Semantic Feature Representations for Normal and Aphasic Language\textsuperscript{1,2}

E. B. Zurif\textsuperscript{3}

\textit{Aphasia Research Center, Boston University School of Medicine}
and
\textit{Boston Veterans Administration Hospital}

A. Caramazza

\textit{The Johns Hopkins University}

R. Myerson and J. Galvin

\textit{Aphasia Research Center, Boston University School of Medicine}

Aphasic and non-neurological patients grouped nouns on the basis of similarity of meaning. These word groupings served as input matrices for hierarchical clustering and multidimensional scaling analyses. The emergent structures suggest that, while the normal adult has a number of levels upon which to organize his lexicon, the adult aphasic's lexicon can be characterized as a set of partial entries that are tied to affective and situational data. The results also suggest that semantic feature representations derived from similarity-of-meaning judgments are of relevance in the study of factors which influence actual language performance.

A number of recent studies on the organization of lexical memory incorporate the notion of an internalized data structure based upon semantic features (e.g., Bower, 1970; Miller, 1967, 1969, 1972; Quillian, 1968). These features, typically mapped as single words (± male) or phrases (± who is married, represent the lexical information available to

\textsuperscript{1} This was a completely collaborative effort between the first and second authors, and the order of first and second listing was decided by a coin toss. A coin toss also decided the order for listing the third and fourth authors, both of whom shared testing and aided in the preparation of this paper.

\textsuperscript{2} This study was supported by NINDS research grants 11408 and 06209 to Boston University School of Medicine and by grant NS 07615 to Clark University. We thank Dr. Warren Torgerson for providing a copy of TORSCA and for his helpful discussions on MDS. We also thank Dr. Howard Gardner for his extensive comments on an earlier draft of this paper, and Drs. Harold Goodglass and Sheila Blumstein as well as Errol Baker and Bud Porter for their helpful and considerable advice throughout this research.

\textsuperscript{3} Reprints may be obtained from E. Zurif. Aphasia Research Center, 14th floor, 150 S. Huntington Ave., Boston, Mass. 02130.
the language user. They are meant to capture the concepts inherent in a word as well as to define the range of semantic relations that a word can enter into with other words.

Aside from questions concerning the possible organizations of such feature representations in memory and methods for their discovery (e.g., Henley, 1969; Miller, 1969; Perfetti, 1972; Rips, Shoben, and Smith, 1973), there remains the issue of whether or not the act of retrieving words in ongoing speech actually exploits componential information of this kind. To the extent that it does, and assuming for the present that the organization of the psychological lexicon can be independently characterized, it might be expected that adults who have a problem in finding words operate from a differently organized or disrupted system of verbal concepts. That is, if word comprehension and production are impaired, the representation of semantic relations in memory should also be disrupted or restricted. We examine this assumption by exploring how lexical information is stored both for normal language users and aphasics.

Although Jakobson applied linguistic concepts to the study of acquired language disorders as early as 1941, the psycholinguistic approach to aphasia has only recently become systematized. Within the past few years, investigations have ranged from careful descriptions of aphasic impairment at the syntactic (Goodglass, 1968; von Stockert, 1972; Zurif, Caramazza, and Myerson, 1972), lexical (Howes, 1964; Whitaker, 1971a) and phonological levels (Blumstein, in press; Lecours and Lhermitte, 1969) to an explicit concern with theoretical linguistic predictions (Blumstein, in press; Green, 1970; Weigl and Bierwisch, 1970; Whitaker, 1971b). There are still many unresolved issues concerning the appropriate basis for classifying the disorders, and as yet, there is no comprehensive neurolinguistic model of aphasia. However, there is general agreement upon one salient point: brain damage is selective in its undermining of the language system.

Moreover, the effect of this selective disturbance is roughly a dichotomous one and bears an intriguing affinity with the distinction between semantic representation and surface structure. Patients with anterior lesions generally produce a nonfluent, agrammatic form of speech relatively well-supplied with words of concrete reference, whereas patients with posterior lesions, although capable of a fairly "fluent" output, are impaired in producing the substantives necessary for an informative utterance. Actually, careful examination shows that posterior patients also make grammatical errors (Goodglass, 1968). But because their speech prosody is normal and because they produce a variety of complex syntactic forms such as embedded clauses and other dependent structures, the clinical impression of grammatical fluency is maintained.
In any event, the starting point of the present study is that although virtually all aphasics have some problem with finding words, whether in spontaneous speech or in formal naming tests, only the most severely stricken show a complete inability to access lexical information. When a moderately impaired aphasic, especially one with anterior damage, searches his memory for the name of an object or event, the process may be slower than normal, but the outcome is more than infrequently successful. Even when the response is not correct, as is often the case with posterior aphasics, it is usually better described as out-of-focus with the target rather than random (e.g., Lhermitte, Derouesne, and Lecours, 1971; Pizzamiglio and Appicciafuoco, 1971).

Given these well-established clinical forms of word-finding difficulties, we have attempted to determine whether or not at least part of the problem, however much embedded in one aphasic syndrome or the other, can be related to the organization of lexical storage. We try to gain leverage on this issue by analyzing some of the structures underlying verbal concepts in the different aphasias and by comparing them with those underlying normal language use. As a corollary, a built-in performance measure exists against which to test the relevance of these analyses and comparisons: the contrast between an aphasic patient’s and a normal speaker’s ability to find words for an utterance.

Our exploration of subjective lexical organization is based on a triadic comparison procedure. A subject is presented with three words at a time, and in each case, is required to choose the two that he feels are most similar in meaning. The use of this procedure involves the assumption that the words to be clustered are conceptually distinct to the subject. This assumption becomes critical when dealing with an aphasic population; and so, special care has been taken here to demonstrate that the aphasic patients “understand” the meanings of the words used, even if they do not produce them easily or correctly in ongoing speech.

For subjects to settle on two words in each triadic array they presumably ignore some of their distinguishing features while attending to others. By this reasoning, the procedure may be used to infer which features are central to subjective lexical organization. But the procedure should not be left unguided; there should be some consideration for the choice of items and their componential structure (Miller, 1969). Such components of meaning need not have some predefined status within any particular linguistic theory (Perfetti, 1972); they need only be plausible, and may be derived as economically as possible from a knowledge of the relations existing between lexical items. For example, the features +human and +male can be applied to many entries in the dictionary and represent relations that are relatively systematic in the language (Miller, 1969).
The question arises, however, as to whether or not the dimensions of the internal lexicon can be characterized solely in terms of these general features. An alternative notion is that subjective lexical organization relies heavily on an undefined residuum of meaning. The latter term is Miller's (1972) and refers to the idiosyncratic details associated with the meaning of any particular word. The features summarizing these details cannot be considered to have any systematic value for the rest of the dictionary. In fact, the residuum may even contain features which are not to be found at all in a standard English dictionary; features which, instead, represent referential and affective information (Miller, 1972).

Thus, as part of our exploration of subjective lexical organization, we have attempted to distinguish between general and residual semantic features. We have chosen a set of 12 concrete nouns which permit the formation of at least two different, and mutually exclusive semantic domains: one domain to be formed by clustering the words on the basis of general semantic features presupposed in their definitions; the other, by grouping the words on the basis of features that fit into the "meaning-residuum" category. The details are supplied in the methods section.

Perhaps an even more central property of the words we have chosen is that they contain features that may be cast into a hierarchical relation. That is, not every word in our array has a value for every feature and certain features dominate others. For example, those items that are classified as +animal may be further classified as +mammal or +reptile, but those that are classified as -animal are undefined for either mammal or reptile. Thus, as one of our analyses, we have applied a hierarchical clustering scheme (HCS) to the triadic comparison data (Johnson, 1967).

The characteristic tree-like structure imposed by HCS, however, can mask internal variation within a class and so be somewhat misrepresentative of the range of semantic relations involved. Therefore, we have also performed a second analysis of the data, which consists of scaling the words by means of a non-metric multidimensional scaling (MDS) program (TORSCA; Young and Torgerson, 1967).

The two forms of analysis, HCS and MDS, have different implications for a theory of the internal representation of the lexicon. HCS gives a solution in $N-1$ dimensions, where the items are taken to be members of mutually exclusive hierarchical classes. MDS, on the other hand, tries to fit the words into a Euclidian space of minimum dimensionality, stressing the possibility that features are not of an "all-or-none" kind, but rather are present or absent to a specifiable degree. Support for this notion stems from recent developments in linguistic theory (Lakoff, in press). Thus, despite our having provided words whose features can be hierarchically ordered, we have hedged on predicting the outcome by
using both HCS and MDS. Furthermore, concurrent use of the two procedures permits an assessment of the extent to which there is internal variation within a class.

METHODS

Selection of Aphasic Subjects

The aphasic patients included in this study were classified on the basis of clinical examination, laboratory data (brain scan and EEG), and the results of the Boston Diagnostic Aphasia Test (Goodglass and Kaplan, 1972). There were five anterior aphasics and five with posterior damage. All but one of these patients had vascular lesions; the one exception had sustained a traumatic injury to the head. Their ages ranged from 39 to 72 years with a mean age of 51 years. All were male and none had been educated beyond the high-school level.

The results of the aphasia examination also indicated that the members of each aphasic group were homogeneous in terms of the syndrome presented. Thus, all of the anterior patients were clearly of the classical Broca’s type. An important feature of this syndrome is dysprosodic, nonfluent speech characterized by frequent pauses, reduced phrase lengths, and agrammatism. Comprehension for spoken language is relatively intact, however, and the ability to name objects and body parts is spared, although finding the name appears to require effort.

Four of the five posterior patients presented the classical syndrome of Wernicke’s aphasia. An essential characteristic here is fluent but empty speech in which indefinite noun phrases such as “thing” or “someone” are substituted for the expected nouns. The style is at best circumlocutory with inappropriate words occasionally substituted (verbal paraphasias). Comprehension is also impaired as is naming ability.

The fifth posterior patient, although “fluent” in his output, could comprehend ongoing speech at a somewhat better level than is typical in Wernicke’s aphasia. This set him apart from the other posterior patients.

The extent to which a patient could be unambiguously classified as being either a Broca’s or Wernicke’s aphasic was only one of the criteria for inclusion in this study. Among the other criteria was the requirement that subjects be able to read single words and recognize their definitions when read aloud. In addition, they had to be able to cope with the metalinguistic demands of the triadic comparison task. The details of our screening procedures are supplied in a following section. What should be noted, however, is that although these criteria served to exclude many more Wernicke’s than Broca’s aphasics, the subjects finally selected could all be considered at least moderately aphasic.

Control Subjects

Five male patients chosen from non-neurological wards of the Boston VA Hospital served as control subjects. The ages of these patients ranged from 24 to 56 years with a mean age of 46 years. None of these patients had been educated beyond the high-school level.

* Clinical tests of speech comprehension range from an informal assessment of the ability to follow social conversation to more structured procedures such as assessing the ability to answer yes–no questions, to point to named objects and body parts, and to respond to spoken commands.
Stimulus Items: Semantic Features and the General-Residual Distinction

There was one array of 12 words. These words and the manner in which they were defined for the patients are presented in Table 1.

We were guided by a number of considerations in choosing the words. First of all, we restricted the category to concrete nouns of relatively high frequency (Thorndike and Lorge, 1944), presuming that such words would be among the easiest for aphasics to handle (Goodglass, Hyde, and Blumstein, 1969; Wepman, Bock, Jones, and van Pelt, 1956). We then included a very basic conceptual feature: ±_human_. We felt that unless the patients could capture this feature in their comparisons, there would be little likelihood of any useful data emerging from this procedure.

We then sought bases for creating subdivisions within each of the animal and human categories. Our principal interest here was whether or not we could form subdivisions based solely on either general or residual semantic features, thus we had to separate conceptually these two components for each word.

For the six human nouns, this separation was carried out by using a guideline suggested by Miller (1969); namely, that negation is likely to deny only the most residual features of a noun. Consider, for example the item HUSBAND (Table 1). If we hear that someone is not a husband, we assume that the person is not married, we do not assume that the person is not a male. For this item, then, the feature male is taken to be the general one, and _who is married_ is considered to form the residuum. By the same reasoning, male is considered to function as a general feature for KNIGHT; female, as general feature for WIFE and MOTHER; and person, as a general feature for COOK and PARTNER. Applying Miller's (1969) terminology, we have labelled these general semantic components, presuppositions.

If presuppositional structure, as defined above, were to guide the comparisons of the human items, the two most compact relations would be formed by the words WIFE and MOTHER, both of which presuppose female, and by the words HUSBAND and KNIGHT, both of which presuppose male. Further, to the extent that these general fea-

<table>
<thead>
<tr>
<th>Noun</th>
<th>Definition given to subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOTHER</td>
<td>A woman who has given birth to a child.</td>
</tr>
<tr>
<td>WIFE</td>
<td>A woman who is married.</td>
</tr>
<tr>
<td>COOK</td>
<td>Someone who prepares food.</td>
</tr>
<tr>
<td>PARTNER</td>
<td>Someone who shares something.</td>
</tr>
<tr>
<td>KNIGHT</td>
<td>A man who has been honored for his bravery.</td>
</tr>
<tr>
<td>HUSBAND</td>
<td>A man who is married.</td>
</tr>
<tr>
<td>SHARK</td>
<td>A fish that is scary and vicious.</td>
</tr>
<tr>
<td>TROUT</td>
<td>A fish that is harmless and good to eat.</td>
</tr>
<tr>
<td>DOG</td>
<td>An animal that is usually kept as a pet.</td>
</tr>
<tr>
<td>TIGER</td>
<td>An animal that is ferocious.</td>
</tr>
<tr>
<td>TURTLE</td>
<td>An animal called a reptile that has a bony shell.</td>
</tr>
<tr>
<td>CROCODILE</td>
<td>An animal called a reptile that is long-tailed and vicious.</td>
</tr>
</tbody>
</table>

*Note. The decision to use animal, rather than mammal, as the superordinate, or general semantic feature in DOG and TIGER was an intuitive one. It should be noted, however, that people generally use these two features synonymously (Rips, Shoben, and Smith, 1973).
SEMANTIC FEATURE REPRESENTATIONS

Features could be cast into a hierarchical relation, the items COOK and PARTNER would be less compactly related to the other terms. COOK and PARTNER presuppose only human. In comparison, HUSBAND and KNIGHT presuppose male in addition to human, and MOTHER and WIFE presuppose female in addition to human. Thus, grouping either COOK or PARTNER with any of the others requires that one general feature as well as the distinctive residuals be ignored. But grouping WIFE with MOTHER or HUSBAND with KNIGHT requires that only the residual component be ignored (Miller, 1969).

Componential similarities were not restricted to the general semantic features, however. The definitions of the human items displayed commonalities among the residual features as well, such as those offered for WIFE and HUSBAND, both of which contain the residual who is married. Thus, if the residual components were to guide the groupings, HUSBAND, WIFE, and perhaps even PARTNER (see Table I) might be expected to form the most compact relation.

The form of the componential analysis differed between the human and animal items. In contrast to the residual features contained in the human-item definitions, those established for the animal items were not intended to capture highly specific semantic details. In fact, for each animal term, the only "definitional" component presented to the patients beyond species membership was one which expressed whether or not the animal could be considered ferocious. But ferocity, although an attribute, can hardly be considered a distinctive feature. The item TIGER, for example, can be much more uniquely specified by its tawny, black-striped coloration.

On the other hand, we felt that ferocity and species membership could be distinguished from each other in terms of the amount of referential information each encompassed; the former feature requires referential knowledge (first or second hand), the latter is less tied to such knowledge and more dependent upon an understanding of the systematic interlexical relations in the language (Bolinger, 1965; Miller, 1972). In this respect, it should be noted that ferocity is not even listed under the dictionary entries for the senses of SHARK and CROCODILE conveyed here; and even TIGER is only "proverbial for its ferocity" (The Oxford Universal Dictionary, 1955). After all, the extent to which sharks, tigers, and crocodiles are ferocious is in large measure a function of extralinguistic setting. Thus, even though it may be argued that all semantic elements ultimately derive from a knowledge of the world, ±ferocity still seems to us to be a much more empirical or perceptual concept than ±mammal, for example. This difference is all we wish to imply by categorizing ±ferocity as a residual component.

Within this framework, in a taxonomic classification based on shared presuppositions—that is, on the class membership of the nouns—the clusters likely to emerge would be SHARK and TROUT (both fish), TURTLE and CROCODILE (both reptiles, and DOG and TIGER (both mammals, or as common usage has it, both animals). In contrast, were the comparison task to be carried out on the basis of non-dictionary residuals, the emergent features would most likely be +ferocity (SHARK, CROCODILE, and TIGER) and −ferocity (TROUT, TURTLE, and DOG).

Procedure

Each subject was first screened on his ability to read aloud single words. For this we simply chose five words at random from the array, and only patients who could read each of the words correctly were included in this study. The awkward articulations of anterior aphasics were discounted in this measure, but literal paraphasias (the substitution of an incorrect phoneme) were not. Whenever any such error arose, the examiner required the subject to try again, and invariably, the second attempt was correct.

We next assessed how well the patients understood what each of the words in the array meant. One word at a time, printed on a 3 × 5 in. card, was placed before the subject. He
was given time to read the word and then instructed to indicate which of four definitions was the correct one for that word. These definitions were spoken aloud by the examiner with sufficient time between them for the subject to signal his choice. The three alternative definitions in each multiple-choice task were always taken from the same major domain of the array as the correct one. For example, when assessing the ability to define a +human item, the incorrect alternatives were also taken from the +human category. When an incorrect recognition occurred, the examiner pointed out the correct choice and later retested the patient on this word. If a patient made more than three or four of these initial errors he was dropped from the study.

As a final screening and training procedure, subjects were instructed in the process of grouping items on a subjective basis. This procedure consisted of having each patient carry out triadic comparisons of geometric forms of various colors. There were three training trials. For the first two, the patient's task was unambiguous: grouping could only be carried out on the basis of color in the first trial, and in the second, shape was the only relevant dimension. The third array of three forms forced the patient to choose between shape and color and only those patients who appreciated the subjective nature of this last task were tested in the study. None of the patients reported on here had any difficulty with the concept that different grouping decisions were possible within one array.

With respect to formal testing, the words were all individually printed on 3 × 5 in. cards and the three words forming each triad were arrayed in a triangular fashion to permit the grouping of non-adjacent words. There were 220 such triads \((C^3_{11})\). These were delivered in the same predetermined quasi-random order to all of the subjects; however, the placement of the words within each triangular arrangement was randomized anew for each patient.

The patients were always required to consider the three possible combinations within each triad before pointing to the two words that they felt to be the most similar in meaning. It should also be noted that the aphasics often verbalized all of the possible groupings within each triad and that this verbalization was spontaneous because they were only required to indicate the possible combinations by pointing. This would not have occurred had the sorting resulted from default; that is, from lowered attention or from spatial neglect.

The patients were tested individually in short sessions spread out over several weeks. At least once in each session, the examiner reassessed the patient's ability to read and to recognize the definitions of the words used.

**Methods of Data Analysis**

As indicated in the introduction, the data were analyzed by both a hierarchical clustering program (HCS) and a non metric multidimensional scaling program (MDS).

HCS operates on a similarity matrix which in the present work takes the form of an incidence matrix, each cell representing the frequency with which a particular word is grouped with another. The procedure itself consists of working backwards from this unclustered matrix to induce the hierarchical organization that may have influenced the sortings. It is an iterative procedure. At each stage the two words that are most closely related are merged and treated as a single entry in a new and smaller matrix. A visual account of this sequence of mergings takes the familiar form of a tree graph.

\(^5\) We were forced to choose the triadic comparison procedure rather than the far simpler method of having subjects create their own similarity-of-meaning partitions. Pilot work had indicated that aphasics were incapable of sorting when the metalinguistic demands were too unstructured.
The decision to interpret the data matrix as a similarity matrix, or equivalently, as a distance matrix, is plausible and can be mathematically supported (Miller, 1969). What must also be considered, however, is how well the data fit a hierarchical structure. In an ideal hierarchical ordering, the two items merged at each stage should be equally related to the other words of the matrix. But in practice when two words are clustered together they are often not equidistant to a third, and the problems arise of, first, determining whether in fact we are dealing with a hierarchical structure, and second, assigning a number relating the merged unit to a third word.

Following Miller (1969), our solution has been to carry out two parallel cluster analyses: one always based on the stronger relations between the cluster and the other words of the matrix (minimum method), the other always based on the weaker relations (maximum method). To the extent that subjects have been sorting on the basis of an implicit hierarchical structure, there should be no difference between the solutions. In practice, if the minimum and maximum solutions produce only numerical differences of the form of different strengths of the same clusters, the data are accepted as being at least minimally compatible with a hierarchical organization. If the solutions yield striking qualitative (topological) differences, such an assumption is usually not supported.

MDS was also applied to the incidence matrices. However, because the rationale underlying the use of MDS is better known among psychologists and because Shepard (1972) and Torgerson (1958) have already provided extensive reviews of MDS and its limitations, there is no need to repeat these details here. What deserves reemphasis is that applying HCS and TORSCA together should permit more insight concerning the structure of subjective lexical organization than would be possible were either technique to be used alone.

RESULTS AND DISCUSSIONS

Hierarchical Ordering of Similarity Judgments

A similarity matrix was constructed for each subject, the cell entries in the matrix representing the number of times the subject had grouped each word with the other. The entries in the five anterior aphasic matrices were then added together as were those in the five posterior matrices, and those in the five control matrices. All subsequent analyses were based on these group data.

A rough measure was taken of the degree of hierarchical ordering imposed by each group on the items of the array. Consistent with the data treatment outlined, this was done by applying both the minimum and maximum methods of analysis to every group matrix and by noting the extent to which they produced qualitatively similar tree graphs. The data were accepted as being compatible with a conceptual hierarchy only when most of the clusters were common to both methods of analysis. As a minimum requirement, any differences between the two outcomes had to be restricted to the upper nodes of the trees; that is, to relations which were based on the smallest number of groupings.

In general the array permitted the patients to produce a rather tractable set of semantic clusters. Of interest, not all the groups showed evidence of having sorted on the basis of an implicit hierarchical organization: the control and anterior aphasic patients did, but not the posterior aphasics. Although the minimum-method outcomes are displayed for
each of the three groups (Figures 1–3), only those for the control and anterior patients are indicative of hierarchical stability.

One other reservation about the posterior group’s tree graph should be discussed, namely the graph is not entirely representative of all the individual matrices generated by the members of this group. Specifically, one posterior aphasic’s sortings more closely resembled the triadic comparisons of the anterior aphasics than those of his own group. This was the posterior patient who had been set apart from the others on a clinical basis because of his better speech comprehension. However, because his “fluent” output and the laboratory evidence indicated posterior damage, and because his comparisons did not appreciably alter the pattern of entries in the group matrix, his data were treated as posterior data.

The problem of intersubject variability did not arise for either the anterior or control groups. The five anterior aphasic matrices clearly resembled each other as did the five control patient matrices.

Figure 1 shows the outcome of applying the minimum-method analysis to the data generated by the non-aphasic patients. The human items were clearly separated from the animal items, and because the minimum and maximum solutions of the data were almost the same, the human feature may be considered as basic and as one which dominates the remaining verbal concepts implied by the structure.

The animal terms seem to have been sorted mostly on the basis of the general semantic features they presuppose. The two strongest relations formed are between SHARK and TROUT and between TURTLE and CROCODILE; the items of the first cluster both presupposing the feature, fish, the items of the second cluster both sharing the presuppositional feature, reptile. Insofar as TIGER and DOG did not get tightly

---

6 We must emphasize, however, that every posterior patient included in this study could recognize the definitions of the 12 words used in the array.
clustered with each other, we may presume that TIGER presupposed *ferocity* as well as *mammal* or *animal* for the patients. Nevertheless, these items were grouped with each other more often than with any other item.

To a much lesser extent, the human item clusters can also be accounted for in terms of a presuppositional structure. Of the three human items most often grouped together, two—WIFE and MOTHER—are related by their class membership. In addition, the relation of WIFE and MOTHER to COOK and PARTNER can be explained by the notion that the presuppositions are hierarchically ordered. COOK and PARTNER both presuppose the feature, *human*, while MOTHER and WIFE both presuppose *female* which in turn presupposes *human*.

However, the control subjects clearly had recourse to classificatory features other than those contained in the presuppositions of the human items. For example, KNIGHT and HUSBAND, the only items directly presupposing *male*, were rarely grouped together. And the item, WIFE was clustered as strongly with HUSBAND as it was with MOTHER.

As Miller (1969) has pointed out, analysis in terms of definitional presuppositions or general semantic features should not be pushed further than warranted. Within the *+human* array, it is quite plausible that the integrated unit of MOTHER, WIFE, and HUSBAND arose because all three terms imply a household relationship. That is, although the union of these three terms might have resulted from a combined use of general and specific features, it is equally likely that this cluster, and its relation to the other items, was effected from a knowledge of social relationships that cannot be captured by a componential analysis of standard dictionary entries.

However, because normal language use reflects a knowledge of referential situations no less than a knowledge of linguistic relations, it should be expected that the control subjects classify some of the human items on an implicational basis. After all, the referents of these words are an important aspect of day-to-day existence. On the other hand, it is of interest to note that these same subjects were capable of supplying a presuppositional grid for the animal terms. This level of analysis might have been applied because of an absence of common referential landmarks. That is, even though the words, themselves, may be familiar, the referents of most of the animal terms are rarely encountered. As it turns out, however, it is precisely this presuppositionally based type of linguistic conceptualization that appears to be most undermined in aphasia.

The tree graph for the anterior aphasic group is presented in Figure 2. As already mentioned, it may be considered to represent a hierarchical structure. In fact, the minimum and maximum solutions applied to the anterior aphasic matrices produced complete qualitative correspondence.
The anterior aphasic graph resembles the control graph in two respects. First, the human–nonhuman feature again appears as the most dominant verbal concept in the array (but note how the position of DOG violates this typal separation); and second, the human items again seem to have been clustered on the basis of something other than pure dictionary information.

Our tentative hypothesis is that in clustering these human items, both the anterior aphasics and the control patients created social categories in terms of distance from self. Within this scheme, the items ranged from closely allied members of a household unit to someone as remote as KNIGHT, a person known to have existed, but with whom no social relations are possible (e.g., Leach, 1964).

Notwithstanding the similarity of their human-item clusters, however, the two groups were guided by very different considerations in judging the animal items. Although the nonaphasic patients combined these items in terms of shared species membership, the anterior aphasics clearly focussed on what we chose to include as residual information. Thus, the aphasic patients generated two major clusters: one consisting of SHARK, CROCODILE, and TIGER, all ferocious, wild, and remote; and the other consisting of TROUT and TURTLE, both partially edible, and both quite harmless.

The outcome may have been partly a function of the form of the definitional statement given to the subjects. Recall that we had intentionally bypassed the most specific semantic features of each animal term in order to maximize the contrast between a semantic domain based on
general semantic features and one based more directly on a knowledge of the world. But even so, the aphasics still had the option of creating a presuppositional structure. That is, the point to be emphasized is not that the anterior patients were guided by the definitional statements, but rather that they seemed constrained by a recognition of the similarities and differences among the attributes of each noun. What this suggests is that, compared to the lexical structure underlying normal language use, the memory code in anterior aphasia is more restricted in its range of conceptual integration. In effect, verbal concepts in anterior aphasia appear to be more tightly tied to affective and situational data.

Consistent with this notion is the finding that DOG was clustered more often with humans than with animals. Insofar as anterior aphasics gain the meaning of a word in terms of the extralinguistic events it signals, this outcome is to be expected. That is, DOG and the human items presumably got grouped together on the basis of something like dog being man’s best friend (e.g., Leach, 1964).

The anterior aphasics and control subjects not only generated different kinds of clusters for the animal nouns, they also differed in terms of how well they integrated the clusters they produced. The measure of this is represented as the height of the node connecting any two items or any two groups of items; and this height may be considered proportional to the degree of noise inherent in the similarity judgments. That is, the less consistent the comparisons are, the higher the connecting nodes will be. Reading the heights of the animal groupings directly from the two tree graphs, it can be observed that the tightest clusters generated by the controls—between SHARK and TROUT and between TURTLE and CROCODILE—are more compact than any of those generated for the animal items by the anterior aphasics.

It should also be noted that the noise in the anterior aphasics’ judgments is restricted to the animal terms. That is, they clustered the human terms together as strongly as did the control subjects. This suggests that a lexical memory structure emphasizing practical or empirical information is relatively less efficient in integrating words whose referents are rarely perceived. However, to anticipate the outcome of the multidimensional representation, some of the noise may have resulted, not because of an impoverished structure, but rather because of an interfering ability to recognize other types of semantic relations. That is, as the multidimensional representations will hopefully make clear, although lexical organization appears to be weakened and even somewhat dislocated by anterior brain damage, the disruption does not appear to be total. However, this will be taken up in greater detail in the following section.

What should be emphasized at this point is that for the words used in
this study the internalized dictionary of the anterior aphasic can be represented as a hierarchically ordered set of features. Moreover, this organization seems intuitively valid in terms of everyday language use. It might be more restricted and less efficient than that underlying normal language behavior, but that is what would be expected for a patient who must search his memory so laboriously before finding the word he needs. To this extent there appears to be a relation between the psychological processes underlying verbalization and the present description of the internal lexicon.

The case for advancing this relation is strengthened further by the results obtained from the posterior aphasics (Figure 3). We selected patients who could recognize the definitions of the nouns we used. Yet, the clusters generated by these posterior patients are as lacking in specific conceptual features as their speech is empty. Even the basic human–animal concept was not convincingly recovered from the comparisons.\(^7\)

Nor, for that matter, is there any strong evidence of hierarchical stability in their clusters. The only clusters that are common to both solutions are MOTHER, WIFE, and COOK; PARTNER and HUSBAND; and TROUT and TURTLE. The fact that the items in each of these first two clusters are related in terms of shared presuppositions should more or less be discounted here, however. Unlike the other two patient groups, the posterior aphasics often sorted on the basis of how

---

\(^7\)Actually, the maximum-method analysis did show a separation of the human and animal terms (with the exception of the item DOG). Our only reason for choosing to present the minimum-method outcome was because of the emphasis it gives to the more compact clusters.
easily any two words could be used in a copula sentence, as for example, "My mother is a good cook." This example is not based on inference; the process was usually carried out aloud.

In any case, it should be noted that the posterior patients clustered the human items more compactly than the animal items. As with the anterior aphasics, this finding suggests a relative inability to integrate items whose referents do not play an important role in terms of environmental requirements. Beyond this finding, however, nothing interpretable seems to emerge from the posterior group's tree graph, which again suggests that at least a part of the word-finding problem in aphasia is rooted in a disruption of the underlying lexical organization.

Multidimensional Representation of Similarity Judgments

Our intention in carrying out multidimensional representations of the similarity matrices was to supply an alternative to the hierarchical feature model. Although we had provided words whose features could be cast into a hierarchical arrangement, the posterior aphasic subjects, especially, did not always follow the dictates of this organization. We felt, therefore, that a spatial representation might add considerable information to that obtained from the taxonomic structures.

The number of dimensions chosen for each representation was determined with respect to two criteria: (1) how well the dimensionality of the space could fit the data points, using Kruskal's stress index and Torgerson's index (Young and Torgerson, 1967); and (2) how well the dimensionality permitted clarity of conceptual representation. On the basis of these criteria, three-dimensional representations were chosen for both the control and anterior aphasic data (for the two representations, stress .013 and index .99). Although an increase in the number of dimensions would have provided a minimally better fit to the data, interpretability would have suffered.

The semantic space derived from the control group's similarity matrix is schematically presented in Figure 4. Numbers have been provided for the horizontal axis only to indicate the extent of the breaks on this axis.

It is readily apparent that the humaness dimension cannot be characterized as a continuum, but really consists of two discrete classes: +human and +animal. This finding reaffirms the mutual independence of the two domains already noted in the hierarchical analyses.

If a class distinction is also to be found along the diagonal dimension, it is certainly less pronounced. However, given the typal separation of human and animal terms, a problem arises in interpreting the diagonal dimension. Can it be considered psychologically common to both the human and animal domains, or is it different for each? Intuitively, the latter possibility seems to be the more valid. Thus, the human items...
It should also be noted that the arrangement of the animal terms along the diagonal axis is quite consistent with the outcome of the hierarchical clustering scheme. That is, despite the spatial structure of the multidimensional representation, there is still a perceptible division based on species membership, violated only by the placement of TURTLE. Other than this exception, however, DOG and TIGER are lined up at one end; SHARK and TROUT are at the other end; and definitely occupying a middle ground is the aquatic reptile, CROCODILE.

In fact, it is only with respect to the vertical dimension that we find an aspect of normal memory organization not apparent from the hierarchical representation. The placement of the animal items along this dimension more or less defines a continuum of ferocity (although slightly...
violated by the position of CROCODILE), whereas the arrangement of the human items define its counterpart, aggressiveness. The case for specifying the latter continuum is obviously not a strong one, being derived solely from the contrast between KNIGHT and the other items, with HUSBAND occupying a neutral position. Nonetheless, the structure is still shown to be richer than one based solely on shared presuppositions, and this is especially so for the animal terms.

The organization of lexical memory in anterior aphasia is also shown to be less restricted than that inferred from the clustering representation. As indicated by the diagonal axis in Figure 5, the rated distances among the animal items line up, at least partially, in terms of a land-adaptedness continuum. The linear ordering is far from perfect, however; in fact, the position of TURTLE is even more out-of-line than on the control group's representation. But the continuum does emerge, and to this extent suggests that the anterior aphasics retained some appreciation of definitional presuppositions or of general semantic features.

However, the internalized lexicon of the anterior aphasic is still shown to be less well-organized than that of the normal language user. First, the item DOG is again shown to violate the typal separation of human and animal terms, and within each of these domains, the items are much

![Figure 5. Multidimensional scaling solution for triadic comparisons of anterior aphasic patients.](image-url)
more loosely related to each other along the humaness dimension than they are in the normal group’s representation. The animal terms, especially, are spread out. Secondly, the social distance dimension apparent in the normal structure cannot be isolated in the anterior aphasic space. In fact, other than the separation of human and animal items, the only dimensions to emerge clearly are the ones relying heavily on extralinguistic knowledge. These we have labelled fierceness and aggression, the latter being along the diagonal axis in this instance.

Thus, although the memory structure underlying anterior aphasia retains some of the conceptual features of the underlying organization in normal language use, the organizational ties are less tight. Further, there is a shift in the data base: the anterior aphasic’s organization emphasizes extralinguistic information; the normal language user’s structure accommodates to a greater degree features representative of the systematic relations among words.

The posterior group’s similarity matrix could only be interpreted in a one dimensional representation. The human terms were clustered fairly tightly, with the placement of the animal terms accounting for most of the variation in the linear ordering. It should also be mentioned that DOG was more closely tied to the human terms than KNIGHT. Thus, the posterior aphasic’s inability to find substantives in ongoing speech again appears to be reflected by the lack of cross-references in his internal lexicon.

CONCLUSIONS

The primary goal of our analyses was to exhibit and compare the organizational features of long-term lexical memory in the different forms of aphasia and in normal language use. Because both typal and spatial analyses were used for this purpose there is a problem in determining which offers the more accurate representation of semantic feature organization. Had one solution or the other been grossly inadequate (in terms of minimum-maximum comparisons for HCS and stress measures for MDS) we would have had an indication of which to choose. The current state of the art, however, permits no easy judgement when, as in the present study, both analyses yield plausible characterizations of semantic features (Fillenbaum, 1972). Although on an a priori basis HCS is the more appropriate analytic technique for the words used in this study, the application of MDS reveals that even for this set of words, the semantic features are not entirely of an “all-or-none” kind. Only by using both procedures do we see that quantitative dimensions are superimposed on a discrete class structure.

Hybrid scaling techniques are currently being devised to determine the extent to which a structure may be both typal and dimensional
(Degerman, 1970). However, until such scaling procedures are more fully worked out, or until performance measures are developed which unambiguously determine the best structure for a given semantic domain, the concurrent use of typal and spatial techniques seems advisable.

The point to emphasize is that both types of analysis differentiated among the groups in a rather consistent manner. That is, whether the solution was typal or spatial, the semantic structures that emerged for the three groups seem to differ in the range of their application. Within the context of our limited data base, normal adults could shift from a categorization partially related to the referential conditions of their immediate social environment (as shown by the human-item clusters) to one based on an explicit notion of class membership (as shown by the animal-item clusters). The aphasic patients, on the other hand, did not show this shift, being restricted, for the most part, to a more concrete-emotional form of word categorization. Thus, the interpretation we favor is that, although the normal adult has a number of levels upon which to organize his lexicon—some referentially practical, others linguistically practical, the adult aphasic mostly retains those features of words that relate to perceived or imagined environmental situations, especially the affective components of such situations.

The notion that an aphasic retains only partial entries for each word and that these entries show predominantly referential ties, suggests that his lexicon resembles the child’s (cf., Anglin, 1970; Clark, 1971; Leontiev and Luria, 1973). If this is so, however, and if the processes underlying word finding are related to how lexical components are organized, it would follow that children show the same word-finding problems in spontaneous speech that aphasic patients show. Because this is not the case, there are obviously other important linguistic and psychological disruptions in aphasia than those described here. For that matter, although the patients we tested could cope with the metalinguistic demands of triadic comparisons and showed no clinical signs of intellectual or attentional deterioration, the possibility cannot be dismissed that general conditions of brain damage, however well localized, exert some effects on “problem-solving” ability.

Nevertheless, the data do indicate a relation between subjective lexical organization and word finding difficulties in aphasia. The structures produced by the anterior aphasics were hierarchically ordered in a meaningful, even if somewhat restricted, manner; those produced by the posterior aphasics showed no such order: Anterior patients very often find the appropriate substantive for an utterance; posterior patients very often use indefinite nouns or nouns that are “out-of-focus.”

In this sense, although the analyses presented here say nothing of the
processes that might be required to exploit the features of lexical organization, they do suggest that describing what a person knows about his dictionary is of relevance to the exploration of factors affecting actual language performance.

REFERENCES


LAKOFF, G. Hedges: a study in meaning criteria and the logic of fuzzy concepts. Papers from the eighth regional meeting, Chicago Linguistic Society. Chicago: University of Chicago Linguistics Department, in press.


