THE DENOTATA OF COLOR TERMS

Eric H. Lenneberg

Paper read before the Linguistic Society of America, August, 1953

Last year I had occasion to run a psycho-linguistic experiment which, I believe is not without interest to linguists. The problem for which I sought an answer was this: Apparently every language selects a certain limited number of sense data from the physical environment for which it provides a name or means of communication. Thus speakers are induced to refer to some data with greater frequency than to other potential sense data. This is a statement of fact. It does not imply that the speakers of a given language could not talk about anything they wanted to. It is one of our basic axioms that anything can be said in any language. Yet the fact remains that in any language at a given time certain states of the world are more easily communicated than others.

The question that I posed to myself is this: Is cognition at large, or is a particular cognitive process such as recognition or perception affected by the selectivity displayed by all languages?

To answer this it is first of all necessary to know objectively what the language selects. In other words, we have to be able to describe exactly what certain items refer to in the "real world." Thus we are immediately caught in a piece of semantic research which seems almost insoluble for the majority of word-classes and many aspects of language.

To avoid this difficulty, it is customary to take a short cut, that is translate a foreign language into the investigator's own language. While this is in many instances and for many purposes a necessary and sufficient method, it is important to realize that this method cannot help but distort our picture of the semantic structure of a foreign language. Translation does not give us the complete extent of physical objects or events to which a given word refers to. This situation can be illustrated

with kin-terms; in a given context the correct translation of a word may be uncle. Yet the original term may also cover that relative whom we call father, or it may perhaps be restricted to father's brother but not cover mother's brother. These distortions are avoided nowadays in anthropological research because a highly specialized metalanguage has been created with a one-to-one correspondence with the biological "real" structure. The biological structure is in a sense, a reality meter as Professor Voeglin has recently called it. Reality meters, or, as I would prefer to call it, extra-linguistic reference points can be established only in certain limited realms of the lexicon. The world of colors is such a realm (probably the simplest and most universal one, this being the reason why I chose it for my research) but there are a few other fields, especially where sense perceptions are described:

> shapes
>
> tastes
>
> sounds
>
> time-perception
>
> tactual perception
>
> space
>
> points
>
> continuity/discontinuity
>
> several million color distinctions

This then is the logical arrangement of the perceptual data sensed by us as colors.

Our experiment had two distinct phases: 1) one in which we ascertained the linguistic properties of the colors, i.e. we determine objec-

tively how speakers deal in their language with the physical correlates; and a second phase in which we use a new set of subjects who have to manipulate the physical correlates without, however, talking about them. The manipulation in our experiment consisted of recognition of the colors. Other experiments are in progress in which the manipulation is concept formation and learning. In phase two, we study whether subjects handle colors of a particular language property significantly differently from the way they handle colors with which their language deals differently. Let me describe phase one in greater detail.

The first task is to make a collection of color terms without any reference to the denotata. By surveying the corpus of material on a language or by eliciting color terminology from a sufficient number of informants, we compile that list of terms which is common knowledge to every speaker. Thus we obtain lists ranging from three to about 40 terms, the number varying from language to language. This is the preliminary linguistic work. On the side of the referents, the preliminary work consists of preparing a choice of colors in such a way that any color is likely to fit one of the color terms. Since the number of colors we can discriminate exceeds ten million, it is obvious, of course, that a collection of this magnitude is impractical and not likely to be necessary, inasmuch as no language has a million color terms. If, on the other hand, we had but a couple of dozen colors to fit to the compiled terms, we would force our informants to fit his terminology into rigid categories which is apt to do violence to his usage. Furthermore, if we wanted to operate from the start with, say 24 colors, how are we going to choose these colors from the wealth of possible colors? Our choice couldn't

help but being prejudiced by the color terminology and usage of our own mother tongue. In short, a compromise has to be made.

(Blackboard) Turn to color space.

Suppose we divide up this space into a number of equal, little volumes and that we took the midpoint of each volume as a color sample of all the color points in this volume, then we would have reduced considerably the population of colors, and could operate with a representative sample of all the colors. Fortunately, this job has been done for us by a scientific organization which is semi-officially attached to the Johns Hopkins University. It is the Munsell Color Company. These people produce color samples for scientific purposes, and they have also developed in the form of an atlas the type of sample that I have just described. The entire atlas (it is called the Munsell Book of Colors) comprises about 40,000 colors. There is a huge psychophysical literature on these Munsell colors, and several charts and tables are available with the help of which we may assign each Munsell color a specific point within the color space just described. In field work, we do not have to use the entire color atlas but could simply use

(Blackboard) a few slices out of this solid;

e.g. black to grey

show others

show my chart

With such a large variety of colors, subjects feel freer to describe colors in greater detail and will readily point out to you that you don't even have a good representative of color so-and-so. With the help of this type of color chart, it is usually easy for the informant to extrapolate from

this chart, and he can tell you whether color so-and-so is browner, greyer, darker, brighter, etc. than anything that is on your chart.

With the help of this type of chart, it is now possible to determine fairly accurately what particular color words refer to. What we have determined so far is merely a number of points in the color space for which there is a name. We do not know yet over what area each name extends itself, or in other words, how the entire space is structured in a particular language. This and a number of other problems can be answered by observing the following proceedure: Suppose we have compiled 10 of the most common color names and have obtained 10 points that have names in the color space; plotting these points in the coordinate system we obtain a certain distribution, which for English presents a cluster here and a rare area there. Since these clusters are unavoidable for reasons into which I cannot go now, it is important for us to know what happens to other parts of the space. Given a certain density of points in any part of the color space, let us interpolate so many points in the rare region of the space that we finally obtain an even and homogeneous distribution of approximately equidistant points throughout the space. For each point, we manufacture a cardboard card with the color mounted on it.

Show card.

Next we draw a sample of the speech-community of not fewer subjects than twice the number of cards which we have prepared, and administer the following test. The subject is tested for colorblindness and asked a few questions on his experience in color-descrimination. Then he is shown a chart containing all the colors for which there are cards.

Show random chart.

After removing the random chart each color card is presented to him separately and he is instructed to give the name of the color. The sequence of color presentation is randomized for each subject and a record is kept of the subject's entire response. Next, subject is given one of these charts.

Chart and vanilin

This is the end of the first phase of my experiment. As a sample of possible results let me tell you what happens in English.

(Describe on blackboard) Names--distribution

Maps--size and oscillation

Foci

Structure

These data help us to piece together the entire structure of the color space as conceptually and linguistically dealt with by English subjects.

I have described this phase of my experiment in such detail because of its relevance particularly to semantic theory, a few remarks on which shall conclude my paper. But before that, let me roughly outline what happened in the second phase of the experiment. Here subjects had to remember colors. They were shown some colors for a second and then, after half a minute they were required to pick out the colors they had been shown from a large chart containing 250 colors, systematically arranged. The subjects referred to the colors by numbers--no names were used in this phase. The results of this clearly show that colors which have a "good name" in English are recognized with greater facility by English speakers than colors with a "bad name." The many controls used and the psychophysical detail will be described in a paper to be published in one of the

psychological journals.

And now to the semantic implications. Without taking position on any side of the many guerrillas that are being fought in this field, let me comment on two of the most common approaches to the descriptions of meaning. In the first approach, the meaning of a given segment of speech is specified by linking it to some other segments or classes of segments of the same speech. This is the principle underlying Harriss's discourse analysis, most of the logicians' semantics, and even Osgood's semantic differential.

The medium of my colors were color cards. Suppose I had used instead of cards a number of women each of whom has a different hair color. In this case, I might have elicited some of the color terms elicited with my cards, yet the physical correlates would have changed considerably. The best example of this is _red_. The optical properties, in terms of my variables, of red hair, are significantly different from the optical properties of a red card in my sample. There are several other ways in which I could produce a change in the denotata, holding the significata constant. I have done this experimentally; I have given speakers only three cards: a reddish brown, a yellowish brown, and a muddy blue. When I asked my subjects some irrelevant questions about the three cards in their hands (about size, weight, etc.) they would invariably refer to them simply as the red one, the yellow (sometimes brown) one, and the blue one, yet in the earlier set-up these same cards did not once elicit the words red, yellow, or blue. I could go on with other examples of situations in which something is called _red_ which, in a different situation, would not be called _red_.

Thus it is clear, that even the most basic sense-words for which it is all too often assumed that their meaning could easily be established by ostention, i.e. by immediate relation to the physical environment, must, in almost all circumstances, also be defined contextually.

Thus the model (blackboard) of a closed system which has a few points of anchorage by which that system is coordinated to the physical environment is far too simple. The fine relation that exists between a language system and the real world to which it is synchronized, has not yet been explained satisfactorily. Whatever working model we use in our semantic theories, we should beware of postulating any fixed relationships that link points of the language system to points of the physical or social environment.