DOCUMENT RESUME

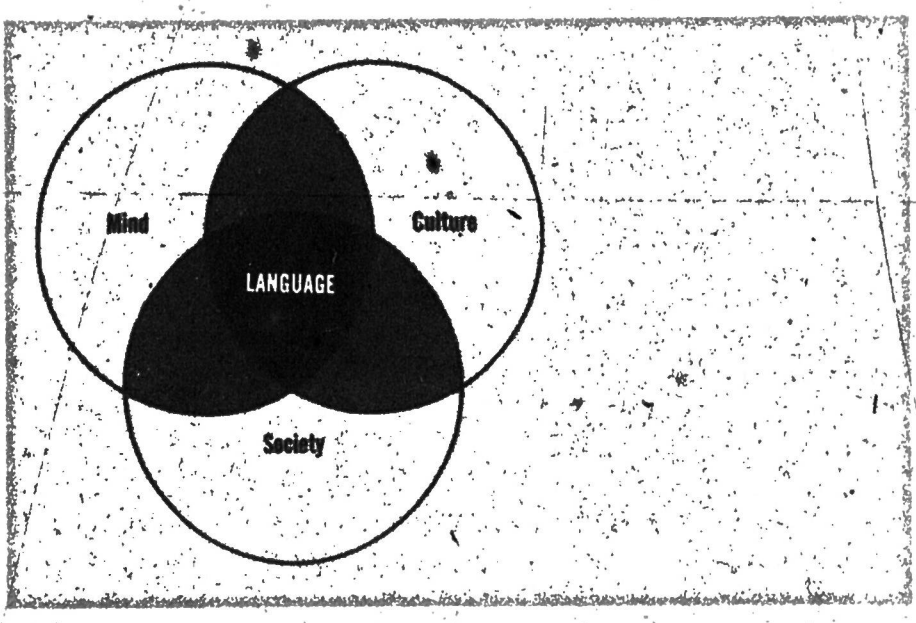
ED 140 680 FL 008 698

AUTHOR	Kay, Paul
TITLE	Tahitian Words for Race and Class. Working Papers of the Language Behavior Research Laboratory, No. 40.
INSTITUTION	California Univ., Berkeley. Language and Behavior Research Lab.
SPONS AGENCY	National Inst. of Mental Health (DHEW), Rockville, Md.
PUB DATE	Aug 75
GRANT	MH-25703
NOTE	40p.
EDRS PRICE	MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS	*Anthropological Linguistics; Ethnology; Folk Culture; *French; Language Patterns; Language Research; Language Variation; *Lexicology; *Linguistic Theory; Malayo Polynesian Languages; *Portuguese; *Semantics; Sociocultural Patterns; Sociolinguistics; Vocabulary
IDENTIFIERS	Ethnographic Semantics; *Tahitian

ABSTRACT

This volume is based on field work conducted in 1960 in Papeete and in a rural district of Tahiti, under the guidance of Douglas Oliver. Section two, which is based on a Ph.D. thesis (Kay 1963), develops the hypothesis that Tahitian words for social classification and the common French translations are semantically equivalent for most native speakers of Tahitian. This means that when a French word is used, the Tahitian meaning is intended. Section three, which attempts to relate the data in section two to more recent anthropological research, develops the hypothesis that words designating racial/social types in Brazilian Portuguese do not constitute a shared semantic system which permits native speakers to communicate effectively in this domain. This semantic ambiguity may be due to a desire to blur class and race distinctions. Section four discusses the implications of these hypotheses for semantic theory, and presents research on other sets of words whose meanings seem to imply underlying notions of quantity and statistical distribution. (Author/AM)

Working Papers of the
Language Behavior Research Laboratory
University of California, Berkeley



The Language Behavior Research Laboratory is an interdisciplinary research project at the University of California at Berkeley. It is supported by Research Grant No. MH 25703 from the National Institute of Mental Health. This support is gratefully acknowledged.

~~Working Papers of the Language Behavior Research Laboratory may be obtained at cost from the~~

~~University of California~~
~~Language Behavior Research Laboratory~~
~~2220 Piedmont Avenue~~
~~Berkeley, California 94720~~

TAHITIAN WORDS FOR RACE AND CLASS

Paul Kay
University of California, Berkeley

Working Paper No. 40
Language Behavior Research Laboratory
1975

TAHITIAN WORDS FOR RACE AND CLASS

To appear in a special issue of the Journal de la Société des Océanistes (Paris) dedicated to Douglas L. Oliver

Paul Kay

August 1975

I. Introduction

It is a pleasure to contribute to this volume, honoring Douglas Oliver some data and tentative conclusions regarding Tahitians' conceptions of major social categories as reflected in their use of language. The field work on which this essay was based was conducted under the guidance of Professor Oliver in 1959 and 1960 in Papeete and a rural district of Tahiti, and the main body of the essay, Section II, appeared in essentially its present form in a Ph D thesis (Kay 1963) also done under the direction of Professor Oliver. The data reported as well as the theoretical framework employed in Section II are thus a decade and a half old. I cannot therefore vouch for the contemporary ethnographic accuracy of the account given therein. I would hope, nevertheless, that apart from any vestigial ethnographic value, conclusions reached in 1960 regarding Tahitians' conceptions of major social categories may have some application to other ethnographic investigations in the Pacific and other areas, and perhaps also to some more general semantic questions. In this hope I have added Section III, which attempts to relate the data and model discussed in Section II to some more recent research by anthropologists on racial/social categories and Section IV, which considers some recent linguistic work on other sets of words whose meanings seem to imply underlying notions of quantity and statistical distribution. Section IV also presents some more general proposals for semantic theory: (1) that the semantic values of linguistic forms are better understood in terms of indices to cognitive schemata than in terms of sets of features (C. Fillmore 1975), (2) that virtually any sort of formal structure that is readily apprehended by the human mind may serve as a such a cognitive schema and (3) that in particular such schemata may, and often do, involve quantitative continua.

II. The Semantics of Race in Tahitian and Tahitian French

The major terms used for racial classification by native speakers of Tahitian, whether or not they are also native speakers of French, are listed below (page 4) with French equivalents and English glosses. The Tahitian and French forms are, in my opinion, practically identical semantically. This assertion is sufficiently novel to require substantiation. Since I formed the hypothesis only after returning from the field, such evidence as can be adduced in support of it will have to be argumentative and anecdotal rather than systematic and factual.

The semantics of French, as spoken by Tahitians, tends to differ from metropolitan French in key areas, particularly with regard to social roles and behavior patterns, so as to furnish a word for word isomorphism with Tahitian conceptualization of these areas. For example, usage by Tahitians, when speaking French, of the words honte and pitié frequently strikes a metropolitan Frenchman as inappropriate. I would contend that for Tahitians in the semantic contexts where the substitutions are made, these words mean exactly the same things as the Tahitian words haama and arofa, although no one who learned French in France would use honte and pitié in these ways. Haama and arofa are frequently used words in Tahitian. In glossing them, one is inclined to emphasize their "orientation to action" or "attitude to alter" aspect much more than their aspect of expression of an internal state. The Tahitian words describe modalities of social behavior while the roughly corresponding French words are concerned with internal states of individual persons. However, when Tahitians speak French, it is clear--albeit intuitively--that in the vast majority of cases it is the Tahitian meaning which is intended rather than the French. When asked in French why he does not take action to collect a debt

owed him, a Tahitian will often reply "Ça fait honte." I think it is fair to represent the way this utterance strikes a Frenchman by the gloss: 'That is/would be shameful.' However, an appropriate gloss for the corresponding Tahitian "E haama," might be, "That is/would be thoroughly improper." There is doubtless some overlap in the meanings, but they are not identical. The French word is used, not to signify its usual meaning in metropolitan French, but the meaning of haama.

This formulation accounts for an otherwise puzzling fact. For a long-time in Tahiti I was perplexed by the assertion of some metropolitan French people that even Tahitians who appear to speak French fluently and elegantly, "do not really understand the meaning of the words." The statement is perplexing because it is difficult to imagine how a large number of individuals can learn a language well on the plane of expression and yet each independently learn it poorly on the plane of meaning. However, in terms of the argument of the preceding paragraph, the observation is easy enough to account for. The differences between the French spoken by Frenchmen and by Tahitians are due only partially to imperfect leaning on the part of individual Tahitians. More generally, there exists in Tahiti a dialect of French which differs systematically from metropolitan French in some semantic areas (and in some phonetic details) while differing hardly at all morphologically and syntactically.

There are, of course, a few Tahitians, particularly those educated in France, who speak the standard semantic dialect just as there are a great many who speak no dialect of French at all.[1] These facts do not affect the

---

[1] Among Tahitians a reliable shibboleth of speakers of the metropolitan semantic dialect of French is use of the T-V contrast in pronouns and verbs in a way that makes pragmatic sense from a French person's point of view. Such speakers form a distinct minority.

hypothesis that a local semantic dialect exists. The characteristic fact about this local dialect is that in certain areas of meaning, especially those dealing with common social roles and social behaviors, many French words are used as if they were perfect translations of certain Tahitian words. Hence, when a Tahitian discusses social behavior in French, his semantics, and hence the cognitive schemata implicit in his discourse, are apt to remain largely Tahitian. It is according to this argument that I hypothesize that Tahitian words for social classification and the common French translations are semantically equivalent for most native speakers of Tahitian.

MAJOR CATEGORIES OF RACIAL CLASSIFICATION IN PAPEETE

Semantic Category	I	II	III	IV	V
Tahitian	taata Tahiti, taata	afa	popaa	afa-tinito	tinito
French	Tahitien Polynésien	demi	popaa Européen	demi-Chinois	Chinois
English[2]	Tahitian Polynesian	mixed-blood	European White	half-Chinese	Chinese

The list of terms given in the preceding table is not exhaustive. However, a model which takes the five columns of the table as exhaustive of the universe of categories of racial classification is an adequate representation of the native conceptualization of race. (From this point on, I am concerned only

[2]The English glosses are quite deceiving if taken as translations. They refer only to the biological aspect of a classification which is not mainly biological.

with categories I, II and III. For discussion of Chinese and part-Chinese, see Moench 1963).

However, the model is not simply a list of these categories, as the list can be and is applied in two different ways. First and most obvious, the three categories are sometimes used in an ordinary way to refer to empirical classes of individual people. Informants generally agree about the defining attributes of Tahitians, mixed-bloods and Europeans. Theoretically, although not in fact, any two informants will agree about the classification of a mutual acquaintance. (For discussion of informant agreement about diagnostic criteria and disagreement on specific diagnoses see Frake: 1961.) However, I have often found inter-informant agreement on specific assignments of individuals to categories which run counter to the explicit criteria of classification. For example, although pure Polynesian ancestry is an explicit criterion for the designation Tahitian, I found very few speakers who would classify a poor and uneducated individual living "Tahitian style" as anything but Tahitian, despite an almost pure caucasoid appearance. In this respect, the situation is quite different from what Frake describes for Subanun diseases, as speakers are clearly not making individual judgments on the basis of the simple presence-absence variables (e.g. rich vs. poor, educated vs. uneducated, racially "pure" vs. racially mixed) they claim to be using.

The psycholinguistic work of N. Cliff proposes a semantic theory which offers a plausible explanation for this situation.[3] The aspect of Cliff's

[3] Cliff's theory was tested on a sample of nine English adverbs of intensity (e.g., somewhat, very, extremely) and fifteen evaluative adjectives (e.g., good, bad, contemptible). The theory holds that each adjective in the set has a numerical value representing its degree of favorableness and that each adverb has a multiplicative value, analogous to a scalar in vector multiplication.

work that is suggestive in the present context is the general notion that the variables defining semantic schemata may be quantitative, in the usual sense of having some of the non-trivial properties of the real numbers, and that these variables may be combined in ordinary algebraic formulae to define the semantic values of lexical categories. With regard to the first kind of racial classification in Tahiti, the inference would be that what Tahitian informants present as discrete, presence-absence type diagnostic criteria for racial classification (such as rich vs. poor) may in fact be continuous variables.

As previously noted, Tahitian informants can assign individual persons to racial categories in such a way as to produce a high degree of agreement among informants, despite the fact that the assignments are acknowledged to violate one or more of the explicit criteria of classification. It is possible that in addition to using continuous variables as the basis of their judgments, in contrast to the dichotomous variables they profess to be using, speakers are using some particular, but at present unknown, algebraic combination of these scale values, in producing judgements. This combination would constitute a continuum of Tahitianese-Europeaness along with individuals are placed on the basis of an algebraic combination of a number of quantitatively perceived properties of the individual. The lexical categories Tahitian, mixed-blood, European would then index regions on this continuum.

Hence, the meaning of an expression like "very bad" may be represented by a number and this number is the product of the number assigned to the adjective "bad" and the "scalar" assigned to the adverb "very." The 135 possible adjectives-adverb pairs were administered to three large groups of undergraduate subjects and the basic scale values for each pair were obtained by the successive intervals technique. The major lines of the theory were very strongly confirmed; for details see Cliff (1959).

Of course, this does not mean that speakers are aware of using such a schema any more than Cliff's subjects were aware that by "very bad" they meant something just about exactly one and one half times as unfavorable as "bad."

Evidence was not systematically collected to substantiate this hypothesis as it was formulated only after field work was completed. Nevertheless, I can think of no alternative hypothesis which will account adequately for the following incidentally gathered facts and observations:

1. Informants agree on the criteria for racial classification.

2. These criteria are presented as dichotomous variables, but they are all readily, and perhaps more naturally, conceivable as continuous variables (e.g. rich vs. poor, educated vs. uneducated, pure Polynesian ancestry vs. mixed ancestry).

3. The conscious model for racial classification presented by many informants is empirically inoperable as it (1) claims to be able to classify any individual and (2) defines the classes in terms of a proper subset of the set of logical possibilities of presence and absence of the diagnostic criteria. Hence it cannot classify those individuals whose description in terms of these criteria falls outside the set of descriptions which are assigned to a racial category. Reconsidering the example given above, there is no theoretical classification for a person who is poor, uneducated, and follows a typically Tahitian pattern of economic consumption but is at the same time of largely European ancestry. However, such individuals definitely exist.

4. There is considerable agreement among informants on certain individual classifications which are impossible by the conscious model.

So far, I have discussed only one of the ways the racial classification dimension is used in Papeete. I have perhaps over-emphasized certain specu-

lations about the mathematical details of this way of classifying people. Whatever validity these speculations may contain, the general nature or "purpose" of the classification deserves more emphasis. This usage, like most classification schemes were are familiar with, assigns every object to one and only one class. It partitions a set of objects.[4] Used in this way, the racial dimension has three regions to one and only one of which every person may be assigned once and for all. According to this usage of the racial dimension, at a given time there exist in the non-Chinese population of Papeete three distinct collections of individuals: one Tahitian, one mixed-blood, and one European. Each collection is felt to have its own membership, customs, attitudes, typical style of life, etc.

However, this is not the only way the racial dimension is used. It is also true that speakers, at different times and in different social or conversational contexts, assign the same person to different racial categories. The form this phenomenon most often takes is for the speaker to contrast the same individual (often but always himself) at one time with "the Tahitians" and at another with "the mixed-bloods." Generally this occurs when some derogatory aspect of the stereotype of the group is question is prominent in the discussion. For example, I have heard many Papeete residents at one time contrast themselves with the "uneducated natives" and at another with the "avaricious mixed-bloods." In effect, a speaker appears to locate the same person, including himself, in different regions of the racial dimension on different occasions. Used in this way the racial dimensions does not determine membership in a set of fixed classes of people.

---

[4] Exception is made for hierarchial taxonomies. Such taxonomies do not themselves partition the set of objects but contain within themselves such a partition.

The question naturally arises of the relation between the two apparently conflicting uses of the racial dimension. Do they operate independently of one another; is the usage to be employed by a given person at a certain time determined randomly or according to some decision function of which even the arguments are unknown? Or is one of these usages basic and the other an alternative which operates only under certain specifiable conditions? I would incline toward the latter view. Here again the data are only anecdotal, but they are better accounted for by the second explanation.

An approximation to the unconscious model of social classification employed by Tahitians might be something like this: In general, an individual is assigned to the social category corresponding to his position on the racial dimension as determined by the perceptions of the classifier of his measures on the appropriate scales. However, there are only two cutting points on the dimension, determining three lexical categories. A situation may, therefore, arise in which the speaker wishes to distinguish the degree of nativeness-Europeanness of two individuals both of whom he would usually place in the same region. Since both individuals belong by definition to the same absolute racial category (as they are in the same region of the scale), one or the other must be assigned to a category other than his usual one.

As far as may be judged, this hypothesis is substantiated by the data. Every case I recorded of a surprising racial designation turns out upon later consideration to involve contrast of two actors in a dyadic interaction situation usually involving economic competition. The general pattern is to attribute the behavior of the actor whose part the speaker is not taking to some unpleasant aspect of the stereotype of the racial category to which that actor is assigned. The rule determining which actor is assigned to a category

other than his usual one does not appear to be connected with whose side the speaker is on. Evidently if the two actors being classified are both in either the lowest or the highest region, there can only be one choice for the unusual classification; the actor nearest the boundary of the middle region is displaced to that region. If both actors are in the middle region, one might suspect that the choice for unusual classification is made on the basis of which actor is closest to a boundary; in a sense, the choice is made which requires the least "stretching" of the racial dimension.*

The following anecdote will illustrate the kind of data relevant to the general hypothesis governing alternative usages of the racial dimension. One informant, with whom I had many conversations, explicitly classified himself as a mixed-blood. Also in normal conversation he implicitly and naturally assigned himself to this category in explaining to me the style of life of the "aborigines" in contrast to his own. Subsequently an unpleasant situation arose between this man and a European tenant of his concerning a rented property. Throughout the rather long and acrimonious dispute the informant reported each development to me, invariably putting a low value on the conduct of his tenant and attributing the tenant's immoral behavior to his race. So far, no individual's absolute racial classification had necessarily been overriden.

However, later the same informant became involved in a somewhat similar situation with an individual he classified as a mixed-blood. In this case, my informant did not hesitate to classfy himself as a Tahitian, in contrast to his mixed-blood antagonist, and explain the other's behavior in terms of the avariciousness and immorality of "the mixed-bloods."

Tahitian words for race and class appear to take their meaning from a

---

*[Footnote on page 10a]

(FOOTNOTE)

* Stephen Palmer has suggested to me that apparently inconsistent classification of a given individual across contexts might be due to a particular context's according greater weight to a given variable than another context gives to that variable. Thus, if the context of discussion is primarily economic, greater weight might be given to economic variables in determining racial classification than if the context of discussion involved, say, Biblical exigesis or athletic ability. According to this view, there is not a unique function from the underlying variables to the racial continuum but a number of such functions, the selection of which on a particular occasion of speaking would depend perhaps on both semantic context (e.g., topic of discussion) and pragmatic context (e.g., relations of interlocutors to each other). This seems a plausible suggestion; I have at present no way of evaluating it empirically.

quantitative dimension cut into three regions, each of which determines a racial category. By and large, individuals agree on the boundaries of the categories, but the exact extent and possible patterning of individual variation with respect to the placing of boundaries could only be determined by direct investigation specifically designed for that purpose. Since this model was constructed after the field work was concluded, no attempt was made to design or perform such an investigation. There probably exists variation among individual Tahitians with respect to the particular boundary points between regions of the racial continuum and perhaps with respect to other details of racial classification. The general framework presented here seems, however, to be widely shared.

III. Words for Race and Class in Brazilian Portugese

M. Harris has studied the words for race and class in Brazilian Portugese and reported his findings in an article entitled "Referential Ambiguity in the Calculus of Brazilian Racial Identity" (Harris 1970). The title expresses Harris's conclusion that Brazilian Portugese words such as branco, preto, negro, mulato, etc. are ambiguous. By "ambiguous" Harris does not mean what is usually meant by that term. Rather what Harris wishes to demonstrate is that these words are either (a) vague, (b) subject to wide interpersonal variation in meaning or (c) both. The distinction between (a), (b) and (c) is not important to Harris as he is not interested in language as a psychological phenomenon. His interests are purely social and he wishes only to show that words designating racial/social types in Brazilian Portugese do not constitute a shared semantic system which permits native speakers of this language to communicate effectively in this domain. In his concluding paragraph Harris speculates "...there may be a positive, conservative structural reason for

maintaining and maximizing the noise and ambiguity [in this semantic domain]... Objectively, there is a correspondence between class and race in Brazil...; the more negroid the phenotype the lower the class. Prevention of the development of racial ideology may very well be a reflex of the conditions which control the development of class confrontations... In Brazil racism and caste formation would unite the lower class" (Harris 1970:12). Harris seems to be saying that since clear talk about race and class might lead to class consciousness on the part of the oppressed black majority (and thence perhaps to revolution), the society or someone or something creates in the language a systematically confused semantic system regarding race and class that prevents such clear talk. Harris does not speculate on the agent or the process that might act upon the language acquisition process of each Brazilian child to bring about the unusual situation he claims to have discovered.

Surprisingly, such peculiar situations, in which certain social actors achieve a deliberate vagueness in language, are not entirely unknown. I. Zaretsky (1969) has demonstrated that systematic vagueness exists in the argot of Spiritualist churches.

> A particularly interesting example of an ethnosemantic analysis... of unquestionable sociological significance is that performed by Irving Zaretsky on the religious argot of San Francisco Spiritualist Churches (Zaretsky 1969). Zaretsky found that many common English words were being used in ways that were clearly not the normal meanings of those words. An ethnosemantic analysis of all such words revealed a subset with the following interesting properties. Words belonging to this set have unmistakable although vague connotations of supernatural "spirit forces" at work in the world, and in particular through the agency of the medium. However, the referential meanings of these forms are extremely nebulous. This is not a question of the investigator's being unable to discover the referential meanings but rather that Zaretsky has shown through detailed comparison of the contexts in which these expressions are employed that the referential meanings are in fact very vague. He then goes on to show two important ways in which the very referential vagueness of these words contributes to the perpetuation of the social organization in which they are employed. First, referential vagueness permits the medium to transmit acceptable messages from the spirits to a parishioner who has written down a specific question that the medium has not seen. For example,

Ethnographer: Can you recall any message you received in church recently?
Informant: Well, let's see...She told me I was going to get drafted...
Ethnographer: Did she actually come right out and tell you you will get drafted, just like that?
Informant: Well, not exactly. She kept muttering about how I have upset conditions around me and how she saw the color green in my vibration...She just went on like that for a while...Well, I know what she was talking about because I waited to hear from my draft board all this time...But I certainly didn't want her to tell me that I will be drafted...
Ethnographer: How did you know what she meant by condition, vibration, environment?
Informant: Well I didn't...No, actually it was obvious...I guess she could tell psychically what bothered me...Well, I don't know, but I think it was pretty obvious what she meant...[Zaretsky 1969:124].

This referential vagueness of the terms employed permits the ardent parishioner to endow them with whatever particular referential interpretation he wishes on each occasion of use. Secondly, this same referential vagueness protects the medium from the possible legal charge that she is engaged in the illicit activity of forecasting specific future events for money (Kay 1970:25f).

In the case of the Spiritualist churches it is clear who is causing the pathological linguistic vagueness and how they are doing it. The Spiritualist argot is essentially the property of the mediums and they teach it to neophyte parishioners. But who or what is teaching each little Brazilian child to use words for race and class in ways that are (a) excessively vague, or (b) different one from the other, or (c) both? Granting for the sake of argument that if such vagueness in fact existed in Brazilian racial terms it might in some sense support the political status quo, by what plausible mechanism could it have been brought about? The ruling classes, who are the ones Harris suggests are the beneficiaries of this alleged vagueness of language, are not in charge of the process of language acquisition of the black peasants and workers.

In short, lacking any mechanism that might bring about such a peculiar linguistic condition, Harris's assertion that systematic vagueness and/or uncontrolled interpersonal variation is the rule amongst Brazilian Portugese

terms for racial types is implausible and leads one to examine the method he used to arrive at such a startling conclusion. Harris's procedure was first to make "a deck of 72 full face drawings constructed out of the combination of three skin tones, three hair forms, two lip, two nose, and two sex types" (p. 2). Each subject was exposed to the thirty-six picture of his or her own sex, and was required to give the qualidade, typo, raca, or cor of the fictional person depicted. No indication is given of who thought up the stimuli nor of whether any checks were made to determine if the subjects thought the pictures looked like people. Of the eight examples reproduced in Harris's article some look to me, speakingly strictly intuitively, more like real people than others. It is clearly Harris's intention that each series of thirty-six pictures give some sort of representation of the diversity of Brazilian physiognomic types, but there is no evidence or argument that they in fact do so. In short, the stimuli lack face validity (in every sense) and Harris seems unaware that in studies of this type, where proxy stimuli are used--in this case artifically constructed pictures alleged to represent the domain of racial types rather than a careful selection of real persons or pictures of real persons--it is generally expected that a demonstration be given that the stimuli really are from the subjects' point of view proxies for what the investigator wishes to so interpret them.

One might guess the next step would be to elicit the words for racial types without using the pictures and then ask subjects which picture(s) best represent each word. Such was not the case. Rather Harris presented all the pictures to each subject and required the subject to assign a racial-type name to the picture. So if, as appears to be the case, many of the stimulus pictures are not even convincing renderings of Homo sapiens, much less good

exemplars of any particular Portugese racial word, the subjects are forced to assign them a racial word anyway.

Under these conditions, it is not surprising that subjects were less than unanimous in naming the pictures. It is precisely the lack of agreement among subjects in this picture naming task, however, that leads Harris to conclude that the Brazilian Portugese words for racial types are "ambiguous" and spurs him on to the socio-political speculations mentioned above.

So far we have considered only the extent to which Harris's stimuli represent real phenotypic diversity in the Brazilian population. But even if we were to grant for the sake of argument--and against plausability--that these stimuli do in fact satisfy this criterion, the study suffers from a more serious conceptual defect. Harris himself is aware intuitively that the meanings of the words in question have something to do with physical appearance but also a lot to do with socio-economic status. Harris says, "Many observers have pointed out the partial subordination of 'racial' to class identity in Brazil exemplified by the tendency for individuals of approximately equal socio-economic rank to be categorized by similar "racial" terms regardless of phenotypic contrasts and by the adage, 'money whitens'" (1970:1). But of course if the semantic domain in question contains a complex combination of factors regarding physiognomy, motor habits, facial expression, dress, income, speech, etc. and the stimulus materials provide information only on the first factor, it is distressingly easy to see why the subjects cannot apply the words more consistently to the stimuli than they do, and of course this lack of interspeaker agreement in classifying such impoverished stimuli allows no conclusion regarding vagueness or ambiguity in the words under study.

The study contains several other errors of method and interpretation, not

all of which need be considered here. One of these is pointed out by R. Sanjek (1971) who performed a follow up study using the same stimulus materials but supplemented by some other procedures. A student of Harris's, Sanjek, while curteously citing Harris's conclusion of inherent "ambiguity," found that inter-speaker variability was dramatically reduced when speakers from a single speech community were studied by the same technique. Harris had pooled the data of subjects from the states of Bahia, Alagoas, Pernambuco, Ceará, Brasilia, and São Paulo, in itself an explanation for the lack of intersubject agreement. Sanjek comments with commendable caution

> I am reluctant to claim that my analysis extends beyond my sample or, at most, beyond Sitio. Studies within the state of Bahia report terms which I did not find at all (Kottak 1963; Hutchinson 1957), even though I asked deliberately several informants if they knew them. The term mulato, which has both low salience and low level of agreement in Sitio, is no doubt of high salience in other parts of Brazil as Harris' data (1970) and a few tests I conducted with the picture set in Salvador and Rio de Janeiro suggest. I should add that several informants in Sitio mentioned in conversations a term which is used for preto in Vila do Conde, the nearest community, but which, they said, "we do not use in Sitio" (Sanjek 1971:1139).

On the whole, the data Sanjek presents on Brazilian racial terms are consonant with the kind of model proposed above for comparable terms in Tahiti, although Sanjek does not reach this conclusion. There appears to be a racial-social semantic continuum whose end points might be roughly glossed "Black/poor/uneducated" and "white/rich/educated." A particular individual's position on this continuum is calculated by means of a complex function taking a series of physical appearance variables and social variables - some discrete valued and some continuous - onto the continuum, which then represents some weighted average, loosely speaking, of all the constituent variables. I would suggest that this is the basic situation reflected by the common Brazilian metaphor "money whitens." "Racial" terms represent regions on this continuum, and they are no more ambiguous or vague than other words.

There is, moreover, evidence in the Sanjek article that Brazilian Portugese contains the same kind of double barreled usage of race terms as Tahitian. That is, these words may be used not only to designate a member of the class of people having a value within a certain range on the continuum corresponding to that word but also to indicate a contrast between two individuals whose absolute values fall within the same region.

> In presenting a cognitive map which I claim is shared in a modal sense (by at least five-sixths of my informants), I want to be clear that such competence does not have a one-to-one correspondence with verbal behavior. I believe rather that the expression of the cognitive classification is altered by environmental (situational, sociological) variables which are essential for an understanding of why any term is actually uttered. Such variables would include at least the economic class, the dress, personality, education and relation of the referrant [sic] to the speaker; the presence of other actors and their relations to the speaker and referrant [sic]; and contexts of speech, such as gossip, insult, joking, showing affection, maintenance of equality or of differential social status, or pointing out the referrant [sic] in a group[5] (Sanjek 1971:1128, italics added).

And Sanjek continues directly in a footnote,

> [5]In terms of this last context, on the basis of her fieldwork in Chile where a similar but less complex system obtains, Sister Jennifer Oberg has pointed out to me that identifying one actor as, say, moreno, may indicate merely that he is more "moreno" in appearance than others (Sanjek 1971:1142).

It appears that the model posited for the Tahitian data may apply in some degree to Brazilian words for race and class and possibly in the semantics of comparable domains in other languages as well. It would not be surprising if the model constructed for the Tahitian data did not apply point for point in the Brazilian situation, however. In particular, it is apparent that Brazilian Portugese, even if one takes a single speech community at a time, has a much larger number of racial terms than Tahitian. One possible difference in the underlying semantic model is that there may be more involved than a single race-class continuum. Nevertheless several of Sanjek's observations suggest

that central to this semantic domain is such a continuum and that an individual's value on this continuous semantic variable is a function of a number of other variables, some physical, some socio-economic, and some probably cultural, e.g., style of dress, manner of speaking, paralinguistics, body motion, and so on. It also seems clear that racial terms are used sometimes to locate a person on this continuum and at other times to express the relative position of two people on the continuum, leading to apparent inconsistencies in application of the terms. I do not deny the possibility that there may be some terms whose significance is more purely a matter of physical type than others, but I suggest that further empirical research on Brazilian terms for race and class might well be informed by this sort of model. In general in doing semantics we are not forced to chose between componential analysis on the one hand and on the other a claim that the domain contains a lot of referential ambiguity, whatever that may mean. There are probably many ways for a language to structure a semantic domain. Almost any sort of structure that is easily and naturally apprehended by the mind may perhaps serve as the underlying schema for a lexical domain. This hypothesis will be developed in the next section, particularly with regard to schemata that involve continuous quantities.

IV. Implications for Semantic Theory

The model developed in Section II to explicate the meanings of Tahitian words for race and class represents an approach to the semantics of words that differs from the standard, structural one. The standard approach to semantics characterizes the meaning of a word as a set of semantic features. This approach is exemplified in Bloomfieldian structural linguistics by Lounsbury (1964) and in generative linguistics by Katz and Fodor (1963). The set of

semantic features comprising the meaning of a word is usually interpreted as a set of necessary and sufficient conditions on the application of the word. (For a critical but generally favorable appraisal of the standard feature method, also known as componential analysis, see Lyons 1969. For a negative appraisal see Fillmore 1975.)

The explication given to the meanings of the Tahitian words discussed in Section II does not correspond to a set of semantic features. Rather what was presented was (1) a cognitive schema for conceptually organizing some part of the real world, i.e., the conceptual continuum of "race" in Tahiti, constructed from a weighted average of values on a variety of physical, social, and cultural dimensions, (2) a set of lexical categories, 'Tahitian', 'mixed-blood', 'European', and (3) a body of rules saying how the lexical categories may be applied to various parts of the schema.* Explication of word meanings in terms of schemata, sets of lexical categories, and rules for applying the latter to the former is quite distinct from the traditional semantic feature theory.

One may ask whether the analytical device adopted in the case of Tahitian words for race and class is merely an ad hoc convenience or whether it exemplifies a general approach that constitutes a serious alternative to the feature theory. I believe the latter is the case, that the cognitive-schema-plus-lexical-set-plus-rules-of-application model is a generally preferable alternative to the feature model. I will not, however, argue the issue in its

---

*Less detail was given about the rules, but it was indicated that such rules may utilize pragmatic information. Thus a mixed-blood person fairly low on the scale might characterize himself as a 'Tahitian' in order to distance himself socially from another mixed-blood hearer, but if the same speaker were to try to characterize himself as a Tahitian to an audience all of whose members were clearly more Tahitian than he, it would surely count as an attempt to ingratiate himself rather than as a ploy to distance himself from his interlocutors.

broadest terms here, as those arguments have been made by others (principally Fillmore 1974a,b, and particularly 1975 and the references cited there; also D'Andrade 1971 and H. Gladwin 1971). I will rather assume a framework in which a semantic account of a conceptual domain consists of (1) a cognitive schema (2) a set of lexical categories and (3) a body of rules specifying the conditions for felicitous application of the latter to the former.

Fillmore, to whom I am chiefly indebted for this view, uses the term 'scene' in roughly the way I am using 'schema'. The difference in terminological choice apparently has to do with the fact that Fillmore has in mind as paradigm examples the mental representations of prototypical action sequences, e.g., a commercial transaction, while the examples I have considered in detail--color kinship, ethno-biological categories--are based on schemata that lack the narrative quality that is connoted by 'scene.' I would propose 'schema' as the more general term, apt both for designating dynamic schemata, 'scenes', and non-dynamic schemata of the kind considered in this paper.*

Fillmore uses the term 'frame' for the set of lexical categories, emphasizing that these categories are related, not only to the schema that they index, but also to each other by virtue of the fact that use of any one of them activates the entire schema. Thus, for example, as soon as I mention a purchase, the prototypical commercial transaction schema is activated and questions or comments regarding the price, buyer, seller, etc. are in order.

I will retain Fillmore's use of 'frame'. Frames, on this view, constitute the relevant sets of lexical categories for semantic analysis. In some cases they will correspond to 'lexical domains' or 'semantic domains' as specified

---

*L. Bloom, P. Lightbown and L. Hood (1975:22ff.) find a notion of schema essentially the same as the one used here useful in modeling the acquisition of semantics by the child.

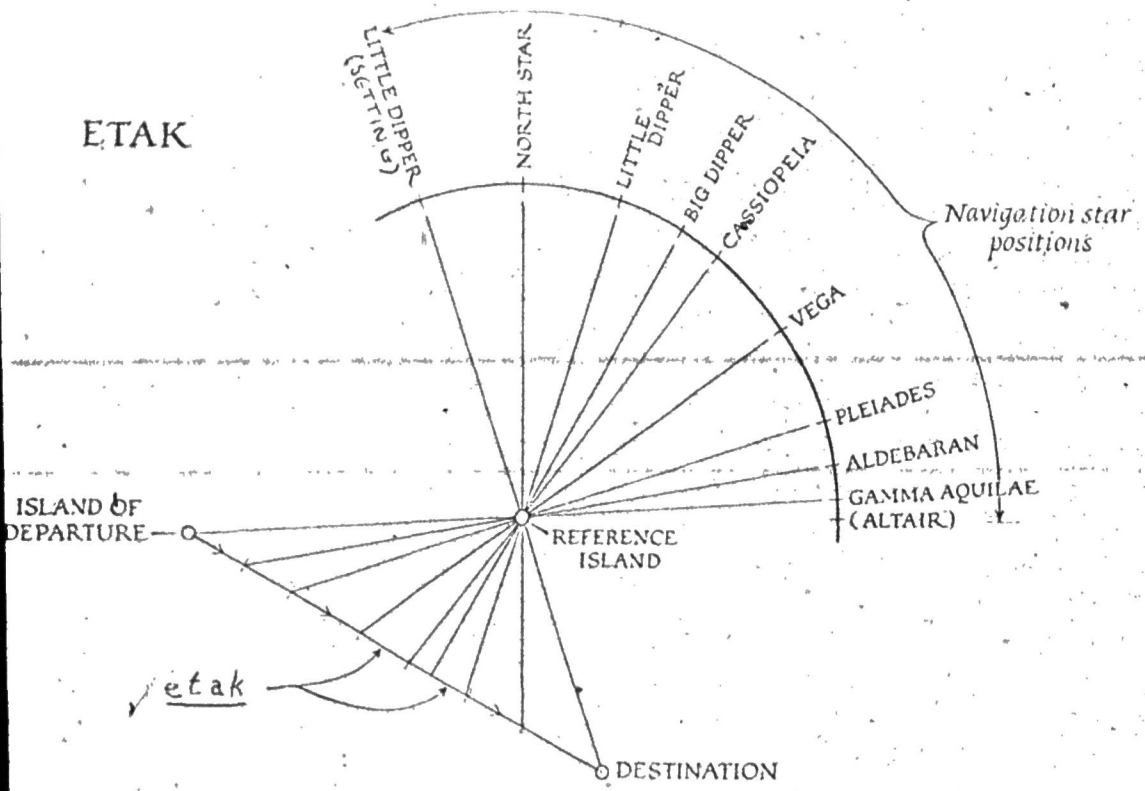
under the feature theory and in some cases they will not. For example, the kinship frame constitutes a semantic domain under the feature theory since its members may be thought of as sharing a features of meaning (Lounsbury 1964;1073), but the writing frame, containing as it does lexical terms like pencil, paper, language, message would not constitute a semantic domain under, say, Lounsbury's definition (cf. Fillmore 1975:125-6) unless one were to invent on entirely ad hoc feature like 'involved in the writing schema'.

Within this assumed framework I would like to address a more particular question: may a cognitive schema on which lexical meanings are based contain continuous quantities or scales? That is, may continuous quantities play a role in the meanings of words? I think there is evidence available from a variety of conceptual domains indicating an answer in the affirmative.

We turn now to some examples of quantitatively based schemata underlying the meanings of words. The first example is from T. Gladwin's interesting study of navigation on Puluwat in the Central Carolines (T. Gladwin 1970). One particularly important schema employed by these sophisticated navigators is represented in Figure 1 and is based on the local star compass, which in turn is based on the rising and setting positions of sixteen prominent stars distributed at unequal distances around the celestial horizon. One of the words defined by this schema is etak, a unit of distance of a sea voyage. In a particular voyage, a reference island is chosen so that lines of sight from the rising or setting positions of compass stars through the reference island to the boat's position will subtent roughly equal intervals on the line of the voyage. Each such interval is called an etak. Since the stars are not equally spaced around the celestial horizon and since it is in general not possible to find a reference island equidistant from the point of departure and the

destination, all etak are not equal in miles for a given voyage, and the etak of different voyages are not comparable to one another at all.

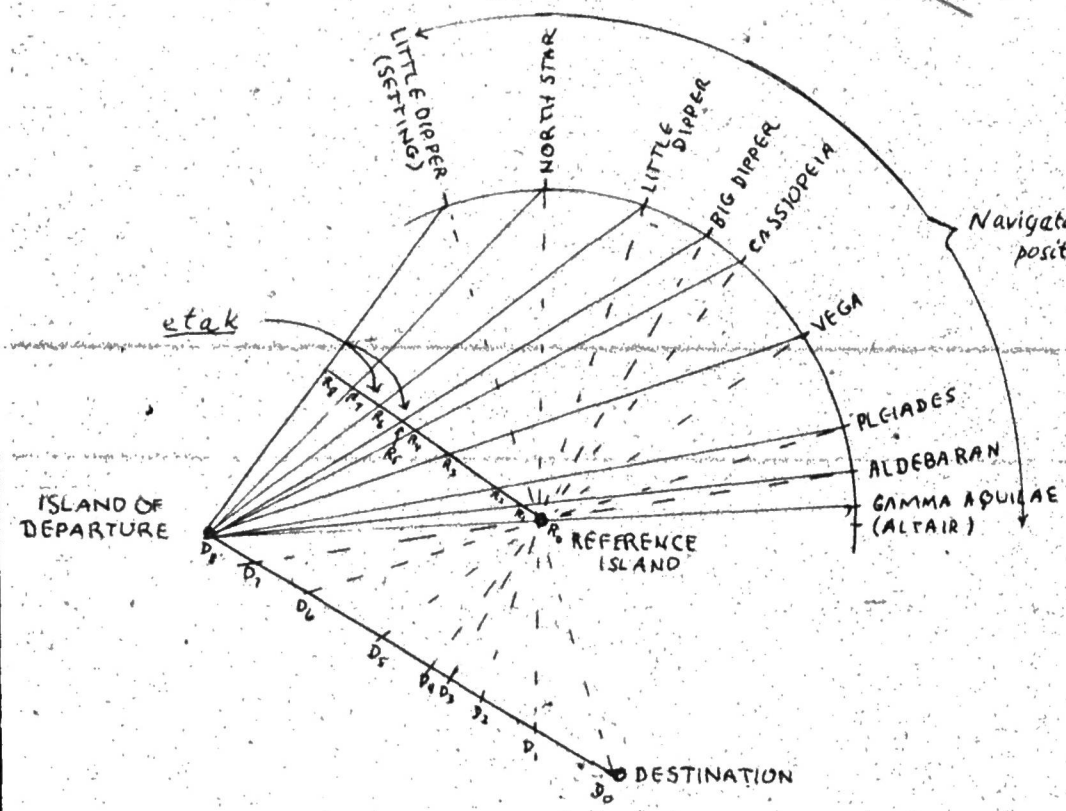
Figure 1, although it conveniently represents the content of etak in Western terms, does not correspond to the image Puluwatese navigators describe in explaining the meaning of etak to their native apprentices. Whereas in our image, the stars and islands are stationary and the boat moves, according to T. Gladwin (1970:181-9) the Puluwatese describe a related schema in which the stars and boat are stationary and the islands move. Gladwin assures us that the Puluwatese do not really think the islands move, but it appears more natural to them to represent the situation in terms of the imaginary movement of the islands than in terms of the real movement of the boat. Gladwin does not speculate on why this is the case.



A Western Version of the Schema Underlying the Puluwat Concept etak [1]

1. Source: T. Gladwin (1970: 185)

Figure 1



A Putatively More Puluwatese Version of the Schema Underlying the Puluwat Concept etak[1]

1. $D_0,\ldots,D_n$ represent successive positions of the moving destination.
   $R_0,\ldots,R_n$ represent successive positions of the moving reference island.

Figure 2 .

In Figure 2, I have drawn a picture that I believe accords better with the description in Gladwin's text of the Puluwatese schema underlying etak. Note that in this schema, with moving islands, the destination island has to move faster than the reference island. Gladwin does not say whether or not he interrogated Puluwatese navigators on this point. etak so calculated are fractions (though not equal fractions) of a voyage, in this case unequal "eighths." This seems sufficient to satisfy the Puluwatese conceptually and from other things Gladwin has to say about how the Puluwatese think about etak in particular and navigation in general it seems possible that the question whether the reference and destination islands move at different speeds may never have occured to them. Gladwin emphasizes the abstractness of this schema, which he calls a cognitive map, unifying several abstract concepts of the navigational system of this preliterate people. It is clearly a quantitatively based semantic schema.

Examples of quantitative schemata underlying lexical frames need not be sought in exotic languages. J.R. Ross (1970) has pointed out that implicitly comparative English adjectives such as great, large, big, tall, wide, thick, many, much, often, fast, and so on imply not only the underlying notion of a continuum but also the notion of a distribution over that continuum about some standard, perhaps mean, value and having a known measure of dispersion, perhaps variance or standard deviation. For example, if it takes me on the average thirty minutes to bicycle from my home to the university with a standard deviation of one minute, ninety-nine percent of the time I will make the trip in twenty-seven to thirty-three minutes (assuming the distribution of times is normal, which is not a matter of substance here). In any case, if I make the trip one day in twenty-five minutes, I am surely justified in saying, "I did it

fast today." If on the other hand a teenage boy takes anything between five seconds and an hour on the phone to work up to asking a girl for a date, averaging thirty minutes (and with, say, a standard deviation of twenty minutes), and one evening he accomplishes this feat in twenty-five minutes, I am not justified in saying, "He did it fast tonight." As Ross says, "The point is that such words [e.g., fast] presuppose a knowledge of the scattering around the mean" (1970:365). In our terms, the lexical frames in which such words participate index cognitive schemata which contain not only the notion of an underlying continuum but also a distribution over this continuum with parameters of central tendency and dispersion specified.

The use of this schema in common talk about ambient temperature furnishes a nice illustration of the frequently elusive distinction between semantic information (that which constitutes the meaning of linguistic expressions) and factual information about the nature of the world (see, for example, Fillmore 1974a:IV-5). This contrast is sometimes stated as the distinction between those items of information that should appear in the dictionary and those which should appear only in an encyclopaedic ethnography. The distinction is also roughly equivalent to what is expressed in semantic feature theory, misleadingly I think, as that between criterial and non-criterial features.

It has been my experience in living in places that differ not only in mean ambient temperature but in the variances of those distributions that the number of complaints one hears about the weather in these places is about the same. For example, upon moving from Berkeley to Honolulu I was at first struck by what seemed to me unjustified claims of the form "It's hot today" or "It's cold today" when the temperature seemed to me not sufficiently extreme to justify such comment. After living in Hawaii for a while, I apparently learned

unconsciously the factual information about the variance to be expected in daily temperatures so that I could accurately predict when people would complain about heat or cold and, if I wished, myself complain in a way that would not elicit contradiction from permanent residents. I would suggest an explanation along these lines. The use of hot in "It's hot today" means roughly 'more than some number n of standard deviations above the mean temperature of days in this place at this time of year'. (The precise number of standard deviations is not of course germane to the argument.) When I learned to complain about the heat in the way local residents of Hawaii do, I had not learned new semantics for hot and cold but had learned new information about the world. The semantics of hot and cold in the relevant contexts specify deviations from the mean in terms of numbers of standard deviations (or something comparable), but not in terms of absolute distance. To apply the words to the world correctly we have to know the relevent facts about the world, for example the number of degrees of the standard deviation of temperatures, but information about the numerical value of any particular standard deviation of temperatures is not part of the meaning of hot. As the seasons change or we move from place to place the meanings of the words hot, cold, warm, cool do not change; they remain defined in terms of the degree of variance from the mean of the relevant distribution. What changes is our factual knowledge about the relevant distributions.

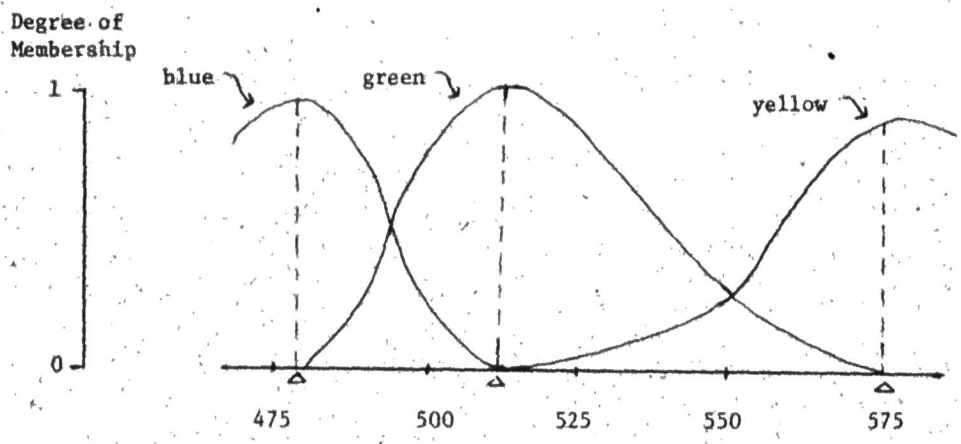
In the above account I have oversimplified somewhat, but I think the oversimplification does not invalidate the argument. It is probably true that people complain more about heat in the summer and about cold in the winter. Assuming, as is probably the case, that this does not reflect an asymmetery of daily temperatures about their seasonal means, it cannot be that the reference point about which variation is judged is simply the mean temperature

for the relevant time and place. Similarly it is quite possible to say felicitously, "Well it's hot again today, just like every day," if for example we are in Panama City. Clearly such a locution does not mean that the temperature is several standard deviations above the mean daily temperature for Panama City in Panama City every day. Probably what is meant in these cases is either that it is hot today relative to some point of ideal comfort rather than relative to the mean temperature, or that it is hot relative to some distribution of temperatures over a wider range of places and seasons, or some combination of the two. The point is that in each way in which the original model of temperature distribution, with its reference point and measure of dispersion is defective, the remedy is to propose some other distribution of temperatures as the contextually appropriate one. So the point remains that hot (or cold) are defined in terms of a schema that locates the temperature in question on a continuum of temperatures as further above (or below) a reference point than a certain number of standard deviations. Which is the relevant distribution is to be determined by the hearer from the context. That hearers are not always certain of the contextually appropriate distribution is evidenced by the following sort of dialoque, in which I have participated more than once. "Boy, it's cold today!" "Do you mean it's cold for Berkeley in the summer or just that we have cold summers?" The second speaker is asking implicitly for the correct distribution in which to interpret the word cold, yielding further evidence that it is in terms of a schema involving a statistical distribution that the meaning of hot, cold, warm, cool must be interpreted.

A third example of a quantitative schema underlying a lexical frame is words for colors. It has been argued in detail (Kay and McDaniel 1975) that the meanings of color words are best represented as functions that map points

of a three dimensional space of hue, brightness and saturation into the real interval [0,1]. This mapping is interpreted as a fuzzy set (Zadeh 1965, 1971a, 1971b). Thus a word like green is a fuzzy set, and the higher the number a percept is assigned by the relevant function, the better example of green that percept is. Figure 3 depicts schematically the green function and parts of the adjoining blue and yellow functions, where for convenience brightness and saturation are held constant. This model is supported by the available neurological evidence (DeValois et al. 1966; DeValois and Jacobs 1968) and psychophysical evidence; Kay and McDaniel 1975; McDaniel 1972, 1974, forthcoming; Sternheim and Boynton 1966).

A fourth example of a group of words defined in terms of underlying quantitative schemata are hedges and intensifiers such as kind of, sort of, somewhat, very, extremely, and so on (Zadeh 1972; G. Lakoff 1972). Zadeh and Lakoff treat hedges and intensifiers as predicate modifiers which operate on a fuzzy set changing the shape of the function. For example, the predicate modifier corresponding to the word very "squeezes" the function, as exemplified in Figure 4.



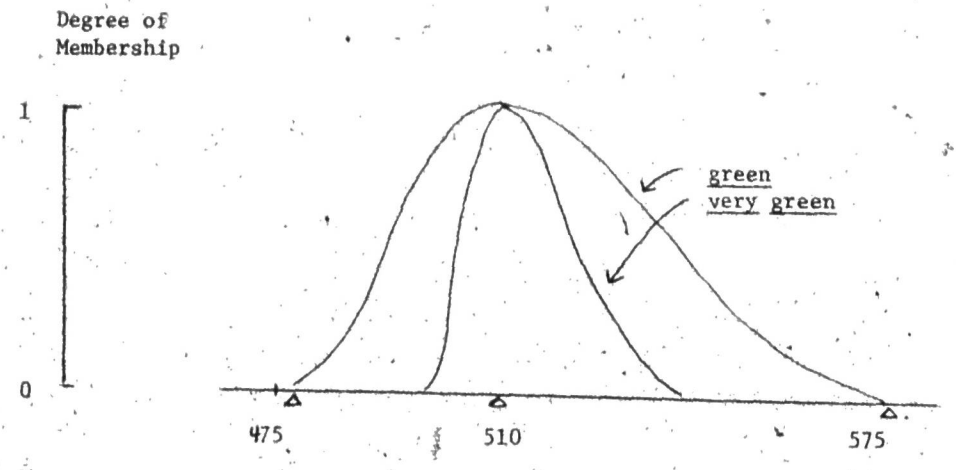
NOTE: (1) The detailed shape of the curves are not intended literally. Only the maxima and minima of the curves as drawn here correspond to established psychophysical parameters.

(2) The only pure green sensation(s) are those having a dominant wavelength of ca. 510 nm.

(3) Any sensation of dominant wavelength greater than 475 nm. (unique blue) and less than 575 nm. (unique yellow) is to some positive degree green.

Green

Figure 3



Green and Very Green

Figure 4

We have considered several examples of quantitative schemata underlying lexical frames. The examples considered, although few in number, came from both exotic and familiar languages (Tahitian, Puluwatese, Brazilian Portugese, English) and represent lexical domains that are both perceptual (color words) and highly conceptual (words for race/class). I have not argued here for the schema and frame theory of semantics as such, but only that the relevant schemata may in some cases involve continuous quantities. But if the examples I have given are correct even in their general outlines, then the feature theory, which is based on the presumption that basic semantic units are always discrete, cannot deal successfully with the lexical domains considered here.

References

Bloom, Lois, Patsy Lightbown and Lois Hood
1975 Structure and Variation in Child Language. ms. 58 pp.

Cliff, N.
1959 "Adverbs as Multipliers," Psychological Review, V. 66.

Code Civil
1959 Paris. Dalloz.

D'Andrade, Roy G.
1971 Modifications of the Feature Model. Paper read at the Annual Meeting of the American Anthropological Association. New York.

DeValois, R.L., I. Abramov and G.H. Jacobs
1966 Analysis of Response Patterns of LGN cells. Journal of the Optical Society of America 56:966-977.

DeValois, R.L., and G.H. Jacobs
1968 Primate Color Vision. Science. 162:533-40.

Fillmore, Charles
1974a The Future of Semantics. In C. Fillmore, G. Lakoff, R. Lakoff (eds.) Berkeley Studies in Syntax and Semantics. Dept. of Linguistics and Institute of Human Learning. Berkeley.
1974b Pragmatics and the Description of Discourse. In C. Fillmore, G. Lakoff, R. Lakoff (eds.) Berkeley Studies in Syntax and Semantics. Dept. of Linguistics and Institute of Human Learning. Berkeley.
1975 An Alternative to Checklist Theories of Meaning. In C. Cogan, H. Thompson, G. Thurgood, K. Whistler, T. Wright (eds.). Proceedings of the First Annual Meeting of the Berkeley Linguistic Society. Berkeley.

Frake, C.
1961 "The Daignosis of Disease among the Subanun of Mindanao," American Anthropologist. 63, 1.

Gladwin, Hugh
1971 Semantics, Schemata, and Kinship. Paper read at Annual Meeting of American Anthropological Association. New York.

Gladwin, Thomas
1970 East is a Big Bird. Harvard. Cambridge, Mass.

Harris, Marvin
1970 "Referential Ambiguity in the Calculus of Brazilian Racial Identity" Southwestern Journal of Anthropology. 36:1-14.

Hutchinson, Harry N.
1957 Village and Plantation Life in Northeastern Brazil. University of Washington Press, Seattle.

Katz, Gerald and Jerry Fodor
1963 The Structure of a Semantic Theory. Language. 39:19-29.

Kay, Paul
1963 Some Aspects of Social Structure in Manuhoe. Unpublished Ph.D. Dissertation. Harvard.
1970 "Some Theoretical Implications of Ethnographic Semantics". Current Directions in Anthropology (Bulletins of the American Anthropological Association, Vol. 3, No. 3, part 2).

Kay, Paul and Chad McDaniel
1975 Color Categories as Fuzzy Setts. ms.

Kottak, Conrad
1963 Race Relations in Arembepe. Columbia-Cornell-Harvard-Illinois Summer Field Studies Program. Mimeo.

Lakoff, G.
1972 Hedges, A Study in Meaning Criteria and the Logic of Fuzzy Concepts. Papers from the Eighth Regional Meeting Chicago Linguistic Society (ed.) P.M. Permanteau, J.N. Fevi, G.C. Phares. Chicago.

Lounsbury, Floyd
1964 The Structural Analysis of Kinship Semantics. In H.G. Lundt (ed.) Proceedings of the Ninth International Congress of Linguists. The Hague. Mouton.

Moench, R.
1963 Economic Relations of the Chinese in the Society Islands. Cambridge, Mass.: Unpublished Ph.D. Thesis (Harvard).

Ross, John R.
1970 "A Note on Implicit Comparatives" Linguistic Inquiry. 1:363-660.

Sanjek, Roger
1971 "Brazilian Racial Terms: Some Aspects of Meaning and Learning" American Anthropologist. 73:1126-43.

Sternheim, C.E. and R.M. Boynton
1966 Uniqueness of Perceived Hues Investigated with a Continuous Judgemental Technique. Journal of Experimental Psychology. 72:770-76.

Zadeh, L.A.
1965 Fuzzy Sets. Information and Control. 8:338-53.
1971a Quantitative Fuzzy Semantics. Information Sciences. 3:159-76.
1971b Fuzzy Languages and Their Relation to Human Intelligence. Memo. No. ERL-M302. Electronics Research Laboratory. University of California at Berkeley.

Zaretsky, I.I.
1969 The Message Is the Medium: an Ethnosemantic Study of the Language of Spiritualist Churches. Unpublished Ph.D. dissertation. University of California at Berkeley.