QUANTITATIVE LINGUISTICS

Vol. 13

Empirical Semantics II

A Collection of New Approaches in the Field

edited by

Burghard B. Rieger

Studienverlag Dr. N. Brockmeyer
Bochum 1981

Rieger, B. (Ed.): Empirical Semantics, Bochum: Brockmeyer 1981

THE STRUCTURE OF THE SUBJECTIVE LEXICON

Francis Jeffry Pelletier

Department of Philosophy,
University of Alberta,
Edmonton, Alberta, Canada

In the folklore from many different areas of personal advice and in the annals of academic wisdom of many different fields, one can find this anecdote:

> Three blind men were set the task of determining what entity was placed before them. The first blind man grasped ahold of a long, thin, pliable object which moved of its own accord. "It's a snake," he decided. The second blind man wrapped his arms around a very large, unmovable object. "It's a tree," he smugly affirmed to himself. The third blind man was in the meanwhile patting a large, flat, pliable object. "Beyond a doubt," he confidently decided, "it's a palm leaf." The three blind men argued about the true identity of the mysterious object until the elephant trampled them to death. You see, the first blind man had grasped the elephant's trunk, the second its leg, and the third its ear.

The morals drawn from this story vary from area to area. I intend to draw this one: One cannot understand the part without first understanding the whole. In the area of the present study, namely the area of subjective meaning of terms, this moral becomes: One does not fully understand the meaning of a term (word, morpheme) unless one has a firm grasp on the whole complex network of which that term is a part. One can perhaps have a partial understanding -- particularly when the term names some entity (e.g., the term 'Cleopatra's needle') or denotes a well-known group of physical object (e.g., 'dog'), but the more the term has no immediate correlate, the more its meaning must be understood solely in terms of the position it holds in (what shall be called) its immediate semantic network. More on this concept will be given below, but perhaps this much is suggestive.

One preliminary thing should be made clear. The present study does not attempt to define or characterize or elucidate the concept of meaning. The account to be given below of "semantic network" is not inconsistent with any of those theories of semantics and meaning which view as primary any of the following: (a) a relation between a word and a "thing" (object, set, idea), (b) a relation between a sentence and a communicative intention, (c) a relation between a sentence and a stimulus (or a response), (d) a relation between a sentence and its truth conditions, or (e) a relation between a sentence and a disposition to behave. Instead it is the concept of the subjective lexicon which is being here characterized. In this study, no view is taken as to whether the characterization of this psychological structure is to be seen as basic in an account or description of the semantics of language, or whether it is to be somehow "defined," "reduced," or explained by recourse to, say, an interaction between (b) and (e) of the above.

1. THE SUBJECTIVE LEXICON

As I mentioned above, I view an individual's subjective lexicon as a complicated system or network of related structures. By a structure, I have in mind something like this. The meaning of a given word or morpheme is located at an arbitrary (psychological) point. Some other words or morphemes are located fairly closely -- the closer the two words or morphemes are located, the more similar the individual perceives their meanings. Of course, "similarity of meaning" is not a simple concept; it is to be expected that word A can be more similar in meaning to word C than B is with respect to property P, but B might be more similar in meaning to C than A is with respect to property Q. So, to characterize this (psychological) distance which represents similarity of meaning, it is conceivable that we need

to account for an arbitrarily large number of possible properties along which words might differ in meaning. We shall call these properties "dimensions of meaning."

Given a group of n terms, there need be no more than (n-1) dimensions of meaning to characterize the total distances amongst terms. (And that would be the case when each pair of terms differs from each other pair along a unique or orthogonal dimension). Of course, it is not to be expected that this maximal characterizing number of dimensions of meaning will be required, since it is expected that speakers will (at least in part) compute the distance between A and B as a function of the distances that they each have to C. If they always do this, then the number of dimensions of meaning required to completely characterize the distances will be less than (n-1).

It is to be expected that, given a group of terms which they perceive as similar, speakers will have some fairly small number of dimensions of meaning along which they evaluate the similarity (or difference) of all the terms. This last holds for a group of terms that speakers think of as belonging to the same "field," that is, terms which are seen as (in some way) belonging together. In such a case, we shall say that the terms belong to the same "semantic domain" or "semantic field". It is also to be expected that psychological structures defined by two distinct semantic domains may be related by some dimension(s) of meaning, but it will either be that (a) the relationship is along one of the dimensions used in one of the structures and the dimension measures the distance between the two structures as very great compared to within-structure measurements, or (b) the relationship is to be measured along a dimension which plays no role in measuring the distances internal to each of the structures. Put spatially, each structure will be tightly knit, but the two structures are only loosely connected to one another. Considered together, the two structures form a superstructure or portion of a semantic network; the total semantic network is one's subjective lexicon.

Under this conception, the (psychological) meaning for a person will be a complex consisting of what a person thinks the

term denotes together with a specification of its position in his subjective lexicon. It should be noted, since the issue will arise again, that the meaning of a term for some arbitrary person can be independent of its meaning for any other person or its "meaning in language as a whole", except insofar as these latter may help condition the term's position in this person's subjective lexicon.

The theoretical position taken here is superficially similar to various approaches in modern linguistics. (These are recounted in Appendix I). But let me state a major methodological difference between the present study and these more purely linguistic theories of Katz, Bierwich, the generative semanticists, and the componential analysts. The methodology pursued by these writers is introspective in nature; that is, to discover what features a word (in one's own language) has, one consults one's one network. To discover what features a word in another language has, one goes through the procedure outlined in Appendix I for the Choctaw example. But even this last ultimately rests on introspection, for the final test is the hypothetical "if adopted, would these features lead English speakers to use English kinship terms in the same way that Choctaw speakers use Choctaw kinship terms?" And that question is resolved by the investigator consulting his own intuitions. When, one might ask, would such a procedure be acceptable? One case would be (for the same-language case) if all speakers had identical networks for their lexicon. Broadening this, the different-language case would require that there be a "universal semantics". However, we think it seems a plausible assumption (1) that one learns one's concepts (features) and their relationships solely from experience, and (2) that different people have qualitatively different experiences which are relevant to this learning. One can deny this assumption. In fact some transformational linguists seemingly do deny (1), even in the case of lexical structures, and claim that linguistically relevant concepts are innate (see LEECH 1974, FODOR 1977, KATZ 1966, 1972). One might instead deny (2) and claim that, so far as linguistically relevant experience

goes, it is qualitatively the same for us all (and that individual biological differences are irrelevant to language learning).

Both of these denials are seemingly difficult positions to maintain; nonetheless, one or the other denial seems to be required by the methodology of all these theorists who think that they can get empirical truth about the lexicon by introspection. The trilemma which faces all theorists who view word meaning as being somehow explicable in terms of semantic features or primes is this: Either (a) It is possible that different speakers have different semantic networks and the proper methodology for investigating the lexicon is to devise an empirical test to determine the differences, or else (b) Every speaker's language-learning-relevant experience is identical and there is some (language-independent) empirical method of discovering this, or (c) There is a "universal semantics" in the sense that there is a "fixed set of features common to all languages" and each speaker is innately determined to acquire that semantics.

The major methodological assumption of this study is that (a) of this trilemma is correct. The goal is to find a suitable empirical test; and a condition on the adequacy is that there be an explanation of how some people can sometimes communicate. In this assumption and goal we are at one with ANGLIN (1970) and FILLENBAUM & RAPOPORT (1971).

Given the preceding theoretical framework, and the assumption that a person A understands a person B's use of a term t only if the meaning of t for A is the same as the meaning of t for B, it ought to strike everyone as incredibly surprising that anyone ever understands anyone else. For, that would require that this incredibly complicated system of interrelated networks is the same for us all. In fact, many linguists, philosophers, and anthropologists have taken precisely this position (although sometimes for different reasons).

The usual way the argument for this position runs is this: (1) Everyone admits that (even within one's own idiolect, and surely within "language as a whole") words are ambiguous, have

different shades of meaning-relationships with other words, etc. (2) Because of our varied backgrounds, it would be an impossible coincidence if we all had exactly the same shades of meaning-relationships in mind whenever we used a given word -- even a fairly technical word. Therefore, communication does not even ordinarily succeed. Not only is there sometimes "lack of communication" and "misunderstanding" in spite of sincerity of intention, but this is by far the normal state of affairs.

One despairs at the picture presented by such thinkers. "Pop culture" tells us that there is sometimes "failure to communicate" in spite of sincere attempts on both sides; but the present outlook would have us believe that we always "fail to communicate". International misunderstandings we (hope we) can live with, but the idea that we do not "really" understand even our best friends ever, is depressing. The theoretical position presented by the given line of reasoning is clearly at odds with what we take to be social fact -- that language is an instrument we use for communication.

Is there any way to find out? Of course, merely pointing to the supposed fact that people talk with one another, of that they appear to learn, is not sufficient; for our present theorists tell us that this is all illusion. What is needed is some complex task which will give us an independent measure of the presumed network, so that we can see whether different people's networks are alike. I believe that the multi-dimensional scaling analysis (MDS, to be described below) does this. As a further side-benefit, the particular tests carried out in the present study can inform us of the effects of "education" on one's subjective lexicon.

Finally, consider once again the concept of a semantic network. Any "semantic domain", and therefore any member of it, can be viewed as being characterized by different "dimensions of meaning". As a simple example, consider a set of phrases of the form "colour word + shape word", so that we have black circle, red square, etc. Rather obviously, there are two prominent dimensions of meaning in these phrases: a colour dimension, and a

shape dimension. A group of temperature terms might be expected to have one dimension of meaning: degree of heat. Similarly, a group of kinship terms might be expected to have at least two dimensions of meaning: a (geneological) nearness one, and a sex one. But these are only the obvious dimensions. It is envisioned that, say, in a group of emotion terms, various (perhaps unexpected) dimensions could be used: a pleasant/unpleasant dimension, an active/passive dimension, a rejection/attraction dimension, etc. And of course, it is also likely that some of these dimensions are much stronger than others; further, different people may have different dimensions, or different strengths of the dimensions appearing.

If, now, a term is applied in a novel situation, it could be because one of its dimensions of meaning applies well, but its other dimensions may apply only because of a drastic attempt at forcing. Sometimes this forcing results in new insights and even genuine conceptual breakthroughs, as for example Einstein's use of simultaneity (of events) or Freud's use of motive (as applied to the unconscious).

What of the original problem: how is it that people can ever communicate, given that it is essentially a problem of coding (a portion of) one's semantic network and that no two semantic networks are the same? I view the problem to be essentially this. The following four propositions lead to the unacceptable conclusion that we never communicate with one another: (1) Our subjective lexicon is a semantic network, (2) This semantic network is acquired solely through experience, (3) No two people have precisely the same experiences, (4) True communication consists in coding a portion of one's subjective lexicon and having it uncoded into precisely the same network. From (1), (2), and (3) it follows that no two people have the same semantic networks, which, together with (4), yield the conclusion that even with sincere intent no one ever understands anyone else. Since this conclusion is absolutely unacceptable, there must be some false premise among these four. (1) and (3), at least for the present study, are not controversial, and so we shall not worry ourselves about them.

Notice that the intermediate conclusion, that it is not the case that people have the same semantic networks, is due entirely to premisses (1), (2), and (3). Given, as we are, the truth of (1) and (3), if we can show that the intermediate conclusion is true, we will have evidence to justify premise (2). Granted, it is not conclusive, but it is hard to see what other empirical predictions could be relevant.

As we shall see in what follows, this intermediate conclusion is supported, and hence we have good grounds for belief in premise (2). Now, since we cannot believe the conclusion, we must deny the only remaining premise, premise (4). It is the hypothesis of this work that if two people's semantic networks are "close enough" there can be genuine communication.

I do not wish this to be taken as a truism: I intend to show that there are large, definable groups of people, each group having essentially the same networks (as judged by MDS), and within these groups communication can carry on normally. More generally, we shall show that people who have more-or-less the same educational background in a rather specific area of word-meaning will form one of these groups. We also shall show that when tested in an area in which these same people have distinct backgrounds, they will not fall into the same groups. We take this to show that it is similarity of (learning) experience which determines the groups and not some "innate structure" -- not even "innate structures which are evoked by experience".

"Failures of communication" are therefore predicted to occur between people in different groups. And these "failures" can be documented on independent grounds -- they are not merely vacuous predictions. People within a group are to be thought of as "truly communicating" with one another; and the degree to which people between groups "fail to communicate" is a function of how unlike their semantic networks are -- that is, as a function of how dissimilar the meaning of a term they wish to use is for them.

There are two groups in the present study: (a) a group of professional philosophers, and (b) a group of university students

with no philosophical training. This study uses two groups of terms (or "semantic domains") common in philosophy: for the purposes of this report (and also to indicate the specific area of philosophy in which each domain is commonly employed) the domains are called 'action terms' and 'science terms'. With respect to group (a), the use of these terms should make the variablility of networks less susceptible to variation because of educational backgrounds and insure that the subjects have given some reflective thought to the meaning-relationships among the terms, since these are professionals dealing with a group of more-or-less technical terms in their discipline. The use of such terms ought to at least partly ameleoriate the charge of ULLMAN (1962, 249), who points out:

> That the neatness with which words delimit each other and build up a kind of mosaic, without any gaps or overlaps, has been greatly exaggerated. This is true only of specialized and rigidly defined systems such as army ranks; in ordinary language, vagueness, synonymy, ambiguity, and similar factors will produce a much less tidy picture.

It is not envisaged that even professional philosophers who have carefully thought about their meanings can use "technical" terms in a rigid manner. After all, the terms are borrowed from ordinary discourse and will still have vestiges of the original fuzziness about them; furthermore, even in technical discourse (except perhaps with respect to army ranks) terms are "extended" in meaning in various ways. Part of being a useful term is being usable in novel circumstances in such a way as to capture a part of the new circumstance, and bring the rest of the circumstance under the general purview of the term used.

2. METHODOLOGICAL REMARKS

There have been a number of studies of the sturctures of various semantic domains. Anthropologists have investigated the domains of pronouns and kinship terms (HAMMEL, 1965; ROMNEY and D'ANDRADE, 1964; WALLACE and ATKINS, 1960). But these kinds of studies have all been within the framework of componential analysis, with the kind of inadequate methodology mentioned in Appendix I. This is the case even in those works that are avowedly concerned with "psychological reality" as WALLACE (1965). Osgood and his associates (see OSGOOD, et al., 1957) have used antonymous rating scales to obtain "similarity of meaning" data, and also have investigated "semantic interaction" in an attempt to discover (or characterize?) possible vs. anomalous word combinations. DEESE (1965), CLARK (1968), HENLEY (1969), HENLEY, et al. (1968) have used association techniques to "map out" (association) relationships that hold for some given term. Miller has argued for the use of hierarchical clustering techniques in analyzing these sorts of data. But more closely related to the present approach is that of FILLENBAUM and RAPOPORT (1971).

The present report rejects the approach of DEESE and associates on two grounds. First, I desire to maintain a fixed semantic domain and discover the relationships that hold among the terms within that domain. Deese's method is open-ended in the sense that the subjects are allowed to "freely associate" any term and then a "mapping" is made of the "kinds" of associations performed (subordinate, coordinate, location, etc.). Secondly, and even more fundamentally, association techniques do not in any straightforward sense tell us about the meaning relationships which hold between terms for a subject. If one is asked to "freely associate" whatever he wants to a given term, it is obvious that the responses may be the result of many factors other than what the individual has in mind by the "meaning" of the stimulus word. (After all, this is why the technique is used in psychonanalysis.) And ignorance often enters into such associations: Bagdad has as its most common associate India, but it surely cannot be part of the meaning of Bagdad for anyone (see "Dictionary" in DEESE, 1965).

The present report follows FILLENBAUM and RAPOPORT (1971) rather than OSGOOD in characterizing the notion of a semantic space or network. Osgood required the subjects to associate pairs of bipolar adjectives (the scales) with selected nouns; a matrix of interconnections among the scales was computed and factor-analyzed to yield a semantic space of small dimensionality. The present report does not require the use of bipolar scales to determine the location of a term. Rather it asks the subjects to give similarity ratings of pairs of terms and uses multi-dimensional scaling (MDS -- to be described below) methods to determine a minimum dimensionality; the location of terms on these dimensions may then be analyzed to determine what bipolar adjectives might be associated with these dimensions. That is, rather than "promoting" the subjects by stating in advance what the dimensions might be, the subjects unwittingly construct their own dimensions. (It is assumed that the scales determined in this manner will yield the aforementioned "dimensions of meaning".)

Fillenbaum and Rapoport used three procedures to obtain data: (1) tree construction, (2) construction of complete, undirected graphs, and (3) direct grouping or classification. I reject the third procedure as it imposes (perhaps unjustifiably) a hierarchical structure on the data. The first two procedures are very tedious for the subject. They involve the assigning of a rank ordering to every possible pairwise combination of terms and also to every possible combination of groups of terms. This involves an immense number of judgments of relative similarity of meaning ((n-1)! judgments on n terms), leading to subjects' inability to make meaningful judgments, especially in the middle parts of the experiments. Furthermore, subjects are not allowed to assign the same rank to any two comparisons, even if they want to.

The present procedure differs first in only insisting on pairwise comparisons of meaning-similarity, thus reducing the number of judgments to n(n-1)/2. This is done by presenting the terms as an "upper triangular matrix" (minus diagonal). Secondly, subjects were first to scan the list and pick out the most similar pair and the least similar pair, and assign the values 1 and 9 respectively. They then had "anchoring points" to compare their later judgments against (since they had been forced to make choices for extremes); they were allowed to assign to different comparisons the same rating which, together with the small number of different scale values, make their comparisons meaningful; they had their attention focussed on elements of the given domain (instead of constructing "new" members of the domain by considering combinations -- partial trees -- of the original list); and they did not have to go back and change any ratings (as they might if they were to simply start making judgments and later come across a pair which was "more different" (or "more similar") than any they had encountered up to that point.

In analyzing their data, Fillenbaum and Rapoport pooled data from their subjects. This seems suspicious, in spite of their considerations (pp. 25-29), since it would be presumptuous to assume that the same features (dimensions of meaning) are equally prominent for different subjects. And given that the judgments are not unidimensional, even if all subjects agreed on what meaning-dimensions were important, they might very well disagree on the relative importance of them. In fact, it is a part of the purpose of this study to investigate the cogency of the assumption that people actually do agree on these dimensions.

At an intuitive level the MDS method can be understood like this. Suppose you want to locate, in space, a point. No dimensions are required to measure its distance from itself. Now suppose you add another point to the space. There are two possibilities: either it coincides with the first point or it doesn't. In the former case, one still requires no dimensions to measure the distance between the two points. In the latter case the distance can be measured in one dimension, call it A. If we add a third point, we again have two possibilities. Either it will fall on dimension A (the line connecting the first two points) or it won't. If it does fall on A, no further dimension is needed to measure the distances amongst the three points. If it doesn't fall on A, a new dimension B is called for; and the resulting measurements of distance can be described as falling on the plane AB. If a fourth point is now added to the space, it might fall on the plane AB, or not. If not, a third dimension C is required to measure all distances, and ABC will determine a three-dimensional space. A fifth point may or may not fall in the ABC space; if not, a fourth dimension D is required to measure all the distances, and ABCD determines a four-dimensional space.

An MDS analysis of n points starts by postulating the (n-1) maximal number of dimensions required to describe the distances amongst the n points and searches to see if they are all really necessary: Whenever two dimensions overlap -- that is, whenever the points being measured can be measured accurately without recourse to one of the dimensions -- the analysis collapses two dimensions (i.e., deletes one of them). Furthermore, when the deletion of a dimension will introduce very little measuring error, the dimension is deleted and the amount of error is recorded. This is done by deleting the dimension whose deletion introduces least error first; deleting the dimension whose deletion introduces next-most error second; etc. The error is cumulative and measured as stress (KRUSKAL 1964), which indicates how close the new interpoint distances are as compared to the original. Stress is a normalized sum of squared deviations (an analogue of the standard error of estimate in regression analysis).

Let $p(i,j)$ be a proximity measure on the points i and j, and $d'(i,j)$ be a function on $p(i,j)$ such that

(1) $d'(i,i) = 0$

(2) If $p(i,j) = p(k,m)$, then $d'(i,j) = d'(k,m)$

(3) If $p(i,j) < p(k,m)$, then $d'(i,j) > d'(k,m)$.

Let d(i,j) denote the distance between points i and j. Then for a specified number of dimensions r, the stress, S, for N points can be computed by (this account mirrors FILLENBAUM and RAPOPORT, 1971: 21)

$$S = \sum_{i=1}^{N} \sum_{j=1}^{N} (d(i,j) - d'(i,j))^2 \; / \; \sum_{i=1}^{N} \sum_{j=1}^{N} d(i,j)^2$$

Not surprisingly, stress increases with increase of N and decreases with increase of r. When the stress required to delete another dimension becomes "too great" we do not delete and treat that as our "optimal dimensional representation".

The present study uses the recommendations of STENSON and KNOLL (1969) in this regard. Further, the results reported in YOUNG (1970, Table 3) can tell us when the amount of stress is "too great" to permit a sufficiently high degree of metric determinancy in a given solution, and it turns out that using the value we have in this study does give a very high degree of metric determinacy. This is so because "in most standard cases of interest, the data so heavily overdetermine the solution that the particular measure of goodness of fit that (one chooses to) optimize generally has only a negligible effect on the result" (SHEPARD, 1969: 30, as quoted by FILLENBAUM and RAPOPORT 1971: 22).

The distance between any two points (as used in the formulae above) can be measured in a variety of ways. The above discussion rather implicitly assumed that the measurement will be Euclidean -- that is, that the distance between points i and j, d(i,j), can be measured from a third point k by the formula:

$$d(i,j)^2 = d(i,k)^2 + d(j,k)^2 - 2 \cos A \; d(i,k)d(j,k)$$

where A is the angle made at the juncture of the lines from i to k and j to k. One might also use non-Euclidean measures of this distance. However, it seems to follow from SHEPARD's (1969)

work that the Euclidean metric is very robust so long as one can assume a continuous underlying space of some well-defined dimensionality. As the theoretical discussion of a subjective lexicon implied, we take this assumption to be part of the concept of a subjective lexicon. With this assumption, "the data so overdetermine the representation that erroneous assumption of a Euclidean metric will still permit a satisfactory recovery of the true underlying structure and, indeed, even a determination of the unknown metric" (SHEPARD, 1969: 34; as quoted by FILLENBAUM and RAPOPORT, 1971: 24).

A multi-dimensional scaling analysis which allows for the assessment of individual differences (e.g., the INDSCAL program) can intuitively be understood as being a way of finding an "average" common-word-space for all the subjects rating the same group of terms. The words can then be placed in this common word space and then an analysis of the dimensions used can be given which shows how those terms relate to this common-for-all-subjects word space. The subjects are assigned scores which amount to a measure of the salience of each dimension for them. (Of course, this intuitive explanation is not strictly speaking correct. Actually, the within-subject normalized similarity scores are all used to construct an overall word space for all terms in which they can be located. Then a weighting factor is assigned to each subject for each dimension; this weight measures the contribution of that dimension in accounting for the scores of that subject).

The possibility of using hierarchical clustering analyses (HCP -- described below) on the data obtained will also be investigated. However, it is not expected that this method will be necessarily appropriate, since hierarchical clustering is most appropriate with data that can naturally be assumed to be grouped into superordinate/subordinate classes; and the present domains are not expected to be structured this way. It is possible, though, that some new information can be forthcoming about the underlying structure of these domains if the hierarchical clustering results are consistent with the MDS results. Roughly

speaking, this means that given a low-dimensional representation of the data, the clusters indicated by the hierarchical grouping are also "natural, simple and compact" in the dimensional representation (cf. FILLENBAUM and RAPOPORT, 1971: 39).

However, an HCP analysis of the weightings given to each subject by the INDSCAL analysis can show us groups of subjects who view the dimensions in a similar way. This use of HCP can be very useful in describing those subjects who perform the same in their judgments of similarity.

An HCP analysis of the similarity-of-meaning data is a way of converting it into a taxonomic system. Intuitively what happens is this: as input we are given a distance measurement from any term to any other. We search for two terms which are most similar to each other, that is, they are most similar to each other when one compares their distances from all the other terms. After these two terms are found they are grouped together and considered a new term with a specified distance from all other terms, while the original two individual items are deleted. This procedure is repeated until there is but one term (group) left. The output of such a procedure is a hierarchical classification of terms into groups ("clusters") and sub-groups, and, in the method used in this thesis, with a measure of how much "error" is introduced by deleting the terms in favour of the construction of a new term, the "cluster".

The method must specify for a newly-constructed cluster, a distance from it to each of the remaining terms or clusters. There are at least three reasonable ways to do this. The distances from the cluster constructed out of i and j to some other element k might be taken to be the minimum of the distances from i to k and j to k, the maximum of the distances from i to k and j to k, or the average of the distances from i to k and j to k. That is, where d(i,j) is the distance from i to j, and d((i,j),k) is the to-be-constructed distance from the cluster (i,j) to k, we can choose any of these:

$$d((i,j),k) = \min(d(i,k), d(j,k))$$

$$d((i,j),k) = \max(d(i,k), d(j,k))$$
$$d((i,j),k) = (d(i,k) + d(j,k))/2$$

In general, choice of constructing new distances will yield different hierarchical solutions for the same data. (See FILLENBAUM and RAPOPORT, 1971: 31-32 for an example demonstrating that the "min" and "max" choices yield distinct hierarchies). In the present study the method of maximizing was used. (See JOHNSON, 1967, for reasons to choose this so-called diameter method rather than the minimizing ("connected") and averaging methods).

MDS and hierarchical clustering procedures represent distinct structural models according to which one analyzes the similarity-of-meaning data. Each procedure analyzes the data according to a different goal, and it is clear that choosing to do one rather than the other amounts to a commitment that the underlying, true structure of the data has a certain form. Any set of ordinal procimity data can be subjected to a hierarchical clustering analysis, and a certain hierarchy will be output. However, the interpretability of the resulting hierarchy is not only a function of which hierarchy is the output of the analysis, but also of the a priori plausibility of the assumption that the underlying, true structure is hierarchical. So, from a mathematical point of view, hierarchical clustering analyses make very weak assumptions about the data -- only that they are ordinal; however, the method makes a very strong, and somewhat implausible (from the point of view of the structure of the subjective lexicon), assumption about the underlying structure of the data -- that it is class-inclusive in nature. MDS, on the other hand, makes comparatively strong assumptions about the data: that there is a continuous underlying space of some well-defined dimensionality, and that the distances between points are "additive" in the same dimension (indeed, in the present study, Euclidean additive); however, these mathematically strong assumptions have a high degree of a priori plausibility about the true underlying structure of the subjective lexicon.

For any individual and domain, the term which has the lowest mean score is that term which the subject sees as being the most similar in meaning to the other terms in that domain. The term with the smallest standard deviation will be that term which is viewed as the most "stable" in meaning of the terms in that domain. So analyses of these more traditional sorts can also yield valuable information about how the subjects perceive the semantic domains. By pooling these similarity judgments, we can arrive at an overall semantic space for all subjects (that is, a low-dimensional representation of the "average" low-dimensional representations). Against this space we can plot subjects to see whether they form distinguishable groups. We can therefore also discern whether these groups correspond to the various schools of philosophy (for those judgments of the professional philosophers), and whether the "educated" philosophers' groups correspond to those of the naive subjects.

3. SUBJECTS AND PROCEDURE

The terms used in this study are given in Tables 1 and 2. The primary sample consisted of 10 professional philosophers, all of whom had received their Ph.D. in philosophy and all of whom were native speakers of English. There were 33 students in the secondary sample, all native speakers of English and none having taken a philosophy course. (Further details, including raw scores, are available from the author.)

It should be noted that the subjects were given instructions which call for judgments of similarity of meaning without further clarification of that concept. Many subjects, especially among the philosophers, requested clarification of this. None was given other than "Use whatever criteria seem appropriate to you for this task."

Table 1: Action Terms

1. ACCIDENTAL	7. INADVERTENT
2. CAUSED	8. INVOLUNTARY
3. COMPELLED	9. MECHANICAL
4. DETERMINED	A. SPONTANEOUS
5. EFFORTLESS	B. THOUGHTLESS
6. FREE	C. UNCAUSED

Table 2: Science Terms

1. ANALYTIC	C. NECESSARY
2. A PRIORI	D. OBJECTIVE
3. CONCEPTUAL	E. OBSERVABLE
4. CONTINGENT	F. PROBABLE
5. CONVENTIONAL	G. REASONABLE
6. EMPIRICAL	H. REGULAR
7. ESSENTIAL	I. SUFFICIENT
8. GIVEN	J. THEORETICAL
9. HYPOTHETICAL	K. TRUE
A. LAWLIKE	L. VALID
B. NATURAL	

The philosophers were asked various questions designed to place them within one or another tradition in, or school of, philosophy. This was ascertained by direct questioning and comparison of this answer with their dissertation topics and published research areas. The results of this are shown in Table 3.

Table 3: Classification of Philosophers and their Fields of Specialization

Subject No.	Fields of Specialization
1	Analytic: Phil.Mind, Phil.Educ., Phenomenology
2	Analytic: Phil.Mind, Ethics
3	Analytic: Marxism, Aesthetics
4	Analytic: Phil.Mind
5	Analytic: Aesthetics
6	Analytic: Mediaeval Phil
7	Analytic: Phil.Mind, Phil.Lang.
8	Analytic: Marxism, Phil.Lang.
9	Analytic: Phil.Religion
A	Analytic: Phil.Religion, Metaphysics

As mentioned above, the present task is considerably simpler than various other ways of measuring "similarity of meaning". On n terms, this task requires $n(n-1)/2$ judgments; undirected tree-construction requires $(n-1)!$ judgments; and sorting requires something between these two numbers (depending on the sorting strategy involved). Furthermore, these instructions require the

subject to first pick out a "maximally similar" and a "maximally dissimilar" pair first, thereby giving "anchoring points" against which to make the remaining judgments -- something which is lacking in both sorting and tree-construction tasks. As an added benefit, the present procedure, unlike the others, allows subjects to rate two distinct pairs as "equally similar". And finally, it seems to be marginally preferable, all other things being equal, to ask subjects for their judgments directly, rather than by the devious, indirect methods invoked in the other two procedures mentioned.

The results of the similarity judgments were fed into an INDSCAL program on the University of Alberta AMDAHL 470V/6 to yield a common-to-all-philosophers space for terms (for each of the two groups of terms). Against this space, the philosophers are mapped, so that comparisons amongst philosophers can be perceived. An HCP then took their individual weighting factors on such a space and formed groups of philosophers (for each of the two sets of terms) in accordance with whether they had the same judgments of similarity of meaning of those terms or not. A typical member of each group was then chosen for individual study. This individual study was done rather "clinically" -- such a typical philosopher had a MDS program run on his responses, and his subjective lexicon was intuitively characterized (by considering properties of his "dimensions of meaning"). Furthermore, the means, standard deviations, etc., of the similarity ratings of the individual terms of these typical philosophers were studied in order to more fully characterize his view of that set of terms. The same procedure was followed with the group of students, although in this report only the primary sample of philosophers will be discussed. Finally, each of the "typical philosophers" was examined by comparison with the HCP groups formed from the students' INDSCAL weighting factors to determine whether the students' groups are the same as the philosophers' groups.

4. RESULTS AND DISCUSSION

The results of the questions asked of the philosophers to determine their fields of specialization are summarized in Table 3.

As discussed above, an INDSCAL analysis is a form of multidimensional scaling where a common space of low dimensionality (here: three dimensions) is found for the terms being judged, and the subjects are located in that space. The INDSCAL analysis therefore takes the raw scores and computes a space containing (say) three dimensions according to which all subjects are judging the terms. This space can be constructed from the x-weightings (i.e., the coordinates of the terms being analyzed in their object space) of the terms in three dimensions. The subjects have a score, called a w-weighting, which is a measure of the extent to which they attach importance to each of the three dimensions.

The degree of reconstructability of the original scores from the normalized x- and w-weightings can be seen from Table 4. With the exception of Subject 10 on Science Terms, the correlations are quite high. These normalized scores yield Figures 1 to 12: the first three are plots of the philosophers in the common word space for Action Terms, the next three are plots of the philosophers in the common word space for Science terms, the next three are the positions of the Action Words in the common space, and the last three are the positions of the Science Words in their common space. Superimposed on the first six plots are the groups indicated by the HCP (described below).

From Figures 7 to 9, a rough characterization of the three dimensions for the semantic space of the Action Terms can be given. Dimension 1 for Action Terms arranges the words from a high score for COMPELLED, DETERMINED; slightly less high for CAUSED, MECHANICAL; to low scores for FREE, SPONTANEOUS, UNCAUSED; and slightly less low for ACCIDENTAL. It is clear that this (primary) dimension is a measure of <u>necessitation</u>. Dimension 2 for Action Terms arranges the words from a high score for

Table 4: Correlation Between Original Data and Scores Computed from x- and w-Weightings

Subject	Action Terms	Science Terms
1	0.859	0.798
2	0.839	0.631
3	0.800	0.700
4	0.741	0.668
5	0.815	0.797
6	0.800	0.698
7	0.805	0.740
8	0.833	0.621
9	0.686	0.869
A	0.779	0.442
Average	0.796	0.696

FREE, EFFORTLESS; less high for SPONTANEOUS; still less high for CAUSED; to a low for INVOLUNTARY; slightly less low for ACCIDENTAL, THOUGHTLESS, UNCAUSED, INADVERTENT. This (secondary) dimension is a measure of the purposefulness of (human) action. Dimension 3 for Action Terms places UNCAUSED high, COMPELLED, DETERMINED less high; it places THOUGHTLESS, INADVERTENT low, ACCIDENTAL, EFFORTLESS slightly less low. This (ternary) dimension is difficult to interpret, but seems to run from "out of one's control for some external reason" to "within one's control but not attended to". (So it seems that EFFORTLESS has two vectors of meaning illustrated by its placement on dimensions three and two. In the second dimension it means: easy vs. difficult to do. In the third dimension it amounts to: done without exerting any purposeful effort of any sort vs. done without exerting any effort at accomplishing that particular effect -- although some effort may be involved in trying to do maybe something else.) Perhaps both high and low scores are excuses for an action, ranging from "done because of some external reason" vs. "done unawares". Middle scores would then be terms that 'excuse' does not apply to.

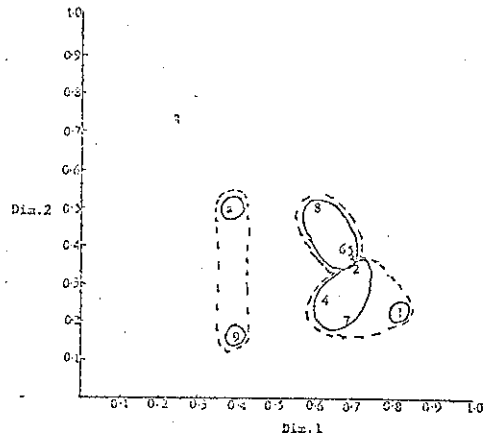

Fig. 1. INDSCAL Analysis: Philosophers in Common Action Term Word Space, Dimensions 1 vs. 2. 5-group HCP solution in solid outline. 3-group solution in broken outline.

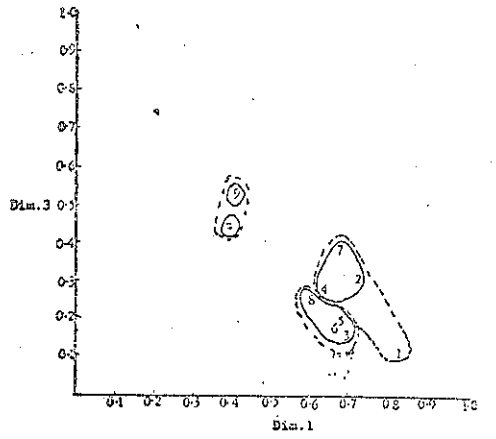

Fig. 2. INDSCAL Analysis: Philosophers in Common Action Term Word Space, Dimensions 1 vs. 3. 5-group HCP solution in solid outline. 3-group solution in broken outline.

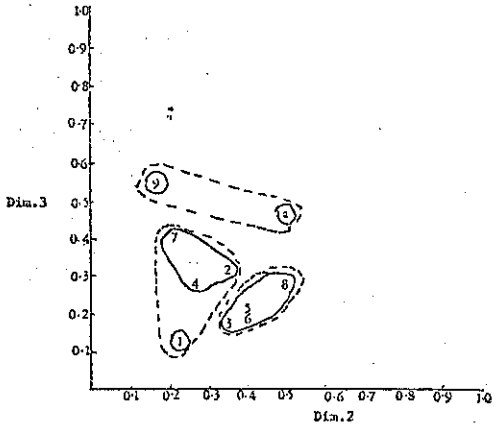

Fig. 3. INDSCAL Analysis: Philosophers in Common Action Term Word Space, Dimensions 2 vs. 3. 5-group HCP solution in solid outline, 3-group solution in broken outline.

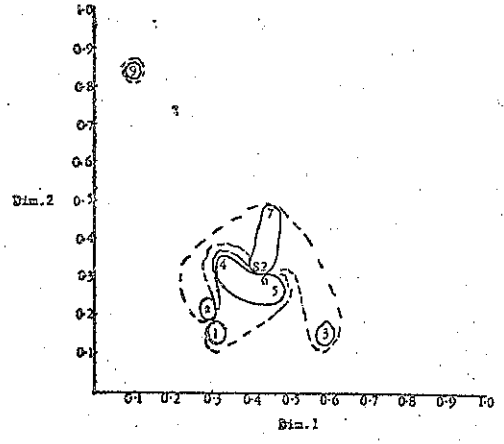

Fig. 4. INDSCAL Analysis: Philosophers in Common Science Term Word Space, Dimensions 1 vs. 2. 6-group HCP solution in solid outline, 3-group solution in broken outline.

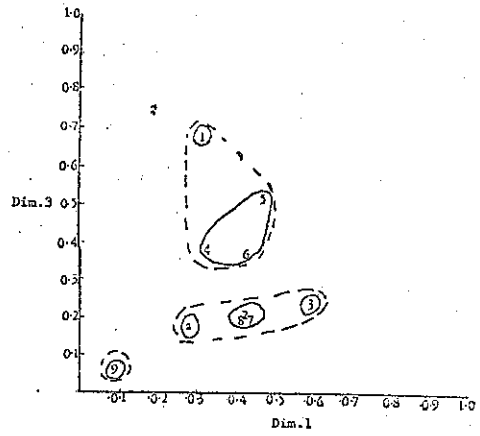

Fig. 5. INDSCAL Analysis: Philosophers in Common Science Term Word Space, Dimensions 1 vs. 3. 6-group solution HCP in solid outline, 3-group solution in broken outline.

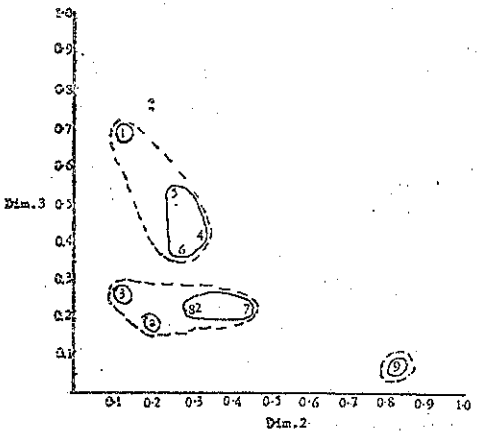

Fig. 6. INDSCAL Analysis: Philosophers in Common Science Term Word Space, Dimensions 2 vs. 3. 6-group HCP solution in solid outline, 3-group solution in broken outline.

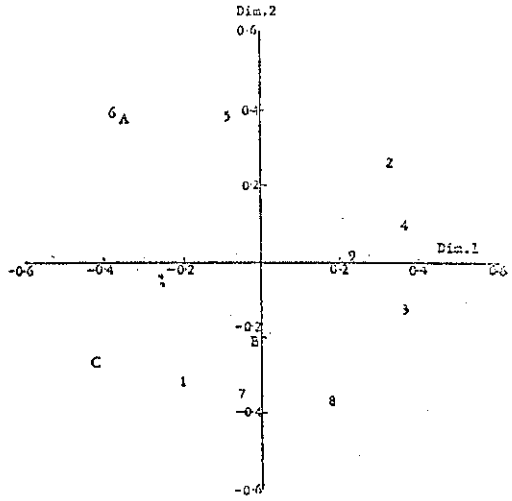

Fig. 7. INDSCAL Analysis: Action Terms Placed in Common Action Term Word Space, Dimensions 1 vs. 2.

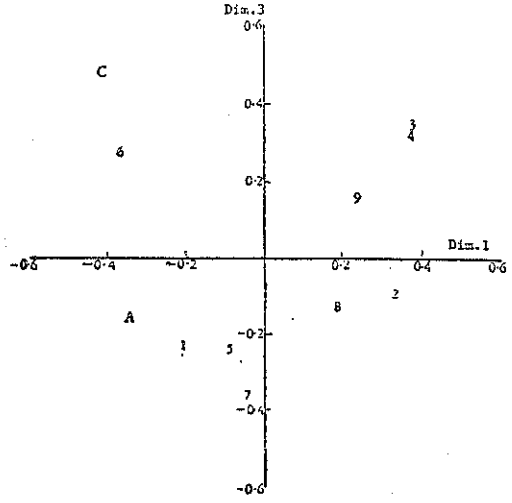

Fig. 8. INDSCAL Analysis: Action Terms Placed in Common Action Term Word Space, Dimensions 1 vs. 3.

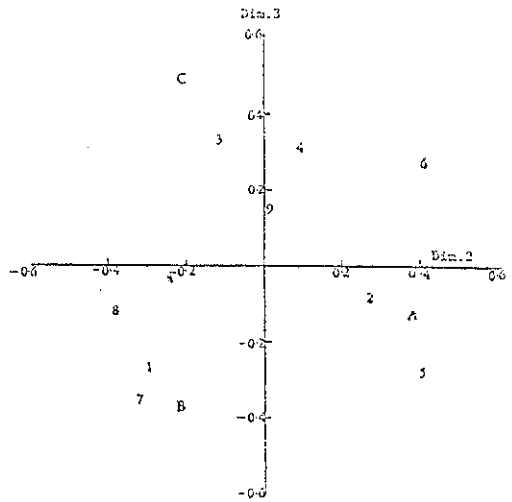

Fig. 9. INDSCAL Analysis: Action Terms Placed in Common Action Term Word Space, Dimensions 2 vs. 3.

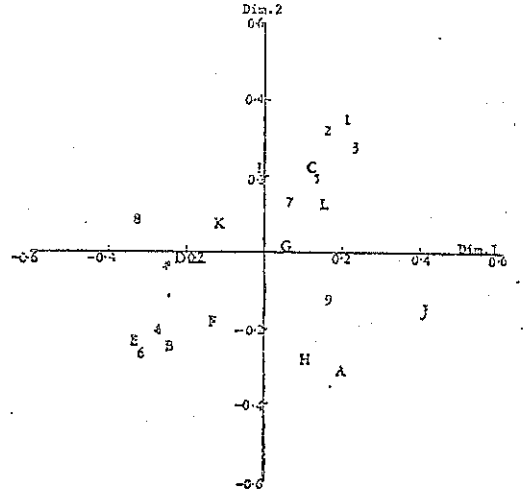

Fig. 10. INDSCAL Analysis: Science Terms placed in Common Science Term Word Space, Dimensions 1 vs. 2.

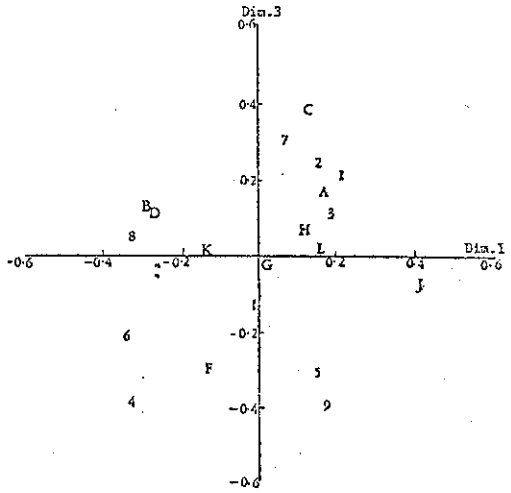

Fig. 11. INDSCAL Analysis: Science Terms Placed in Common Science Term Word Space, Dimensions 1 vs. 3.

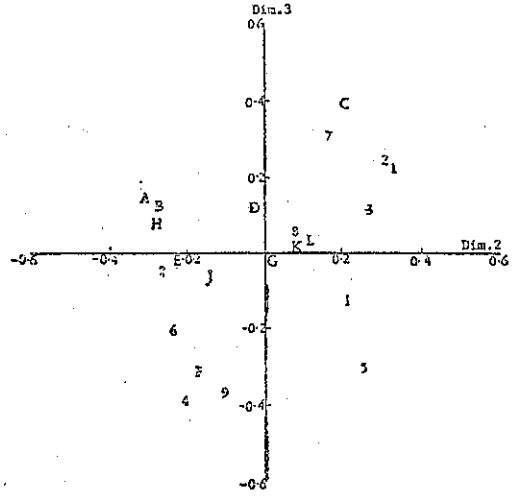

Fig. 12. INDSCAL Analysis: Science Terms Placed in Common Science Term Word Space, Dimensions 2 vs. 3.

It can be seen from Figures 1 to 3 that the different philosophers accord more importance to some one (or two) of the dimensions than to the others. For example, Subject 1 gives little importance to either dimension 2 or dimension 3. By far the most important for this subject is dimension 1. On the other hand, Subject 9 accords little importance to dimension 2, somewhat more to dimension 1, and by far the most to dimension 3. Subject A accords approximately equal weight to all three dimensions. The two groups (indicated on the plots of Figures 1 to 3) consisting of Subjects 2,4,7 and Subjects 3,5,6,8 differ in that the latter views dimension 2 as more important than the former does (on the whole) and the former group views dimension 3 as more important than the latter group does (on the whole). This is most clearly shown in Figure 1.

From Figures 10 to 12 we can get a rough idea of the three dimensions which characterize the philosopher's semantic space for Science Terms. Dimension 1 ranks terms from a high THEORETICAL, lower for CONCEPTUAL and ANALYTIC, still lower for HYPOTHETICAL, A PRIORI and REGULAR; to a low EMPIRICAL, OBSERVABLE, GIVEN (and slightly less low CONTINGENT). Clearly, this primary dimension rates terms as to how much "inference" is involved, or how far removed from empirical, observable reality. Dimension 2 ranks A PRIORI, ANALYTIC very high, CONCEPTUAL slightly less high, NECESSARY, CONVENTIONAL, SUFFICIENT less, ESSENTIAL still less; it ranks LAWLIKE lowest, REGULAR, NATURAL, EMPIRICAL slightly higher, CONTINGENT, OBSERVABLE slightly higher, and PROBABLE still higher. This dimension seems to go from a high of "logically foundational" to a low of "empirically foundational". Those terms in the middle will be terms which are not foundational at all; on the negative side but in the middle will be non-foundational terms having to do with empirical matters (e.g., PROBABLE). Dimension 3 has NECESSARY very high, ESSENTIAL high, ANALYTIC, A PRIORI somewhat lower; it has HYPOTHETICAL, CONTINGENT very low, CONVENTIONAL and PROBABLE low, EMPIRICAL slightly higher. This dimension seems to be a measure of logical certainty attached to the term.

Again, it can be seen from Figures 4 to 6 that different philosophers attach different importance to the different dimensions. For example, Subject 9 appears to attend only to dimension 2. Subject 3 attends mainly to dimension 1, and much less to dimension 3, still less to dimension 2. Subject 1 attends mainly to dimension 3, much less to dimension 1, and still less to dimension 2. Subject A attends about equally to all dimensions. (Subject A's normalized w-scores do not very well represent his actual ratings, as Table 4 shows. This suggests that a three-dimensional solution provides a poor fit for his data in this field). The other groups indicated on these plots, 4,5,6 vs. 2,7,8, differ mainly on how much importance they attach to dimension 3.

The question naturally arises: why should we choose a three-dimensional solution over any other -- e.g., over a two-dimensional or four-dimensional solution? The answer to this can be put in terms of the correlation between the original ratings and the reconstructed ratings made on the basis of the low dimensional solution. If, in the Action Term case, every dimension were equally important, we would expect each of the 11 dimensions to contribute equally to the placement of the terms; i.e., we would expect each dimension to account equally for the variance and correlations. Instead, the first dimension accounts for most, the second dimension for next most, etc. When we reach a point where the addition of a dimension does not add more information (improve correlation) than would be expected by chance, then we say we have an optimum solution. Thus, we get a rough estimate as to whether we have an optimum solution in n dimensions by considering whether adding the (n+1)st dimension accounts for more of the improved correlation than one expects by chance addition of another dimension. What do we expect from a chance addition of a dimension? For the Action Terms, given a solution in n dimensions with a correlation of r, we expect each of the remaining 11-n dimensions to account equally for the 1-r correlation not yet accounted for, i.e., for (1-r)/(11-n).

Our three dimensional solution for Action Terms has a correlation of 0.796, so if none of the higher dimensions are important, we would expect that they each would contribute 0.204/8 (= 0.025) improvement in correlation. So, we expect an unimportant forth dimension to have an average correlation of 0.821. The four dimensional solution was done and it in fact has an average correlation between the original data and the reconstructed data of 0.823 (see Table 5). We therefore judge that the addition of a fourth dimension to a solution is not important. Similar reasoning tells us that the third dimension is important. From Table 5 we see that the correlation between the original scores and the reconstructed scores for the two dimensional solution is 0.704, so if a third dimension is not important, we would expect it to contribute 0.296/9 (= 0.033).

Table 5: Average Correlations Between Original Scores and Reconstructed Scores in Different Dimensional Solutions (Philosophers)

No. Dimensions	Action	Science
4	0.823	0.725
3	0.796	0.696
2	0.704	0.607

That is, we would expect a three dimensional solution to have a correlation of 0.737. However, the three dimensional solution has in fact a correlation of 0.796, and so we consider the third dimension to be important. Carrying the same reasoning over to the Science Terms, the three dimensional solution correlates 0.696, so we expect each higher dimension to add 0.304/17 (= 0.018), so we expect a four dimensional solution to correlate 0.714. In fact, as Table 5 shows, it correlates at 0.725. A case can perhaps be made out for this being of some importance (the distribution properties of the MDS and INDSCAL correlations are not known), but we view it as not of interest and stay with a three dimensional maximum. (In fact, this highly improved correlation comes mainly from an improvement in Subject A's correlation, from 0.442 to 0.545). The two dimensional solution correlates 0.607, so we expect the higher dimensions to contribute 0.393/18 (= 0.022) to the correlation, so we expect the three dimensional

solution to correlate 0.629. The fact that it correlates 0.696 shows that the third dimension is important.

The normalized w-weightings give us information about how important each subject views each dimension (in comparison with the other subjects). We might therefore try to group subjects according to how much alike they are in their perception of the similarity of these dimensions. For this end, the three-dimensional w-weightings of the philosophers for each semantic field were entered as data for an HCP.

The results (diameter method) are presented as Tables 6 and 7. It should be noted that the program forces a grouping regardless of how much "error" is introduced. For Table 6 we should consider two possible solutions as to how many groups of philosophers there are for the Action Terms. The first solution would be to consider these groups: {A}, {9}, {3,5,6,8}, {2,4,7}, {1}. The second solution would be to consider these groups: {9,A}, {3,5,6,8}, {1,2,4,7}. Table 6 gives the error scores (measure of within-group variance introduced) for the successive groupings. It will be noted that the attempt to form four groups (by adding Subject 1 to {2,4,7}) makes the error score jump from 0.195 to 0.300; the attempt to form two groups (by adding {1,2,4,7} to {3,5,6,8}) makes the error score jump from 0.349 to 0.470. For Table 7 we should consider these two solutions: there are six groups {A}, {9}, {2,7,8}, {4,5,6}, {3}, {1}; or there are three groups {2,3,7,8,A}, {1,4,5,6}, {9}. The attempt to form five groups (by joining subject A to {2,7,8}) makes the error score jump from 0.189 to 0.287; the attempt to form two groups (by joining {2,3,7,8,A} to {1,4,6}) makes the error score jump from 0.361 to 0.573, as Table 7 shows. These groups are the ones which have been indicated by concentric closed figures in Figures 4 to 9. The fact that these figures are "well defined" -- without discontinuities and relatively simple -- shows that the INDSCAL and HCP analyses support one another. (Apparent overlap of clusters and non-simple shapes is an artifact of the two-dimensional representation of the three dimensions. Consideration of the other projections will confirm this.)

Table 6: Error Scores for Successive Groupings of Philosophers on Action Terms

Groups Formed	Error Score
{5,6}	0.018
{3,5,6}	0.064
{2,4} {3,5,6}	0.108
{2,4,7} {3,5,6,5}	0.163
{2,4,7} {3,5,6,8}	0.195
{1,2,4,7} {3,5,6,8}	0.299
{1,2,4,7} {3,5,6,8} {9,A}	0.349
{1,2,3,4,5,6,7,8} {9,A}	0.470
{1,2,3,4,5,6,7,8,9,A}	0.595

Table 7: Error Scores for Successive Groupings of Philosophers on Science Terms

Groups Formed	Error Score
{2,8}	0.010
{4,6} {2,8}	0.107
{4,6} {2,7,8}	0.145
{4,5,6} {2,7,8}	0.189
{4,5,6} {2,7,8,A}	0.287
{4,5,6} {2,3,7,8,A}	0.345
{1,4,5,6} {2,3,7,8,A}	0.361
{1,2,3,4,5,6,7,8,A}	0.573
{1,2,3,4,5,6,7,8,9,A}	0.950

What are the groups indicated by the HCP? Is there any reason to choose either the five-group or three-group solution over the other (for Action Terms)? Is there any reason to choose either the six-group or the three-group solution over the other (for Science Terms)? Let us start this discussion by a consideration of the Action Terms. One rather interesting feature of this study is that action theory is a (central) branch of the philosophy of mind, and a number of our philosophers have as their field of specialization the philosophy of mind. In fact, the group {1,2,4,7} of the three-group solution is precisely those philosophers. This is certainly a reason for preferring the three-group solution. But what of the five-group solution? This solution separates {1} from {2,4,7} -- a feature which becomes understanda-

ble when one notices that Subject 1 has had some training in Continental philosophy. It therefore seems that more-or-less uniform training in a field (here: philosophy of mind) will cause people to perceive semantic elements from that field in the same way as one another; and that moderate differences in training (e.g., Continental philosophy superimposed on Analytic philosophy of mind) will cause pertubations in this perception.

What of the other philosophers? The three-group solution joins {9,A}, both of whom have philosophy of religion as an area of specialization. We might note in passing that many of the terms of action theory (e.g., FREE, INADVERTENT, UNCAUSED, etc.) have special uses in philosophy of religion, and it is this which perhaps accounts for why {9,A} are separate from the others {3,5,6,8}. The five-group solution separates Subjects 9 and A; this is perhaps recognition that Subject A does, but Subject 9 does not, have other specializations. Both solutions form the group {3,5,6,8}. This group consists of analytic philosophers whose specialization is neither philosophy of mind nor philosophy of religion. It therefore seems that the three-group solution is preferable, since it is more easily interpreted, and since there is a natural way to interpret and understand the five-group solution in terms of the three-group one.

It furthermore seems to follow from my analysis of the Action Terms that a general training in a field (as all analytic philosophers have some general training in philosophy of mind) when not augmented by specific training in that field or in some field which also uses the concepts of that field (as philosophy of religion uses some of the terms used in philosophy of mind), will produce a homogeneous group regardless of any other differences within this group. Thus, for example, the members of the group {3,5,6,8} have specialities as distinct as Marxism and Mediaeval philosophy, or as philosophy of language and aesthetics, yet their common, very general training in philosophy of mind directs them into a single group; and this group is distinct from that group of philosophers who specialize in philosophy of mind, even though individual philosophers chosen from the two groups may have other specializations which are identical (e.g., philosophy of language).

The explanation of the results of the Science Terms is not to be organized in the same way as that of the Action Terms, because none of the philosophers has philosophy of science as one of their specialities, and only three of the subjects (2, 7, and 8) reported any training at all in the field. Instead, we view it as a method of discovering "accidental" sameness of training among people who have no special or general training in the area (except that we might predict 2, 7, and 8 to form a group). In fact, in the six-group solution {2,7,8} do form a group, thus confirming the conclusions formed on the basis of the analysis of the Action Terms. In this six-group solution, the only other non-singleton group is {4,5,6}. There is no obvious reason why these subjects should form a group. There is also no obvious reason why Subjects 3 and A should group with {2,7,8} in the three-group solution, nor why Subject 1 should group with {4,5,6}, nor why Subject 9 should not cluster with some other group. I take this to be a discovery that hidden and unexpected factors can bring about difference or similarity of subjective semantic networks even when some overt factors might suggest otherwise. (E.g., Subjects 3 and 5 both have aesthetics as a speciality, but they fall in different groups; and Subjects 4,5, and 6 group together even though they have no field in common).

5. MDS ANALYSIS OF TYPICAL PHILOSOPHERS

One problem with the preceding INDSCAL analyses of the philosophers' and students' dimensions of meaning is that it does not quite prove the strong claim that different people might have entirely different dimensions of meaning. For, the INDSCAL technique (like any statistical technique) makes certain assumptions about the source of the data. Of interest here is INDSCAL's assumption that every subject has access to the same dimensions, and it then goes on to measure the salience of each of these dimensions for each subject. But we would like to show

that they are not even aware of the same dimensions. For this end, we should look at the MDS analyses of individual philosophers and compare their dimensions of meaning. (Admitting, of course, that these will be very gross comparisons, as there is no recognized statistical test to be used in such comparisons between separate MDS analyses. Nonetheless, there are such striking differences among our philosophers that these intuitive comparisons will carry considerable weight.)

"Typical philosophers" are defined by their deep embeddedness in one of the groups given by the HCP (see Figures 13 and 14). Thus, for the Action Terms, Philosophers 5 and 6, Philosophers 2 and 4, and Philosophers 9 and A, will be typical. For the Science Terms, Philosophers 2 and 8, 4 and 6, and 9 will be typical. We pick one philosopher from each of these groups and use him to represent the group as a whole. The individual MDS analyses of these philosophers' dimensions of meaning are given in Tables 8 and 9 for the Action Terms and Science Terms respectively. It can be seen from Table 8 that Philosophers 2 and 5 differ almost entirely in the fact that their third dimensions are different. In the discussion above of the Dimensions of Meaning for Action Terms, Dimension 1 was called a measure of "Necessitation" and Dimension 2 was called a measure of the "Purposefulness of human actions". We here see that both philosophers 2 and 5 individually use these dimensions for their first two vectors of meaning in characterizing these terms. The third dimension was disjunctively characterized as: "Out of one's control for an external reason vs. in one's control but not attended to" or maybe "Excusable for some external reason vs. excusable because it one's control vs. done because of some external reason". Philosopher 5 ranks the terms from high SPONTANEOUS, EFFORTLESS to low ACCIDENTAL, INADVERTENT, THOUGHTLESS -- a dimension ranking "easily performed but responsible for vs. done unawares".

This illustrates the truth of the remark made earlier that philosophers who specialize in philosophy of mind (like philosopher 2) have learned some subtle thing which is encoded as a ternary vector and which other analytic philosophers have not learned.

Table 8: Dimensions of Meaning for Some Typical Philosophers (Action Terms)

Philosopher 2

Dim. 1	Dim. 2	Dim. 3
COMPELLED	DETERMINED	EFFORTLESS
INVOLUNTARY	FREE	.
MECHANICAL	CAUSED	.
DETERMINED	.	.
.	.	.
.	.	.
.	.	.
.	.	.
SPONTANEOUS	UNCAUSED	DETERMINED
FREE	COMPELLED	.
UNCAUSED	INVOLUNTARY	MECHANICAL
		UNCAUSED

Philosopher 5

Dim. 1	Dim. 2	Dim. 3
COMPELLED	CAUSED	SPONTANEOUS
CAUSED	FREE	EFFORTLESS
DETERMINED	EFFORTLESS	.
INVOLUNTARY	MECHANICAL	.
.	.	.
.	.	.
.	.	.
SPONTANEOUS	INVOLUNTARY	THOUGHTLESS
FREE	COMPELLED	INADVERTENT
UNCAUSED	UNCAUSED	ACCIDENTAL

Philosopher A

Dim. 1	Dim. 2	Dim. 3
COMPELLED	INADVERTENT	UNCAUSED
INVOLUNTARY	INVOLUNTARY	DETERMINED
DETERMINED	ACCIDENTAL	.
CAUSED	UNCAUSED	.
.	CAUSED	.
.	.	.
.	.	.
.	.	.
ACCIDENTAL	MECHANICAL	EFFORTLESS
UNCAUSED	DETERMINED	ACCIDENTAL
FREE	FREE	INADVERTENT
		THOUGHTLESS

Philosophers 2 and 5 differ almost entirely in their third dimensions.
Philosopher A has a third dimension more like 2's than like 5's.
Philosopher A's second dimension differs from both 2's and 5's.

Table 9: Dimensions of Meaning for Typical Philosophers (Science Terms)

Philosopher 5

Dim. 1	Dim. 2	Dim. 3
ANALYTIC	HYPOTHETICAL	PROBABLE
NECESSARY	THEORETICAL	
A PRIORI	.	
CONCEPTUAL	.	LAWLIKE
ESSENTIAL	.	ESSENTIAL
.	.	.
.	.	.
.	.	.
.	.	OBJECTIVE
.	.	CONTINGENT
.	LAWLIKE	VALID
PROBABLE	OBJECTIVE	ANALYTIC
EMPIRICAL	NECESSARY	
CONTINGENT	OBSERVABLE	CONVENTIONAL

Philosopher 8

Dim. 1	Dim. 2	Dim. 3
CONCEPTUAL	REGULAR	HYPOTHETICAL
A PRIORI	THEORETICAL	
ANALYTIC	SUFFICIENT	
THEORETICAL	LAWLIKE	CONTINGENT
.	HYPOTHETICAL	VALID
.	.	TRUE
.	.	.
.	.	.
PROBABLE	.	.
REGULAR	REASONABLE	OBJECTIVE
CONTINGENT	TRUE	ANALYTIC
NATURAL	EMPIRICAL	ESSENTIAL
OBSERVABLE	CONTINGENT	NECESSARY
EMPIRICAL	GIVEN	
	VALID	

Philosophers 5 and 8 have the same first dimension. Perhaps 5's second dimension is the same as 8's third dimension?

Philosopher A has a first dimension very similar to philosophers 2 and 5, but his second dimension is different from their similar second dimension. His second dimension is extraordinarily difficult to interpret: it perhaps ranges from a high "events for which a definite responsibility cannot be ascribed" to a low "events for which a definite responsibility can be ascribed" (although the responsible agent might sometimes be the person, sometimes be an outside force). This kind of dimension might be seen as a result of A's philosophy of religion. A's third dimension is more like 2's than 5's.

Table 9 shows the differences between the typical philosophers 8 and 5 on Science Terms. It will be noted that they have (pretty much) the same first dimension. It runs from a high ANALYTIC, A PRIORI, CONCEPTUAL, less high NECESSARY, THEORETICAL, ESSENTIAL to a low EMPIRICAL, CONTINGENT, PROBABLE less low OBSERVABLE, NATURAL. This appears to be the same as the first dimension given by the philosophers' INDSCAL analysis. A strong similarity might be noted in 5's second dimension and 8's third dimension, although it would be difficult to say exactly what it is: perhaps "the amount of empirical certainty attached" (but this does not fit philosopher 8's third dimension very well). It is clear that our two typical philosophers use the Science Terms quite differently. (Which fact explains why the INDSCAL analysis has such a lower correlation than the Action Term analysis).

6. A LOOK AT THE SEMANTIC NETWORKS OF THE TYPICAL PHILOSOPHERS

Table 10 gives the mean scores and standard deviations of the individual Action Terms for our typical philosophers 2, 5 and A, and Table 11 gives them for the individual Science Terms for the typical philosophers 5 and 8. A term which has a low mean score will be one that the philosopher sees as being rather similar to all the rest of the terms, and one which has a high mean score will be one that he views as being dissimilar to the rest. In other words, those terms with the lowest mean scores will, in some sense, define the center of his semantic network,

Table 10: Analysis of Philosophers 2, 5, and A (Action Terms)

Term	Phil. 2		Phil. 5		Phil. A	
	mean	s.d.	mean	s.d.	mean	s.d.
ACCIDENTAL	3.83	1.99	4.58	2.11	7.33	2.84
CAUSED	4.92	2.47	4.67	2.10	6.75	3.28
COMPELLED	4.92	2.84	4.83	2.59	6.83	3.01
DETERMINED	5.17	2.72	5.00	2.00	7.08	2.97
EFFORTLESS	5.33	2.50	4.67	1.83	7.75	2.99
FREE	5.75	2.45	4.67	2.19	7.42	2.79
INADVERTENTLY	4.50	2.32	4.42	1.88	6.92	3.29
INVOLUNTARY	4.58	2.84	4.75	2.18	6.83	2.95
MECHANICAL	4.67	2.77	4.50	1.73	7.33	2.96
SPONTANEOUS	5.08	2.39	4.67	2.10	7.25	3.25
THOUGHTLESS	3.42	1.78	4.25	1.76	7.00	3.25
UNCAUSED	6.33	2.77	5.17	2.21	7.83	2.55
r(2,5)=0.78;	r(2,A)=0.54;		r(5,A)=0.32.			
MOST CENTRAL:	THOUGHTLESS ACCIDENTAL		THOUGHTLESS INADVERTENTLY MECHANICAL		CAUSED INVOLUNTARY COMPELLED	
MOST PER-IPHERAL:	UNCAUSED		UNCAUSED		UNCAUSED	
			DETERMINED		EFFORTLESS	
MOST STABLE:	THOUGHTLESS ACCIDENTAL		MECHANICAL THOUGHTLESS EFFORTLESS INADVERTENTLY		UNCAUSED FREE ACCIDENTAL	
MOST VARIABLE:	COMPELLED INVOLUNTARY		COMPELLED		INADVERTENTLY CAUSED SPONTANEOUS THOUGHTLESS	

GENERALLY: Philosopher 2 sees the terms neither as being very similar nor dissimilar in meaning (mean of means = 4.88). He has moderate stability in his judgments of similarity (s.d. = 2.50).

Philosopher 5 sees the terms neither as being very similar nor dissimilar in meaning (mean of means = 4.68). He is very stable in his judgments of similarity (s.d. = 2.05).

Philosopher A sees the terms as being very dissimilar in meaning (mean of means = 7.19). He is very unstable in his judgments of similarity (s.d. = 3.00).

while those with highest mean scores will show the outer limits of his semantic network. This gives us an alternate method of characterizing the semantic network -- one which shows it as a group of concentric circles. Admittedly, such a characteriza-

Table 11: Analysis of Philosophers 5 and 8 (Science Terms)

Term	Philosopher 5		Philosopher 8	
	mean	s.d.	mean	s.d.
ANALYTIC	4.29	2.31	3.90	2.34
A PRIORI	4.33	2.03	4.33	2.67
CONCEPTUAL	4.29	2.00	4.10	1.95
CONTINGENT	4.95	2.60	4.52	2.40
CONVENTIONAL	4.43	1.63	4.71	1.90
EMPIRICAL	4.14	2.56	4.14	2.17
ESSENTIAL	4.71	2.08	3.52	1.78
GIVEN	4.05	1.60	4.57	2.34
HYPOTHETICAL	5.05	1.69	4.95	2.16
LAWLIKE	3.86	1.62	4.62	1.80
NATURAL	4.05	1.53	4.19	1.66
NECESSARY	4.57	2.44	4.24	2.32
OBJECTIVE	3.29	1.68	4.10	1.55
OBSERVABLE	4.38	2.09	5.14	2.31
PROBABLE	5.29	2.22	4.43	1.86
REASONABLE	3.81	1.50	3.67	1.28
REGULAR	3.81	1.36	3.95	1.40
SUFFICIENT	4.76	1.37	4.95	1.69
THEORETICAL	4.90	1.92	3.57	1.66
TRUE	4.14	1.49	3.90	1.87
VALID	3.95	1.91	5.43	2.16
MOST CENTRAL:	OBJECTIVE		ESSENTIAL THEORETICAL REASONABLE	
MOST PER-IPHERAL:	PROBABLE HYPOTHETICAL		VALID OBSERVABLE	
MOST STABLE:	REGULAR SUFFICIENT		REASONABLE REGULAR	
MOST VARIABLE:	CONTINGENT EMPIRICAL NECESSARY		A PRIORI	

GENERALLY: Philosopher 5 sees the terms as somewhat more similar than dissimilar (mean of means = 4.34). Very stable in his judgments of similarity (s.d. = 1.88).

Philosopher 8 sees the terms as somewhat more similar than dissimilar (mean of means = 4.33). Very stable in his judgments of similarity (s.d. = 1.96).

tion of the semantic network does not give the detailed internal information garnered from an MDS analysis about the relatedness of the individual terms; but still, it can be of use in charac-

terizing the overall structure and limits of the person's field. As one can see from Table 10, some of the philosophers (here: 2 and 5) see the terms forming their semantic networks as being neither very similar nor very dissimilar (their mean of means are 4.88 and 4.68 respectively). This would mean that, in the overall structure of their networks, the terms are neither "bunched together" nor "spread out". Other philosophers (here: A), on the other hand, see the terms as being very dissimilar (mean of means 7.19). This might mean that they do not view the set as "cohesive" (at least to the extent that 2 and 5 do), and might even argue that, for them, the terms are not really characterizable as falling into any natural grouping (like "action terms"). There is one caveat that must be made here. Since the scales of judgement are not absolute, the difference among the philosophers on this point may be nothing more than response bias for one or the other end of the scale. (I.e., one philosopher may see 5 as the "middle" value meaning "about average" while another philosopher may see 7 as the "middle" value). A bias like this might, I think, account for some of the disparity between our philosophers, but I think it cannot account for it all. After all, they were all given the same instruction set and example, and they were all told to first mark the "1's" and "9's". This surely will give a strong tendency for stabilization among the different philosophers, and cannot account for the dramatic differences we have recorded. However, it must be recognized that we are here making an inference about this -- the data we have do not allow that this is being measured directly.

The standard deviations of the terms might be taken as giving a measure of "how stable" the philosopher is in judging similarity. More accurately, it is a measure of how the philosopher views each of the terms. So, if he thinks that a term has the same distance from each of the other terms, the term will get a low standard deviation score. And if he does this for all the terms, he will get a low overall standard deviation score. However, in these studies, where there is a large number of judgements of similarity to be made, there is a sense in which this really does measure "how stable" the philosopher is in

judging similarity. For this reason we shall continue to call it a measure of the stability of judgment, realizing of course that this is an inference doubly removed from any direct measure of the philosopher's actual stability of judgement. In Table 10, it can be seen that philosopher 5 is very stable in his similarity judgments (s.d. = 2.05), that philosopher 2 is moderately stable in his similarity judgments (s.d. = 2.50), and that philosopher A is markedly unstable in his similarity judgments (s.d. = 3.00).

CONCLUSION AND ACKNOWLEDGMENTS

The results are, in themselves, interesting, I think. But of broader interest are the possibilities opened up by MDS and the related techniques mentioned in this report for investigating various areas of study. Some of these are discussed in Appendix II.

I would like to take this opportunity to thank especially William J. Baker of the Linguistics Dept. of the Univ. Alberta for guidance, encouragement, and much help in doing this study. Gary D. Prideaux, Lois M. Markworth (of the Univ. Alberta Linguistics Dept.), and William W. Rozeboom (of the Centre for Theoretical Psychology at the Univ. of Alberta) also have made helpful comments on earlier drafts. And I should especially thank my colleagues and friends who answered the questionnaires and performed the (very tedious and time-consuming) task reported herein.

APPENDIX I: A COMPARISON WITH OTHER APPROACHES

The theoretical position taken in this study is superficially similar to various approaches taken in modern linguistics. In KATZ & FODOR (1963), and as adopted by CHOMSKY (1965), KATZ (1966, 1972), BIERWICH (1969, 1971) among others, the lexicon is represented as containing various parts. One of these parts is a set of features which represent, inter alia, the "concepts" which are "parts of a meaning" of a lexical item. These writers argue that the lexicon of a linguistic theory must contain these features (and other things also), in order to adequately account for both semantic phenomena and syntactic phenomena. A lot has been written by philosophers in criticism of this conception of semantics (see VERMAZEN 1967, LEWIS 1972, PELLETIER 1977), but in linguistics it has been commonly adopted (sometimes with minor alterations). One kind of alteration is given by WEINRICH (1966) who argues about the precise formulation of these features and how they are to be combined when producing the meaning of longer phrases containing the words. Another kind of alteration is to claim that the lexicon does not contain morphemes plus features associated with them, but rather only the features (or "semantic primes", as they are then called), which are later "lexicalized". (That is, at some point in the (transformational) derivation of a sentence, the co-occurrence of these features triggers some grammatical rule which alters the features into a morpheme). This is the view first given (I think) by GRUBER (1965), followed by generative semanticists, and studied under the title of 'lexical decomposition'. See ZWICKY (1976) for a summary and bibliography.

One difficulty (according to some) with this general approach is that it seems to commit one to a "universal semantics". That is, it requires that the set of features or semantic primes be a "fixed, universal vocabulary" (CHOMSKY 1965: 160; see also McCAWLEY 1968) which holds for all speakers of all languages. A consequence of this is that it requires the theory to account for why not all words in all languages are directly inter-translatable, i.e., why not every (logically) possible combination of features corresponds to a word in each language. To explain this can be seen as one part of the goal of the account of "exceptions" given in LAKOFF (1965). One might attempt to avoid this difficulty by denying a "universal semantics" of this sort, either along the lines taken by WHORF (1946) or along the lines taken by practitioners of "componential analysis". (A summary of componential analysis and thorough bibliography is given by NIDA, 1974; see also WALLACE, 1965).

According to these latter kinds of writers, there is no "universal semantics", at least not in the sense that every feature in any language is taken from some fixed vocabulary. Rather, it is envisaged that each language has its own stock of features, and that these features may only occasionally be identical to those of another language. It would seem to follow from this view that people from distinct linguistic cultures can never "truly understand one another"; and indeed, this is precisely the conclusion reached by WHORF (1946). There are a variety of difficulties associated with a strong form of this theoretical position; but this is not the place to go into them. (For discussion and bibliography see, for example, PENN (1972) and COOPER (1973).) Most componential analysts do not seem to be aware of this apparent consequence of their position, and seem to believe that the semantic network determined by a group of terms in one language as defined by the relevant semantic features of that language can be explained or approximated in another language by indicating features of this other language which, if adopted, might lead its speakers to use their terms in the same way as the corresponding terms are used in the language under study. Consider the following example: The semantic network determined by the kinship terms of Choctaw as defined by the relevant semantic features of Choctaw are viewed as explained or approximated in English when features of English are assigned to the English kinship terms in such a way that, if adopted, English speakers would use the English kinship terms in the same way that Choctaw speakers use their kinship terms.

When the methodology of componential analysis is stated this way, its conceptual inadequacies become transparently obvious. For one thing, even if one could have adequate evidence for each of the other steps in the above example, it still would not follow that the analyst could be justified in asserting that the features of Choctaw kinship terms were in fact the ones he uses to approximate hypothetical usage in English. For, after all, it was postulated at the outset that the set of features in the two languages need not be the same; indeed, the whole theoretical rationale of componential analysis makes it impossible to ever get adequate evidence to settle the issue, since we cannot get "outside language". (For a summary of similar criticism, see BURLING (1964); some of these issues are taken up in WALLACE (1965).)

I wish here to point to another unpleasant theoretical consequence of componential analysis. It seems to have gone unnoticed by componential analysts that, for the very same reasons one might want to deny the necessary existence of a "universal set of features", one ought to deny the necessary existence of an interpersonal set of features common to a given language. After all, if one holds to the view that an individual acquires concepts (including these features) through experience, one should recognize the possibility (if not likelihood) that individual learning experiences will differ and lead individuals to (a) have subtly different features in their subjective lexicon, and (b) associate features with words in subtly different ways. If now, we add one more seemingly plausible assumption, namely that communication takes place when the relevant protion of the speaker's semantic network is replicated by the hearer (an assumption that componential analysts seem to find agreeable, see NIDA, 1975, Chap. 1), we are led to the radical conclusion that our theory decrees that we are never in a position to have adequate justification for the hypothesis that individual speakers of even the same language ever communicate. That is, the Whorfian claim can, according to their theory, be made for idiolects as well as for different languages. This is an intuitively unacceptable consequence and, if we wish to embrace some theory like componential analysis, we must find some way to prevent this conclusion from being entailed.

The real problem here is that in order to have adequate evidence for any theory which views word meaning as explained by "features", "primes", or "dimensions" (including the present study) we must be able to specify some method to discover them. And no amount of introspection can succeed, because this merely brings the "features", "primes", "dimensions", etc., back to those in whatever language(s) the investigator knows -- thus being circular. One advantage of the present study is that it provides a way out of this introspective circle.

APPENDIX II: EMPIRICAL SEMANTIC APPLICATIONS OF MDS

I think it is clear that MDS is to be recommended as an investigative tool -- superior to the introspective analyses now used -- for componential analysts. It has application also in the study of "pseudo agreements/disagreements" among professionals of various social science fields. (For an explanation of "pseudo agreement/disagreement" together with an extended study of one amongst businessmen and politicians, see TENNESSEN (1959) and the studies reported therein). For example, it seems that politicians might very well appear to agree (or disagree) on such terms as 'conservative', 'liberal', 'socialist', and the like, when in fact there is strong disagreement (or agreement) in their second and third dimensions. Not only is the study of subjective meaning of use in the "practical" fields like politics, but it might also be useful in the study of the subjective lexicons of such social scientists as anthropologists, sociologists, political scientists, and so on.

Another area that the method might be of use is in psychoanalytic evaluation. Suppose, for instance, that a patient were given a set of emotion terms such as LOVE, HATE, SEX, GREED,

etc. One might conjecture that it could be quite revealing of important psychological traits that a person has, for example, LOVE and HATE close together on some dimension, whereas other people have LOVE and HATE rather far apart on all dimensions.

However, from the broadest theoretical point of view, of paramount interest is the issue of innatism with respect to word meaning or word connotation. I think this preliminary investigation lends considerable support to the view that (if this conception of the subjective lexicon is accepted) a strict innatism of the form adopted by such linguists as Katz and Leech is empirically false.

The conception of one's subjective lexicon which allows the possibility of two person's having completely different mental structures associated with the same lexical item, and never being able to tell, is called 'privatism'. (For example, privatism allows that we associate distinct subjective impressions with the word 'red', but yet all our behaviour with respect to 'red' is identical -- we both call the same things "red".) Privatism is bolstered by a plausible account of language learning called 'empiricism': the view that, at childhood, the mind is, as it were, a blank tablet and one's experiences etch upon this tablet to form the concepts, connotations, and meanings he will associate with various words. And since we each obviously have distinct experiences we each therefore will have distinct subjective associations with the same term. But this kind of empiricism with respect to language learning, and the conception of successful communication which holds that communication amounts to "sharing mental structures", leads naturally to the view that we rarely (if ever) have true communication since our childhood experiences are so varied; and even if we do, we would never know it.

One popular point at which to challenge this view and its undesirable conclusion is at its account of language learning. We therefore distinguish a view of the human mind according to which it is not a blank tablet waiting for experience to etch upon it, but rather is a block of veined marble upon which ex-

perience chiseled. Even though our childhood language learning experiences might be somewhat different, so long as they are fairly similar, the fact that we have the same veins in our mental marble will guarantee that we will arrive at the same concepts or connotations or meanings associated with all our ordinary terms. This view is innatism, and in the area of lexical semantics has been explicitly championed by such philosophico-psychologico-linguists as KATZ (1966, 1972), FODOR (1977), and LEECH (1974). Other linguists have offered arguments for the innateness of syntactic structures; such arguments, it seems, could easily be adapted to lexical semantics. It thus seems that innatism with respect to the lexicon -- the meanings or connotations or concepts that an individual associates with individual terms -- is a view that has, or should have, many advocates.

Which view is correct? MDS can help us decide; for innatism says that people with similar but non-identical experiences will have the same subjective lexicons, while empiricism says that there will be a monotonic gradation of differences -- the more distinct the experiences, the more distinct the subjective lexicons.

Our primary sample, a group of ten professional philosophers, have experiences relevant to the philosophic terms being considered. And these experiences can independently be judged according to degree of similarity to one another. It was found that (a) even very similar experiences produce somewhat dissimilar networks, (b) the more similar two peoples' experiences are, the closer their networks are to one another, and consequently, (c) groups whose members have similar experiences tend to "clump together" and be quite different from other groups all of whose members have similar experiences.

It is therefore shown that innatism of lexical structures, when taken as a theory that has empirical reflexes, is false. In their belief that it is needed to account for language learning and communication, such theorists as Katz, Fodor, and Leech are not just misguided (as some critics think innatism

to be), but conclusively wrong. However, innatism need not be taken as a theory that has empirical reflexes of this nature. We have not investigated this non-empirical theoretical innatism in this study; nor is a work in the sciences, even a social science, the proper place to carry out such an investigation. Such a task can best be left for the philosophers.

REFERENCES

ANGLIN, J.M. (1970): The Growth of Word Meaning, Cambridge: MIT Press.

BIERWISCH, M. (1969): On Certain Problems of Semantic Representations, Foundations of Language 5, 153-184.

BIERWISCH, M. (1971): On Classifying Semantic Features, in: Steinberg & Jacobvits (1971, 410-435.

BURLING, R. (1964): Cognition and Componential Analysis: Hocus Pocus or God's Truth?, American Anthropologist 66, 20-28.

CHOMSKY, A.N. (1965): Aspects of the Theory of Syntax, Cambridge: MIT Press.

CLARK, H.H. (1968): On the Use and Meaning of Presupositions, Journal of Verbal Learning and Verbal Behaviour 7, 421-431.

COOPER, D. (1973): Philosophy and the Nature of Language, London: Longmans.

DEESE, J. (1965): The Structure of Associations in Language and Thought, Baltimore: John Hopkins.

FILLENBAUM, S. & RAPOPORT, A. (1973): Structures in the Subjective Lexicon, New York: Academic Press.

FODOR, J. (1977): The Language of Thought, New York: Academic Press.

GRUBER, J. (1965): Studies in Lexical Relations (MIT dissertation).

HAMMEL, E.A. (Ed.) (1965): Formal Semantic Analysis, Special issue of American Anthropologist 67:5, pt. 2.

HENLEY, N.M. (1969): A Psychological Study of the Semantics of Animal Terms, Journal of Verbal Learning and Verbal Behaviour 8, 176-184.

HENLEY, N.M., NOYES, H.L. & DEESE, J. (1968): Semantic Structure in Short-Term Memory, Journal of Experimental Psychology,

JOHNSON, S.C. (1967): Hierarchical Clustering Schemes, Psychometrika, Vol. 32, 241-254.

KATZ, J.J. (1966): The Philosophy of Language, New York: Harper & Row.

KATZ, J.J. (1972): Semantic Theory, New York: Harper & Row.

KATZ, J.J. & FODOR, J. (1963): The Structure of a Semantic Theory, Language 39, 170-210.

KRUSHAL, J.B. (1964): Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis, Psychometrika 29, 1-27.

LAKOFF, G. (1965): Irregularity in Syntax (Ind. Univ. Diss.), published 1970, New York: Holt, Rinehart & Winston.

LEECH, D.K. (1974): Semantics, Harmondsworth: Penguin.

LEWIS, D.K. (1972): General Semantics, in: Davidson, D. & Harman, G. (Eds.) Semantics of Natural Language, Dordrecht: Reidel, 169-218.

LYONS, J. (1970): New Horizons in Linguistics, Harmondsworth: Pelican.

McCAWLEY, J. (1968): Lexical Insertion in a Transformational Grammar Without Deep Structure, Chicago Linguistic Society IV, 71-80.

MILLER, G.A. (1967): Project Grammarama, in: Miller, G.A. The Psychology of Communication, New York: Basis Books, 125-187.

MILLER, G.A. (1969a): A Psychological Method to Investigate Verbal Concepts, Journal of Mathematical Psychology 6, 169-191.

MILLER, G.A. (1969b): The organization of Lexical Memory: Are Word Associations Sufficient?, in: Talland, G.A. & Waugh, N.J. (Eds.), The Pathology of Memory, New York: Academic Press, 223-236.

MILLER, G.A. (1972): English Verbs of Motion: A Case Study in Semantics and Lexical Memory, in: Melton, A.W. & Martin, E. (Eds.), Coding Processes in Human Memory, Washington: Winston, 335-372.

MILLER, G.A. (1973a): Communication, Language, and Meaning: Psychological Perspectives, New York: Basic Books.

MILLER, G.A. (1973b): Psychology and Communication, in: Miller (1973a), 3-12.

NIDA, E. (1974): Exploring Semantic Structures,Munich: Wilhelm Fink Verlag.

NIDA, E. (1975): Componential Analysis of Meaning, The Hague: Mouton.

OSGOOD, C.E., SUCI, G.J. & TANNENBAUM, P.H. (Eds.) (1957): The Measurement of Meaning, Urbana: Univ. of I11. Press.

PELLETIER, F.J. (1977): ([How/Why]) Does Linguistics Matter to Philosophy?, Southern Journal of Philosophy 15, 393-426.

PENN, J.M. (1972): Linguistic Relativity versus Innate Ideas, The Hague: Mouton.

ROMNEY, A.K. & D'ANDRADE, R.G. (1964): Cognitive Aspects of English Kin Terms, in: Romney, A.K. & D'Andrade, R.G. (Eds.) Transcultural Studies in Cognition (Special issue of American Anthropologist 66:3, pt. 2, 146-170).

SHEPARD, R.N. (1969): Some Principles and Prospects for the Spatial Representation of Behavioural Science Data (unpublished).

STEINBERG, D.D. & JACOBVITS, L.A. (1971): Semantics, Cambridge: Cambridge U.P.

STENSON, H.H. & KNOLL, R.L. (1969): Goodness of Fit for Random Ranking in Krushal's Non-Metric Scaling Procedure, Psychological Bulletin 71, 122-126.

TENNESSON, H. (1959): 'Free Enterprise', an Empirio-Semantic Analysis of a Slogan, Synthese 11, 72-84.

ULLMAN, S. (1962): Semantics: An Introduction to the Sciences of Meaning, New York: Barnes & Noble.

VERMAZEN, B. (1967): Review of [Katz & Postal 1964] and [Katz 1966], Synthese 17, 350-365.

WALLACE, A.F.C. (1965): The Problem of the Psychological Validity of Componential Analysis in: Hammel (Ed.) (1965), 220-231.

WALLACE, A.F.C. & ATKINS, J. (1980): The Meaning of Kinship Terms, American Anthropologist 62, 58-79.

WEINRICH, U. (1966): Explorations in Semantic Theory, The Hague: Mouton.

WHORF, B.L. (1956): Language, Thought, and Reality, Cambridge: MIT Press.

YOUNG, F.W. (1970): Nonmetric Multidimensional Scaling: Recovery of Metric Information, Psychometrika 35, 455-473.

ZWICKY, A. (1976): Linguistics as Chemistry: The Substance Theory of Semantic Primes, in: Anderson, S.R. & Kiparsky, P. (Eds.) A Festschrift for Morris Halle, New York: Holt, Rinehart & Winston, 467-485.