

LA THÉORIE DU PROTOTYPE: DE LA PSYCHOLOGIE EXPÉRIMENTALE À LA LINGUISTIQUE

Jean-Michel Fortis
Université Paris Diderot
UMR 7597 “Histoire des Théories Linguistiques”

Le contexte d'émergence

- 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 (= le groupement de couleurs dissemblables dans une catégorie, par ex. "vert", et la dénomination des couleurs).
- Cette théorie naît au sein d'un débat théorique précis, celui entourant l'hypothèse dite de Sapir-Whorf.

Le contexte d'émergence

OBJECTIFS:

- Montrer comment le principe de relativité de Whorf est devenu une hypothèse (“l’hypothèse de Sapir-Whorf”) qu’on peut “opérationnaliser” (rendre testable).
- Montrer comment Rosch prolonge et infléchit les recherches qui l’ont précédée, en distinguant 2 aspects:
 - 1) sa prise de position sur l’hypothèse de Sapir-Whorf.
 - 2) l’outil descriptif que constitue sa théorie du prototype.
- Décrire chronologiquement le développement de cette théorie, qui s’est construite par agrégation successive de concepts et d’outils d’origines diverses.
- Montrer pourquoi et comment cette théorie a été importée par les linguistes (elle est aujourd’hui un pilier de la linguistique cognitive).

N.B.: Cet exposé reprend un article publié en 2010 dans la revue en ligne **Corela**.

Le contexte d'émergence

- 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, principalement dans les travaux d'Eric Lenneberg, Roger Brown et John Roberts.
- But : 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.

Whorf et le “principe de relativité”

“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).

Quelques précautions sur Whorf

- Whorf ne dit pas que la langue contraint intégralement la perception ou l'expérience intime (il y a des universaux cognitifs / expérientiels: asymétrie figure / fond, conscience de l'avant / après temporel).
- Whorf ne dit pas que les langues sont des univers qui ne peuvent communiquer.
- Whorf ne dit pas que la langue détermine la culture (il dit qu'elles sont en "consonance").
- Whorf s'intéresse d'abord aux cryptotypes (catégories linguistiques latentes qui déterminent des phénomènes de surface), et aux "linguistic patterns" qui contraignent notre conceptualisation du temps, de l'événement etc. (au-delà de la couche perceptive / expérientielle universelle).
- NB: "relativité" est une allusion à Einstein (et Jung) déjà présente chez Sapir [1949] (1924): 159, "*relativity of the form of thought*".

Quelques précautions sur Whorf

- Le principe de relativité est dans un mouvement historique, celui du *linguistic turn*, dont un des aspects est la défiance à l'égard du langage (*the metaphysical garbage view*, in Joseph 2002).
- La linguistique a une valeur thérapeutique (de décentrement; en accord avec *the magic key view* in Joseph 2002). Sapir semble joindre ces deux visions. La science moderne a aussi cette valeur aux yeux de Whorf.
- La linguistique participe au questionnement des cadres de pensée hérités (par ex. substantialisme vs relations, principe de contradiction etc.) qu'on trouve par ex. chez Cassirer (une réf. implicite de Rosch), et chez des auteurs plus marginaux comme Korzybski (connu de Whorf).

Du principe à l'hypothèse

- Selon Koerner (2002 : 39), Hoijer (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).
- La notion d'hypothèse convient mieux aux recherches présentées ici, puisqu'il s'agit de la soumettre à l'épreuve expérimentale.

Du principe à l'hypothèse

- C'est Eric Lenneberg qui (re)lance le débat sur la catégorisation des couleurs. (relance car cf. Saunders, 1992 et 2007, *The debate on colour naming in 19th century German philology*; voir aussi Lucy 1992, *Language diversity and thought*).
- Sa version de l'hyp. de Sapir-Whorf:
“La structure d'une langue donnée affecte-t-elle les pensées (ou la capacité de penser [*thought potential*]), la mémoire, la perception, la capacité d'apprentissage de ceux qui parlent cette langue?” (Lenneberg 1953 : 463).
- Paradoxalement, cette version est dirigée *contre* l'approche de Whorf.

Lenneberg critique de Whorf

- Whorf est trop littéraliste: il glose littéralement des énoncés de langues amérindiennes pour les rendre étranges.
- Whorf néglige le fait que la construction de la signification n'est que partiellement fondée sur les signes linguistiques.
- 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.

Lenneberg: opérationnaliser l'hypothèse de S-W

- Les couleurs sont un domaine référentiel qui joue le rôle d'un pur donné indépendant de ces contextes culturel, social, individuel, et dont le découpage par la langue ne laisse pas de latitudes au locuteur (contrairement à *empty* chez Whorf).
- La relation entre ce “pur donné” et la langue est appelée *codification*.
- Mettre en évidence un effet whorfien = établir une relation entre la codification et un comportement testé en laboratoire.
- Idée simpliste de base: un “terme” de couleur dénomme une catégorie, une catégorie est un groupement [*grouping*] de couleurs.
- La naïveté est totale. Brown (1976: 125), le collègue de Lenneberg, nous apprend incidemment que c'est Chomsky qui les invitera à distinguer *meaning* et *reference*

Lenneberg: opérationnaliser l'hypothèse de S-W

- Notion psychologique de catégorie: classe de stimuli traités comme équivalents du point de vue de certains de leurs attributs.
- Par ex. chez Bruner et al. (1956): on soumet à des sujets des cartes portant un motif variant sur plusieurs dimensions (par ex. forme et couleur) et on leur demande de trouver le “concept” (= catégorie), c’est à dire le principe de classement déterminé à l’avance par l’expérimentateur (par exemple ‘carré et rouge’).
- Les attributs de la couleur sont: teinte, saturation et brillance.
- Ces attributs ne sont pas mis en doute: le protocole n’impose pas de quitter l’environnement culturel familier; ces attributs sont des grandeurs mesurables; il existe des matériels standardisés (des nuanciers). Pour une critique, Saunders (1992) et nombreux articles, Dubois & Cance (2009).

Brown & Lenneberg (1954): l'expérience intraculturelle

- Objectif: mise en relation d'une variable dépendante, le rappel d'une couleur, avec un indice dit de "codabilité" de la couleur.
- **Protocole:** les couleurs sélectionnées 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 dépend de : la longueur du nom (plus il est court, plus la couleur est censée être codable) ; le temps mis pour la dénommer ; le degré d'accord inter- et intra-sujetif 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 non.

Brown & Lenneberg (1954): l'expérience intraculturelle

- L'inspiration vient ici de Zipf (1935), qui mettait en relation la longueur d'un mot avec sa fréquence. Lenneberg pense que la fréquence d'un mot est un bon index de la disponibilité de la catégorie à laquelle il renvoie.
- Les résultats confirment l'hypothèse "relativiste" de départ.

Lenneberg & Roberts (1956): l'expérience interculturelle

- Problème: comment savoir si les noms de couleur sont la cause d'une meilleure reconnaissance ou sont seulement l'expression d'une meilleure discriminabilité?
- La solution est de faire varier la langue.
- Protocole: sujets anglophones et zunis (ethnie du Nouveau Mexique, langue 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 (320 pastilles, 40 teintes x 8 niveaux de brillance) à saturation maximale.

Lenneberg & Roberts (1956): l'expérience interculturelle

- Le zuni ne distingue pas le jaune et l'orange. De fait, ils ne sont pas reconnus non plus dans la tâche de mémorisation.
- Certaines de couleurs les plus typiques sont identiques de langue à langue, même si les frontières des catégories diffèrent (mais pas de conclusion anti-relativiste!).
- "...la meilleure caractérisation des concepts, est d'en faire des zones de typicalité croissante et décroissante sur un continuum de stimulation" (Lenneberg 1957: 2).
- Des aspects essentiels anticipant sur Rosch sont en place: le protocole interculturel, les couleurs focales, la typicalité.

Lantz & Steffire (1964): la précision communicationnelle

- Surmonter les contradictions en proposant un nouvel index remplaçant la codabilité: la Précision Communicationnelle (*communication accuracy*).
- La P. C. = efficacité avec laquelle la description d'une couleur par un locuteur permet à un autre sujet d'identifier ladite couleur dans un nuancier.
- Le nouvel index prédit à la fois les résultats de Brown et Lenneberg et ceux de Burnham et Clark.
- MAIS cet index n'emporte pas l'adhésion. Brown (1976: 145) explique pourquoi: il porte sur l'efficacité individuelle à communiquer, pas sur la langue... L'index de P. C. sera également rejeté par Rosch.

Le retournement universaliste

- Burnham et Clark (1955) contredisent la thèse d'une corrélation positive entre codabilité et reconnaissance, en utilisant cette fois des couleurs moyennement saturées.
- Lenneberg, pour des raisons qui tenaient à ses recherches sur les rapports entre maturation cognitive et le développement du langage chez l'enfant, et à une mise en question de l'empirisme (en partie inspirée par Quine), s'éloignait de son relativisme initial pour adopter une position beaucoup plus biologisante, qualifiée de "néokantienne" (Lenneberg 1962).
- Le contexte épistémologique est en train de changer (sciences cognitives, psychologie du raisonnement et de la catégorisation, modèle computationnel, universalisme chomskyen, approches référentialistes en sémantique et anthropologie, réseaux d'A.I. etc.)

Berlin & Kay (1969): les *basic color terms*

- Enquête sur 98 langues. Pour 20 langues, les données sont de première main (mais souvent recueillies auprès d'un seul informateur, de plus connaissant l'anglais) ; les matériaux des 78 autres langues sont glanés dans des sources littéraires diverses, parfois anciennes (et donc B & K connaissent les travaux du 19^{ème} s. sur la dénomination des couleurs, notamment ceux de Magnus, 1880, *Untersuchungen über den Farbensinn der Naturvölker*).
- 1^{ère} phase: 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.
- Protocole: 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.

Berlin & Kay (1969): les *basic color terms*

- Les critères: 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 Hickerson 1971 et Saunders 1992 pour un examen critique). Le mot *terme* signale que Berlin & Kay ne s'intéressent pas à la catégorie lexicale des désignations qu'ils obtiennent.

Berlin & Kay (1969): les *basic color terms*

- Comme Lenneberg et Roberts, B & K observent une 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).
- 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.
- L'évolution est la suivante (Berlin & Kay 1969 : 4):
blanc / noir > rouge > vert *ou* jaune > vert et jaune > bleu > marron > violet [*purple*], rose [*pink*], orange, gris (ou une combinaison de ces termes)

Berlin & Kay (1969): les *basic color terms*

- Cette évolution refléterait un développement économique et technologique, les systèmes les moins différenciés étant propres aux peuples “primitifs” (sic), tandis que les systèmes plus complexes (dont, évidemment, l’anglais) seraient caractéristiques des “nations les plus civilisées du monde” (Berlin 1970: 14).
- Accusée de redécrire les données disponibles pour les conformer à l’hypothèse de départ (cf. Hickerson 1971), l’étude de Berlin et Kay n’en a pas moins reçu un accueil globalement favorable (Saunders 1992). Prestige amplifié par le succès de la théorie du prototype et sa (presque) réduction à un modèle neurophysiologique (Kay & McDaniel 1978)*, deux développements postérieurs qui ont contribué à la légitimer (mais l’universalité des couleurs focales reste en débat).

Dès 1972, McDaniel montrait que la réponse neuronale à des couleurs atteignait un pic pour certaines couleurs focales (jaune, bleu et vert).

K & McD. (après De Valois et al. 1966, 1968): 1 couleur = 1 **ensemble flou** de réponses neuronales à des longueurs d’onde. Pensent que la séquence de B & K correspond à la différenciation progressive de super-couleurs (*composites*, union d’ensem. flous) en couleurs *primaires* (n, blc, ble, v, j, r = réponses neuronales fondamentales) et couleurs *dérivées* (orange, déf. modifiée de l’intersection de r et j).

Rosch: les débuts

- Formation initiale de philosophe (mémoire sur Wittgenstein), puis part étudier la psychologie à Harvard.
- Thèse (1969) dirigée par Roger Brown (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). Influence théorique de Jerome Kagan (psy. du développement).
- En 1969, accompagne son mari, l'anthropologue Karl Heider (fils du psy. autrichien Fritz Heider), chez les Danis (ouest de la Nouvelle Guinée). Etudie les interactions mère-enfant et surtout la catégorisation des couleurs et des formes. Ekman, croisé à Berkeley, propose aux Heider d'expérimenter sur sa théorie des expressions faciales d'émotions universellement reconnues (joie, tristesse, dégoût, surprise, crainte, et, mais avec des restrictions, colère; Ekman 1971). Rosch a des doutes sur cet universalisme, mais ses réticences seront surmontées....

Rosch: les débuts

- Premier article sur les enfants, suite à Berlin & Kay: les enfants choisissent préférentiellement de désigner les couleurs focales de B & K lorsqu'on leur demande de montrer la couleur correspondant à un nom, ou simplement de choisir une couleur parmi d'autres. Il leur est plus facile de retrouver une couleur donnée parmi un choix d'échantillons si la couleur testée est focale (Rosch Heider 1971b).
- On retrouve le problème de Brown & Lenneberg: on ne peut pas exclure un biais linguistique.

>> d'où l'intérêt des Danis: ils ont "essentiellement" deux termes, associés surtout à la brillance (l'un pour les couleurs "chaudes", l'autre pour les couleurs "froides"), pas de noms pour les couleurs focales de Berlin et Kay. S'il y a un meilleur traitement des couleurs focales, la raison doit être cognitive.

Recherches sur les Danis

- Problème: si on lit Karl Heider et des articles moins connus de Rosch (1972b, 1972c), on voit que le lexique des couleurs dani n'est pas si appauvri:
“...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 contourne le problème en choisissant les sujets qui n'emploient que les deux “termes de base”, assimilés à blanc / noir de B & K.

Recherches sur les Danis

- Les résultats obtenus aux tâches de mémorisation par des anglophones et des Danis sont plus proches que ceux obtenus entre l'épreuve de dénomination et l'épreuve de mémorisation dans une même culture (Rosch & Olivier 1972)
>> le statut privilégié des couleurs focales a des raisons cognitives (problème: le terme pour les couleurs chaudes a 2 foci, un rouge sombre et un rose pâle).
- Tâche d'apprentissage à partir de couleurs et de formes: faire apprendre aux Danis à former des catégories "naturelles" (construites autour de couleurs focales ou de bonnes formes) ou "non-naturelles" (construites autour de couleurs internominales, par ex. vert-jaune, ou à partir de bonnes formes bruitées, par ex. un triangle dont un côté est interrompu). Dans chaque catégorie de formes définie par Rosch, le prototype est une sorte de module générateur.
>> les catégories "naturelles" sont acquises plus facilement.

Le “prototype”

- Apparition du terme de “prototype” (Rosch 1973a) pour couvrir les couleurs focales et les bonnes formes.

- Pourquoi ce terme?

A l’origine, la notion de prototype a d’étroites affinités avec celle, en psychologie, de *schéma*.

Grossièrement, la notion de schéma (qui remonte au schème de Kant) aboutit au 20^{ème} s. à celles de *schéma superficiel* / *schéma postural* chez le neurologue anglais Henry Head.

La notion de schéma

- Le schéma postural est 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).

La notion de schéma

- Le psychologue anglais Bartlett reprend cette notion à Head dans une célèbre monographie sur la mémorisation (*Remembering*, 1932).

La mémorisation: une reconstruction active, par laquelle un stimulus est assimilé à une forme familière qui, lui préexistant, est aussi appelée *performed scheme*. Assimilation et déformation sont des processus affectant tout matériau mémoriel, et ne sont pas donc pas restreintes aux modalités perceptives (ou proprioceptives / kinesthésiques).

Théorie complexe, “clinique“, anti-Ebbinghaus (composantes culturelle, “tempéramentale“, affective), simplifiée ensuite.

La notion de schéma

- Woodworth (1938): la mémorisation de formes est 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). Ce schéma deviendra une structure en fonction de laquelle les déviations sont reconnues, et que les déviations modifient à leur tour constamment.

La notion de schéma

- En partie par Hebb, la notion se transmet à Attneave (1957) avec l'objectif 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" (but ultime: s'opposer à une théorie purement associationniste de l'apprentissage).
 - Chez Rosch, le schéma d'une bonne forme est générateur; le prototype permet la reconnaissance; et le prototype est aussi un stéréotype (forme la mieux reconnue), avec une dialectique entre reconnu / qui permet de reconnaître
- >> traits partagés avec le schéma-prototype.

La notion de schéma

- Exemples de matériel utilisé par Attneave (1957):

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



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

La notion de schéma

- A la suite d'Attneave, un certain nombre d'études reprendront l'idée du prototype générateur d'une série. Certaines expériences 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).

Les catégories “sémantiques” et la typicalité

- Les couleurs focales et les bonnes formes permettent-elles de penser les autres catégories? Y a-t-il des prototypes ailleurs?
- Rosch (1973b) pense aux catégories “sémantiques”, c-à-d aux ensembles de termes (généralement des noms) référant aux membres d'une classe d'entités et qui n'appartiennent pas à une dimension perceptive. Elle dispose d'un outil récent: les listes de Battig & Montague (1969): 56 catégories avec leurs membres (élicités) classés par fréquence et rang de mention.
- Catégories choisies par Rosch: les fruits, les sciences, les sports, les oiseaux, les véhicules, les maladies, les légumes, les crimes.
- L'expérience est un ballon d'essai: voir si, aux yeux des sujets, la notion de typicalité fait sens lorsqu'elle est appliquée aux catégories “sémantiques”.

Les catégories “sémantiques” et la typicalité

“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” 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. (...) 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” etc.” (1973b : 131-2)

[suivent les instructions pratiques]

Les catégories “sémantiques” et la typicalité

- Résultats: corrélation entre fréquence et typicalité (la pomme est un fruit jugé plus typique que l'olive, et *pomme* est plus fréquent qu'*olive*).
- Fréquence plutôt que typicalité? Rosch pense que non:
 - (1) Est-ce la fréquence qui conditionne la saillance des couleurs focales et des bonnes formes ?
 - (2) Elle met sa foi dans l'idée que les instructions, par leur précision, ont déterminé les sujets à fonctionner en mode “catégorie roschienne”, et non en mode “fréquence”.

Les catégories “sémantiques” et la typicalité

- Autre test classique: amorçage (*priming*; Rosch 1975b, 1975d):

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.

On fait de même avec les catégories sémantiques (par ex. avec “oiseau” comme amorce).

Les catégories “sémantiques” et la typicalité

- 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” (*internal structure*) et des contours flous. Les membres des catégories “sémantiques” ne sont pas encore des complexes d’attributs.

>> image gestaltiste de la catégorie

Prototype et Gestalt

- Ce rôle de la Gestalt est confirmé par la comparaison établie entre le prototype et les notions de *point de référence* (*reference point*) et *point d'ancrage* (*anchoring point* ; Rosch 1975a).
- Comparaison suggérée par Wertheimer (1938), qui pointait l'asymétrie de certains jugements de comparaison: 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° .

Prototype et Gestalt

- Rosch reprend Wertheimer avec comme matériel des lignes, des nombres et des couleurs.
- Les sujets doivent compléter 'x est quasiment y' au moyen des stimuli *physiques* et ajuster physiquement x p/r à y (prototype fixe / instance mobile, ou l'inverse; elle attend que x soit plus près de y quand y est fixe et est le prototype).
- Les contextes linguistiques sont soigneusement choisis (pour éviter au maximum les problèmes; Fortis 2010 pour une discussion).

>> idée d'une isomorphie entre langage et structuration perceptive et d'une nature "indexicale" des noms de couleur, de nombres, de lignes. Le test linguistique est issu d'un allée-retour entre Lakoff et Rosch.

Lakoff et le prototype

- Lakoff venait 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 1965 (Lakoff 1973).
- Il s'appuyait lui-même sur une présentation où Rosch (1971c) avait exposé des résultats préliminaires sur les degrés de typicalité de membres de catégories "sémantiques".
- Dans l'étude de Lakoff, les contextes du type 'x est quasiment y', ou 'x est un vrai y' étaient analysés comme des manifestations linguistiques de l'existence de degrés d'appartenance catégorielle. Lakoff appelait *enclosures* (*hedges*) ces contextes.
- Il 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*.

Lakoff et le prototype

- L'étude de Lakoff sur les enclosures est révélatrice d'une perspective encore logiciste, pré-linguistique cognitive, que signale 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.
- A ma connaissance, il s'agit de la première prise en compte de la théorie roschienne du prototype en linguistique.

NB: idées similaires (l'appartenance à une catégorie est floue) chez Labov (1973) *The boundaries of words and their meanings*, dans une perspective contextuelle, variationniste de la dénomination (travaux initiés en 1964).

Lakoff et le prototype

- A cette époque, Lakoff, en partie sous l'influence de John Ross, s'intéressait aux phénomènes de gradience non seulement dans le lexique proprement dit mais dans la grammaire (*Nouniness* date de 1973).
- La 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). Par ex. la faculté pour un adv. de se préposer dépend d'une hiérarchie des adv. et d'une hiérarchie de fermeture des îlots:
 - a. Tomorrow it's likely that Sam will leave town.
 - a'. ? In Berkeley it's likely that there will be a riot. [adv. de temps > adv. de lieu]
 - b. Tomorrow I think John will leave.
 - b'. * Tomorrow I'm surprised that John will be in his office. [*I'm surprised that* est un îlot plus fermé que *think that*]

Le prototype, cosa mentale

- Qu'est-ce qu'un prototype de catégorie "sémantique"? Un membre très typique, comme 'pomme' pour 'fruit'? Ou bien n'y a-t-il que des jugements de typicalité?
- Si le prototype est une représentation, de quelle nature est-elle? Imagée, verbale, amodale?
- La notion de prototype est-elle identique, qu'on parle de couleurs, de formes, de la dénomination *fruit*?
- Rosch oscille entre deux conceptions, réifiante et non-réifiante (Rosch 1977, 1978, 1999) et la nature mentale du prototype ne paraît pas très claire (plutôt une "image" dans Rosch 1975b).

La formation des catégories sémantiques

- Les catégories sémantiques \neq couleurs, formes.
- Elles n'ont a priori pas de base neurophysiologique.
- Elles doivent être acquises par apprentissage à partir des membres de la catégorie et d'un processus de traitement de l'information (Rosch & Mervis 1975: 574).

>> Comment et pourquoi s'opère cette formation des catégories sémantiques?

Qu'est-ce qui fait tenir ensemble les membres d'une même catégorie sémantique?

Solidarité de la catégorie et ressemblance de famille

- S'inspirant de Wittgenstein, Rosch considère 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*.
- Wittgenstein (§66-67 des *Investigations Philosophiques*, ex. de la catégorie 'jeux'):

“Je ne puis caractériser mieux ces ressemblances que par les mots *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.”

La notion de ressemblance de famille

- Pourrait être un écho de Brentano (*Von der mannigfachen Bedeutung des Seienden nach Aristoteles*, 1816: 60): faisant allusion à la question classique évoquée par Aristote dans *Métaphysique* Γ, 2, qui est de décrire les rapports entre les sens de *sain* dans *médecine saine*, *homme sain*, *urine saine* (analogie *pros hen*, c-à-d par référence à une chose, ici la santé), Brentano (1816 : 60) parle de *noms de famille* (*Familiennamen*).

Pourquoi des catégories: une vision pragmatiste

- Rosch (1975c: 197)

“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”

- Bruner (1957: 126).

“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, elle tient le docteur à distance si on en mange une chaque jour. (...) Le sens d'une proposition, comme Peirce l'a remarqué dans sa célèbre étude sur la théorie pragmatique de la signification, est l'ensemble des jugements hypothétiques qu'on peut formuler sur les attributs ou conséquences associés à cette proposition.”

Pourquoi des catégories: une vision pragmatiste

- Une catégorie “utile” est telle que les attributs de ses membres tendent à être interprédictibles.
- Encore une fois, même lien entre interprédictibilité et typicalité chez Bruner (chez qui le jugement de typicalité est gestaltiste):
“...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*]. (...) 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.” (Bruner et al. 1956)

La ressemblance de famille: expérimentation

- Rosch et Mervis (1975): 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.
- Protocole: sujets listent attributs et on vérifie si les exemplaires prototypiques identifiés dans une autre expérience ont les propriétés (1) et (2).

La ressemblance de famille: expérimentation

- La propriété (3) pose problème: qu'est-ce qui contraste avec *chaise*? Les rép. des sujets sont "déraisonnables" ou trop "créatives" (Rosch & Mervis 1975: 584).
- Solution: employer des catégories artificielles, des suites de lettres (par ex. PHMQB ou XPHMQ) où une lettre vaut un attribut et pour lesquelles les ressemblances sont *perceptives*. Résultat: les sujets apprennent plus facilement des catégories dont les membres ressemblent moins à ceux d'autres catégories.
- Ce problème de la dimension contrastive des catégories touche un point sensible: si on ne parvient pas à donner les catégories contrastives des catégories "sémantiques", c'est peut-être parce qu'un item peut appartenir à plusieurs paradigmes et que le choix du paradigme est contextuel (Rastier 1991)...
- La dimension contrastive sera en grande partie évacuée de la LC.

Le prototype vu du côté de l'attribut: la validité d'indice

- Les exemplaires d'une catégorie ne sont pas égaux, mais les attributs non plus.
- Un attribut peut prédire l'appartenance à une catégorie mieux qu'un autre.
- Pour quantifier cette force prédictive, Rosch reprend Beach (1964) et son concept de *relative amount of evidence*: la somme des 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.
- Le *relative amount of evidence* est une notion que Beach hérite du psychologue d'origine autrichienne Egon Brunswik (disciple de Bühler, membre du Cercle de Vienne, installé aux E.-U. depuis 1937). Sa doctrine: le fonctionnalisme probabiliste.

Le fonctionnalisme probabiliste

- Un organisme repère des couplages causaux entre des indices, par exemple, la disparité des images rétiniennes, et une propriété du monde, 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 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.

Le fonctionnalisme probabiliste

- Rosch renomme le *relative amount of evidence* de Beach en *cue validity*.
- Elle intègre aux attributs, comme Brunswik et Tolman, les aspects manipulatoires / moteurs et utilitaires de l'objet (l'objet comme chose-à-mastiquer ou chose-apaisant-la faim chez B & T; origine: Kurt Lewin).
- Un indice n'a pas besoin d'être définitoire ni propre (on peut utiliser la taille comme indice de l'intelligence). Idem pour un attribut à la Rosch:

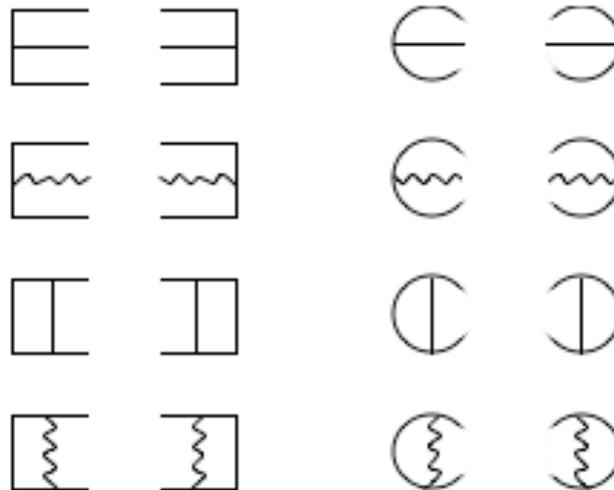
“la distinction entre sauvage et domestiqué est sans pertinence pour une définition rigoureuse d'*oiseau* puisque les oiseaux comme les non-oiseaux peuvent être sauvages ou domestiqués ; cependant, les oiseaux sauvages sont considérés comme plus centraux que les oiseaux domestiqués” (Rosch, 1973b : 141-2).

Indice et information

- Rosch retraduit la validité d'indice en termes informationnels: des attributs sont *redondants* lorsque de l'un on peut prédire la présence des autres.
- Les prototypes ont la validité d'indice la plus élevée = les prototypes maximisent la redondance de l'information.
- La redondance reflète quelque chose de réel dans l'environnement, que Rosch, à la suite de Garner (1974), appelle la *structure corrélationnelle*.

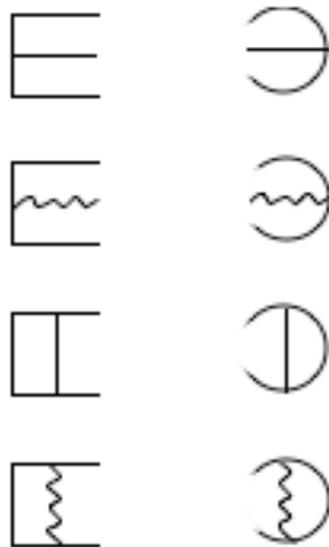
La structure corrélacionnelle chez Garner (1974)

- Soit l'ensemble total de 16 stimuli à 4 dimensions (forme de la figure, ouverture, orientation de la ligne, forme de la ligne) de 2 valeurs chacune:



La structure corrélacionnelle chez Garner (1974)

- On prend la moitié de l'ensemble (avec dimensions et valeurs équiprobables):



- 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: la *structure corrélacionnelle* de ce sous-ensemble entraîne que l'une de ses dimensions de variation est *redondante*.

Convergences entre Rosch et Garner

- Garner insiste 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.
- Il souligne que le contraste entre catégories semble guider la catégorisation: “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).
- Rosch retient le *réalisme* de Garner: la structure corrélacionnelle et la redondance sont des propriétés de l'environnement qui se reflètent dans la structure des catégories.
- MAIS il y a des divergences (voir Fortis 2010).

La question des niveaux de catégorisation

- Les catégories “sémantiques” ont introduit un nouveau problème dans la théorie du prototype : les noms qui identifient les membres d’une catégorie sont à différents niveaux d’abstraction (*quadrupède* est plus abstrait que *chien*).
- Le problème est de comprendre pourquoi nous catégorisons les mêmes référents à différents niveaux d’abstraction.
- Dans le milieu de Rosch, 2 approches sont disponibles: l’approche pragmatique / utilitariste de Brown (1958a/b, 1965, *middle level of abstraction*) et l’approche taxinomique de Berlin. C’est surtout cette dernière qui semble avoir influencé Rosch.

Les niveaux de catégorisation: Berlin et les *folk genera*

Berlin (1973, 1978):

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*).

Les autres niveaux (règne ou variété par ex.) sont plus rares.

Les niveaux de catégorisation: Berlin et les *folk genera*

- Selon Berlin, le niveau générique est privilégié : il inclut un plus grand nombre de membres, qui sont généralement monotypiques, et qui tendent à être désignés par des termes inanalysables ou peu productifs (comme les *basic color terms*).
- Le genre populaire correspondrait à des **discontinuités** perceptives du monde biologique (Berlin 1978 : 11). 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).

Genre et niveau générique

- Pendant un bref moment, Rosch reprend l'idée de genre de Berlin sous le nom de *niveau générique* (Rosch 1975c).
- MAIS Rosch place 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* (comme Berlin elle a une définition conceptuelle du genre, mais la taxinomie lui importe moins).

Genre et niveau générique

- Le niveau générique est exprimé en termes de *corrélation*:
“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.” (Rosch 1975c : 201)
- Rosch retrouve ainsi ses principes de structure corrélationnelle, de redondance, de validité d’indice et de contrastivité.

Le niveau de base

- Rosch renomme le niveau générique en “niveau de base” (Rosch 1976; Rosch et al. 1976).
- Le niveau de base fait l’objet d’une expérimentation, à partir d’une liste “épurée” de catégories.
- Protocole: lister des attributs, mesurer le taux de chevauchement des objets à un niveau donné, identifier des formes moyennées.

L'épuration des catégories

- L'expérience exigeait que les membres de catégories de base et subordonnées soient encore identifiables à partir d'un tracé; des catégories ont disparu (les fruits par exemple).
- 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.
- 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. 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. (Rastier 1991 pour une critique).

Résumé

Les “catégories sémantiques” sont organisées comme les catégories “naturelles”, autour de prototypes (tous les membres ne sont pas également typiques de la catégorie, ils sont associés par une ressemblance de famille). Leur structuration obéit au principe de maximisation de la *validité d’indice* (*cue validity* ; maximisation de la redondance des indices et de la contrastivité). Au niveau de base, la validité d’indice est maximale et la discontinuité morphologique par rapport au niveau supérieur est maximale.

recherches sur les
couleurs (Lenneberg
et al.)
couleurs focales de
Berlin & Kay

Gestalttheorie,
bonnes
formes, points
de référence

notion de schéma /
prototype (Head >
Bartlett > Attneave)

ressemblance
de famille,
Wittgenstein

niveau de
catégorisation, *folk
genera* ap. Berlin,
middle level ap.
Brown

**THEORIE
DU PROTOTYPE**

notion pragmatiste et
psychologique de la
catégorie, typicalité,
Bruner, Lenneberg

Fonctionnalisme
probabiliste,
validité d'indice,
Egon Brunswik,
Beach

Théorie de
l'information,
redondance, structure
corrélacionnelle
(Garner)

L'importation de la théorie en linguistique

- Thème du flou (logique floue, ensemble flous, grammaire floue, appartenance floue): voir Lakoff sur les enclosures.
- La théorie du prototype illustre une convergence vertueuse entre sciences (psychologie, neurophysiologie, anthropologie), à une époque où cette convergence est promue par le conglomérat *cognitive science* et où le mentalisme (pour parler comme Bloomfield) n'est plus tabou.
- Renouveau de la sémantique, dans un contexte d'acculturation relative à la tradition de sémantique lexicale (sémantique générative).
- Rejet progressif du formalisme chez les futurs tenants de la linguistique cognitive.

L'importation de la théorie en linguistique

- Avec la défaite de la sémantique générative, certains linguistes américains se trouvent sans affiliation théorique. La théorie du prototype remplit un vide, en leur permettant de faire de la sémantique lexicale.
- Date jalon: 1975, *Summer Institute* à Berkeley.

Lakoff: “Plusieurs exposés historiques ont été présentés cet été-là. D’abord, il y a eu le premier exposé par Eleanor Rosch de ses résultats sur les catégories de niveau de base. En second, il y a eu le premier exposé de Talmy sur ses travaux qui montraient que les concepts associés aux relations spatiales avaient des primitives de nature topologique et orientationnelle [ex. *front / back*]. En troisième lieu, Charles Fillmore, qui, pour la première fois, défendait la nécessité d’une approche sémantique en termes de cadre (*frame*). Et en quatrième, Paul Kay présentait son travail réalisé avec Chad MacDaniel sur le fondement neurobiologique de la catégorisation des couleurs” (Huck & Goldsmith 1995 : 117).

“Et donc, face à toutes ces preuves, durant cet été de 1975, je me suis rendu compte que la grammaire transformationnelle et la logique formelle étaient désespérément inadaptées et j’ai arrêté de faire de la sémantique générative” (Ruiz De Mendoza Ibáñez 1997 : 39).

L'importation de la théorie en linguistique

- L'approche est sémasiologique: les attributs roschiens deviennent des sèmes et la catégorie un lexème (Coleman & Kay 1981 sur le verbe *lie*; Brugman 1981 sur *over*; sur cette transition, voir Kleiber 1990).
- La théorie permet de traiter facilement la polysémie, en se fondant sur l'idée que les acceptions d'un lexème sont comme les membres d'une catégorie roschienne, unis par une simple ressemblance de famille.
- La théorie du prototype envisage la catégorisation comme flexible. Elle permet donc de récupérer directement le thème de la formation par *analogie*, et indirectement de réévaluer le rôle de l'usage et de la fréquence. Cf. la notion de *schéma* chez Bybee & Slobin (1982): verbes irréguliers anglais, réduits à des schémas du type
 ...æŋ(k)]_{verb / past}, où le schéma est un prototype (admet des déviations).

L'importation de la théorie en linguistique

- La théorie permet d'exprimer les propensions empiristes et pragmatistes qui sont "latentes": les acceptations prototypiques sont généralement les acceptations concrètes (p. ex. spatiales pour les adpositions).
- Elle converge ainsi avec la théorie des métaphores (Lakoff & Johnson 1980), la métaphore devenant un procédé de dérivation d'une acception non prototypique à partir d'une acception prototypique.
- Lakoff (1987) et Lakoff & Johnson (1999, *Philosophy in the Flesh*), font de cette convergence un argument en faveur d'une théorie globale de la connaissance qui prétend dépasser le rationalisme occidental en réévaluant le rôle du corps et de l'expérience vécue dans l'appréhension de domaines "abstraites".

L'importation de la théorie en linguistique

- MAIS au cours du transfert vers la linguistique certains aspects de la théorie sont abandonnés: la dimension contrastive et le calcul de la validité d'indice sont ignorés (les traits sémantiques ne sont pas pondérés).
- Les interactions avec la psychologie sont paradoxalement rares et une fois l'adoption de la théorie effectuée, les linguistes se désintéressent de l'évolution des théories de la catégorisation en psychologie.

Texte complet et références disponibles à

<http://corela.edel.univ-poitiers.fr/index.php?id=1243>

ou via

<http://htl.linguist.univ-paris-diderot.fr/jmfortis.htm>

Merci!