THE NOUN-VERB INTERSECTION METHOD FOR THE STUDY OF WORD MEANINGS

by

L.G.M. NOORDMAN
(Groningen)

and

W.J.M. LEVELT
(Nijmegen)

The aim of the present study is to investigate the semantic structure of concepts. The question is whether concepts can be characterized with respect to semantically significant features and how these features can be detected. This study is part of a cross-lingual research project initiated by Osgood.

The meaning of a word is assumed to be composed of a finite set of discrete parts or "features". The psycholinguistic interest is in recovering these features from experimental data, resp. testing their psychological reality. It is supposed that the sets of features of two words (in the present case noun phrase and verb phrase) determine in some way the acceptability of the word combination, e.g. *he made a machine* is an acceptable combination, whereas *he made a horse* is not, because a horse is not an artifactual object. The analytic problem, then, is to arrive at a specification of the feature structure of particular words from a set of judgments on their combinatorial possibilities.

This study concerns the categorization of nouns. The sample of nouns that has been selected was Osgood's Atlas of Affective Word Meanings. This Atlas has been adapted and translated into the Netherlands language by Jansen and Smolenaars (see also Jansen en Smolenaars, 1966). The combinations used are noun-verb compounds. Osgood has stated rules for Atlas categorization by noun-verb intersection for American English. These rules concern the selection of 10-15 of the most frequently used verbs that make as many differentiations as possible among nouns. The combinations of these verbs with the Atlas nouns are judged with respect to the acceptability or nonacceptability of the combinations. These judgments yield the experimental data.

In order for the analysis of these data to be linguistically relevant, a psychological model is needed that relates the semantic feature structure to the acceptability judgments. One can conceive of several such models. The one proposed by Osgood is a linguistically inspired model. However, a measurement technique appropriate to this model has yet to be developed. Another model, the distance model, which leads to adequate measurement, is linguistically not very attractive.

The first part of this study deals with the models relating meanings to acceptabilities. First, the distance model will be presented, as well as the techniques of analysis appropriate to this model. Then the model proposed by Osgood is presented and a procedure for the analysis is suggested which, however, is not quite appropriate to the model.

The second part of this study concerns the experiment. The rules proposed by Osgood to select 10-15 verbs that make as many differentiations as possible among nouns will be presented in some detail. The procedure followed in the present study differed only in some minor points from Osgood's one. Along with these modifications the whole process will be discussed that leads to the selection of 13 verbs to be used with the Atlas nouns. The results of the analysis of the noun-verb intersection matrix will be presented. The results will be discussed in terms of the respective models.

TWO MODELS RELATING ACCEPTABILITIES TO MEANINGS

The first model is a distance model. It assumes that the meaning of a word can be represented by a point in a Euclidian space. Semantic features are directions in this space. These directions are not necessarily orthogonal. The projections of a point on these features define the feature structure of the word. Psychological relations between meanings can be represented by distance relations among the points. The distance between two points is an inverse monotonic function of the acceptability of the combination formed by the two words.

The techniques of analysis appropriate to this psychological model are the usual multivariate analysis techniques. The problem, however, is that this psychological model is linguistically not very attractive. First, semantic features are – at least partially – hierarchically ordered. Therefore, many of the directions will be irrelevant for a number of nouns. These nouns will have zero projections on these dimensions; but a zero projection is different from no projection on a given dimension. Secondly, semantic features often have discrete values e.g. [+ living] vs. [– living]. Therefore, they cannot be represented by continuous distances.

In spite of the fact that a distance model is a rather questionable one, it has been used because it allows for the following straightforward analyses of the experimental data:

(a) centroid analysis.

In this analysis only the distances between nouns or groups of nouns will be dealt with. The 13 verbs are considered as variables which span a 13 dimensional space in which the nouns are located; the acceptability judgments are considered as the scores of the nouns on the variables. The nouns are divided into a number of a priori groups. On the basis of the acceptability scores, the centroids of these groups are computed. The distances between these centroids are analysed with Kruskal's MDSCAL algorithm and with the hierarchical clustering scheme of Johnson.

(b) vector analysis.

This part also deals with the categorization of the nouns only. An adapted form of factor analysis is applied to the noun x verb matrix, the nouns being considered as the variables. The multiple correlations of certain a priori features with the factor scores of the nouns are computed.

(c) distance analysis.

A procedure designed to unfold and scale both nouns and verbs has been applied. Both nouns and verbs are represented as points in the same meaning space.

The second theoretical model relating the feature structure to acceptabilities is the one proposed by Osgood (1970). The meaning of a word is assumed to consist of a finite set of discrete features, as it is in the previous model. However, the codings on these features are considered to be either + or 0 or –. A combination of two syntactically related concepts (e.g. noun/verb; verb/adverb; adjective/noun) can be apposite, permissible or anomalous. A combination is assumed to be anomalous if the codings of the words on one or more features have opposite signs (coding + and –). A word combination is assumed to be permissible if for each feature the coding of one or both words is zero. A word combination is called apposite if the codings of the two words on at least one feature are the same and if the codings on the other features are not opposed in sign.

The properties of the model can be described in the following way: the coding of features is discrete, as opposed to continuous; the interaction within features is all-or-nothing (the fusion on one feature must represent the dominant sign or be zero) as opposed to algebraic (same signs summate and opposed signs cancel); the relations between features is segregate (number of shared or antagonistic codings being irrelevant) as opposed to aggregate (final solution depending on e.g. the ratio of shared to antagonistic codings).

It is clear that the multidimensional data analysis techniques are not appropriate to this model. A feature discovery procedure appropriate to this model has yet to be developed.

However, as a check of this model one could start with the above mentioned procedure (c), then compute the multiple correlations for the nouns between their coordinates and their codings on a priori features and remove those features for which correlations are low. Projections of all nouns on the remaining "strong" features are computed, and subdivided into three categories: high, medium, low. In this way each noun gets a +, 0 or – coding on each feature. Also for the verbs these codings can be computed.

Using Osgood's rules concerning the interactions within features and the relations between features, one can postdict the original acceptability judgments on the basis of the computed discrete codings on the features.

This procedure leads to postdictions that are different from those made according to the distance model. This fact is illustrated by Figure 1. The distances A-B, C-D, and E-F are all equal. According to the distance model, the combinations are, therefore, equally ac-

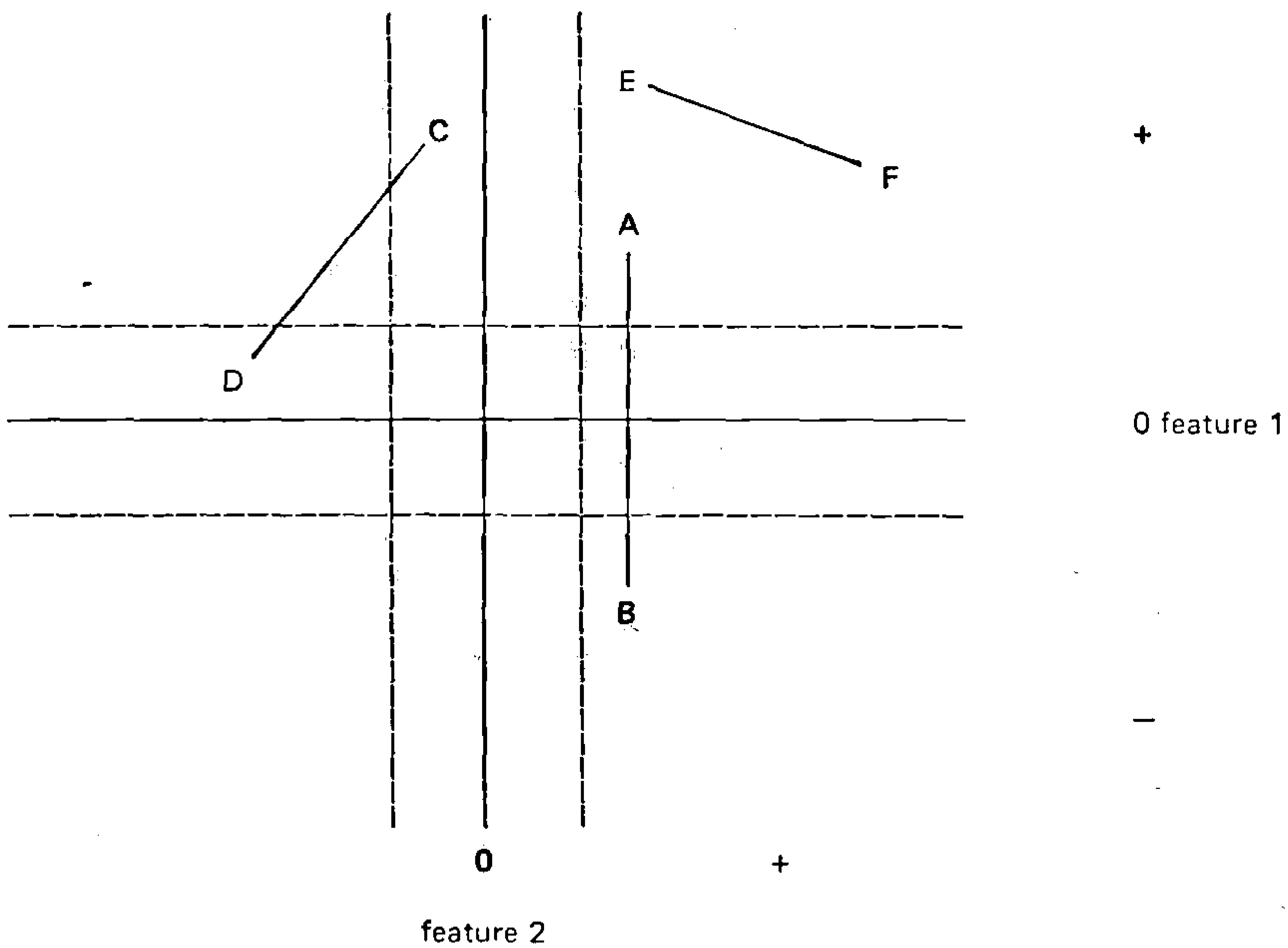

Figure 1. Comparison of the prediction of acceptability scores according to Osgood's model and according to the distance model.

ceptable. According to the model proposed by Osgood, however, A-B is an anomalous combination, D-C is a permissible one and E-F is an apposite one. Because Osgood's model is linguistically more attractive, it is mainly this model and the computation procedure mentioned above that has been used in the postdiction task.

This computation procedure, however, is not adequate to the model; it has all the shortcomings of a spatial model mentioned above. The use, moreover, of multiple regression on discrete codings is questionable. However, the point of the analysis is to obtain, by whatever procedures, codings on features such that the original acceptability data can be predicted according to Osgood's rules. If the analysis, in spite of its shortcomings, succeeds in yielding such a list of feature codings, this can be taken as a confirmation of Osgood's model. If, however, no such result obtains, then this can be due to either inappropriateness of the analysis or inadequacy of Osgood's model.

THE EXPERIMENT

Summary of Osgood's rules for Atlas categorization by noun-verb intersection

(1) Selection of generalized concepts.

Construct a hierarchy of basic features and select "exemplar" concepts (10-15 nouns) for termini of branches. See Figure 2.

(2) Selection of verbs.

Select 40-60 high frequency verbs. The verbs often will be multisense, which is appropriate for the present purpose. Clear homonyms, however, have to be dropped, as well as verbs that function as auxiliaries and verbs that are part of metalanguage. Each verb is assigned to either the transitive or intransitive category. The 40-60 verbs define the columns, the 10-15 nouns define the rows in the intersection matrix. Leave two rows between each of the "exemplar" concepts for later additions.

(3) Assignment of determiners (articles) to nouns

English uses articles to distinguish senses of nouns. It is necessary to be consistent where determiners make a difference in sense. The general rule in assigning articles to nouns is that the dominant meaning of the nouns has to be adopted. If a specific determiner is given in the Atlas translation, then use that form and its implied sense, e.g. A CUBE.

Use rules appropriate for the probably dominant meanings for teen-age subjects in assigning determiners.

Assign determiners to "exemplar" nouns selected in step (1) and to the entire set of Atlas concepts.

(4) Making acceptability judgments.

Acceptability means that the sentence formed by the noun and verb can be said and readily understood in the ordinary (not rare or poetic) use of the language by teen-age native speakers. Acceptability should be indicated by +, non-acceptability by – in the noun x verb matrices. The verbs used

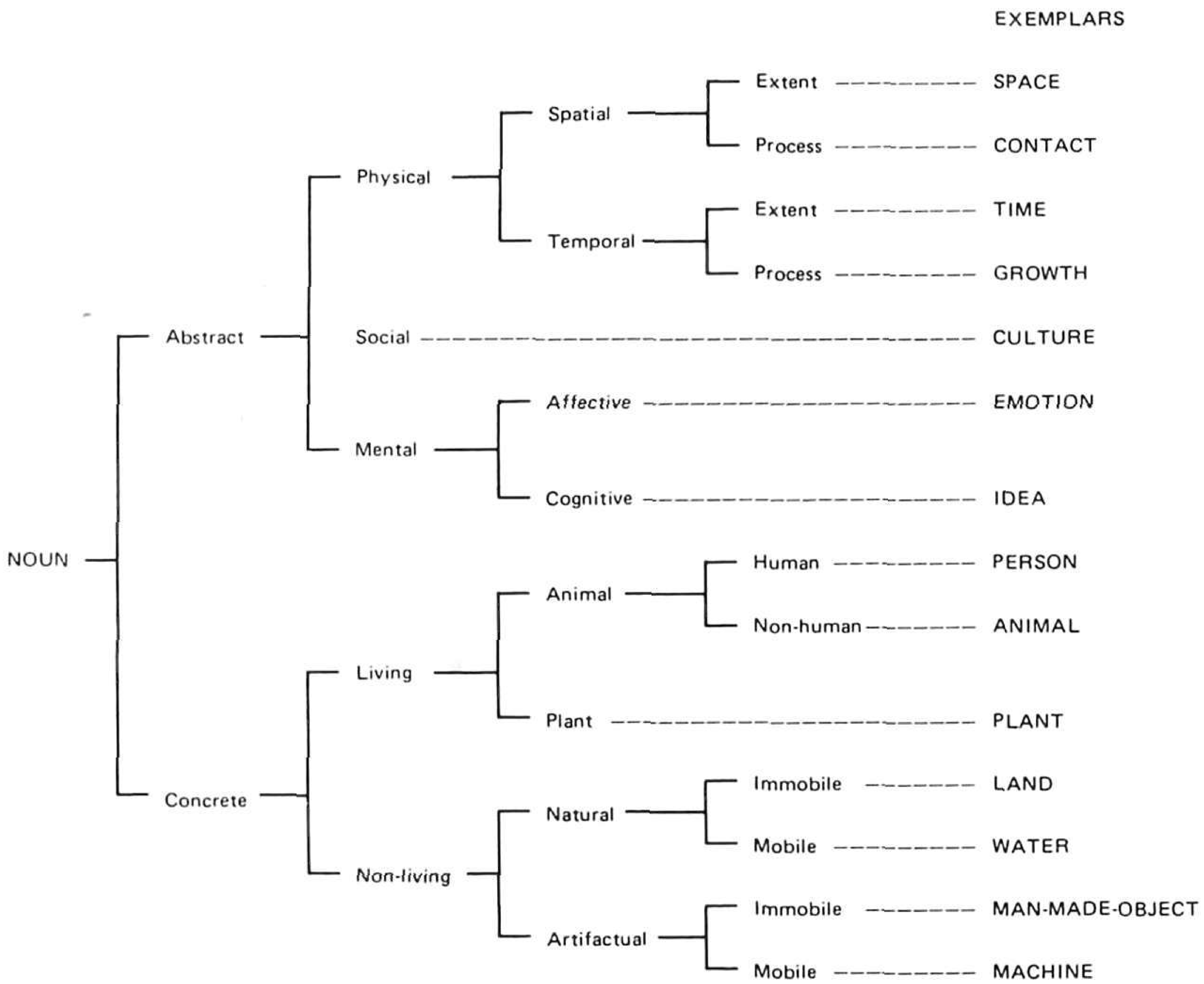

Figure 2. Basic feature hierarchy.

in this procedure are deliberately multi-sense words; therefore, be liberal in accepting sentences in any common sense of usage. In American English the past tense is consistently used, since the present tense is ambiguous and often requires auxiliaries. One may shift pronouns to make better sense with certain nouns; for a few nouns the tense must be shifted (*tomorrow will come*). Always go completely across noun rows (all verbs) since noun senses, once "determined" are more stable than the multi-sense verbs.

(5) Preliminary verb reduction.

Record the +/– ratio for each verb and the ratio for each noun. Provisionally eliminate all verbs with extreme ratios (ratios more extreme than 2/12 or 12/2 in the English example). Assign the remaining verbs to categories on basis of making identical discriminations among the 10-15 "exemplar" nouns. Among pairs of nouns at the lowest level of the tree, select in order that one of each pair having the greatest imbalance in ratio. Change its codings on all verbs to whichever sign (+ or –) is dominant for that noun. Verb categories will then collapse. The nouns at higher levels will be eliminated in the order of the imbalance of their +/– ratios, but a higher level noun may not be eliminated before the nouns under this one are eliminated. Continue this process until there are between 8 and 12 or so categories of identically coded verbs. By this procedure the verbs are categorized in a verb "tree". Select one verb from each category using the following criteria: (a) balanced +/– ratio, (b) difference in coding from other verbs being selected, (c) ease of coding judgments, (d) making interesting distinctions among the nouns.

(6) Expansion of the noun sample.

In the two empty rows on each side of each "exemplar" noun in the original noun x verb matrix, insert two additional nouns drawn from the Atlas concepts. The two added concepts for each "exemplar" should have the same higher order features as the "exemplar" but added (or different) lower order features. The two added concepts should be as different from each other as possible. Code the 20-30 added concepts on the 8-12 verbs selected under step 5.

(7) Refinement of verb sample.

For each noun triplet (each "exemplar" and its two satellites) list those verbs in the 8-12 reduced set which make distinctions in coding. Record the frequencies with which the verbs make distinctions; record the number of "critical" distinctions (i.e. a distinction made by only one verb). Note those nouns within triplets that are relatively undifferentiated by these verbs (in American English: triplets with fewer than two distinctions). Look for additional verbs in the total set of 40-60 which will distinguish among the under-differentiated nouns. List for each underdifferentiated noun pair within triplets those new verbs which would increase differentiation. Sum for each new verb the number of such differentiations, as well as the number of "critical" ones. The verbs with the highest values are "good bets" for addition.

The final verb selection involves simultaneous use of several criteria: (a) the frequency with which a verb makes distinctions, particularly "critical" ones, (b) independence of distinctions made by a verb from those made by others, (c) the contribution of some unique form of discrimination by a verb, (d) judged ease or difficulty in using verb with nouns and making judgments of sentence acceptability, (e) just plain good intuitive "feel" about the verb!

The following procedure has been devised for selecting the American English verbs:
(a) beginning with the original set of 8-12 verbs and then continuing with new verbs, locate each in the verb tree, (b) moving upward in this tree to its middle levels, contrast each test verb against others in the same categories in terms of all criteria just mentioned, (c) select the best in each such set, (d) continue this process until about 10-12 verbs have been selected, those being "representative" of the major limbs of the verb categorization tree. (The 8-12 original verbs do not have to be selected), (e) add a few verbs which seem to make unique discriminations.

The total number of final verbs should be no less than 10 and probably no more than 15.

The selection of the verbs used in the Netherlands noun-verb intersection matrix

(1) Selection of the generalized concepts.
The basic feature tree and the 14 generalized concepts, suggested by Osgood, have been used (Figure 2).
(2) Selection of the verbs.
For the selection of the 40-60 highest frequency verbs, the frequency count of (van Berckel, Brandt Corstius, Mokken and van Wijngaarden, 1965) has been used, summing each verb over tenses. Seventeen verbs have been assigned to the intransitive category; 43 to the transitive category.
(3) Assignment of articles to nouns.
The Netherlands translation of the Atlas nouns has been used. Each noun has been used in the meaning that is dominant in ordinary speech.
Assigning determiners seems to be a very important, but sometimes arbitrary task. The rules given for American English are not quite sufficient, since the meaning of a noun is not always unequivocally determined by a specific article. Especially the generic use of nouns caused some problems (e.g. *beavers build dams*; *the hand is quicker than the eye*). A noun, when used categorically, was considered to have another meaning than when not used categorically (e.g. *they have the atom bomb*; *they saw the atom bomb*), but this difference is not expressed by a difference in the articles.
The categorical use of nouns perhaps deserves special attention, because it changes drastically the acceptability of the noun-verb combinations (e.g. if one uses *the atom bomb* in its generic sense, one cannot say *they saw the atom bomb*). The question is whether the acceptability of the combinations of these nouns with the verbs is more determined by the special use of the nouns than by the meaning the article confers on the

nouns. It is worthwhile to investigate whether the categorically used nouns have a rather similar +/– pattern, in spite of rather different meanings.

(4) Making acceptability judgments.

It was not considered necessary for the acceptability of a noun-verb combination that the sentence, isolated from a larger situational context, could always be understood as a meaningful sentence. As in American English, the past tense was used. The acceptability judgments were made by 3 independent judges.

(5) Preliminary verb reduction.

There were 23 verbs with an acceptability ratio more extreme than 2/12 or 12/2. These were provisionally eliminated.

At the lowest level of the tree, the order of elimination of the nouns was: *groei* (growth), *contact* (contact), *land* (land), *persoon* (person), *machine* (machine). The other nouns were eliminated in the order: *tijd* (time), *emotie* (emotion), *beschaving* (culture), *ruimte* (space), *water* (water), *dier* (animal), *idee* (idea), *plant* (plant), and *object* (object).

In spite of the fact that the noun *object* had the most balanced +/– ratio (29/31), the verb tree is rather asymmetric. This is because the imbalanced verbs are eliminated. Twenty-one of the 23 eliminated verbs were [-object]verbs. The tree is not only asymmetric with respect to [+ object] and [– object]. There are also many more transitive than intransitive verbs: 43 and 17 resp. The verb count of van Berckel e.a. (1965) indicates that among the most frequent verbs, the majority of the verbs are transitive ones. Parenthetically, the verbs that can be either transitive or intransitive have been assigned to the category in which they function most naturally. Among the [– object] verbs there are more intransitive verbs than among the [+ object] verbs (p =.05 one tail test). This accounts for the relatively low number of intransitive verbs in the final list (3 out of 13).

If the concepts *growth, contact, land, person, machine, time, emotion, culture, space* and *water* are eliminated, the verbs are subdivided into groups based on the distinctions made by the concepts *animal, idea, plant,* and *object.* There are 11 groups of verbs. The following verbs were selected: *geven* (to give), *houden* (to keep), *horen* (to hear), *verklaren* (to explain), *trekken* (to pull), *maken* (to make), *leiden* (to lead), *begrijpen* (to understand), *veroordelen* (to condemn), *beginnen* (to begin) *and vaststellen* (to assess).

(6) Expansion of noun sample.

The nouns added from the Atlas in the empty rows were: *vierkant* (square), *lijn* (line) under "exemplar" *ruimte* (space); *aardbeving* (earthquake), *reis* (travel) under "exemplar" *contact* (contact); *toekomst* (future), *jaar* (year) under "exemplar" *tijd* (time); *leven* (life), *ziekte* (disease) under "exemplar" *groei* (growth); *huwelijk* (marriage), *wet* (law) under "exemplar" *beschaving* (culture); *pijn* (pain), *hartstocht* (passion) under "exemplar" *emotie* (emotion); *bedoeling* (purpose), *probleem* (problem) under "exemplar" *idee* (idea); *moeder* (mother), *boer* (farmer) under "exemplar" *persoon* (person); *vogel* (bird), *vis* (fish) under "exemplar" *dier* (animal); *bloem* (flower), *appel* (apple) under exemplar" *plant* (plant); *Afrika* (Africa), *zon* (sun) under "exemplar" *land* (land); *rivier* (river), *vuur* (fire) under "exemplar" *water* (water); *schilderij* (picture), *dijk* (dam) under "exemplar" *object* (object); *atoombom* (atom bomb), *vliegtuig* (airplane) under "exemplar" *machine* (machine).

(7) Refinement of verb sample.

Twenty-one of the 42 noun pairs within triplets were differentiated only once or not at all by the 11 verbs. Seventeen of the 49 verbs that previously had been dropped from the set of 60 verbs made three or more differentiations between noun pairs that were not differentiated by the 11 verbs. The other 32 verbs made less than three, but mostly no distinctions. Among the 49 verbs there were none that made a critical distinction. This is not surprising at all, since for a distinction to be critical at this stage a verb must make a distinction not previously made by a much larger group of verbs.

The 32 verbs that made less than three differentiations among undifferentiated noun pairs within triplets were eliminated. As a partial check of the procedure, the total number of distinctions made by each of these 32 verbs was counted. These frequencies were in general very low. If these verbs had made many distinctions, one would not have been justified in eliminating them, in spite of the fact that they did not differentiate the noun pairs underdifferentiated by the 11 original verbs. Another interesting point is that of these 32 verbs all but two had an extreme +/– ratio and were therefore eliminated in the first stage of the procedure. There were only two imbalanced verbs that made three differentiations among underdifferentiated noun pairs within triplets. It can be concluded that the criteria: balance of +/– ratio; large number of distinctions among underdifferentiated noun pairs within triplets and large total number of distinctions select approximately the same group of verbs.

The final verb selection process dealt with the 11 original verbs plus the 17 added ones, a total of 28. Using the original tree of the 37 balanced +/– verbs to determine which were the major branches, about 10-12 verbs had to be selected. Because the number of selected verbs in each branch had to be proportional to the importance of the branch – the importance being determined by the number of verbs in a branch – a branch of 3-4 verbs had to be represented by one verb in the final list. The verbs had to be selected from the 28 candidates, mentioned above.

The process of selection was started at the top of the tree as follows: the branch [+ object] had 27 verbs, so that eight verbs had to be selected; only 19 of these 27 verbs were candidates for the final verb list: the eight verbs had to be selected from these 19 verbs. Other things being equal, it was preferred to choose verbs from groups that were different at a subordinate level. Thirteen verbs were selected in this manner: *geven* (to give), *krijgen* (to receive), *houden* (to keep), *nemen* (to take), *zien* (to see), *volgen* (to follow), *maken* (to make), *leiden* (to lead), *begrijpen* (to understand), *vaststellen* (to assess), *komen* (to come), *staan* (to stand), *beginnen* (to begin). The verb *vaststellen* was finally dropped because of difficulty in making acceptability judgments. The verb *voorkomen* (to prevent) was added – which figured in the tree in the same major branch as *vaststellen* – because it seemed to make interesting distinctions (social events).

There is a difficulty with additions like this. One eventually will find a distinction, [+ social events] [– social events], emerging out of the final noun-verb intersections. If so, it cannot be concluded that this distinction is an important and frequently used one in the language of teen-age native speakers, because it has not been found by the rather objective selection process, based on the use of the language by teen-age speakers. However, it might be a very interesting one.

The data

The selected verbs and the nouns of the Atlas formed a 13 x 480 matrix (20 out of the 500 Atlas concepts were dropped, because they were not nouns). The acceptability of each combination was judged by two independent judges. When the whole judgment task was finished, the judges discussed the discrepancies in their judgments. Any noun-verb combinations they continued to judge differently, were judged by a third person, eventually by a fourth and a fifth person. This concerned about 200 noun-verb combinations.

The order of presentation of the nouns was such that similar words were equally distributed over the whole list. However, it was impossible for a single judge to make his judgments completely independent from those made by the other judges. This increased the consistency of the judgments, but, in a clustering procedure of the nouns this could be a cause of partially artificial clustering.

The analysis

In the first section the number and size of groups of identically coded nouns will be described. Some groups can easily be interpreted. In the following sections, the data will be analysed according to the distance model.

Number and size of groups of identically coded nouns

Dichotomous coding on 13 verbs maximally yields $2^{13} = 8192$ different + – patterns. However, only 142 different patterns were found. Table 1 lists the number and size of groups of identically coded nouns. It appeared that 400 nouns were grouped together in only 62 groups. Some groups can easily be identified: there is a group of articles of furniture (4 nouns), a group of vehicles (5 nouns), a group of colors (5 nouns), a group of seasons, month, day, night (18 nouns). There are three groups of persons. Not all persons are found in these groups. One consist of 24 nouns indicating relatives and infirm people. This group is distinguished by the verb *leiden* (*they led the N*, N indicating a noun), from a group of 15 nouns indicating professions. This group is distinguished by the verbs *krijgen* (*they got the N*), *houden* (*they kept the N*), *nemen* (*they took the N*) from a group of 4 nouns also indicating professions. There are two groups (5 and 12 nouns) of articles of use and articles of dress. These groups are mutually differentiated by the verb *komen* (*the N came*). There are two groups (22 and 11 nouns) of mass nouns (e.g. butter, iron). These are nouns that were used generically in their combinations with the verbs. The two groups are mutually differentiated by the verb *maken* (*they made N*). There is a group of five animals; they are differentiated by the verbs *begrijpen* (*they understood the N*) and *leiden* (*they led the N*) from two other groups of animals.

TABLE 1. Number and size of groups of identically coded nouns

	Number of groups	Number of nouns in a group
	80	1
	22	2
	13	3
	9	4
	4	5
	14	more than 5
total	142	

Centroid analysis

This part deals only with the categorization of the nouns. It is supposed that the five a priori features Concrete/Abstract, Living/Non-living, Physical/Non-physical, Mental/Non-mental and Social/Non-social play a role in the categorization of the nouns. Furthermore it is hypothesized that the five a priori features are hierarchically ordered (see Figure 2).

Each noun was judged with respect to the five a priori features by three judges. On the basis of these judgments the nouns are divided in five a priori groups: Concrete-Living, Concrete-Non-living, Abstract-Physical, Abstract-Mental, Abstract-Social (hence-

TABLE 2. Percentage of nouns from each noun group that formed an acceptable combination with each verb (left) and Z values for the distinctions Concrete/Abstract and Living/Non-living made by each verb; if $|Z| > 1.96$ then $p < .05$ (right).

	CL	CN	AP	AM	AS	Concrete/ Abstract	Living Non-living
geven gave N	22	62	8	10	13	8.2	−1.6
krijgen received N	32	74	19	29	26	6.8	−2.1
houden kept N	30	43	8	15	14	6.6	1.4
nemen took N	26	67	10	0	8	10.1	−0.6
zien saw N	100	99	22	25	32	16.8	9.1
volgen followed N	68	17	4	8	17	7.5	11.3
maken made N	0	62	13	25	13	4.6	−7.0
leiden led N	40	6	1	0	15	4.9	9.2
begrijpen understood N	61	3	3	44	26	2.1	10.0
voorkomen prevented N	0	1	9	15	22	−5.9	−3.4
komen N came	70	23	26	4	6	7.5	10.8
staan N stood	66	7	0	0	0	9.8	15.4
beginnen N began	68	2	29	6	20	2.8	11.5

forth indicated by CL, CN, AP, AM, AS). The group size was 158, 125, 142, 21 and 34 resp. These five categories correspond rather closely to those found by Miller (1969) in a sorting experiment: Living Things, Non-living Things, Quantitative Terms, Psychological Terms and Social Interactions.

If one tabulates for each of the five groups of nouns the percentage of acceptable combinations with each verb it appears that the features Concrete/Abstract and Living/Non-living are very strong ones (Table 2). Each one of the 13 verbs differentiates the Concrete nouns from the Abstract nouns. For each verb the proportion of acceptable combinations with Concrete nouns is significantly different from the proportion of acceptable combinations with Abstract nouns. Furthermore the Living nouns (CL) are very well differentiated from the Non-living nouns (CN, AP, AM, AS). For 10 out of the 13 verbs the proportion of acceptable combinations with Living nouns is significantly different from the proportion of acceptable combinations with Non-living nouns.

In order to investigate whether the five a priori features play a role in the categorization of the nouns, the distances between the centroids of the different groups of nouns are computed and analysed in the following manner.

The 480 nouns x 13 verbs matrix of + and – (coded as 1 and 0) was standardized by columns, so that every verb contributed equally to the total variance. This matrix can be represented by 480 points in a 13 dimensional space. The centroids of the two main groups (Concrete and Abstract) and of the five subgroups (CL, CN, AP,

TABLE 3. Distances between the centroids of the main groups and the centroids of the subgroups.

A	2.64					
CL	2.06	3.53				
CN	1.71	2.98	3.77			
AP	2.77	0.63	3.66	3.08		
AM	2.81	0.75	3.74	3.06	1.28	
AS	2.59	0.48	3.43	2.99	1.05	0.84
	C	A	CL	CN	AP	AM

AM, AS) were computed as well as the distances between the centroids of the subgroups and the centroids of the main groups (Table 3). It appears that the distances of the centroids of the concrete subgroups to the concrete main group are smaller than those to the abstract main group. The distances of the centroids of the abstract subgroups to the abstract main group are smaller than those to the concrete main group. The groups are very well differentiated according to the Concrete/Abstract dimension.

In order to gain a rough impression of the overlap of the five groups, the distances of each noun to the centroids of the five subgroups were computed. The percentage of nouns of each of the five groups, for which the distance to the centroid of the "own" subgroup is smaller than the distance to the other centroids, is 66, 74, 71, 50 and 25 resp. for CL, CN, AP, AM and AS resp.

Although there is a certain overlap among the abstract groups, the centroids will be considered for the moment as points representing their groups. The five centroids are scaled with the Kruskal MDSCAL procedure. Table 4 represents the coordinates of the centroids in a three, two and one-dimensional Euclidian space. It appears that the three dimensions are three hierarchically ordered features. They can be identified as Concrete/Abstract, Living/Nonliving for the Concrete groups, and Human/Non-human for the Abstract groups. (The stress is necessarily 0,00% in a three dimensional Euclidian Kruskal analysis for five points.) If the five centroids are scaled in a three dimensional Euclidian space, taking into account the absolute values of the distances instead of only the ordinal

TABLE 4. Coordinates of the five centroids in a three, two and one-dimensional Euclidian space (Kruskal's MDSCAL-analysis).

	I	II	III	I	II	I
CL	–1110	– 985	– 015	–1148	–1014	–1998
CN	– 647	1020	008	– 668	1051	577
AP	573	– 045	448	592	– 046	451
AM	643	010	– 292	664	010	516
AS	541	000	– 148	558	000	455
stress		0.00%			0.76%	1.19%

properties of the distances, the results are very similar to those of Kruskal's MDSCAL analysis. The results are represented in Table 5. The dimension of the one-dimensional space corresponds to a Living/Non-living feature.

Because the features seem to be hierarchically ordered, a Johnson analysis was applied to the same material (Figure 3). The diameter and the connectedness method yield the same result. Therefore, it can be concluded that the presented hierarchy reflects the underlying psychological organization of the material. This hierarchy is apparently psychologically more appropriate than the logically required hierarchy according to which first the distinction Concrete/Abstract is made and then the Living/Non-living distinction for concrete material. This conclusion is confirmed by the fact that the feature of the one-dimensional space of the MDSCAL solution is the same Living/Non-living feature.

Summarizing, the analysis up to this point leads to the identification of three hierarchically ordered features: Living/Non-living, Concrete/Abstract and Human/Non-human.

Vector analysis

This part deals only with the categorization of the nouns. In order to investigate the semantic features of the nouns, one can factor analyse the 480 nouns x 13 verbs matrix, considering the nouns as variables. The nouns are then represented as points in a meaning space. Semantic features are directions in this space. The

TABLE 5. Coordinates of the five centroids in a three dimensional Euclidian space (the metric properties of the distances begin taken into account).

	I	II	III
CL	−1.59	−2.05	0.00
CN	−1.15	1.70	0.00
AP	1.35	0.02	0.63
AM	1.36	0.13	−0.74
AS	1.20	0.09	−0.46
stress		5.95%	

feature structure of a noun is represented by the projections of the point on the different directions. A factor analysis of intersection data however, yields a space of one more dimension than necessary to account for the data (Coombs, 1964; Levelt, 1967). The exact factor analytic solution which eliminates the spurious factor is not feasible. Levelt (1967), however, suggested a procedure – the co-variance method – which is an approximation of the exact solution.

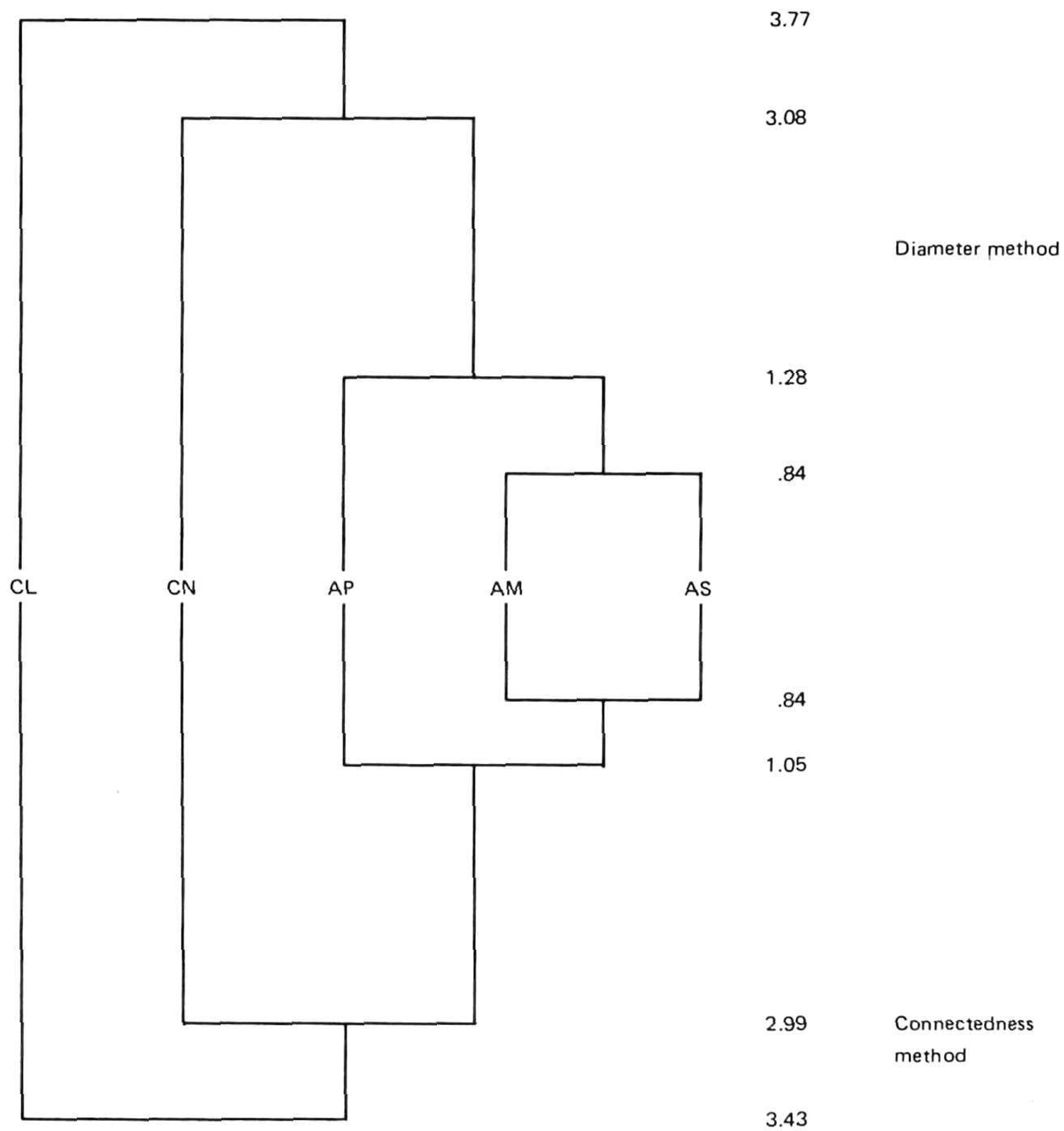

Figure 3. Hierarchical clustering scheme solution (Johnson) for the five centroids.

From each cell of the matrix, the appropriate column mean is subtracted. The factor analysis has to be applied to the matrix of covariances between the variables. This procedure was applied. The factor analysis program for an unlimited number of variables by Hofstee (1970) was used. This program avoids the difficulty of computing the covariances without being a Q-analysis: the computed factor loadings are loadings of the variables (nouns), not of the "elements" (verbs).

Although the factor loadings were rather low (no single factor loading was significant at .05 level) three factors could easily be identified (Table 6). Almost all the Concrete-Living nouns had a positive loading on the first factor; the Concrete-Non-living nouns had negative ones; the loadings of the Abstract words being close to zero. Factor 2 and factor 3 differentiated the Concrete and Abstract nouns.

The factor pattern was rotated according to the Varimax criterion. The first factor was similar to the first factor of the unrotated analysis (Table 6); the second factor was similar to the second and third unrotated factors. The interpretation of the other factors was not possible.

Because there is no reason to suppose that semantic features correspond to orthogonal axes in a meaning space, several oblique factor transformations were applied to the matrix of factor loadings. (Simple loadings solution, promax solution and independent cluster solution or a solution according to the quartimax criterion). These transformations yielded factors that were almost orthogonal. The results were quite similar to those of the Varimax solution.

TABLE 6. Factor pattern and eigenvalues for the 480 nouns

	I	II	III	IV	V	VI	VII	VIII
CL	+	–	–					
CN	–	–	–					
AP		+	+					
AM		+	+					
AS		+	+					
	24.6	7.4	6.9	4.7	4.0	3.6	3.0	2.8

TABLE 7. Multiple correlations between the factor loadings and the E, P and A scores of 476 nouns.

Factor	R	% variance accounted for	F
E	.29	8	**5.3
P	.20	4	2.5
A	.27	7	**4.7

**p < .01

No other features than the two a priori ones were detected by the factor analysis procedures.

With the semantic differential technique three dominant factors or features of word meaning were discovered regularly: an Evaluation factor, a Potency factor and an Activity factor. Because the E.P.A. system is a reliable and valid characterization of at least part of the human semantic system, it is worthwhile to investigate whether these features play a role in the categorization of the nouns. The multiple correlations of the factor loadings with the E scores, P scores and A scores of the 476 nouns were computed (the A, P and E scores for four nouns were not available). The results are represented in Table 7. Although two of the three multiple correla-

TABLE 8. Multiple correlations between the factor loadings and the scores on each of the a priori features of the 480 nouns.

Feature	R	% variance accounted for	F
Concrete /Abstract	.80	64	**119.4
Living /Non-living	.77	59	** 96.4
Physical /Non-physical	.51	26	** 23.8
Mental /Non-mental	.42	18	** 14.4
Social /Non-social	.45	20	** 16.9

**p < .01

tion coefficients are significant, only a very low percentage of the variance is accounted for by the E, P and A scores. If, as Osgood suggests (1970), the E.P.A. system is a characterization of the metaphorical use of the language, this result is not surprising.

In order to test the psychological reality of the five a priori features, each noun was scored 1 or 0 with respect to these features (see above). The multiple correlations were computed between these scores and the factor loadings for the 480 nouns (Table 8). All correlations are highly significant. It should be clear that these a priori features are not independent e.g. a noun that represents an abstract entity also represents a non-living entity if the scoring is dichotonous. A much larger percentage of the variance is accounted for by these a priori scores than by the E, P and A scores.

Distance analysis

This part deals with the semantic space for both nouns and verbs. The noun-verb intersection matrix were analysed by the MINI-RSA program of Roskam (1970). MINI-RSA is an unfolding and scalogram analysis program for a partial ordering of a number of column elements by each of a number of row elements. The program yields a representation of each element as a point in a k-dimensional space in such a manner that – for each row element – the ordering of the distances with the column elements is in correspondence with the ordering in the corresponding row of the data matrix. A measure of stress is defined for the rows; the program minimizes the root-mean-squared stress.

Sixty-one nouns did not form an acceptable combination with any of the 13 verbs; these nouns had to be dropped. The resulting data matrix consisted of 419 row elements (nouns) and 13 column elements (verbs).

The MINI-RSA program yields coordinates for both the nouns and the verbs in the same meaning space. Because the solution in a two dimensional space was a slightly degenerated one, the coordinates of the verbs and the nouns in a three dimensional space are used in the following analysis. The iterative process was stopped at a stress $\leqslant .03$.

In order to test whether an a priori feature had been an actual feature in the acceptability judgments, the earlier described procedure was applied: the multiple correlation was computed between the scores of the nouns on such a feature and the coordinates of the nouns in the three dimensional meaning space. Besides the five a priori features discussed above, some other features were tested.

Three different features Human/Non-human were defined: (1) only human beings are judged as Human and scored 1 while all the other nouns are scored 0, (2) only abstract human concepts are judged as Human, (3) human beings and abstract human concepts are judged as Human. Another feature: Generic/Non-generic is a syntactical one; the nouns which had been used generically in the noun-verb combinations are scored 1; the other nouns are scored 0. These judgements were made by one judge. The multiple correlations are represented in Table 9. All the features except Physical/Non-physical are highly significantly different from zero at the 1% level: $F(3, \sim) = 3{,}78$. It should be remembered that some features are highly correlated with each other. In the next section some other features will be discussed.

TABLE 9. Multiple correlations between codings on a priori features and coordinates (MINI-RSA) of 419 nouns.

Feature	R	% variance accounted for	F
Concrete/Abstract	.45	21	35.78**
Living/Non-living	.65	43	102.53**
Physical/Non-physical	.09	1	1.20
Mental/Non-mental	.30	9	13.95**
Social/Non-social	.37	13	21.25**
Human/Non-human (1)	.69	48	126.29**
Human/Non-human (2)	.52	27	52.15**
Human-Non-human (3)	.61	37	79.69**
Generic/Non-generic	.43	19	31.95**
Object/Non-object	.64	41	96.19**
Artifactual/Non-artifactual	.67	45	115.24**
Temporal/Non-temporal (1)	.22	5	6.99**
Temporal/Non-temporal (2)	.32	10	16.12**

**p < .01

Postdiction of the original data according to Osgood's model

Because the MINI-RSA program yields coordinates for both nouns and verbs in the same meaning space, one can try to postdict the original acceptability scores according to Osgood's rules. Therefore, discrete codings (+ 1, 0, – 1) have to be obtained for both the nouns and the verbs on the semantic features. The following procedure seems appropriate.

A significant multiple correlation between noun coordinates and codings on an a priori feature is assumed to indicate the psychological reality of that feature. The scores of the nouns on each of these features can be computed with the multiple regression technique, the dimensions of the meaning space being considered as the predicting variables and a feature as the predicted variable. This technique then yields continuous scores for the nouns on these features.

The scores of the verbs on the features can be obtained in a similar way. They are estimated by a multiple regression equation using the same regression coefficients as have been used for the nouns. The scores for the nouns and the verbs on the different features must be divided in one way or another into three groups: high, neutral and low, scored as + 1, 0 and – 1. On the basis of these discrete codings the acceptability scores of the noun-verb combinations can be postdicted using Osgood's rules. The apposite and permissible noun-verb combinations in the predicted matrix are then considered as acceptable combinations (+); the anomalous noun-verb combinations are considered as non-acceptable combinations (–).

There are two kinds of erroneous predictions: a predicted acceptable noun-verb combination when the combination is, in fact, a non-acceptable one. This kind of error is called a patchable error by Osgood; in this model a patchable error will be corrected if a feature is found on which the codings of the noun and the verb have opposed signs. A postdicted non-acceptable noun-verb combination, when the combination is in fact an acceptable one, is called an unpatchable error by Osgood; an unpatchable error cannot be corrected in this model.

The difficulty of this procedure is how the range of the codings of the nouns and verb on each feature has to be divided in a plus, zero and minus "area". One can reduce the number of patchable errors by narrowing the zero area, only a few nouns and verbs will be coded zero; this will increase the number of unpatchable errors. One can reduce the number of unpatchable errors by extending the zero area of the range; many nouns and verbs will then be coded zero; but this will increase the number of patchable errors.

The best postdictions have been obtained with only three features: Concrete/Abstract, Living/Non-living and Human/Non-human (3). After some trial and error it appeared that the best way of dividing the range of the codings was to take as division points the mean plus approximately a quarter of the standard deviation and the mean minus approximately a quarter of the standard deviation. On the basis of these features the acceptability of all the noun-verb combinations had been predicted. The number of predictions is 5447. The number of unpatchable errors was only 226 i.e. 4.2%. The number of patchable errors, however, was 1334 i.e. 24.5%. Narrowing the zero area of these features increased the number of unpatchable errors considerably. The features Mental/Non-mental, Social/Non-social, Human/Non-human (1), Human/Non-human (2) and Generic/Non-generic – the range of the codings on these features being divided in the same way as was done for the three features mentioned above – did not improve the postdiction of the acceptability judgments. An inspection of the kind of patchable errors suggested a number of features that eventually would improve the postdiction of the acceptability judgments. Object/Non-object and Artifactual/Non-artifactual were added. Two different features Temporal/Non-temporal, were defined: (1) only periods of time are judged as Temporal, (2) periods of time and temporal processes are judged as Temporal. Only the feature Temporal (1) slightly improved the postdictions: the postdiction using the four features yielded 4.4% unpatchable errors and 23.3% patchable errors.

An inspection of the 226 unpatchable errors does not support the hypothesis that these errors are due to failures in the original acceptability judgments of the noun-verb combinations.

It might be interesting to compare the results obtained using Osgood's model with the results obtained using the distance model. In the meaning space common to both nouns and verbs, – yielded by the MINI-RSA program – a distance *d* had been defined such that two points separated from each other by a distance greater than *d* corresponded to an noun-verb combination which was postdicted as an unacceptable one. A noun-verb combination was postdicted as an acceptable one when the two corresponding points were separated from each other by a distance less than *d*. The best result was obtained for $d = .60$: 1619 out of the 5447 acceptability judgments (i.e. 29.7%) were predicted erroneously. This number is not very different from the total number of erroneous predictions in Osgood's model: 1560 i.e. 28.6%; but there the percentage of unpatchable errors is only 4.2%.

The postdiction using Osgood's model can eventually be improved. One can conceive of several rather specific semantic features which, added to the four features mentioned above, would eventually reduce the number of patchable errors. Although it will be hard to restrict the percentage of unpatchable errors to less than 5%, this study suggests that Osgood's model may be considered as an attractive possibility that deserves further study.

Summary

It is assumed that the acceptability of a word combination depends on the meaning of the words. In the present study an attempt is made to characterize word meanings starting from acceptability judgments of word combinations. Acceptability judgments have been obtained for the combinations of 480 nouns with 13 verbs. Two models are discussed that relate acceptability judgments to the structure of word meanings. The model, which, from a linguistic point of view, is the more attractive, is more problematic than the other model with respect to the statistical analysis of the data. A method is presented for the analysis of the data according to the

linguistically more attractive model. It appears that a very small set of semantic features accounts quite well for the acceptability judgments.

REFERENCES

Berckel, J.A.Th.M. van, Brandt Corstius, H., Mokker, R.J. & Wijngaarden, A. van. *Formal properties of newspaper Dutch.* Amsterdam: Mathematisch Centrum, 1965.

Coombs, C.H. *A theory of data.* New York: Wiley, 1964.

Hofstee, W.K.B. Factoranalyse over onbeperkte aantallen variabelen. *Nederlands Tijdschrift voor de Psychologie*, 1970, XXV, 377-379.

Jansen, M.J. & Smolenaars, A.J. Kort verslag inzake een interlandelijk gestandaardiseerde semantische differentiaal. *Nederlands Tijdschrift voor de Psychologie*, 1966, XXI, 211-216.

Levelt, W.J.M. *Semantic features: a psychological model and its mathematical analysis.* Center for Comparative Psycholinguistics, Report, University of Illinois, 1967.

Miller, G.A. A psychological model to investigate verbal concepts. *Journal of Mathematical Psychology*, 1969, 6, 169-191.

Osgood, C.E. Interpersonal verbs and interpersonal behavior. In J.L. Cowan (Ed.), *Approaches to Language and Thought.* Tucson: Univ. Arizona Press, 1970, 133-228.

Roskam, E.E.Ch.I. Data theorie en metrische analyse. *Nederlands Tijdschrift voor de Psychologie*, 1970, XXV, 15-54 and 66-82.

The authors wish to express their gratitude to Professor C.E. Osgood for reading an earlier draft of this paper well and critically; to Dr. E.E.Ch.I. Roskam who executed the MINI-RSA program; to Drs. A.J. Smolenaars who made available the E.P.A. scores, and to Mr. H. Camstra for the computation of the distances in the centroid analysis.