

Categories  
versus  
*rules*

LING 421 Project ~ Tom Cunningham.

# 1. Contents

1. Contents .....	2
2. Introduction.....	3
3. Classical Categorization .....	4
3.1 Caricature .....	4
3.2 Application.....	4
4. Labov (1973).....	5
4.1 Intentions .....	5
4.2 Experiments.....	5
4.3 Results .....	5
5. Rosch (1975).....	6
5.1 Intentions .....	6
5.2 Experiments.....	6
5.3 Results .....	7
6. Armstrong, Gleitman, and Gleitman (1983).....	8
6.1 Intentions .....	8
6.2 Experiments.....	8
6.3 Results .....	8
7. Against Classical Categories .....	9
7.1 Labov.....	9
7.2 Rosch .....	9
7.3 Armstrong .....	9
8. Categories vs. Rules .....	10
8.1 Introduction.....	10
8.2 Rules .....	10
8.3 Rule explanations .....	11
8.4 Rule predictions.....	11
9. Influences on Word Application .....	13
9.1 Objects of application.....	13
9.2 Disagreement .....	13
9.3 What is Rosch measuring?.....	14
9.4 Inconsistency .....	14
9.5 Are Rules Just A Way of Implementing Categories? .....	15
10. Summary .....	16
10.1 Reservations .....	16
11. Conclusion.....	16
12. Bibliography.....	17

## 2. Introduction

This essay is concerned mainly with concrete semantic categories, such as CUP and BIRD and FURNITURE. First with the traditional approach to category description, second with influential experiments concerning category description, third with interpretation of these experiments, and finally a discussion of the justification for imputing categories, with a new non-categorical theory to explain the use of categorical words.

I shall use small caps to refer to categories, thus: DOG.

### 3. Classical Categorization

Many discussions of prototype categorization contrast the new theory to classical or Aristotelian or checklist categorization (Taylor, 1995:21-37; Geeraerts, 1989:588; Rosch, 1973:111; Rosch, 1975:193). Aristotle, in his *Metaphysics* (cited in Labov, 1973: 368), distinguishes between an object's essential and accidental properties.

The theory is not commonly articulated except when criticised because it has often not been thought to be a theory at all, rather as simply an explanation of what categories are or what they should be. Here I present a caricature, the theory in its most extreme form:

#### 3.1 CARICATURE

A category is a set of all objects which meet some criteria; so it is a division of all the world's objects into those which meet the criteria and do not, those which are included and excluded by the category. There are no degrees of membership.

The criterion of membership is a conjunction of features, those objects which possess the features will be admitted.

Each feature is binary, an object either exhibits it or not. A feature picks out a category itself, so each category can be seen as the intersection of other categories.

But if this were true of all categories then there would be an infinite regress. Thus some features must be basic or uncomposed, these are atomic features. Atomic features cannot be learnt as composed of other features, so they are innate and universal.

#### 3.2 APPLICATION

In psychology the classical approach has been most relevant through concept formation experiments which typically use categories with necessary and sufficient conditions for membership (cited in Rosch, 1975:193).

In linguistics componential semantic analysis is strongly modelled along classical lines (Katz and Fodor, 1963, cited in Geeraerts, 1989:588).

## 4. Labov (1973)

### 4.1 INTENTIONS

In this study Labov used a novel method to investigate meaning. Instead of introspection or systematic probing of his intuitions, he conducted psychology-style experiments to determine the influences on word use in normal speakers.

He assumes that meaning is best represented by an algorithm which predicts the average application of a word based on an object's properties. Labov calculates which algorithm would correctly predict speakers' use of terms about containers (mug, cup, etc.), and makes some general points about the patterns of use.

The set of traditional assumptions about the nature of the algorithm he calls the 'categorical view'; it asserts that each category (which corresponds to an algorithm) is discrete, invariant, qualitatively distinct, conjunctively defined, and composed of atomic primes. Labov defines each of these properties; applying to semantic algorithms they mean that the algorithm will predict applicability just if the object exhibits all of a set of features, and that the features are very basic and themselves unanalysable. The experiments are designed to test these traditional assumptions.

### 4.2 EXPERIMENTS

The general pattern of these semantic experiments was to give the participants (subjects) a carefully controlled amount of information (collectively, the stimulus) about some object, then for the subject to return a noun phrase which named the object. In this case subjects were given a line-drawing of a container, sometimes with information about its use or the material it is made from, they then name it as - for example - a funny cup with a stem. The head noun (in the example, cup) is recorded as the subject's classification.

Having carried out a number of these experiments each stimulus can be paired with its typical classification. For example when presented with a drawing of a curved container with one handle and a width:height ratio of 1.0, 100% of Labov's subjects used the word cup to describe it; but if the container's width:height ratio is 1.5 then cup is used only about 85% of the time, with other words like mug and bowl taking up the slack.

### 4.3 RESULTS

Averaging his subjects' responses Labov found a high degree of variability, i.e. only a few types of receptacle were unanimously given the same name, most receptacles were given a consistent variety of names.

Secondly, none of the categories could be well characterized in terms of necessary and sufficient conditions. E.g. it is not necessary for an object to have a handle for it to be a cup, nor for it to be used for drinking tea, or for it to have a saucer, but each of these contributes to the likelihood it will be called a cup.

In conclusion Labov offers a definition of cup, describing both an invariant core and variable skirt which predicts categorization to be proportional to a weighted sum of typical cup features (having a handle, used to drink tea, with a saucer, etc.).

## 5. Rosch (1975)

### 5.1 INTENTIONS

Rosch's 1975 experiments tested subjects' behaviours on being presented with the name of a category and names of category members (eg. category: BIRD, with members ROBIN, SPARROW, etc.). The experiments were primarily carried out to investigate the mental representation generated upon hearing a categorical term (like bird), rather than a meaning or category. Rosch claimed no particular theory of the mental representations she was investigating, rather she was trying to answer a few general questions about them by carefully investigating their influence on subject behaviour. In earlier and later papers Rosch allows herself more speculation.

The general technique of investigation was priming: subjects were told the name of a category (primed) then two seconds later a card was flashed before them. The card displayed either two object names or two pictures, the subject had to decide whether the two stimuli were the same then immediately press a key indicating their choice. Sometimes they were told to only judge the stimuli the same if they were physically identical (e.g. robin and robin), but sometimes if the named or depicted objects just belonged to the same category (robin and sparrow).


Sometimes priming with the name of the category to which the two stimuli belong improves or retards judgement time. The key to interpretation of these experiments is that if the prime (category name) improves the judgement of A-identity more than B-identity (where A and B are subcategories) then the prime's representation must be more relevant or similar to the representation of A than of B.

### 5.2 EXPERIMENTS

Rosch used various criteria to select just ten high-level concrete category terms: fruit, bird, vehicle, vegetable, sport, tool, toy, furniture, weapon, and clothing. For each category lists of about sixty member names were compiled, chosen by frequency of subject naming.

Nine experiments were carried out, each with different groups of university students as subjects. The first experiment gathered data on subjects' conscious judgement of internal structure or representativity. Instructions explained how a category could exhibit differential membership, then subjects were asked to rate each instance on a 1-7 scale according to how good an example of the superordinate category it is.

The second experiment tested the psychological validity of these ratings, by measuring the effect of priming with the category name on judging identity of members, using for each category four high-, medium-, and low-rated members. Sometimes

subcategories were represented by pictures instead of names, e.g.  instead of clock.

The third experiment was performed along similar lines to rule out other possible interpretations of the second experiment, for example that the subjects' behaviour with respect to poor members of categories might be explicable as the poor members being not members at all.

The fourth, fifth, sixth, seventh, and eighth experiments varied elements of the experimental set-up to test various hypotheses about how representations were being processed: whether representations were more similar to the pictures or words, and whether priming improved the perception of stimuli or just the decision of identity.

Finally, the ninth experiment repeated the experiment-2 priming procedures on four subjects twice a day, five days a week, for two weeks. Rosch had carried out previous similar repetition experiments with colour categories and found little increase in performance.

### 5.3 RESULTS

Rosch's main conclusion was that categorical structure exists. The good-exempleness rankings given in experiment one were consistent across subjects. The second experiment confirmed these rankings, the well rated members were also the members whose recognition was improved by category priming, and therefore whose representation is more similar to the category's representation.

It was also found that correct primes for poor members inhibited subjects judgements more than incorrect primes. For example, subjects judged the identity of two instances of vase - a marginal member of furniture - more slowly when primed with FURNITURE than when primed with BIRD.

Varying experimental conditions appeared to show that the mental representation generated by the prime is mostly abstract but perhaps closer to a picture than a word.

When the priming experiments were repeated in experiment nine subjects well learnt the task and all prototype effects disappeared. Similar prototype effects in identifying colours could not be destroyed by repetition. The difference is probably best explained by the physiological basis of colour perception being fixed, whereas our knowledge of concrete categories is plastic.

## 6. Armstrong, Gleitman, and Gleitman (1983)

### 6.1 INTENTIONS

Eleanor Rosch's 1973 paper On the internal structure of perceptual and semantic categories reported similar experimental results to those of 1975 (previous section), but her interpretation was slightly stronger: that the experiments demonstrated family resemblance. Family resemblance is not clearly defined in this paper or in Rosch's original one, but it is taken to typically involve lack of necessary and sufficient conditions, fuzzy boundaries, degrees of membership, and judgement of membership by similarity to a prototype. Armstrong et al. contest the interpretation of consistent exemplariness as showing family resemblance.

### 6.2 EXPERIMENTS

The experiments were modelled on Rosch's 1973 ones, and are very similar to those described in the last section. Eight superordinate categories were used, four of Rosch's original categories (fruit, sport, vegetable, vehicle) and four chosen especially to not exhibit degrees of membership (even number, odd number, female, plane geometry figure). Rosch's 1973 experiments were replicated: first judging subjective exemplariness of members in categories, second testing verification time for statements of the form member is a category. The verification experiment is similar to Rosch's 1975 priming experiment, shorter verification times are thought to show semantic or informational similarity of the member and category.

As a third experiment the subjects were asked directly whether, for each of the eight categories, "it make[s] sense to rate items in this category for degree of membership in the category?" (p286).

### 6.3 RESULTS

Rosch's 1973 results were well confirmed for the four categories common to both experiments. For the four new, flat, categories similar results were obtained: subjects gave a wide range of exemplariness rankings (though not as wide for the first four categories), they agreed on their rankings, and their rankings were corroborated by the verification experiment.

Almost all subjects thought that the four new 'flat' categories did not display degrees of membership. The four categories borrowed from Rosch produced disagreement, about two thirds of the subjects thought fruit, vegetable, and vehicle could display degrees, and about one third thought so of sport.

As has already been mentioned, Armstrong et al. used this data to contest interpreting prototype effects as betraying a family resemblance structure.

## 7. Against Classical Categories

Introductions to and discussions of prototype theory (Taylor 1995; Geeraerts 1989; Lakoff 1987) generally take the experiments of Labov and Rosch to vindicate the approach, and those of Armstrong as a challenge, to be explained in prototype-theory terms.

### 7.1 LABOV

Taylor (1995: 40) portrays the Labov experiment as demonstrating the fuzzy boundaries between categories. He goes on to draw three conclusions about attributes, where attributes are the things on which entity categorization is based. Attributes are not binary, attributes are high-level and cultural, and no set of attributes are necessary and sufficient for category membership. Each of these conclusions contradicts an element of the classical categorization theory.

### 7.2 ROSCH

Taylor (1995: 43-46) unambiguously interprets Rosch's 1975 experimental results, at least for the category FURNITURE, as reflecting degree of membership, thus contradicting the classical model's assumption of binary membership.

Geeraerts (1989) divides prototypicality (a property of categories) into four elements, one of which - clearly derived from Rosch's experiments - is: "Prototypical categories exhibit degrees of membership; not every member is equally representative for a category." It is unclear what the intended relationship is between the clauses either side of the semicolon, since Geeraerts remarks that a category can exhibit them independently and this "implies that a strict distinction has to be made between degree of membership and degree of representativity" (1989:596).

Maclaury ascribes a similar interpretation to Rosch: "each member belongs to the category to the degree that it shares attributes with the prototype" (1991:57), but Rosch herself seems to be careful not to talk of degree of membership.

### 7.3 ARMSTRONG

The Armstrong results, for the category ODD, obviously test any interpretation of Rosch's experiments as showing degrees of membership, and force a strong conclusion either to be watered down or qualified. One of Rosch's categories, 'bird', actually makes almost as good an example as 'odd', since it displayed clear prototypicality effects and yet most people would agree that 'bird' is an all-or-none category.

Taylor explains the apparent conflict between having and not having degrees of membership by saying that there are actually two aspects or characterizations or types of category ODD (1995:68-74). One is the folk category - defined by prototype, the other is the expert category - defined by schema. The first is the cause of the experimental results, so degree of membership may be fairly attributed to it; but it is just our workaday conception of odd, when asked to reflect on odd numbers generally we think of the definition taught in mathematics class, i.e. the expert definition, which does not exhibit degrees of membership.

Geeraerts takes a different approach (1989:597). He says there is just one category ODD, and it exhibits degrees of representativity but not fuzzy boundaries or degrees of membership.

## 8. Categories vs. Rules

### 8.1 INTRODUCTION

In this section I advance a sketch of a different approach to explaining peoples use of everyday concrete nouns, which I call a rule theory. This theory is just a possible manifestation of more general concerns about the relation of category and use. I will try to explain the general concern in this introductory note, and my reasons for it should become apparent as I elaborate and demonstrate the rule theory in subsequent sections.

The classical theory of categorization has always been a mix between being normative and descriptive. Clear thinking books (e.g. Wilson 1963) tell us that all proper categories have necessary and sufficient criteria, and if the words we uses are not associated with such criteria then we are not thinking clearly. When we come to psychology or linguistics and want to study people's actual categories rather than the categories they should have, it is unclear what status the necessary and sufficient criteria have.

Labov's and Rosch's experiments make a necessary and sufficient condition explanation of how we use concrete nouns seem far-fetched, but most proponents of prototype theory - especially in linguistics - retain the assumption that the concrete nouns pick out categories. It is not simply a matter of definition that a noun picks out a category of instances, some nouns seem far from doing any such thing, such as onset and profile, these nouns are only used in certain contexts and have meaning in those contexts; treating them syntactically as categories results in dubious expressions: ?all onsets and ?all profiles. By assuming that the use of all concrete nouns is on the basis of category membership we need to grossly deform what it means to be a category: we allow degrees of membership, we allow a lack of common features, we allow new categories to be spawned by context, and membership to be skewed by hedges. What is left, I want to suggest, is hardly worthy of the name category.

### 8.2 RULES

First a disclaimer: this is admittedly somewhat vague, and is missing many elements, and I imagine similar but better defined theories have been articulated and perhaps rejected in other fields. I advance this mainly to illustrate how in principle word use could be explained without talking about categories.

The rule theory accounts for word application in two parts: the use of a rule, plus a categorical aspiration. This theory is related to Hebb's explanation of thought (1980), the principles of artificial neural networks (Hassoun 1995), Wittgenstein's suggestions that learning language is much like learning rules for the use of words (1967:38)<sup>1</sup>, and Quine's naturalised semantics (1960).

Each word is associated with a rule. A rule is a psychological mechanism, which outputs a degree of certainty. When occasion comes to put a name to something every rule competes and the one which wins is the one with the greatest output. Labov's experiments, then, show which are the optimal conditions for different rules, e.g. for a tapered, handled, porcelain container as wide as it is high, the cup rule outputs the highest degree of certainty so the word cup is used.

A rule's output depends on its input, which is from any other mental event or information - from low level perceptual to high level abstract thought. For a rule associated with a simple word like cup the input need not be just the perceived attributes of the object which is to be named, the input can include any other mental context - such as who you are speaking to, where you are, what you have just said and thought, and how you feel.

---

<sup>1</sup> In paragraphs §80 and §81 Wittgenstein mentions the relation of meaning and rule, and whether logic is an ideal to which language aspires.

The relationship between the input and the output is established by training: when we learn a simple concrete word such as 'cup' we simply begin saying the word and some uses are approved of, others disapproved (c.f. Quine, 1960: ch1). The mental events associated with correct (i.e. approved) uses and not associated with incorrect uses become the inputs of the rule, and when they are activated the rule itself is activated. For example, the mental event it-has-a-handle is an input of the rule it-is-a-cup.

What gives a rule semantics is that we may reflect on its use, and this is its categorical aspiration. Most people unreflectively assume concrete nouns to unambiguously and consistently pick out a set of objects and for all the objects to have something in common. Concrete words often fail to live up to this standard, as shown by Labov's experiments, but when we are made aware of this failure we will be at least embarrassed and quite commonly we change our use by intentionally interfering with the natural input/output relation of our rule. For example, if someone talking about sport generally were asked whether horse racing is a sport (one Rosch's examples of marginal membership) they will most likely decide by fiat on the spot, that it is or is not a sport for the present purposes, than be content with just saying it is a sport to degree 3.18 (Rosch, 19175:232). The categorical aspiration, then, is an individual and social process which interferes with rules to try to make their output more closely approximate a categorical ideal of binary membership and common properties.

Finally one more thing to note about rule theory is that it is reductive: it explains semantics, or at least semantic behaviour, in terms of non-semantic events and processes, i.e. neural association and activation. Moreover the input/output mechanics of rule operation is modelled on well known properties of real (Hebb, 1980) and artificial (Hassoun 1995) neural nets, making it an interdisciplinary and psychologically realistic theory. Within the prototype category, and more broadly the cognitive linguistics framework, almost all explanations of meaning are just in terms of other meanings hoped to be somehow more basic.

### 8.3 RULE EXPLANATIONS

How would the use of rules, accompanied with a categorical aspiration, explain the experimental results we have already encountered?

1. The consistent degrees of representativity in Rosch's experiments: they show the degree of psychological association, particularly to what degree the psychological aftermath of perceiving the superordinate category word is an input into each subordinate category word.

2. Fuzzy boundaries in Labov's experiments: Due just to people having had their rules trained slightly differently, I shall explain this in more detail in a subsequent section.

3. Interdependence in Labov's experiments: The rule model is explicitly interdependent about the influences on the use of a word: as inputs accumulate the output (or certainty) is increased and is more likely to outweigh the output from other words.

### 8.4 RULE PREDICTIONS

How will rules and categories differ in their predictions?

- (1) Objects are what may be members in a category, and membership is based on an object's properties, so if the use of a word is governed by a category we would expect that the same object will be categorized in the same way if all its properties can be equally perceived. A rule, on the other hand, does not just depend on the object's properties, it is explicitly dependent on any mental event, so we would not expect the same object to be consistently categorized the same way on different occasions.
- (2) A category categorizes objects on the basis of their properties, but a rule just goes off on the basis of input from other mental processes. This I think allows the possibility of a rule to be activated without any object being present, just because of associated mental events.
- (3) Since a rule is a neural process, rather than semantic, we would not expect a speaker to have much reflective insight into the conditions of its use.
- (4) Similarly, a rule is learnt to correctly predict the singular use of a terms (an X), but not general terms (all Xs), so we would expect only a rough correspondence. A

mental category, on the other hand, seems to suggest an important commonality between using a word to pick out a single member or the entire category.

## 9. Influences on Word Application

### 9.1 OBJECTS OF APPLICATION

It may seem trivial, but neither Rosch's nor Labov's subjects actually judged membership. Labov's subjects categorized drawings, not receptacles. Rosch's subjects categorized subcategories, not objects. Labov's results would probably be not too different had he used real containers (Kempton 1981 (cited in Taylor, 1995: 76), sometimes used real objects in similar experiments). Rosch's certainly would have been different, her subjects gave the word chair a furniture-ness score of 1.04 out of 7; but given actual chairs, kitchen chairs, upholstered chairs, and deck chairs, the scores would surely diverge<sup>2</sup>.

The ability of subjects to use a categorical word given only subcategories or drawings of members confirms rule-theory prediction (2): that a word could be used in the absence of any actual object to which it properly applied, but just on the basis of associated concepts.

The experiments were in another way unrepresentative of actual categorization, because they were momentary. If subjects are confronted with a cup-shaped container full of flour then some tend to call it a bowl, but if they were originally confronted with an identical container filled with coffee, asked to name it, then watched the coffee drunk and replaced with flour, I think that many less would categorize the object as a bowl.

Here the influence on the term's use is how the object was previously conceptualized. Only by a fair stretch of the imagination is previous conceptualization a property of the object, so it is unclear how this extra influence could be accommodated to a categorical view. Previous conceptualization is, on the other hand, certainly a salient mental event which may contribute to a rule.

More generally, a categorical view has a problem with separating the influences of word application into those which are clues to membership and those which are decisive. For example, both having long hair and having two X chromosomes can influence application of the word 'woman' to a person, but the first is just a clue to her being a woman - the second is definitive. The division is clear in this case, but if we turn to categorizing cups it is not so easy. The angle of depiction and the lighting and the colour and the use of a container may all influence peoples' use of names, but we would hesitate before incorporating them all into the definition of the category 'cup'. Armstrong et al. (1983: p295) more closely detail the difficulties of separating criteria of membership and mere heuristic clues.

A rule makes no such distinction, it simply suggests a degree of certainty about whether a word applies. The division between essential and typical features only makes itself known in the process of categorical aspiration.

### 9.2 DISAGREEMENT

Taylor says of Labov's experiments: "Contrary to the expectations of classical theory, there was no clear dividing line between CUP and BOWL; rather, the one category merged gradually into the other." But one could as easily say this was contrary to the expectations of any categorization theory at all, because the variance in membership only appeared on averaging the subjects' responses. The fuzziness measured was not individual indeterminacy but group disagreement, and if their use of terms was based on categories then Labov's subjects all had quite different categories. The fuzziness could only validly be shared among individual representations if their use was approximately random within an envelope of fuzzy meaning, so if a container was a 90% cup then they would call it 'cup' nine times out of ten.

I think the rule theory has some interesting things to say on the disagreement of Labov's subjects. In his range of all possible receptacles there were a few regions in which

---

<sup>2</sup> Though of course my three examples: kitchen-, upholstered-, and deck-chairs, are only subcategories themselves. I can hardly include any actual chairs as examples with this essay.

most subjects agreed on the appropriate word to use, the central areas of 'cup', 'mug', and 'bowl'. I predict, on the basis of Labov's results and the rule theory, that those areas of agreement correspond to the most common receptacles in the subjects' shared environment.

Imagine that the only receptacles we had were one-handed china cups for drinks and large plastic bowls for soup, and nothing in between. In this case devising a rule to predict application of cup and bowl would be simple, anything with one handle or made of china or used to drink from can be called a cup, anything plastic and used for soup is a bowl. In fact, a number of different rules would do equally well in such an environment: one which predicted being a cup just on the basis of having a handle, or just on the basis of being made of china, or just on the basis of being drunk from. Each of these rules would serve equally well in this restricted environment, but when something is introduced into the gap between cups and bowls, e.g. a plastic one-handed mug for drinks and soup, the different rules conflict. Generally people will devise rules to correctly categorize common objects, but these rules may diverge in their categorizations of rare objects. This is the reason for the enormous disagreement among Labov's subjects: they disagreed just on receptacles which were rare and so receptacles that their rule had not been shaped to deal with. Without any practice on these unusual containers the subjects extrapolated their rules, but rules which agree on some cases may disagree when extrapolated, and this is just what happened in Labov's experiments.

### 9.3 WHAT IS ROSCH MEASURING?

There is a lot of ambiguity in the interpretation of Rosch's 1975 internal category structure experiments. The word 'chair' and 'table' have a differential relationship with respect to the word 'furniture' which can be reproduced in a number of experiments, even substituting pictures of a chair and table for the words. The respect in which the two words differ, relative to their superordinate, is sometimes called:

- (1) 'degree of membership' (Taylor, p43; Geeraerts p593; Maclaury p57),
- (2) 'goodness of membership' (Rosch 1978, p36; Rosch 1973 p135),
- (3) 'representativity' (Geeraerts, p593),
- (4) 'exemplariness' (Rosch, 1973, p139; Armstrong et al., p293)
- (5) 'good exempleness' or 'goodness of example' (Taylor, p43; Rosch 1973, p130),
- (6) 'prototypicality' (Taylor p45; Armstrong et al., p269),
- (7) 'how clear a case something is' (Rosch 1978, p36).
- (8) 'focalness' (Rosch, 1973, p134)

Also the respect is commonly described as degree to which the members resemble the core meaning or focal examples or prototype(s).

These different terms used to name the effect have quite different connotations. Something can be a complete member but unrepresentative (penguin in bird), and something can be representative but not a member (a good forgery).

A rule theory would interpret the value simply as degree of mental association, dependent upon to what degree the mental events caused by the superordinate term are inputs into the rule associated with the subordinate category.

### 9.4 INCONSISTENCY

If speakers used a term on the basis of a category then we would expect them to be consistent in their pairing of objects and categorical terms, but they are not:

1. Taylor (1995: 262) describes two methods of category research by noting conditions of use: onomasiological in which subjects are given stimuli and asked to categorize it, and semasiological in which subjects are given a categorical term and asked to pick out the centre and periphery of its denotation. In experiments on colour it has been found that subjects commonly are inconsistent among these two approaches, i.e. when shown colour c they say 'red', but when asked to choose which of a spectrum of colour instances are red they do not include c. Taylor says that this inconsistency also appears in other domains.

2. To my sister I might call a plastic drinking vessel a 'glass', but to a visitor a 'cup'. What do I really think it is? To suppose that use is based on category membership is to suppose that one of my two namings is false or contextually biased or that I actually operate with at least two categories for each of 'glass' and 'cup' - one informal and one formal. Rosch discusses the effect of context on categorization (1978, p42) but supposes that her results are not affected because subjects in an experiment will be in a kind of basic context.

A rule suggests a word on the basis of the entire mental state of person, not just perception of the object to be named. This explains why encountering the same object in different contexts can lead to the use of different words.

#### 9.5 ARE RULES JUST A WAY OF IMPLEMENTING CATEGORIES?

The rule theory is much more psychologically explicit than most prototype-category theories, are they compatible? I do not think so because on a rule theory what we would call a category would be the ideal towards which people individually and collectively strive, and the ideal has common properties and absolute membership, this seems a very strong disagreement between rule theory and prototype theory - prototypical categories are often defined just as lacking common properties and binary membership.

## 10. Summary

In this essay I cover the experiments and arguments used to support the graded-membership theory of semantic categories, and then add some speculations of my own about how to describe categories and word meaning.

- (1) An ideal classical category divides the world's objects into two exclusive sets depending on their possession of some set of atomic properties; until recently much work in psychology and linguistics has assumed human categorization to be similar to this ideal.
- (2) Labov's 1973 work assumes the meaning of an everyday word is best described as an algorithm which predicts the word's average application based on an object's properties. Experiments show that this algorithm must incorporate interdependence of properties, and it predicts mostly areas of marginal applicability.
- (3) Rosch (1975) used priming techniques, which can crudely measure semantic distance, to confirm that category representations are more similar to the representations of their typical members than the representations of their atypical members.
- (4) Armstrong (1983) challenges Rosch's 1973 interpretation of category experiments by recording categories which obviously lack graduation but which succeed on the tests which Rosch takes to indicate graduation.
- (5) The Labov and Rosch experiments have been interpreted to show deficiencies in the classical model, particularly the independence of criteria and the binary nature of membership. In place of the classical model, degree of membership in a category has been suggested to be congruous to degree of similarity to a mental representation (a prototype). Armstrong's problems pose a problem of accommodation.
- (6) I introduce rules as an alternate means of explaining word application, and rules are augmented with a categorical aspiration.
- (7) Finally arguments from experimental evidence are given for preferring rules to categories.

### 10.1 RESERVATIONS

- (1) The only language I speak is English. I imagine an aspiration towards categories would be more pronounced in European languages and cultures.
- (2) There are many discussions of prototypicality which I have not mentioned nor even seen, probably they cover similar ground to me. In particular I think Lakoff somewhere argues against interpreting effects as structure, and Smith and Medin's *Categories and Concepts* (1981, Cambridge, Mass.: Harvard University Press) considers bipartite theories of categories.

## 11. Conclusion

In essence I have reviewed the experimental evidence motivating the theory of prototype categories, and tried to account for the evidence without reference to psychological categories at all, because I think categories may be seen as the ideal to which actual use aspires.

## 12. Bibliography

- Armstrong, S. L., Gleitman, L. R., and Gleitman, H. (1983). What some concepts might not be. *Cognition* 13: 263-308.
- Geeraerts, D. (1989). Introduction: Prospects and Problems of Prototype Theory. *Linguistics* 27: 587-612.
- Hassoun, M H. (1995). *Fundamentals of Artificial Neural Networks*. Cambridge, Mass.: MIT Press.
- Hebb, D O. (1980). *Essay on Mind*. New Jersey: Lawrence Erlbaum Associates.
- Katz, J. J. and Fodor, J. A. (1963). The structure of a semantic theory. *Language* 39: 170-210.
- Kempton, W. (1981). *The Folk Classification of Ceramics: A Study of Cognitive Prototypes*. New York: Academic Press.
- Labov, W. (1973). The boundaries of words and their meanings. In C.-J.N. Bailey and R. W. Shuy (eds.) *New Ways of Analysing Variation in English*. Washington: Georgetown University Press.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.
- MacLaury, R. E. (1991). Prototypes revisited. *Annual Review of Anthropology* 20: 55-74.
- Quine, W. V. O. (1960). *Word and Object*. Cambridge, Mass.: MIT Press.
- Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (ed.), *Cognitive Development and the Acquisition of Language*, 111-44. New York: Academic Press.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General* 104: 192-233.
- Rosch, E. (1978). Principles of categorization. In E. Rosch and B. B. Lloyd (eds.), *Cognition and Categorization*, 27-48. Hillsdale: Lawrence Erlbaum.
- Wittgenstein, L. (1967). *Philosophical Investigations*. Translated by G. E. M. Anscombe. Oxford: Basil Blackwell.
- Wilson, J. (1963). *Thinking with Concepts*. Cambridge: Cambridge University Press.