

TREE REPRESENTATIONS OF ASSOCIATIVE STRUCTURES IN SEMANTIC AND  
EPISODIC MEMORY RESEARCH

Hervé Abdi  
Laboratoire de Psychologie, Dijon  
Jean-Pierre Barthélémy  
E.N.S.T., Paris

Xuan Luong  
Laboratoire de Mathématiques, Besançon

We expose some research in the area of psychology of memory involving proximity or distance matrices. We propose some ways of building up such matrices. Then we detail an algorithm allowing the representation of proximity matrices by an additive tree, and contrast this new algorithm with previous ones. Finally, we examine some results obtained with this method.

1. INTRODUCTION

The general purpose of this paper is to emphasize the utility and describe the use of additive trees in order to describe data collected in the field of the psychology of memory. This paper is threefold: we first describe some research leading to the construction of distance or proximity matrices; secondly we expose and detail the construction of an additive tree as a representation of the original matrix; finally, we examine the results obtained.

The utilization of clustering methods for attesting the organization of memory or revealing its structure has been strongly advocated recently by some authors in different areas of cognitive psychology (see, among others: Miller (1969), Henley (1969), Friendly (1978), Rosenberg et al (1968), (1972), (1982)). Most of the used methods amount to represent the original matrix by an Ultrametric Tree. Recently, there has been an attempt to build some methods leading to representations less stringent than the classical Ultrametric Tree, i.e. the Additive Tree (see Carroll & Chang (1973), Cunningham (1974), (1978); Sattath & Tversky (1977)). We propose hereafter an (economic) heuristic giving an Additive Tree from a proximity matrix and illustrate it with some examples borrowing from our current research or from classical papers.

## 2. A BUNCH OF EXAMPLES

### 2.1. BARTLETT & "THE WAR OF THE GHOSTS"

In 1932, Bartlett asked a few subjects to read an American Indian folk tale (named "The War of the Ghosts") and to recall the story on several occasions (a method called "repeated reproduction"); in a variant of the method (i.e. "serial reproduction") a chain of different subjects is used, the first being shown the original text and then recalling for the second subject who would pass it to a third and so on. These classical experiments of Bartlett will serve here to illustrate a set of informatic procedures, the aim of which is to build some distances between texts.

#### A) Informatic procedures

For reasons of compatibility, the programs are written in Standard Microsoft Basic (Under CP/M), and at least 64 K-Bytes of RAM and a disk unit are needed. Although these programs include some various possibilities, we will restrict ourselves to the part dealing specifically with the construction of metrics between texts.

It must be clear that when we speak of the text given by a subject, we could speak as well of a set of themes or ideas given by a subject providing an adequate coding of the raw data.

The texts are first transformed in a disk file, then for each text we build the Lexicon associated with it. This Lexicon could be either a Boolean Lexicon (i.e. it merely indicates the Presence or the Absence of the item of Vocabulary) or an integer Lexicon (i.e. it indicates the number of Occurrences of each item). From the different Lexicons (Boolean or Integer) we build - by union - a general Lexicon that defines the Vocabulary shared by the different texts.

#### B) Construction of distances between texts

Depending on the point of view adopted, we could define different distances; as an illustration we examine three ways:

- (i) the texts as subsets of the Vocabulary
- (ii) the texts as Bi-partitions of the Vocabulary
- (iii) a "probabilistic" generalization.

Denote by  $Li$  the Lexicon associated with a text  $Ti$ , the general Lexicon by  $V = \cup Li$ , and by  $\bar{L}T$  the complement of  $Li$  in  $V$ .

(i) Each (Boolean) Lexicon\* is a subset of the Vocabulary and we could use, for example, the well-known distance between sets, the so called cardinal of the symmetric difference:

$$d(Ti, Tj) = |Li \Delta Lj| = |Li \cap \bar{L}j| + |\bar{L}i \cap Lj|.$$

(ii)  $\{Li, \bar{L}i\}$  defines a Bi-Partition of  $V$  (i.e. a Partition with two classes), and so does  $\{Lj, \bar{L}j\}$ . So, we could use some distances between Partitions (cf. Arabie & Boorman (1973)) or Bi-Partitions, e.g. the distance of the symmetric difference between  $i$ -partitions that can be expressed as:

$$\begin{aligned} d(Ti, Tj) &= 2(|Li \cap Lj| + |\bar{L}i \cap \bar{L}j|) (|Li \cap \bar{L}j| + |\bar{L}i \cap Lj|) \\ &= 2(|\bar{L}i \cap \bar{L}j|) (|Li \Delta Lj|) \end{aligned}$$

(iii) In order to take explicitly account of the Integer Lexicons we could look for an extension of (i). With each  $Ti$  is associated a probability measure on  $V$  (i.e. the frequency of the different words); denote the probability of item  $x$  of text  $Ti$  by  $Pi(x)$ ; then we find a family of distances by

$$d_q(Ti, Tj) = \sum_{x \in V} |Pi(x) - Pj(x)|^q$$

## 2.2. FEATURES OF PERSONALITY

This research lays on the border between the work on the organization of the semantic memory and the work upon the "implicit psychology". The purpose is to describe the subjective organization of the qualifiers of the character. As a matter of fact, it has often been noted that we tend to group subjectively some features of character as if we had an "Implicit Theory of Personality" (cf. e.g., Rosenberg et al (1972), Wemer and Vallacher (1977)). In this experiment we select fifty eight qualifiers of the character (using some Thesaurus and a bit of literature,...). These qualifiers are then printed on separate cards and given individually to twenty-eight subjects with the request that he or she sort the cards into piles with the constraint that "the cards in a same pile give the feeling to go together"; subjects were free to choose the number of piles for sorting (for a review of the pro and contra of the sorting method, see Rosenberg (1982)).

\*Notice that one could "fuzzy" the "boolean" distance in (i) and (ii) by taking the fuzzy equivalent of the union and intersection, i.e. Min and Max.

Hence, each subject expresses his opinion by a partition on the set of the qualifiers, and for using as a  $\Omega$ -Methodology (cf. Kerlinger (1973)) the afore evoked distances between partitions could easily be used. The partition given by a subject is associated with a matrix whose rows and columns represent the qualifiers, and where we put a 1 at the intersection of a row and a column if the qualifiers are not sorted in the same pile by the subject. Obviously this is a distance matrix (cf. Miller (1969)), and so will be the matrix defined by the sum of the matrices of the different subjects. In this matrix we simply count the number of subjects who do not put together the qualifiers. Dually we could have defined a matrix of co-occurrences by the sum of the so-called incidence matrix (where a 1 means that the qualifiers are in the same pile), for commodity reasons this is the matrix we give later (cf. Table 3).

It must be noted, in passing, that the Data obtained and consequently the distance matrix depend upon the methods designed for obtaining such Data. In particular, other methods (e.g. word associations, or distances between words in free-recall, etc.) lead to other results (see Abdi (1983)).

### 2.3. OLDIES BUT GOLDIES

It could be useful to compare different methods on well-known data. We will examine here three sets of data (included two "classics").

The first comes from a study by Miller (1969) pioneering the structure of semantic memory. Fifty college students, each sorted forty-eight words according to their similarity of meaning (the subjects were given a brief definition of each noun taken from a dictionary). The co-occurrence matrix (similar to the matrix previously discussed, cf. 2.2) was subjected to hierarchical clustering using Johnson's (1967) connectedness and diameter method. These results are recalled in 4 (cf. Table 4 and Figure 5).

The second is extracted from a study conducted by Henley (1969). She obtained from twenty one subjects an estimation of the distance between twelve animal terms (for each subject she simply counts the number of terms separating the terms of interest in a list given by the subject provided with the instruction to "list all the animals they could"); the matrices are then standardized and the mean for each cell (e.g. across subjects) give the entry of a dissimilarity matrix.

Finally, Friendly (1979), in a paper akin to the two previous ones,

strongly advocate the use of hierarchical clustering to diagram the memory organization. He proposed a proximity matrix obtained somewhat like Henley except that the original data came from free recall experiments. The data were "a priori" organized in three subsets: animal terms, parts of the human body, vegetables. The aim of the study was the recovery of this a priori structure (cf. Bousfield (1953), Tulving (1972)).

### 3. METHODS

Broadly speaking, the methods dealing with proximity data can be divided under two general categories: The Maps and the Graphs.

The Maps encompass - in particular - the different variants of Multidimensional Scaling (cf. Kruskal (1978), Schiffman et al (1981), Young et al, to appear), Optimal or Dual Scaling or Correspondence Analysis depending on the authors (cf. Benzecri (1973), Hill (1974), Nishisato (1980)) or classical scaling or "triple analysis" if the proximity could be judged as distance (Torgerson (1958), Benzecri (1973)).

The general characteristic of the Graphs is to represent the objects as vertices and the proximity (or the distance) between them by arcs (with or without valuation). One could possibly explore the structure of these graphs by looking for the connected components, or its centroid (cf. Abdi (1980)) or by imposing some threshold for drawing the arcs between vertices (cf. The Analysis of Similarity developed by Flament et al (1962), and Degenne and Verges (1973)). One can eventually impose some restrictions on the graph, i.e. being a tree. We detail this point in a moment.

#### 3.1. TREES

A general definition of a Tree can be a cycle free, connected, undirected graph; for convenience assume a valuation on the edges, and define the distance between two vertices as the length of the path from one vertex to the other.

The Ultrametric Tree could be characterized - besides other conditions - by the classical "ultrametric inequality". Three vertices - say  $x, y, z$  - on an Ultrametric Tree verify:

$$d(x, y) \leq \max [d(x, z), d(y, z)] .$$

This property is quite drastic and barely verified by similarity measures obtained from an experiment (see Miller (1969) for a stimulating discussion).

So the problem is to build up the better ultrametric approximation of a given similarity matrix, and a lot of algorithms are available for this job (cf. Sneath and Sokal (1973), Hartigan (1971)). Instead of imposing the ultrametric inequality on the representation of a dissimilarity (or distance) matrix, some authors have proposed to weaken the ultrametric inequality in order to obtain a more general and more natural representation (cf. Carroll and Chang (1973), Cunningham (1974), (1978); Sattath and Tversky (1977) called following the authors weighted tree, free tree, path length tree or unrooted tree (cf. among others: Riordan (1958), Hakami and Vau (1964), Buneman (1971), Dobson (1974)). This kind of a tree is characterized by the following inequality, holding for every four-unlet - say  $x, y, u, v$  - :

$$d(x,y) + d(u,v) \leq \max [d(x,u) + d(y,v); d(x,v) + d(y,u)] .$$

It can be shown that the ultrametric inequality is stronger than the additive inequality which, in turn, implies the triangle inequality. The algorithms for fitting an additive tree to a proximity matrix are less numerous than those designed for the species ultrametric tree. The construction of the additive tree can be separated (for the clarity of the explanation) into two parts:

- 1) the finding of the tree-structure (we dare not use the term skeleton...)
- 2) the estimation of the valuation of the branches of the tree.

The classical approach (as illustrated by Cunningham (1974), (1978); Sattath and Tversky (1977)) makes a direct use of the additive inequality for the finding of the tree-structure, and then estimates the valuation with a least-square method. For our part, we propose another approach that we detail in a moment and contrast with the classical one.

### 3.2. CONSTRUCTION OF THE TREE STRUCTURE

(A) Sattath and Tversky (1977) introduced - in a seminar paper - the notion of "loose cluster".

Let  $E$  be a set and  $d$  a dissimilarity function defined on  $E \times E$ , hereafter denoted as  $\langle E, d \rangle$ . A subset  $A$  from  $E$  is a loose cluster if, for all  $\{x, y\}$  from  $A$  and for all  $\{u, v\}$  from  $E$  (with  $x \neq u, x \neq v, y \neq u, y \neq v$ ), the following inequality is verified.

$$d(x,y) + d(u,v) < \min [d(x,u) + d(y,v); d(x,v) + d(y,u)] . \quad (1)$$

If  $E$  corresponds to the terminal vertices of a tree, then the following

inequality holds for all four-unlet  $\{x, y, u, v\}$ :

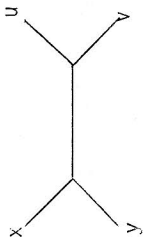
$$d(x,y) + d(u,v) \leq d(x,u) + d(y,v) \leq d(x,v) + d(y,u) \quad (2)$$

(or any of the five inequalities that can be deduced from this one by permutation).

Then if  $\{x, y\}$  is a loose cluster on a tree, from (2) and (1) follows:

$$d(x,y) + d(u,v) < d(x,u) + d(y,v) = d(x,v) + d(y,u) .$$

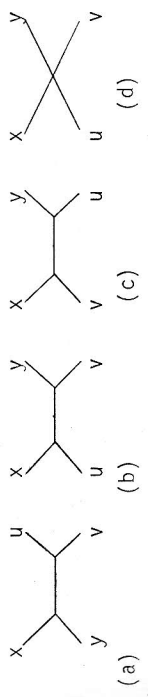
So, in the following tree with terminal vertices  $x, y, u, v$



$\{x, y\}$  on the one hand, and  $\{u, v\}$  on the other are loose clusters, and (2) is then defined. In particular,  $x$  and  $y, u$  and  $v$  are to adopt the following configuration and are said to merge.



Conversely, knowing that  $x, y, u, v$  could be represented by a tree, we get only four possible cases:



Here, the knowledge of a loose cluster is sufficient to characterize fully the representation.

- (1) If at least one pair is loose then:

- $\{x, y\}$  loose  $\Leftrightarrow \{u, v\}$  loose  $\Leftrightarrow$  tree a
- $\{x, u\}$  loose  $\Leftrightarrow \{y, v\}$  loose  $\Leftrightarrow$  tree b
- $\{x, v\}$  loose  $\Leftrightarrow \{y, u\}$  loose  $\Leftrightarrow$  tree c

- (11) No pair is loose, and then we get the tree d.

These previous remarks could be used in order to build the tree structure fitting a  $\langle E, d \rangle$ .

For each  $\{x, y\}$  from  $E$  the Sattath and Tversky's algorithm counts the number

of pairs  $\{u,v\}$  verifying (1); call this number the score of the pair  $\{x,y\}$ . The algorithm then merges the highest score pair - say  $x,y$  - into a new element - say  $z$  -. The dissimilarity between  $z$  and the remaining elements is simply obtained by the average of the dissimilarity of  $x,y$ . The process is then reiterated until all elements have been merged. It can easily be verified that this algorithm gives the exact tree of  $\langle E,d \rangle$  if it exists, providing that the interior vertices of the tree are of degree 3.

(B) We first give some definitions.

(i) Scoring subset and score.

We use the same notations as before. Let  $k$  be an integer less than the cardinality of  $E$ .  $\{a,b\}$  is element of  $E \times E$ , and  $E^* = E - \{a,b\}$ . Then we say that

a  $k$ -subset - say  $S$  - of  $E^*$  is a scoring subset for  $\{a,b\}$  if for all  $x,y$  from  $S$  the following inequality holds:

$$d(a,b) + \text{Max } d(x,y) < \text{Max}[d(a,x) + d(,by)] .$$

The number of  $k$ -scoring subsets for the pair  $\{a,b\}$  is called the  $k$ -score of  $\{a,b\}$  and is denoted  $k\text{-Sc}(a,b)$ .

Let  $\langle E,d \rangle$  be an additive tree and  $k$  equal to 2. The so-called canonical distance gives to every branch of the tree the length 1. Now we can give the:

Proposition: If the canonical distance between two terminal vertices - say  $a,b$  - is two, then  $k\text{-Sc}(a,b)$  is maximal and  $k\text{-Sc}(a,u) = k\text{-Sc}(b,u)$  for every  $u$  of  $E$ .

The proof is left to the reader, but can eventually be found in Luong (1983) with some other properties of the scores.

(ii) Construction of the tree structure.  
Set  $k=2$ . For every pair  $x,y$  we count the number of scoring subsets. We compute then a score matrix. We look for the pair with the maximal score - say  $a,b$  -, then we merge  $a,b$  to give  $c$  and we pose:

$$k\text{-Sc}(c,u) = [k\text{-Sc}(a,u) + k\text{-Sc}(b,u)] / 2 .$$

The process is then reiterated until all the remaining elements have been merged. This procedure is justified by the previous proposition. Notice in passing that we could straightforwardly extend this procedure to any  $k$  (cf. Luong (1983)).

3.3. ESTIMATION OF THE VALUATION OF THE EDGES

(A) Sattath and Tversky (1977), Cunningham (1974), (1978), propose a least square method for estimating the length of the branches of the tree. Precisely, if  $d$  are the distances between terminal vertices on the tree and  $\delta$  the dissimilarity observed,  $d$  is to satisfy:

$$\text{MIN} \sum_{x,y \in E} [d(x,y) - \delta(x,y)]^2$$

This problem is equivalent to the matricial equation:

$$C^t C X = C^t \delta ,$$

where  $\delta$  is the vector of the observed dissimilarity,  $C$  the incidence matrix of the branches and  $t$  denotes the transposed matrix.

Moreover Sattath and Tversky indicate the existence - but without giving it - of a method avoiding the computation of the inverse of a matrix.

(B) The general idea of our procedure is to make a geometrical estimation of the distance; i.e. we embed the dissimilarities in an Euclidean space, and use the geometrical properties of this space in order to obtain an estimation of the distances on the tree.

The tree obtained as described above, is a binary tree. If  $d$  is a distance defined on a binary tree, representing the pair  $\{a,b\}$  by  $z$ , the best possible representation puts  $z$  on the nearest vertex from  $a,b$ . In this case, the distances  $d(z,u) - u$  belonging to  $E$  - are easily estimated:

$$d(z,u) = [d(a,u) + d(b,u) - d(a,b)] / 2 .$$

In the more general case where  $d$  is a dissimilarity, we introduce a "central point" called  $g$ , with:

$$d(g,u) = \left[ \sum_{v \in E} d(v,u) \right] / n ,$$

for all  $u$  in  $E$ , with  $n$  = cardinality of  $E$ .

We then embed the dissimilarities in an Euclidean space. If  $\{a,b\}$  has a maximal score, we build the subtree  $z$  b. The distances on the tree will then be evaluated by the remaining dissimilarities  $d(a,u)$   $d(b,u)$  for all  $u$  from  $E$ . Now  $z$  represents  $\{a,b\}$ ; we make the hypothesis that  $z$  must be brought nearer to  $g$ . According to this, we can determine by the geometry the new dissimilarity  $d(z,u)$  for all  $u$  from  $E - \{a,b\}$ . The processus is then reiterated parallelly with the score procedure described previously.

the original text for ease of comparison. For convenience, we give the distance matrix (Table 1) and the Map resulting of a Factorial Analysis of Distance (F.A.D.) applied to this matrix (Figure 2).

One striking result is the fidelity of the subjects to themselves as could be seen on the additive tree. Moreover, one could easily detect the accurate subject(s) here, for example, the subject LP is fairly accurate both on the same day or 120 days after. These results are worth noting, because the emphasis in the literature is generally put on the gist of the recall for these kind of data instead of Verbatim recall. Nevertheless, the fidelity of the subjects to themselves could be due to the stability of their habits of language as well as to the "quality" or constancy of their memory. The F/D Map gives essentially the same conclusion, although the dimensions revealed by the Map are not easily interpreted. Nevertheless a discrepancy must be noted: the Map indicates L 120 far from ORIG and L1 constringing with the additive tree which gives L 120 near ORIG and L1. In fact the examination of the data matrix reveals that L 120 and L1 are close to each other, so do L1 and ORIG but ORIG and L 120 are quite far from each other. So the two methods differ when a general view is not obvious.

The additive tree obtained from the distance matrix for the "serial reproduction" method allows the recovery of the serial structure (i.e. a semi-order) of the data (although the additive tree model is probably not the better one for this kind of problem). As previously, a simple look at the tree suffices to distinguish the inaccurate "Gaps" (e.g. between R3 and R4, and R6 and R7).

4.2. THE QUALIFIERS OF THE CHARACTER

The data matrix 58 x 58 represents actually our biggest example treated by our methods. The efficiency of our algorithm appears here: we needed seven hours (on a micro computer) to obtain the tree and the valuations on the arcs; in contrast, the classical algorithm would have needed -at the very least - three days in the same conditions.

We have included here in addition to the additive tree, the three-dimensional result of a F.A.D. of the same data. Although there is no major discrepancy between the two approaches, the additive tree is clearly more readable. Probably because the implicit model here is one of cluster rather than one of different dimensions. To be more precise, the sorting instruction induces the construction of clusters and the high degree of agreement.

3.4. EVALUATION OF THE ALGORITHM

The procedure we proposed is clearly heuristic, and so, it could be interesting to have an evaluation of the quality of it. One possible way can be to generate some random distance matrices, to build the additive tree, then to rebuild the distance matrix and to measure the degree of fit between the original matrix and the rebuilt matrix. And so we did. We decided to choose as a measure of fit the classical product moment correlation, but it is well known that r is strongly related to some other measures of fit, e.g. the stress of Kruskal. The results of our simulations are given in table 1.

Number of trials	set size	average r	standard deviation	Min	Max
25	12	.886	.036	.84	.95
20	18	.884	.034	.833	.930
16	24	.861	.029	.811	.914
18	30	.863	.016	.843	.88
10	36	.860	.022	.82	.89
8	42	.855	.018	.81	.905
4	48	.840	.021	.82	.875

Table 1

Goodness of fit measure between random Euclidean distance matrices and rebuilt distance on an additive tree.

It could be seen that the overall degree of fit is quite good. In fact, the results obtained fall closely near the results obtained by Pruzansky et al (1982) with the original ADDTREE of Tversky.

3.5. AVAILABILITY

Our algorithm is translated into standard Basic (CP/M) and a listing of the program is available on request.

4. SOME RESULTS

The different figures and tables are given at the end of the paper.

4.1. THE WAR OF THE GHOSTS

We examine here the results obtained with the distance between texts seen as Bi-partitions (as defined in 2.1.B.ii). Figure 1 gives the additive tree obtained with the "repeated reproduction" method. We have included

between subjects (cf. the data matrix) leads obviously to a "cluster" solution. Nevertheless, some distinct clusters of qualifiers can be identified and be interpreted in terms of "implicit psychology", but this interpretation remains to be verified experimentally, later on.

4.3. OLDIES BUT GOLDIES

For the different examples evoked, we give the original data matrix, the hierarchical solution provided by the authors and the additive tree (except for Henley for which we give a Map resulting from a F.A.D.). The comparison is left to the reader as a passtime....

TABLES AND FIGURES

ORIG	0	ORIG	N1	N15	HP1	HP8	L1	L120	P1	P45	P105	P1145	R1	R15	R45	X180
N1	81	0														
N15	81	32	0													
HP1	119	102	106	0												
HP8	121	102	108	42	0											
L1	78	85	89	107	113	0										
L120	110	91	99	97	99	88	C									
P1	118	83	101	91	97	102	92	0								
P45	109	88	94	100	102	107	95	59	0							
P105	113	92	104	94	100	105	97	49	52	0						
P1145	115	80	84	76	80	107	87	71	80	76	0					
R1	80	59	65	75	83	92	82	80	77	87	53	0				
R15	91	62	68	76	84	93	79	75	78	86	46	23	0			
R45	101	68	68	78	92	103	77	79	86	90	42	37	20	0		
X180	113	78	90	84	94	111	87	83	92	90	58	65	58	56	0	

Table 2

Distance matrix between 15 texts obtained by Bartlett (1932) with the method of "repeated reproduction".

The distance is the distance of the symmetrical difference between Bi-partitions (see text). The subjects are denoted by the letter(s) beginning the Label, the figures following the letter(s) indicate the number of days between the first presentation of "the war of the Ghosts" and the recall.

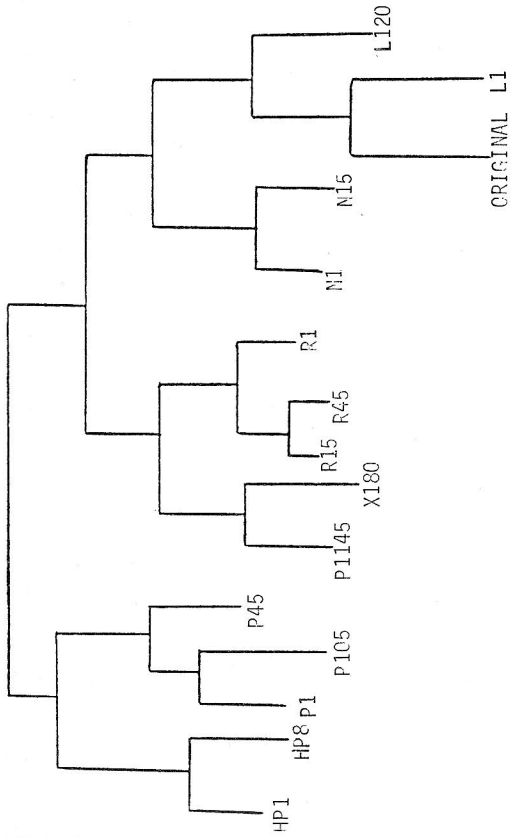


Figure 1

Additive tree obtained from the distance matrix of Table 2. The subjects are denoted by the letter(s) beginning the label, the figure following the letter(s) indicate the number of days between the first presentation of the "War of the Ghosts" and the recall.

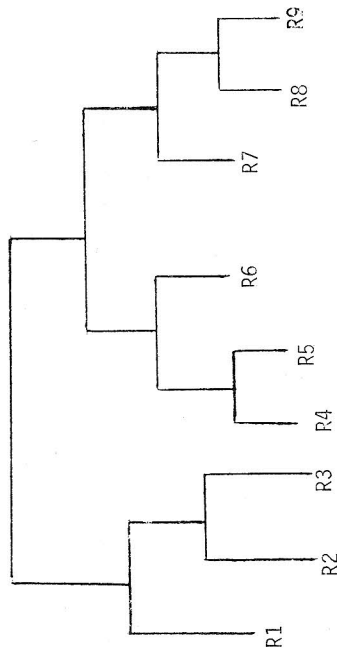


Figure 3

Additive tree obtained from the distance matrix built from the data of Bartlett (1932): "Serial reproduction". The labels give the serial order.

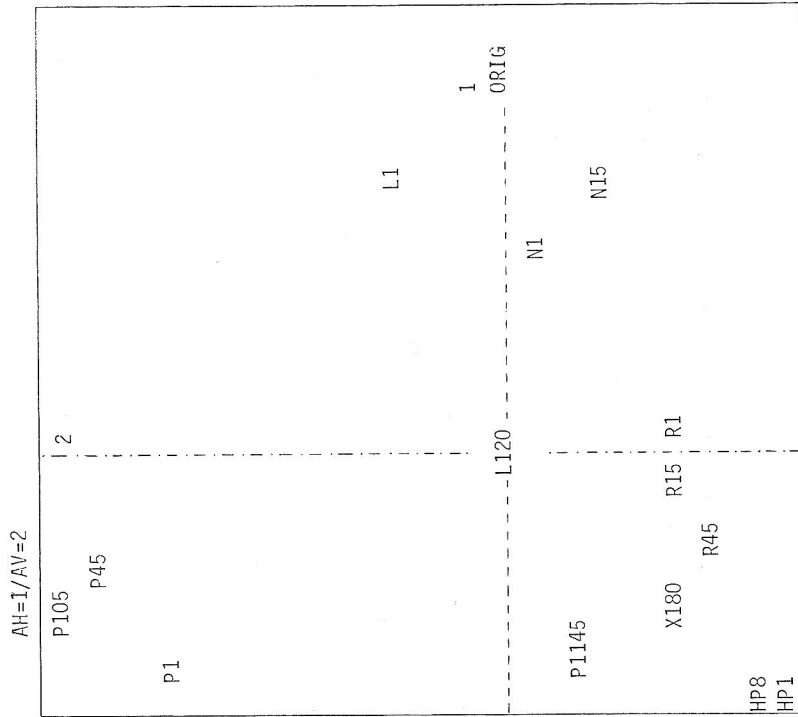


Figure 2

Euclidean representation (factorial analysis of distance) of the distance matrix of Table 2. The subjects are denoted by the letter(s) beginning the label, the figures following the letter(s) indicate the number of days between the first presentation of "the Mar of the Ghosts" and the recall.





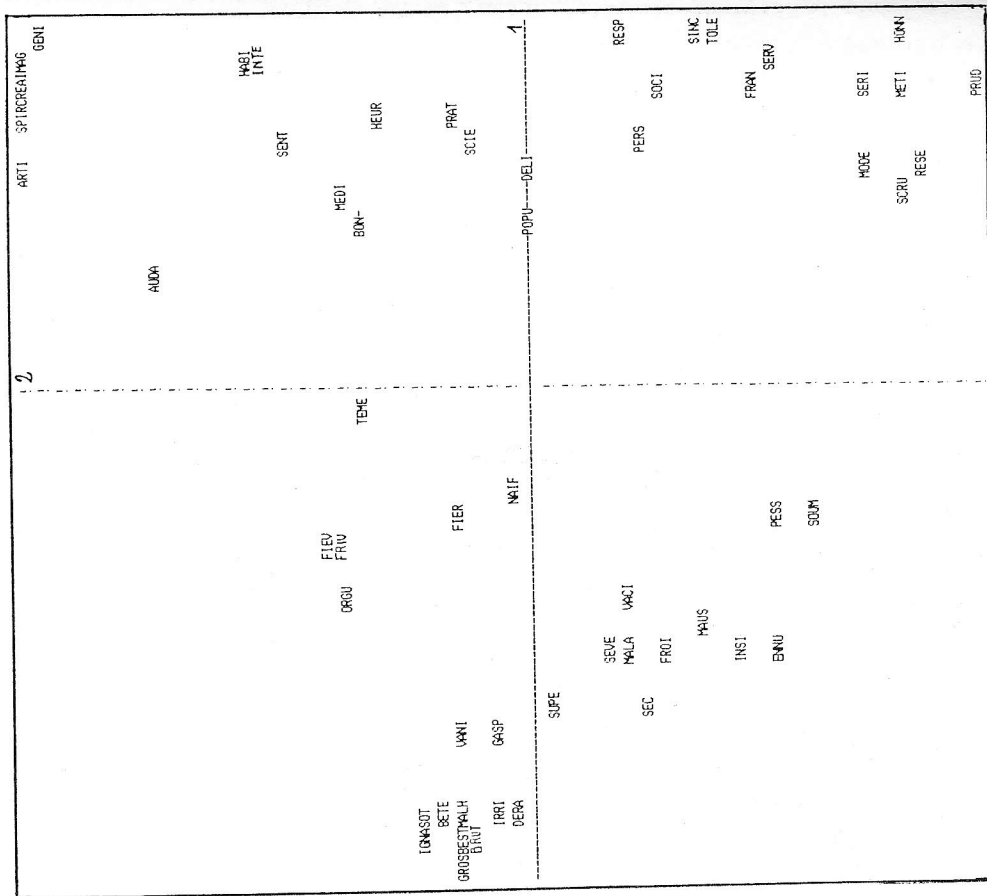


Figure 5

Euclidean representation (factorial analysis of distance) of the distance matrix of Table 3. (qualifiers of the character).  
Axe 1 and Axe 2.

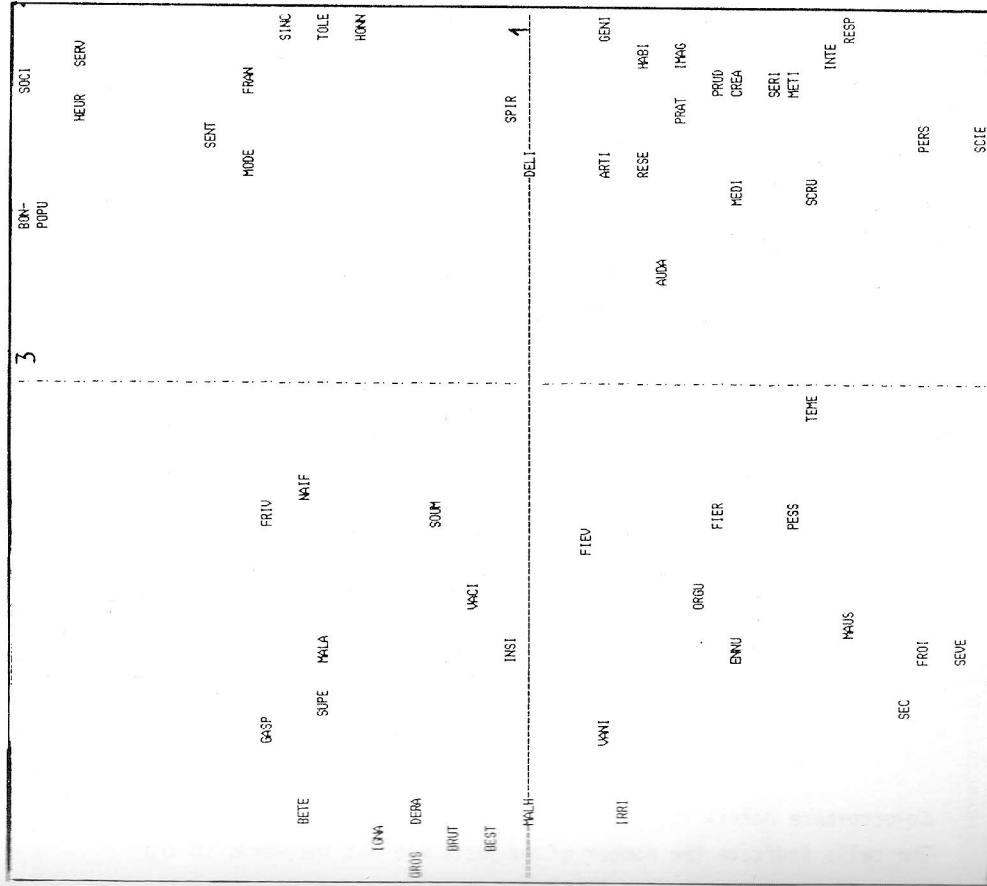


Figure 6

Euclidean representation (factor analysis of distance) of the distance matrix of table 3. (qualifiers of the character).  
Axe 1 and Axe 3.



	Cat	Cow	Deer	Dog	Goat	Horse	Lion	Mouse	Pig	Rabbit	Sheep
Bear	47.2	27.7	40.1	49.6	19.1	29.0	22.6	29.5	21.4	20.3	16.1
Cat		30.9	56.1	2.0	29.0	25.3	24.1	24.8	43.0	41.5	47.1
Cow			43.6	30.2	11.0	7.7	24.5	34.1	17.0	27.9	8.2
Deer				50.9	44.5	43.0	44.7	39.9	41.1	19.9	53.1
Dog					17.0	24.0	26.9	27.5	45.0	39.4	46.8
Goat						7.2	23.1	39.6	19.5	21.8	1.8
Horse							28.6	32.6	25.7	30.1	15.2
Lion								33.2	29.3	33.3	35.0
Mouse									34.9	22.6	51.9
Pig										25.9	19.6
Rabbit											32.5

Table 5

Proximity matrix between animal terms from Henley (1969).

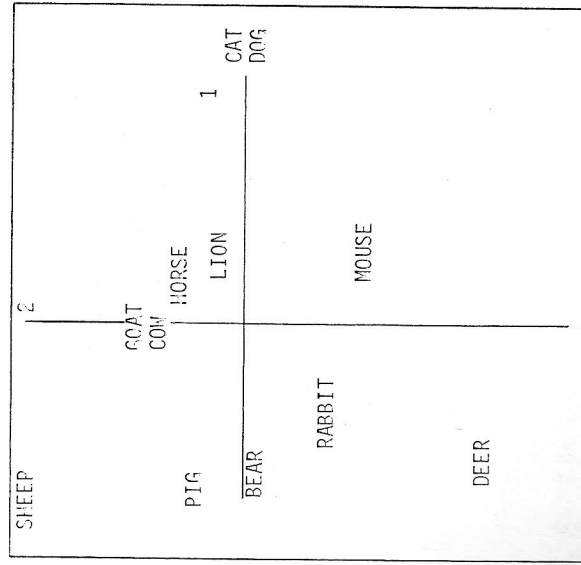


Figure 9

Euclidean representation (factorial analysis of distance) of the distance matrix of Table 5. (Henley's animal terms).

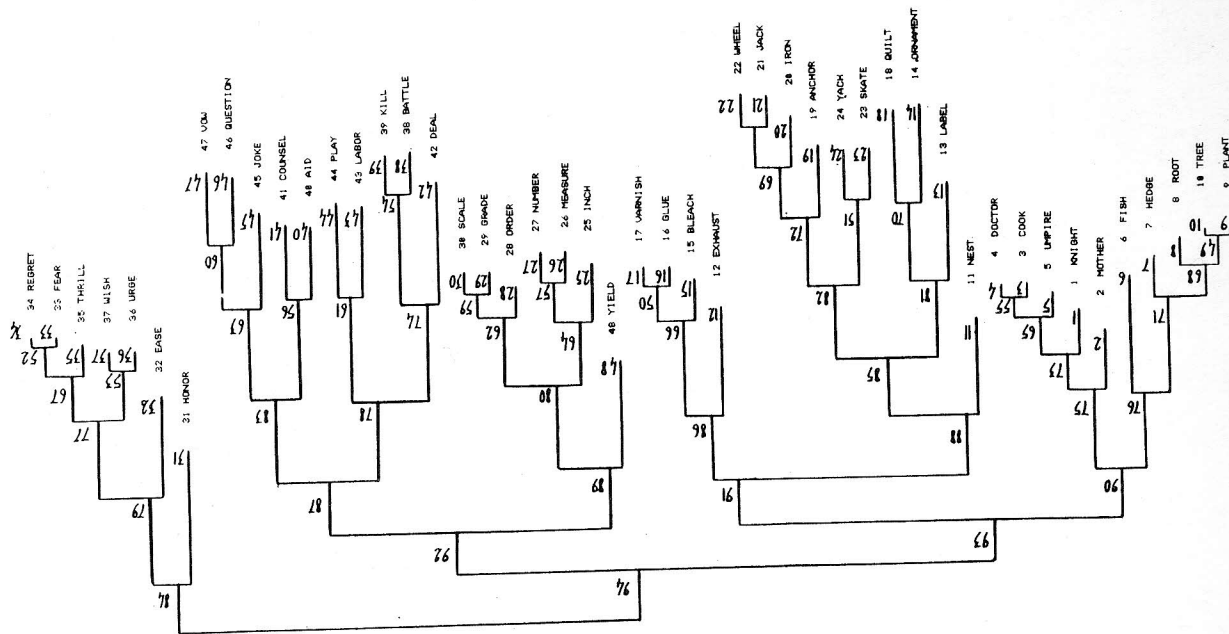


Figure 8

Additive Tree obtained from the data of Table 4.

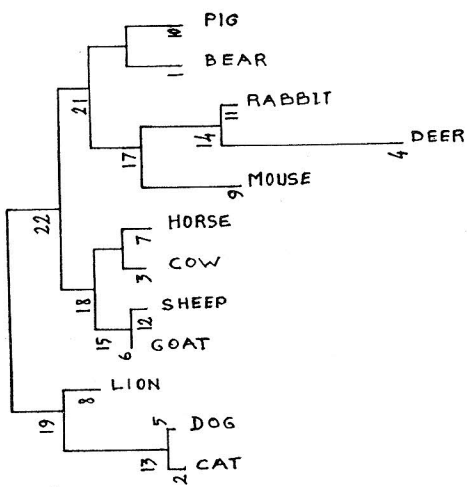


Figure 10

Additive tree obtained from the data of Table 5. (Henley's Animal Terms).

Table 6  
Proximity Matrix between 18 English Nouns obtained by Friendly (1979) in free recall experiments.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
dog	x																		
cat	100	x																	
lion	91	93	x																
tiger	88	91	100	x															
sheep	86	88	92	91	x														
goat	88	89	94	94	100	x													
face	55	57	57	57	52	51	x												
mouth	62	65	62	61	60	56	95	x											
arm	63	66	65	65	61	60	90	92	x										
finger	62	64	63	63	56	58	94	94	93	x									
leg	58	61	62	62	56	57	89	92	96	90	x								
foot	55	57	59	59	59	53	93	90	92	91	98	x							
celery	69	67	56	56	61	63	60	60	57	62	48	49	x						
spinach	69	68	60	58	65	64	61	62	60	62	56	57	96	x					
lettuce	70	69	60	58	64	64	60	61	59	62	54	55	94	97	x				
potato	68	67	60	59	64	65	65	64	67	60	60	60	87	95	93	x			
rice	66	65	54	53	53	58	66	66	65	70	58	60	88	86	85	92	x		
yam	66	65	57	56	60	62	63	64	62	67	58	57	91	93	92	97	99	x	

\*Numbers indicate item numbers.

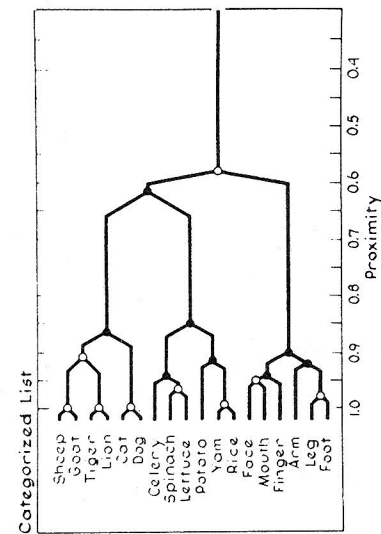


Figure 11

Hierarchical solution for the data of Table 6 (from Friendly (1969)).

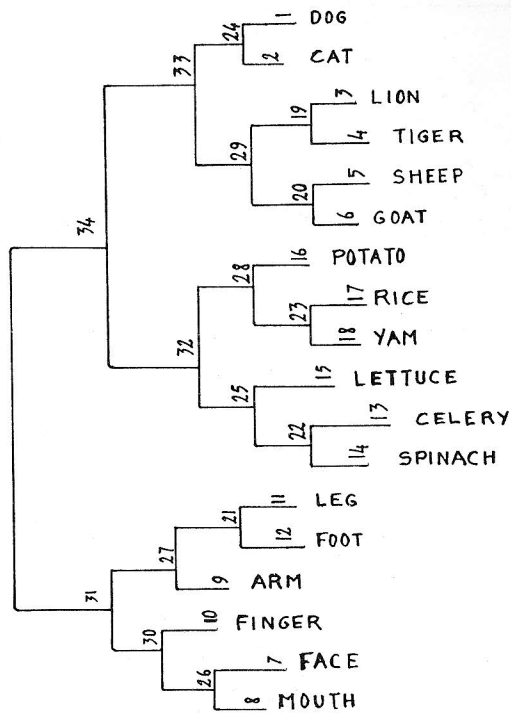


Figure 12

Additive tree obtained from the data of Table 6 (Friendly's data).

REFERENCES

(1) Abdi, H., Approches néoordinales en analyse des données psychologiques, thèse, Aix-en-Provence, (1980).

(2) Abdi, H., Structuration des qualificatifs du caractère, Multigraph, Dijon (1983).

(3) Arabie, P. and Boorman, S., Multidimensional scaling of measures of distances between partitions, Journal of Math. Psychol., 10 (1973), 148-203.

(4) Bartlett, F.C., Remembering, Cambridge University Press, Cambridge (1932).

(5) Benzecri, J.P., L'analyse des données, Dunod, Paris (1973).

(6) Boorman, S.A., Olivier, D.C., Metrics on spaces of finite trees, Journal of Math. Psychol., 10 (1973), 26-59.

(7) Bousfield, W.A., The occurrence of clustering in the recall of randomly arranged associates, Journal of General Psychol., 49 (1953), 229-240.

(8) Buneman, P., A note on the metric properties of trees, Journal of Combinatorial theory, 17(b), 48-50.

(9) Carroll, J.D., Chang, J.J., A method for fitting a class of hierarchical tree structure models to dissimilarity data and its application to some body parts data of Miller's, in Proceedings of the 81th Annual Convention American Psychological Association, 8 (1973), 1097-1098.

(10) Colonius, H., and Schulze, R.H., Tree structures for proximity data, British Journal of Math. and Stat. Psychol., 34 (1981), 167-180.

(11) Cunningham, J.P., Finding an optimal tree realization of a proximity matrix, paper presented at the mathematical psychology meeting, Ann Arbor (1974).

(12) Cunningham, J.P., Free trees and bidirectional trees as representations of psychological distances, Journal of Math. Psychol., 17 (1978), 165-188.

(13) Degenne, A., and Verne, P., Introduction à l'analyse de similitude, Revue Française de Sociologie (1973).

(14) Dobson, J., Unrooted tree for numerical taxonomy, Journal of Applied Probability, 11 (1974), 32-42.

(15) Flament, C., L'analyse de similitude, Cahiers du C.R.O., Brussels (1962).

(16) Friendly, M., Methods for finding graphic representations of associative memory structures, in Puff, C.R. (Ed.), Memory organization and structure, Academic Press, New York (1969).

- [17] Hakami, S.L., and Yau, S.S., Distance matrix of a Graph and its realizability, *Quarterly of Applied Mathematics*, 22 (1964), 305-317.
- [18] Hartigan, J.A., Representations of similarity matrices by trees, *J.A.S.A.*, 62 (1967), 1140-1158.
- [19] Hartigan, J.A., Clustering algorithms, J. Wiley, New York (1975).
- [20] Henley, N.M., A psychological study of the semantics of animal terms, *Journal of Verb. Learn. and Verb. Behavior*, 8 (1969), 176-184.
- [21] Hill, M.O., Correspondence analysis, A neglected method, *Applied Statistics*, 23 (1974), 340-354.
- [22] Johnson, S.C., Hierarchical clustering schemes, *Psychometrika*, 32 (1967), 241-254.
- [23] Kerlinger, F.N., Foundations of behavioral research, Holt, New York (1973).
- [24] Kruskal, J.R., and Wish, M., Multidimensional scaling, Sage, Beverly Hills (1978).
- [25] Luong, X., Voisinae lâche, score et famille scorante, *Multigraph, Besançon* (1982).
- [26] Luong, X., to appear (1983).
- [27] Miller, G.A., A psychological method to investigate verbal concepts, *Journal of Math. Psychol.*, 6 (1969), 169-191.
- [28] Nishisato, S., Analysis of categorical data: dual scaling and its applications, University of Toronto Press, Toronto (1980).
- [29] Pruzansky, S., Tversky, A., and Carroll, J.D., Spatial versus tree representation of proximity data, *Psychometrika*, 47(1) (1982), 3-24.
- [30] Riordan, J., An introduction to combinatorial analysis, J. Wiley, New York (1958).
- [31] Rosenberg, S., The method of sorting in multivariate research, in Hirschberg, N., and Humphreys, L.G., (Eds.), *Multivariate applications in the social sciences*, L. Erlbaum, Hillsdale (1982).
- [32] Rosenberg, S., Nelson, C., and Vivekanathan, P.S., A multidimensional approach to the structure of personality impressions, *Journal of Pers. and Social Psychol.*, 9 (1968), 283-294.
- [33] Sattath, S., and Tversky, A., Additive similarity trees, *Psychometrika*, 42(3) (1977), 319-345.
- [34] Schiffman, S., Reynolds, M.L., and Young, F.W., Introduction to multidimensional scaling, Academic Press, New York (1982).
- [35] Sneath, P.H., Sokal, R.R., Numerical taxonomy, Freeman, San Francisco (1973).

- [36] Torngerson, M.S., Theory and methods of scaling, J. Wiley, New York (1958).
- [37] Tulving, E., and Donaldson, W. (Eds.), *Organization of memory*, Academic Press, New York (1972).
- [38] Wegner, D.M., and Vallacher, R.R., *Implicit psychology*, Oxford University Press, Oxford (1977).
- [39] Young, F.W., De Leeuw, J., Takane, Y., to appear: *Multidimensional Scaling: theory and methods*, L. Erlbaum, Hillsdale.