of conceptual tempo, *Journal of Abnormal Psychology* 71 : 17-24.
- Kay, Paul & McDaniel, Chad K. (1978) The linguistic significance of the meanings of basic color terms, *Language* 54(3) : 610-646.
- Kay, Paul, McDaniel, Chad K., Maffi, Luisa & Merrifield, William (1997) Color naming across languages. In Hardin, Clyde L. & Maffi, Luisa, *Color categories in thought and language*, Cambridge, Cambridge University Press, 21-56.
- Kleiber Georges (1990) *La sémantique du prototype*, Paris, Presses Universitaires de France.
- Kleiber, Georges (1994) Lexique et cognition : y a-t-il des termes de base ?, *Rivista di Linguistica* 6(2): 237-266.
- Kleiber, Georges et Riegel, Martin (1978) Les grammaires floues. In Martin, Robert (dir.), *La notion de recevabilité en linguistique*, Paris, Klincksieck : 67-123.
- Koerner, E.F. Konrad (2002) On the sources of the 'Sapir-Whorf hypothesis'. In Koerner, E.F.K., *Toward a history of American linguistics*, London, Routledge : 39-62.
- Lakoff, George (1973) Hedges : A study in meaning criteria and the logic of fuzzy concepts, *Journal of Philosophical Logic* 2 : 458-508.
- Lakoff, George, 1987. *Women, Fire and Dangerous Things : What Categories Reveal about the Mind*, Chicago, University of Chicago Press.
- Lantz, Delee & Steffle, Volney (1964) Language and cognitive revisited, *Journal of Abnormal and Social Psychology* 69 : 472-481.
- Larsen, Steen F. & Berntsen, Dorthe (2000) Bartlett's trilogy of memory. Reconstructing the concept of attitude. In A. Saito (ed.), *Beyond Bartlett. Interface of cognitive and sociocultural science*, Psychology Press : 90-114.
- Latour, Bruno (1993) Le topofil de Boa-Vista : la référence scientifique : montage photo-philosophique, *Raisons pratiques* 4 : 187-216.
- Leary, David E. (1989) From act psychology to probabilistic functionalism : The place of Egon Brunswik in the history of psychology. In Ash & Woodward (eds), 115-142.
- Lee, Penny (1996) *The Whorf Theory Complex. A critical reconstruction*, Amsterdam & Philadelphia, John Benjamins.
- Lenneberg, Eric H. (1953) Cognition in Ethnolinguistics, *Language* 29(4) : 463-471
- Lenneberg, Eric H. (1957) A probabilistic approach to language learning, *Behavioral Sciences* 2, 1-12.
- Lenneberg, Eric H. (1962) The relationship of language to the formation of concepts, *Synthese* 14 : 103-9.
- Lenneberg, Eric H. (1967). *Biological foundations of language*, New York: John Wiley & Sons.
- Lenneberg, Eric H. & Roberts, John M. (1956) The language of experience: a study in methodology, *International Journal of American Linguistics* 22(2), Memoir 13.
- Lucy, John A., 1992. *Language Diversity and Thought*, Cambridge, Cambridge University Press.

- McDaniel, Chad E. (1972) *Hue perception and hue naming*, BA dissertation, Harvard College.
- Mandler, George (2007) *A history of modern experimental psychology : from James and Wundt to cognitive science*, Cambridge, Mass., MIT Press.
- Peirce, Charles Sanders (1878) How to make our ideas clear, *Popular Science Monthly* 12 : 286-302.
- Posner, Michael I., & Keele, Steven W. (1968) On the genesis of abstract ideas, *Journal of Experimental Psychology* 77(3) : 353-363.
- Rastier, François (1991) *Sémantique et Recherches Cognitives*, Paris, Presses Universitaires de France.
- Reed, Stephen K. (1972) Pattern recognition and categorization. *Cognitive Psychology* 3, 383-407.
- Ritchie, Benbow F. (1964) *Edward Chace Tolman*, Biographical Memoir, National Academy of Sciences, 291-324.
- Rosch Heider, Eleanor R. (1969) *Style and effectiveness of childrens' verbal communications within and between social classes*, doctoral dissertation, Harvard University.
- Rosch Heider, Eleanor (1971a) Information processing and the modification of an "impulsive conceptual tempo", *Child Development* 42 : 1276-1281.
- Rosch Heider, Eleanor (1971b) "Focal" color areas and the development of color names, *Developmental Psychology* 4(3) : 447-55.
- Rosch Heider, Eleanor (1971c) On the Internal Structure of Perceptual and Semantic Categories, unpublished paper, Psychology Dept., U. of California, Berkeley.
- Rosch Heider, Eleanor (1972a). Universals in color naming and memory, *Journal of Experimental Psychology* 93 : 10-20.
- Rosch Heider, Eleanor (1972b) Probabilities, Sampling, and Ethnographic Method: The Case of Dani Colour Names, *Man*, New Series, 7(3) : 448-466.
- Rosch Heider, Eleanor (1973a) Natural categories, *Cognitive Psychology* 4 : 328-50.
- Rosch, Eleanor H. (1973b) On the internal structure of perceptual and semantic categories. In Timothy E. Moore (ed.), *Cognitive development and the acquisition of language*, New York, Academic Press, 111-144.
- Rosch, Eleanor (1975a) Cognitive reference points, *Cognitive Psychology* 7 : 532-47.
- Rosch Eleanor (1975b) Cognitive representations of semantic categories, *Journal of Experimental Psychology : General*, 104(3) : 192-233.
- Rosch, Eleanor (1975c) Universals and cultural specifics in human categorisation. In Brislin, Richard W., Bohner, Salomon & Lonner, Walter J. (eds.), *Cross-cultural Perspectives of Learning*. New York, 177-206.
- Rosch, Eleanor (1975d) The nature of mental codes for color categories, *Journal of Experimental Psychology : Human Perception and Performance* 75(1) : 303-22.
- Rosch, Eleanor (1976) Classification d'objets du monde réel: origines et représentations dans la cognition, *Bulletin de Psychologie*, numéro spécial "La mémoire sémantique" : 242-250.
- Rosch Eleanor (1977) Human categorization. In Warren, N (ed.), *Advances in cross-cultural psychology*, Londres, Academic Press : 1-71.
- Rosch Eleanor (1978) Principles of categorization. In Rosch, Eleanor & Lloyd, Barbara B. (ed.), *Cognition and categorization*, Hillsdale, NJ, Lawrence Erlbaum Associates, 27-48.
- Rosch Eleanor (1983) Prototype classification and logical classification: The two systems. In Scholnick, E., *New Trends in Cognitive Representation: Challenges to Piaget's Theory*, Hillsdale, NJ: Lawrence Erlbaum Associates: 73-86.
- Rosch Eleanor (1987) Wittgenstein and categorization research in cognitive psychology. In M. Chapman, R. A. Dixon (ed.), *Meaning and the Growth of Understanding*, Berlin, Springer : 151-166.

- Rosch, Eleanor (1999a) Reclaiming concepts. In R. Nunez & W.J. Freeman (Eds.) (1999). *Reclaiming cognition: The primacy of action, intention and emotion*. Thorverton, Eng.: Imprint Academic. Published simultaneously as a special issue of *The Journal of Consciousness Studies*, 6, No. 11-12, 61-77.
- Rosch, Eleanor (1999b) Primary knowing: When perception happens from the whole field. Disponible à <http://www.dialogonleadership.org/interviews/Rosch.shtml>
- Rosch, Eleanor & Lloyd, Barbara B. (1978) *Cognition and Categorization*, Hillsdale, N.J.: Lawrence Erlbaum.
- Rosch Eleanor & Mervis C.B. (1975) Family resemblances : studies in the internal structure of categories, *Cognitive Psychology* 7 : 573-605.
- Rosch Eleanor, Mervis C.B., Gray W.D., Johnson D.M. & Boyes-Braem P. (1976) Basic objects in natural categories, *Cognitive Psychology* 8 : 382-439.
- Rosch Heider, Eleanor & Olivier, Donald C. (1972) The Structure of the Color Space in Naming and Memory for Two Languages, *Cognitive Psychology* 3 : 337-354.
- Roth, Emilie M. & Shoben, Edward J. (1983) The effect of context on the structure of categories, *Cognitive Psychology* 15 : 346-78.
- Ruiz de Mendoza Ibáñez, Francisco José (1997) An interview with George Lakoff, *Cuadernos de Filología Inglesa* 6/2 : 33-52.
- Sahlins Marshall D. & Service Elman R. (1960) *Evolution and culture*. Ann Arbor, University of Michigan Press.
- Sapir, Edward (1949) [1924] The grammarian and his language. In Sapir, E., *Selected writings of Edward Sapir in language, culture and personality*, edietd by David G. Mandelbaum, Berkeley – Los Angeles, University of California Press, 150-9 [repris de The grammarian and his language, *American Mercury* 1 : 149-155].
- Saunders, Barbara (1992) *The Invention of Basic Colour Terms*, Utrecht: ISOR.
- Saunders, Barbara (1995) Disinterring basic color terms. A study in the mystique of cognitivism, *History of the Human Sciences* 8(4) : p.19-38.
- Saunders, Barbara (2000) Revisiting basic color terms, *The Journal of the Royal Anthropological Institute*, Volume 6, Issue 1 : 81-99.
- Smith, Edward E. & Medin, Douglas L. (1981) *Categories and concepts*, Cambridge, Mass., Harvard University Press.
- Smith, Edward E., Shoben, Edward J. & Rips, Lance J. (1974) Structures and process in semantic memory, *Psychological Review* 81:214-241
- Sorenson, E.G. (1976) *The edge of the forest*, Washington, Smithsonian Insitution Press.
- Stefflre, Volney, Castillo Vales, Victor & Morley, Linda (1966) Language and cognition in Yucatan: A cross-cultural replication, *Journal of Personality and Social Psychology* 4: 112-115.
- Tolman, Edward C. (1933) Gestalt and sign-gestalt, *Psychological Review* 40 : 391-411.
- Tolman, Edward C., & Brunswik, Egon (1935). The organism and the causal texture of the environment, *Psychological Review* 42(1) : 43-77.
- Tolman, Edward C, Ritchie, Benbow F. & Kalish, D. (1946) Studies in spatial learning II : Place learning versus response learning, *Journal of Experimental Psychology* 36 : 221-9.
- Tulving, Endel (1968) Theoretical issues in free recall. In T.R. Dixon & D.L. Horton, (eds.), *Verbal behavior and general behavior theory*, Englewood Cliffs, Prentice Hall : 2-36.
- Tversky, Amos and Kahneman, Daniel (1974). Judgement under uncertainty: Heuristics and biases. *Science* 185 : 1124-1130.
- Tversky, Amos and Kahneman, Daniel (1982) Judgments of and by representativeness. In D. Kahneman, P. Slovic & A. Tversky (eds.), *Judgment under uncertainty: Heuristics and biases*, Cambridge, UK, Cambridge University Press.
- Wertheimer Max (1912) Experimentelle Studien über das Sehen von Bewegung, *Zeitschrift*

- für Psychologie* 61 : 161-265.
- Wertheimer, Max (1938) Numbers and numerical concepts in primitive peoples. In W. D. Ellis (ed.), *A source book of Gestalt psychology*, New York, Harcourt, Brace, & Co, 265-273.
- Whorf, Benjamin Lee (1950) *Four articles on metalinguistics*, Washington, Foreign Service Institute.
- Whorf, Benjamin Lee (1956) [1940], Gestalt Technique of Stem Composition in Shawnee. In Voegelin, C.F., *Shawnee Stems and the Jacob P. Dunn Miami Dictionary*, Indianapolis, Indiana Historical Society [repris dans Whorf, B., 1956, *Language, Thought, and Reality*, 160-72].
- Whorf, Benjamin Lee (1956) [1941a] The relation of habitual thought and behavior to language. In Spier, Leslie (ed.) *Language, culture and personality, essays in memory of Edward Sapir*, Menasha (Wisconsin), Sapir Memorial Publication Fund [repris dans Whorf, B., 1956, *Language, Thought, and Reality*, 134-159].
- Whorf, Benjamin Lee (1956) [1941b], Languages and Logic, *Technology Review* 43 : 250-52, 266, 268, 272 [repris dans Whorf, B., 1956, *Language, Thought, and Reality*, 233-45].
- Whorf, Benjamin Lee (1956) *Language, Thought, and Reality*, Cambridge (Mass.), The M.I.T. Press.
- Wittgenstein, Ludwig (1961) [1953] *Tractatus logico-philosophicus suivi de Investigations philosophiques*, Paris, Gallimard.
- Wittgenstein, Ludwig (1980) [1969] *Grammaire philosophique*, Paris, Gallimard.
- Zadeh, Lofti (1965) Fuzzy sets, *Information and Control* 8, 338-53.
- Zipf, George K. (1935) *The psycho-biology of language*, Boston, Houghton Mifflin.