HEYMANS BULLETINS PSYCHOLOGISCHE INSTITUTEN R.U. GRONINGEN

NR: HB - 70 - 59 EX

Noun categorization by noun-verb intersection for the Dutch language

L.G.M. Noordman and W.J.M. Levelt
Groningen University

Heymans Bulletins contain computer programs (RP), students' theses (SW), internal notes (IN) and pre-publications (EX). These bulletins, usually with the exception of INternal bulletins, can be obtained from:

Secretariaat Psychol. Instituten "Heymans"
Oude Boteringestraat 34
Groningen, The Netherlands

A request to be placed upon the mailing list should also be sent to this address.

August 1970

NOUN CATEGORIZATION BY NOUN-VERB INTERSECTION FOR THE DUTCH LANGUAGE

L.G.M. Noordman and W.J.M. Levelt
Groningen University

1. INTRODUCTION

This study is a part of a cross-lingual research project. Its aim to investigate the semantic structure of concepts. The question is whether concepts can be categorized with respect to semantically significant features.

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. 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. We used the nouns from Osgood's Atlas of Affective Word Meanings. This Atlas has been adapted and translated into Dutch by Drs. M.J. Jansen and Drs. A.J. Smolenaars. 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 which 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 non acceptability of the combinations. These judgments yield the experimental data.

In order for the analysis of these data to be linguistically relevant, we have to develop a psychological model relating 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. We first present the distance model, which is linguistically not quite tenable, and the techniques of analysis appropriate to this model. Then we present the model proposed by Osgood and suggest a technique of analysis which however is not quite appropriate to the model.

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

2. 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 mostly usual multivariate analyses techniques. The problem however is that this psychological model is linguistically not very attractive. 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.

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 part only the distances between nouns or groups of nouns are 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 five a priori groups: a group of Concrete nouns, further subdivided in Living and Non-living subgroups and a group of Abstract nouns, further subdivided in Physical, Mental and Social subgroups. On the basis of the acceptability scores, the centroids of the five 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 deals only with the categorization of the nouns. 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. The codings on these features however are considered to be either + or 0 or -. A combination of two syntactically related concepts (e.g. N/V, V/AV, ADJ/N)can be apposite, permissable 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 permissable 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.

Resuming the properties of the model we can state the following: 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 resolution depending on e.g. the ratio of shared to antagonistic codings).

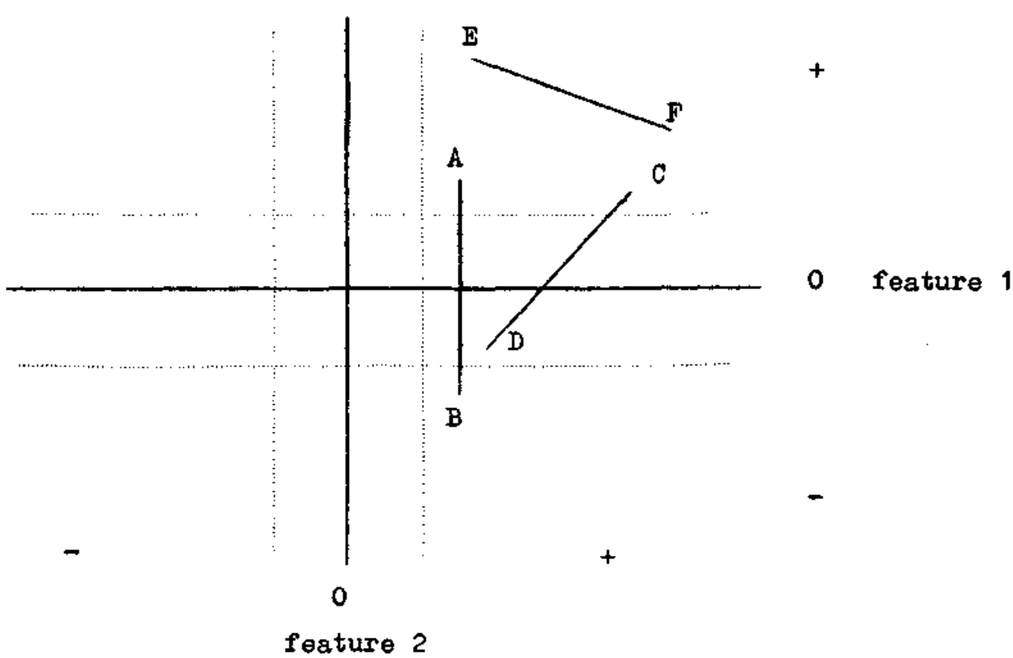
It is clear that the multidimensional data analysis techniques are not appropriate to this model. A feature discovery procedure appropriate to this model has yet to be developed.

However as a check of this model one could start with the above mentioned procedure (c), then compute the multiple correlations between nouns coordinates and codings on a priori features for the nouns and remove those features for which correlations are low. Projections of all nouns on the remaining "strong" features are computed, and subdivided in 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 which are different from those made according to the distance model. This fact is illustrated by fig. 1. The distances A-B, C-D, and E-F are all equal. According to the distance

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



model, the combinations are therefore equally acceptable. According to the model proposed by Osgood, however, A-B is an anomalous combination, D-C is a permissable 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 our analysis, in spite of its shortcomings, succeeds in yielding such a list of feature codings, this <u>can</u> be taken as a confirmation of Osgood's model. If, however, no such result obtains, then this can either be due to inappropriateness of the analysis or to inadequacy of Osgood's model.

3. THE EXPERIMENT

3.1 The rules for Atlas categorization by N-V intersection, as suggested by Osgood.

(1) Selection of 10-15 generalized concepts.

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

(2) Selection of 40-60 high frequency verbs.

If there is no frequency count of verbs available for the language, estimate one by making a verb-type count of 20 to 40 pages of running text. The verbs often will be multi-sense. 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 verbs define the columns, the nouns define the rows in the intersection matrix (40-60 V by 10-15 N). Leave two rows between each of the "exemplar" concepts for later additions.

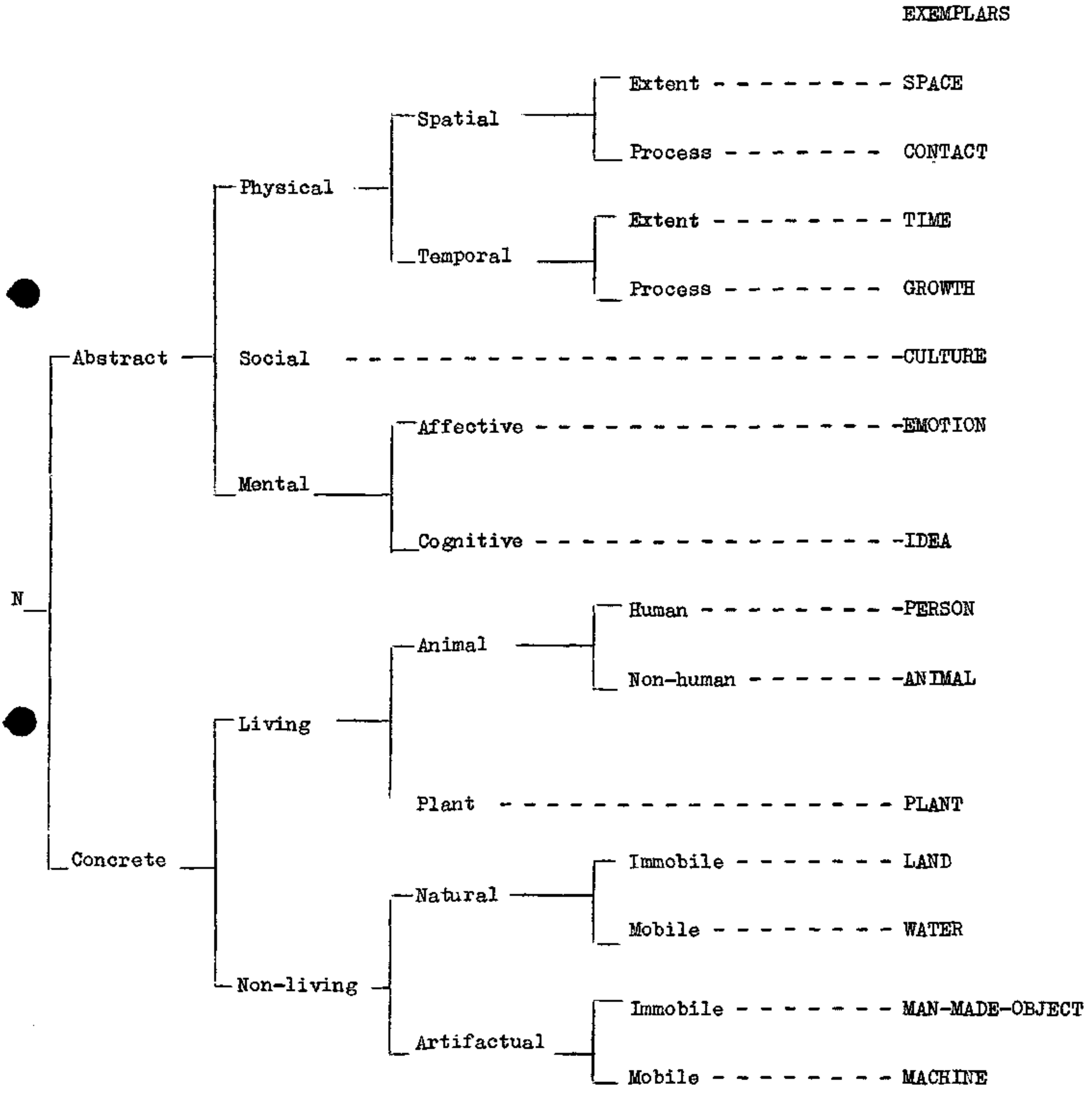
(3) Assignment of determiners (articles) to nouns.

English uses articles to distinguish senses of nouns. It is necessary to be consistent where determiners make a difference in sense. 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 teenage subjects in assigning determiners. For English the following rules were used:

- if N can (or must) be used in singular without any determiner and if the generic (non-determined) sense is the dominant one, use (zero). Ex.: ASIA, CHARITY.
- if N can be either (zero) or (a/the), but the dominant sense is the determined form, use either (<u>a</u> or <u>the</u>) as appropriate to verb being intersected, provided that the use of <u>a</u> vs. <u>the</u> does not change the sense. Ex.: a/the MAN ; a/the STONE.
- if N must be determined and the use of <u>a</u> vs. <u>the</u> does not change the sense, then use either (<u>a</u> or <u>the</u>) as appropriate to verb being intersected. Ex.: a/the BED ; a/the FOREST.

Fig. 2. Basic feature hierarchy



- if N must be determined but the use of a vs. the shifts the sense, then use only that determiner which taps the dominant sense.
  Ex.: the ATOM BOMB (generic sense).

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 N and V 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 N/V matrices. The verbs used 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 was 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 N rows (all V's) since N senses, once "determined" are more stable than the multi-sense V's.

(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" (for an illustration see fig. 3). 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 10-15 noun / 40-60 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 distincions). 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 was used with American English:
(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 categarization 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.

3.2 The selection of the verbs to be used in the Dutch noun-verb intersection matrix

(1) Selection of 10-15 generalized concepts.

We used the basic feature tree and the 14 generalized concepts, as suggested by Osgood. See fig. 2

(2) Selection of 40-60 high frequency verbs.

For the selection of the 40-60 highest frequency verbs, we used the tables of the "Mathematisch Centrum", Amsterdam (van Berckel e.a. 1965), summing each verb over tenses. Seventeen verbs were assigned to the intransitive category; 43 to the transitive category.

(3) Assignment of determiners (articles) to nouns.

The Dutch translation of the Atlas nouns was used. We tried to distinguish senses of nouns and to use each noun in the meaning which is dominant in ordinary speech.

Assigning determiners seems to be a very important but sometimes arbitrary task. The rules given for American English were not quite sufficient, since the meaning of a noun is not always unequivocally determined by a specific article. Especially the generic use of nouns produced problems (E.g., Beavers build dams; the hand is quicker than the eye). We considered a noun, when used categorically, as having another meaning than when not used categorically (E.g., they have the atom bomb; they saw the atom bomb).

The categorical use of nouns perhaps deserve more 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 to the nouns. We will have to investigate whether the categorically used nouns have a rather similar + - pattern, in spite of rather different meanings.

(4) Making acceptability judgments.

A noun-verb combination was judged as being acceptable if the noun and the verb, independently from a larger verbal context, formed a complete sentence which could be said and readily understood in the ordinary speech of teen-age native speakers. However, it was not considered necessary that the sentence, isolated from a larger situational context, could always be understood as a meaningful sentence.

We did not put any restriction on the nature of the relation between N and V. If the combination "They verbed the N" was a meaningful one and if N was an indirect object, then the sentence was judged to be acceptable. This occured only with the verb schrijven (to write) and the noun persoon (person). Because of the imbalance of the +/- ratio, this verb was eventually eliminated.

As in American English, we used the past tense.

The acceptability judgments were made by 3 independent judges.

(5) Preliminary verb reduction.

There were 23 verbs with a +/- ratio of acceptability with the "exemplar" nouns 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).

The process of collapsing the verbs was not started at the bottom, but at the top of the noun tree. The verbs were first separated in a [+ object] group and a [- object] group. These groups were then separated by the noun plant, which gave 4 groups, etc. This procedure leads to the same result as the procedure proposed by Osgood. This procedure seemed to be more practical. For the tree of the 37 verbs, see figure 3.

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. Almost all eliminated verbs were [- object] verbs (21 out of 23). The tree is not only asymmetric with respect to [+ onject] and [- object]. There are also many more transitive than intransitive verbs. The verb count of the "Mathematisch Centrum" indicates that among the most frequent verbs, the majority of the verbs are transitive ones; here: 43 transitive and 17 intransitive verbs. (Some verbs can be either transitive or intransitive. Such verbs are used in 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 explains the relative low number of intransitive verbs in the final list (3 out of 13).

Fig. 3. Preliminary verb reduction; + side of Object
(It was not always possible to give exact
English equivolents for the Dutch verbs)

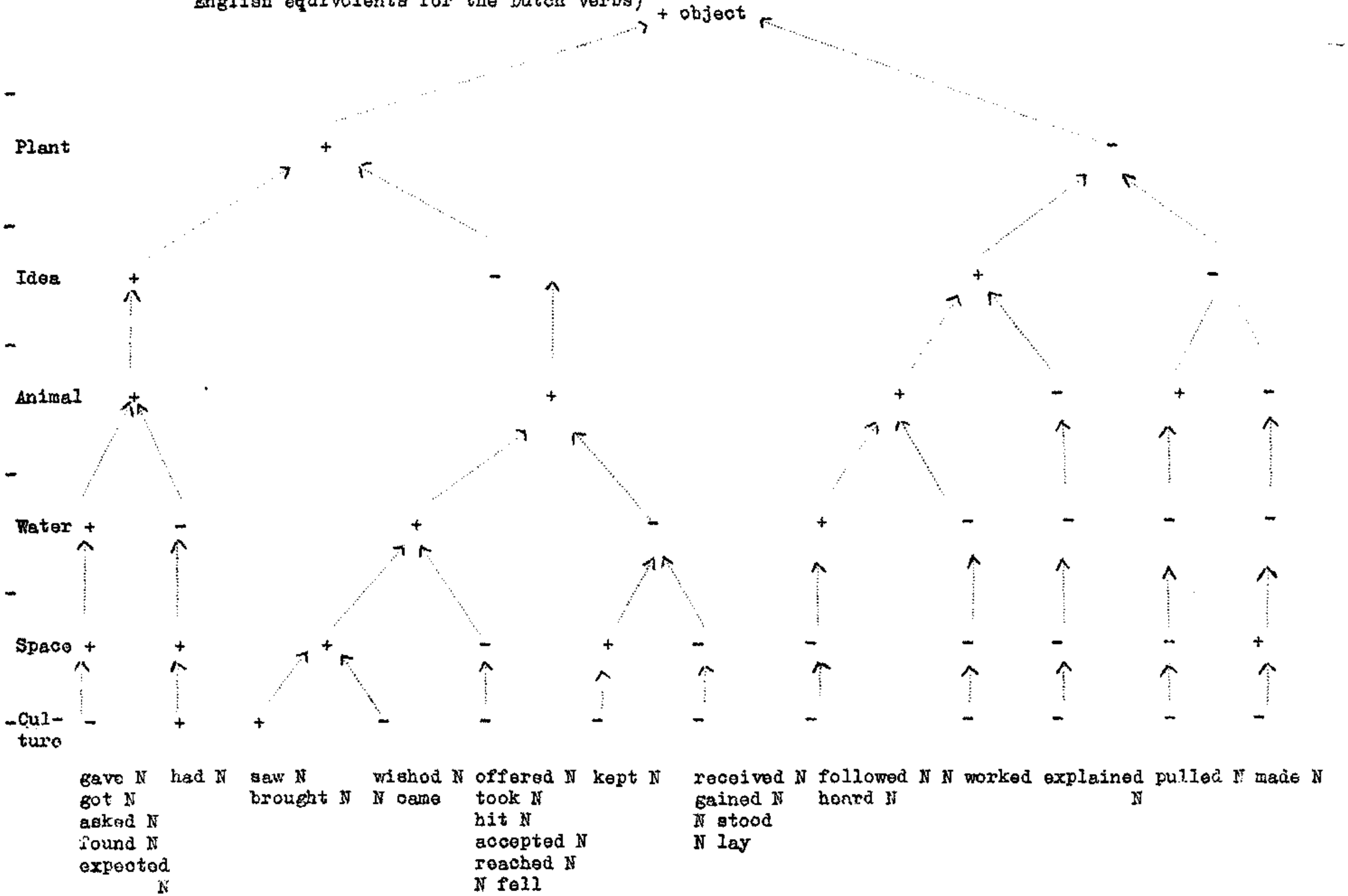
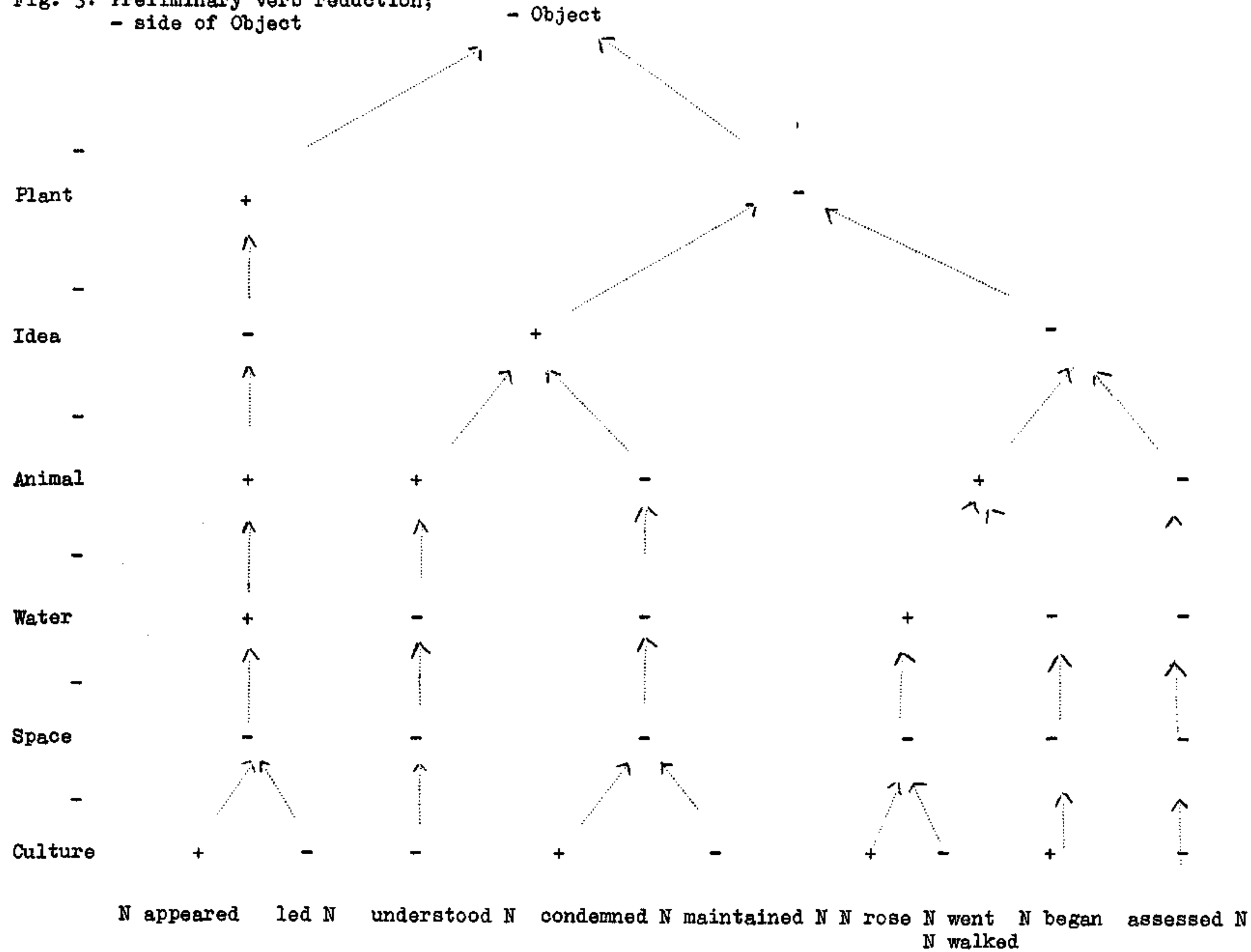


Fig. 3. Preliminary verb reduction;
- side of Object



At the level of dier (animal) there were 11 verb groups. 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).

The procedure up to this point had been executed once before with a group of 14 nouns, of which 2 were different from the group of 14 used here: family instead of culture and vegetable instead of plant (see fig. 2). We got eight verb categories. Six out of the eight selected verbs were also selected in the sample of 11 verbs, mentioned above. The other two verbs could eventually also have been selected in this sample.

(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.

The frequency with which the 11 verbs made distinctions was counted as well as the number of critical distinctions (houden - to keep - and leiden - to lead - each 3; beginnen - to begin -, trekken - to pull - and geven - to give - each one).

Twenty one of the 42 noun pairs within triplets were differentiated only once or not at all by the 11 verbs. We looked in the total set of 60 verbs (among the 49 remaining) for verbs which could differentiate these noun pairs within triplets. We summed for each of these 49 verbs the number of distinctions made among undifferentiated noun pairs within triplets. Seventeen verbs made three or more such differentiations. 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 verbs that differentiated noun pairs within triplets not previously differentiated by any one of the 11 original verbs were also noted.

The 32 verbs which made less than three differentiations among undifferentiated noun pairs within triplets were eliminated. As a partial check of the procedure, we counted the total number of distinctions made by each of these 32 verbs. These frequencies were in general very low. If these verbs had made many distinctions, we 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 which made three differentiations among underdifferentiated noun pairs within triplets. We can conclude 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, we then selected about 10-12 verbs. 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 from 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, we preferred to choose verbs from groups that were different at a subordinate level. We selected 13 verbs 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. We added the verb voorkomen (to prevent) - 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. We eventually will find a distinction, [+ social events] [- social events], emerging out of the final N-V intersections. If so, we can't conclude 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 can be a very interesting one.

3.3 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. N-V combinations they continued to judge differently, were judged by a third person, eventually by a fourth and a fifth person. This concerned about 200 N-V combinations.

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

3.4 The analysis

We start with simply tabulating the number and size of groups of identically coded nouns. Some groups can easily be interpreted. In the next sections the data will be analysed according to the distance model.

3.4.1 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 gives the number and size of groups of identically coded nouns. It appears that 400 nouns are grouped together in only

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
142	

62 groups. Some groups can easily be identified:

There is a group of articles of furniture (4 N's), a group of vehicles (5 N's), a group of colors (5 N's), a group of seasons, month, day, night (18 N's).

There are three groups of persons. Not all persons are found in these groups. One consists of 24 N's indicating relatives and infirm people. This group is distinguishes by the verb _leiden_ (_they led the N_) from a group of 15 N's indicating professions. This groups 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 N's) 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 N's) of _mass nouns_ (e.g. butter, iron). These are nouns that were used genericly 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.

3.4.2 Centroid analysis

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

Each noun was judged with respect to the five a priori features by three judges. On the basis of these judgments the nouns were divided in five a priori groups: Concrete-Living, Concrete-Non-living, Abstract-Physical, Abstract-Mental, Abstract-Social (henceforth indicated by CL, CN, AP, AM, AS). 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 (table 2) it appears that the features Concrete/Abstract and Living/Non-living are very strong ones. The Concrete nouns (categories 1,2) are very well differentiated from the Abstract ones (cat. 3, 4, 5) by each of the verbs. For each verb the proportion of acceptable combinations with Concrete nouns is significantly different from the proportion of acceptable combinations with Abstract nouns (p .05) (Table 3). Furthermore the Living nouns (cat. 1) are very well differentiated from the Non-living nouns (cat. 2, 3, 4, 5). 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 ($p < .05$) (Table 3).

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 N x 13V matrix of + and - (coded as 1 and 0) was standardized by columns, so that each 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, AM, AS) were computed as well as the distances between the centroids of the subgroups and

Table 2. Percentages of acceptable N-V combinations for each of the five groups of nouns

	geven (gave N)	krijgen (received N)	houden (kept N)	nemen (took N)	zien (saw N)	volgen (followed N)	maken (made N)	leiden (led N)	begrijpen (understood N)	voorkomen (prevented N)	komen (N came)	staan (N stood)	beginnen (N began)
CL	22	32	30	26	100	68	0	40	61	0	70	66	68
CN	62	74	43	67	99	17	62	6	3	1	23	7	2
AP	8	19	8	10	22	4	13	1	3	9	26	0	29
AM	10	29	15	0	25	8	25	0	44	15	4	0	6
AS	13	26	14	8	32	17	13	15	26	22	6	0	20

Table 3. Z-values for the distinctions Concrete/Abstract and for the distinctions Living/Non-living

	geven (gave N)	krijgen (received N)	houden (kept N)	nemen (took N)	zien (saw N)	volgen (followed N)	maken (made N)	leiden (led N)	begrijpen (understood N)	voorkomen (prevented N)	komen (N came)	staan (N stood)	beginnen (N began)
Con./Abstr.	8,2	6,8	6,6	10,1	16,8	7,5	4,6	4,9	2,1	-5,9	7,5	9,8	2,8
Liv./Non l.	-1,6	-2,1	1,4	-0,6	9,1	11,3	-7,0	9,2	10,0	-3,4	10,8	15,4	11,5

the centroids of the main groups (table 4). It appears that the distances of

Table 4. 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

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 have 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 given in table 5.

Table 5. Percentage of nouns of each of the five subgroups, for which the distance to the centroid of the "own" subgroup is smaller than the distance to the other centroids.

	%
CL	66
CN	74
AP	71
AM	50
AS	25

Although there is a certain overlap of the abstract groups, we can consider the centroids as points representative for their groups and try to scale these points in a three, two and one-dimensional space. This has been done with Kruskal's MDSCAL algorithm in a Euclidian space and in a city block model. Table 6 represents the coordinates of the centroids

Table 6. 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%

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/Non-living, 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 metric properties of the distances, the results

Table 7. Coordinates of the five centroids in a three dimensional Euclidian space (the metric properties of the distances being 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%	

are very similar to those of Kruskal's MDSCAL analysis. The coordinates are represented in table 7.) The dimension of the one-dimensional space corresponds to a Living/Non-living feature.

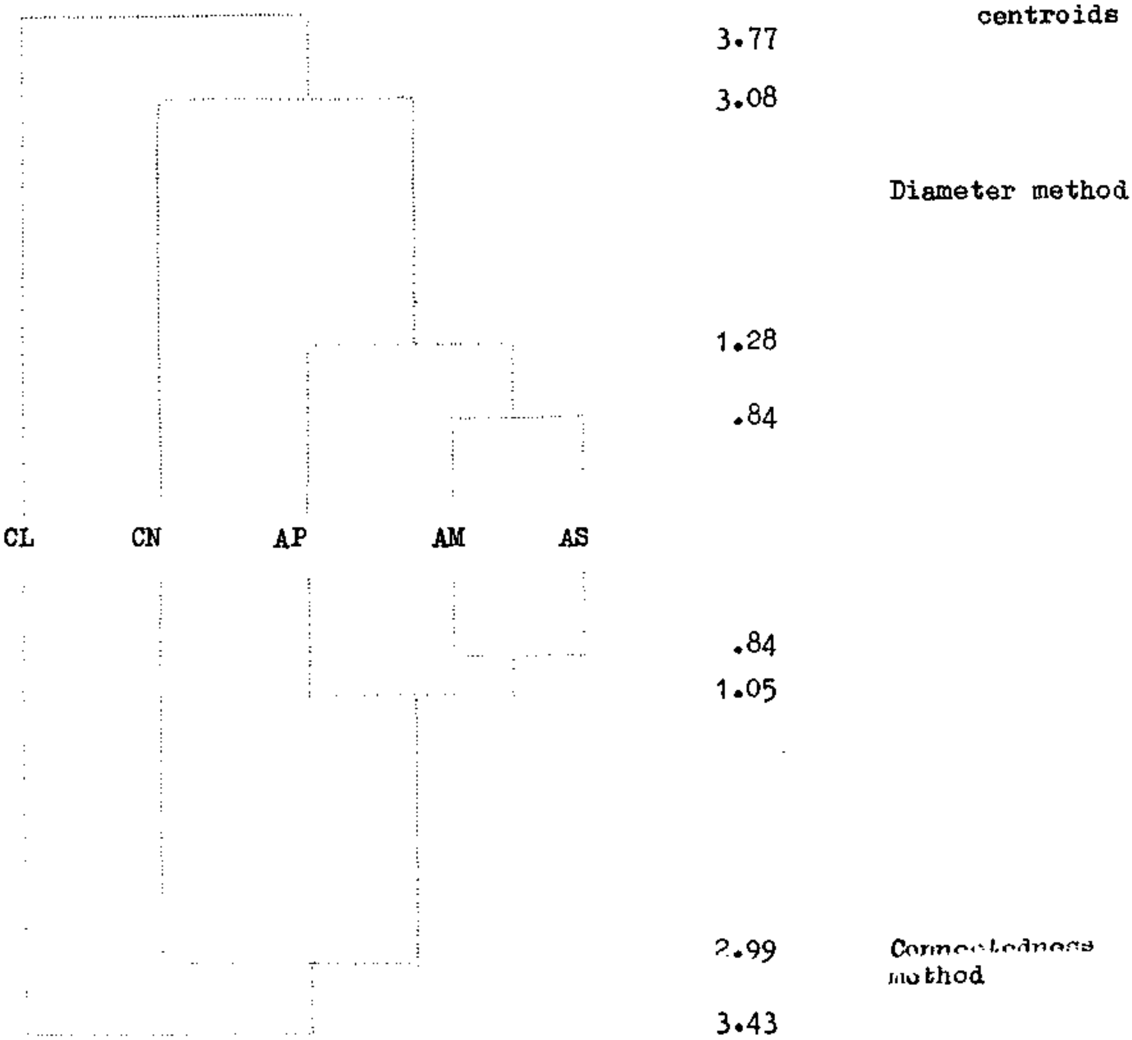
Table 8. Coordinates of the five centroids in a three-two and one-dimensioned city block model (Kruskal's MDSCAL analysis)

	I	II	III	I	II	I
CL	-1569	- 315	038	-1594	- 395	-1820
CN	1034	- 722	- 011	1101	- 731	1257
AP	433	447	316	046	474	190
AM	190	398	- 347	310	374	235
AS	- 088	191	004	138	278	138
stress		1,26%		0,93%		1,58%

Table 8 represents the coordinates for a city block model. Two featues can easily be identified: Living/Non-living and Concrete/Abstract, the former appearing to be the most important one.

Because the features seem to be hierarchically ordered, a Johnson analysis was applied to the same material (fig. 4). The diameter and the connectedness method yield the same result, as was expected.

Fig. 4. Hierarchival Clustering Scheme Solution (Johnson) for the five centroids



The analysis up to this point leads to the identification of the two (hierarchically ordered) features: Living/Non-living and Concrete/Abstract.

3.4.3 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 N X 13 V matrix, considering the nouns as variables. A factor analysis of intersection data however yields a space of one more dimension than necessary to account for the data. The exact factor analytic solution which eliminates the spurious factor is not feasible. Levelt (1967) however suggested a procedure - the covariance method - which is an approximation of the exact solution. This procedure yields a configuration of the nouns, represented as points in a meaning space. Semantic features are directions in this space. The feature structure of a noun is represented by the projection of the point on the different directions. This procedure has been applied. From each cell of the matrix, the appropriate column mean was subtracted. The factor analysis was applied to the matrix of covariances between the variables. 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 9). Almost all the Concrete/Living nouns had a positive loading on

Table 9. 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

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 9); 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.

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 have regularly been discovered: 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, one wanted to investigate whether these features play a role in the categorization of the 480 nouns. The multiple correlations of the factor loadings with the E scores, P scores and A scores of the 480 nouns were computed. (Table 10). Although two of the three multiple correlation coefficients are significant only a very low percentage of the variance is accounted for by the E, P and A scores. (The correlation of the first factor - Living/Non-living - with

Table 10. Multiple correlations between the factor loadings and the E, P and A scores of 476 nouns.

	R	variance accounted for	F
E	.29	8 %	**5,3
P	.20	4 %	2,5
A	.27	7 %	**4,7

** $p < .01$

the A scores was .20; $p < .01$). 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 page 19). The multiple correlations were computed between these scores and the factor loadings for the 480 nouns (Table 11). All correlations are highly significant.

Table 11. Multiple correlations between the factor loadings and the scores on each of the a priori features of the 480 nouns.

	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

** $p < .01$

A much larger percentage of the variance is accounted for by these a priori scores than by the E, P and A scores. Table 12 represents the significant correlations ($p < .01$) between the factor loadings (unrotated) of the nouns and the dichotomous scores on each of the five a priori features. These results correspond very closely to those of the factor analysis (Table 9).

Table 12. Correlations, significant at the .01 level, between the unrotated factor loadings and the scores on the a priori features of the 480 nouns.

	I	II	III	IV	VI
Concrete - Abstract	-.12	-.41	-.64		-.18
Living - Non living	.51	-.42	-.30	-.14	-.17
Physical - Non physical		.31	.24		.30
Mental - Non mental		.27	.25		-.18
Social - Non social		.26	.31	-.14	

3.4.4. Distance analysis

This part deals with the investigation of semantic features common to both nouns and verbs. The MINI-RSA program of Roskam (1970) has been used. 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. These coordinates have been obtained for a six and five dimensional space (the iterative process was stopped at a stress $\leq .05$) and for a four, three and two dimensional space (the iterative process was stopped at a stress $\leq .03$). 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.

To test whether an a priori feature has been an actual feature in the acceptability judgments, the nouns can be scored (1 or 0) with respect to such a feature and the multiple correlation can be computed between these scores of the nouns and their coordinates in the three dimensional meaning space.

Five a priori features had already been defined and the nouns had been judged with respect to them (page 19). The reality of some more features has been tested. Three different features Human/Non-human have been defined: (1) only human beings are judged as Human and scores 1 while all the other nouns are scores 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 genericly in the noun-verb combinations are scored 1; the other nouns are scored 0. These judgments have been made by one judge. The multiple correlations are represented in table 13.

Table 13. Multiple correlations between codings on a priori features and coordinates (MINI-RSA) of 419 nouns.

	R	variance accounted for	F
Concrete/Abstract	.45	21	35,78 xx
Living/Non-living	.65	43	102,53 xx
Physical/Non-physical	.09	1	1,20
Mental/Non-mental	.30	9	13,95 xx
Social/Non-social	.37	13	21,25 xx
Human/Non-human (1)	.69	48	126,29 xx
Human/Non-human (2)	.52	27	52,15 xx
Human/Non-human (3)	.61	37	79,69 xx
Generic/Non-generic	.43	19	31,95 xx
Object/Non-object	.64	41	96,19 xx
Artifactual/Non-artifactual	.67	45	115,24 xx
Temporal/Non-temporal (1)	.22	5	6,99 xx
Temporal/Non-temporal (2)	.32	10	16,12 xx

xx $p < .01$

All the features except Physical/Non-physical are highly significantly different from zero at the 1% level: $F(3,\infty) = 3,78$. It is clear that some features are highly correlated with each other. In the next section some other features will be discussed.

3.5. 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' rules (pag. 4).
We therefore have to obtain for both nouns and verbs discrete codings (+1, 0, -1) on semantic features. These can be obtained by the multiple correlation technique which has already been mentioned.
The procedure is as follows:

A high multiple correlation between noun coordinates and codings on a priori feature is assumed to indicate the psychological reality of the feature. The scores of the nouns on each of these features are 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 yields continuous scores for the nouns on these features.

The scores of the verbs on the features are 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 are divided in three groups: high, neutral and low, scored as +1, 0 and -1. On the basis of these discrete codings the acceptability scores of the N-V combinations are postdicted using Osgood's rules. The apposite and permissable N-V combinations in the predicted matrix are considered as acceptable combinations (+); the anomalous N-V combinations are considered as non-acceptable combinations (-).

There are two kinds of erroneous predictions: a predicted + noun-verb combination when the combination is in fact a non-acceptable one. This kind of errors is called patchable errors by Osgood; in this model a patchable error will be corrected if a feature is found on which the codings of the nouns and the verbs have opposed signs. A postdicted - 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 verbs on each feature is 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 postdiction process was started with the features Concrete/Abstract, Living/Non-living, Human/Non-human (1), Human/Non-human (3), Generic/Non-generic. After some trial and error it appeared that the best way of dividing the range of the codings on the features Concrete/Abstract, Living/Non-living and Human/Non-human (3) was to take as division points the mean plus approximately half a standard deviation and the mean minus approximately half a standard deviation. Using the codings on only these three features, the acceptability of the N-V combinations for 323 nouns had been predicted. The number of predictions is 4199. The number of unpatchable errors was only 146 i.e. 3.5%. The number of patchable errors however was 1002 i.e. 24%. 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 has been 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 which eventually would improve the postdiction of the acceptability judgments. Object/Non-object and Artifactual/Non-artifactual were added. They did not improve the postdiction of the original scores. Two different features Temporal/Non-temporal have been 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 3.7% unpatchable errors and 22.5% patchable errors. (The number of patchable errors could eventually be reduced to 20.5% at the expense of a larger number of unpatchable errors: 5%).

The feature Temporal had been selected in order to correct some patchable errors of time concept-verb combinations. Surprisingly enough this feature did not improve any of these postdictions but mainly postdictions for some concepts indicating Non-living things.

An inspection of the 154 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 N-V combination which was postdicted as an unacceptable one. A N-V combination was postdicted as an acceptable one when the two corresponding points were separated from each other by a distance less than d. Because the fit of the data to the model is obtained by minimizing a root-mean-squared stress for the rows (see 3.4.4), this model will not be expected to yield a very good postdiction of the acceptability judgments. The best result was obtained for d = .60 : 1283 out of the 4199 acceptability judgments (i.e. 30.5%) have been predicted erroneously. This number is not very different from the total number of erroneous predictions in Osgood's model: 1101 i.e. 26.2%, but there the percentage of unpatchable errors is only 3.7%. As has been said earlier however, the comparison of the results obtained by the two models is not a fair one.

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 mentiones 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 which deserves further study.

References

Berckel, J.A.Th.M. van e.a. (1965). Formal properties of newspaper Dutch. Amsterdam: Mathematisch Centrum.

Hofstee, W.K.B. (1970). Factoranalyse over onbeperkte aantallen variabelen. Ned. Tijdschr. voor de Psych., XXV, 377-379.

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

Miller, G.A. (1969). A psychological model to investigate verbal concepts. J. Math. Psych., 6, 169-191.

Osgood, C.E. (1970). Interpersonal verbs and interpersonal behavior. Approaches to Language and Thought (J. Cowan, Ed.) Univ. Arizona Press.

Acknowledgments

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 results represented in table 4 and 5.