

C.E.A. N. 1964

BANQUE DE DONNEES SEMANTIQUES

Mme Evelyne ANDREWSKY, P. NICOLAS (INSERM U 84)
J.P. GRILLO (I.N.S.T.N.)

1er Janvier 1977

Sommaire

On propose un système permettant de déterminer des relations sémantiques entre items lexicaux. Pour ce faire, un descripteur est associé à chaque item lexical; deux types d'algorithmes sont utilisés pour calculer les relations entre descripteurs (relations de "similitude" ou de "prédicativité").

Ce système permet de simuler des expériences linguistiques. Certains résultats ont ainsi pu être prévus puis vérifiés expérimentalement.

TABLE DES MATIERES

	Page
- Introduction	1
- Items lexicaux et propriétés sémantiques associées.	2
- Etude expérimentale des propriétés (ou traits) sémantiques.	3
- Champ sémantique ou traits caractéristiques?	4
- Les traits caractéristiques en Neurolinguistique.	5
- Description de la mémoire (ou "banque de données") sémantique.	6
- Les systèmes de descripteurs	6
- Les programmes de gestion de la mémoire.	8
- Comparaison des entrées lexicales similaires.	8
- Comparaisons prédicatives.	10
- Simulation d'expériences psycholinguistiques	12
- Conclusions.	16
- Bibliographie.	16
- Listings	18

BANQUE DE DONNEES SEMANTIQUES

E. Andreewsky, P. Nicolas (INSERM U 84)
J.P. Grillo (C.E.A.)

I - INTRODUCTION

Les théories de la compréhension des langues naturelles - qu'elles soient issues de travaux de Linguistique, de Psychologie ou d'Intelligence Artificielle - semblent, à l'heure actuelle s'accorder sur un point : l'importance du rôle des connaissances mémorisées dans les processus liés à la compréhension. (Miller G.A., 1976, Winograd T., 1973, Le Ny F, 1975, ...).

Nous proposons, avec ce travail :

- un modèle de "mémoire sémantique", associant à chaque item lexical, un "descripteur sémantique",

- des procédures de traitement de ces descripteurs.

Procédures et mémoire sémantique peuvent être considérées comme des modèles psychologiques, dans la mesure où elles permettent de simuler certains comportements linguistiques. Ce sont aussi des outils de recherche, conduisant à prévoir certains comportements linguistiques spécifiques.

II - ITEMS LEXICAUX ET PROPRIETES SEMANTIQUES ASSOCIEES.

Ce sont les travaux d'Intelligence Artificielle, postérieurs aux années 1970, qui ont amené à mesurer l'importance des propriétés sémantiques mémorisées, dans les processus de compréhension du langage. En effet, des systèmes tels que ceux de question-réponse sur ordinateur, ou de traduction automatique, ne peuvent pas espérer fonctionner sans faire intervenir les propriétés sémantiques mémorisées, encore appelées "la connaissance du monde".

Un exemple des difficultés de la traduction automatique permettra d'illustrer ce que nous venons de dire. Supposons que nous procédions à la traduction, en anglais, des phrases P1 et P2 suivantes :

P1 = Le car a heurté le piéton; on l'a entendu crier.

P2 = Le car a heurté le piéton; on l'a entendu freiner.

Decider que le "l'" doit être traduit par "he" ou par "it", implique l'intervention de connaissances extérieures aux phrases P1 et P2, et relatives aux propriétés sémantiques associées aux entrées lexicales "car", "piéton", "crier", "freiner". En effet, aucune règle d'ordre syntaxique ne permet de résoudre le problème du référent ("piéton" ou "car") du pronom "l'". Sa résolution implique l'évaluation de la pertinence conceptuelle du rapport sujet/action des couples suivants :

piéton - crier

car - crier

piéton - freiner

car - freiner.

Un système traitant les phrases P1 et P2 (que ce soit le cerveau ou non) a besoin, pour procéder à cette

évaluation, d'un ensemble d'informations : les connaissances mémorisées, que nous appellerons "propriétés sémantiques", liées aux items lexicaux de la phrase.

Nous représenterons ici ces propriétés sémantiques au moyen d'un descripteur par item lexical. Les raisons qui nous font choisir ce type de représentation sont dérivées de l'expérimentation psycholinguistique, et nous les exposons dans le paragraphe suivant.

III - ETUDE EXPERIMENTALE DES PROPRIETES (OU TRAITS) SEMANTIQUES.

Deux directions principales caractérisent les études expérimentales concernant les propriétés sémantiques (ou traits) mémorisées, (encore appelées "même" en Psychologie - Le Ny, 1975).

- la première (illustrée par G.A. Miller, 1969) utilise une technique de tris. On demande aux sujets de grouper les items sémantiquement semblables, à partir d'une liste non ordonnée. L'analyse statistique des tris ainsi effectués permet de mettre en évidence quels sont les traits sémantiques qui servent de critères aux regroupements effectués. Cette analyse donne des mesures de proximité entre chaque couple d'items (le nombre de fois qu'il est classé dans le même groupe par l'ensemble des sujets). Un espace métrique est ainsi déterminé; il comporte des nuages d'items bien séparés par des traits tels que "animé" / "objet" ou "humain" / "animal".

- la seconde, que nous illustrerons par un article de H. Goodglass et E. Baker, 1976, utilise une technique classique en Psychologie expérimentale : le relevé des erreurs et des temps de latences, sous diverses conditions. On présente, ici, aux sujets, un objet-cible, tout en leur proposant oralement une suite de mots; les sujets doivent indiquer si le mot proposé est associé ou non à l'objet cible. On trouve des différences remarquables dans les erreurs et les temps de latences selon

que les rapports entre mots et objet-cible sont du type "similitude" (Ex : orange-pomme) ou autres (Ex : orange -fruit, ou orange - manger etc.).

III - 1 Champ sémantique ou traits caractéristiques?

Les expérimentations psycholinguistiques s'appuient sur des théories relatives au stockage des concepts en mémoire, et notamment sur l'une ou l'autre des théories suivantes :

- Champ sémantique :

Les concepts associés aux items lexicaux sont des points dans un espace sémantique. Celui-ci peut se représenter comme un réseau de points correspondant à chaque concept, où les distances (nombre de points permettant de relier deux concepts) sont les mesures de la proximité de ces concepts. On appelle cet espace "champ sémantique":

- Traits sémantiques :

Les concepts sont représentés par un vecteur descriptif, composé de traits sémantiques. Cette représentation permet des descriptions récursives (les traits appartenant à un concepts générique, fruit, par exemple, ne sont pas explicitement présents dans les descripteurs spécifiques comme "orange".)

Nous avons choisi, dans ce travail, le cadre théorique de la représentation des concepts en traits sémantiques. Les raisons qui nous ont conduit à ce choix sont doubles :

- d'une part, les expériences psychologiques telles que celles dont nous avons indiqué les références ci-dessus, s'interprètent plus facilement avec une théorie en traits sémantiques; il faut, en effet, si on adopte une théorie construite au moyen de champs sémantiques, faire des hypothèses "fortes" sur la structure de ces champs, qui ne semblent pas utilement fondées.

- d'autres part, la pathologie du langage (Aphasie)

conduit à des observations comportementales qui ne peuvent pas s'interpréter sans faire intervenir les traits sémantiques en tant qu'objets traités par les processus de compréhension. Le paragraphe suivant a pour but d'indiquer les caractéristiques de ces comportements.

III - 2 Les traits caractéristiques en Neurolinguistique.

Un des postulats tacite d'une grande partie de la littérature sur la psychologie du langage est la notion de phénomène unitaire reliant un objet (ou un item lexical) à un concept mémorisé lui correspondant. Certains comportements pathologiques ne permettent pas de conserver ce postulat. Deux exemples, détaillés dans E. Andreevsky, 1975, en donnent les raisons.

Le premier concerne un cas d'agnosie visuelle (erreurs dans la dénomination des objets). Soit m l'objet à dénommer; nous désignerons par P_m la verbalisation du nom de cet objet. Dans certains cas d'agnosie visuelle, où le patient perçoit parfaitement les objets (qu'il peut dessiner avec précision), ce n'est pas P_m qui est verbalisé, mais $P_{m'}$. Les différences entre m et m' sont souvent difficile à relever; mais, il est tout à fait remarquable que m et m' peuvent ne différer que par un seul trait sémantique.
Exemples :

pour m = sandale, botte, chaussette, on obtient les dénominations suivantes :

$P_{m'}$ = coiffure d'été, passe-montagne, bonnet en laine.

(ici, le trait "pour les pieds" étant remplacé par "pour la tête")
On peut essayer de comprendre, et même de formaliser relativement simplement les processus de dénomination conduisant à ce type de comportement pathologique dans le cadre d'une représentation des concepts en traits sémantiques.

Le second concerne un cas d'alexie totale, c'est à dire un sujet qui ne peut ni comprendre, ni verbaliser aucun item lexical présenté sous forme écrite (le patient correspondant à un niveau de culture supérieure; il a subi une lésion hémisphérique, conduisant à cette alexie comme seul trouble du langage).

Le patient, qui n'a "aucune idée" de la signification des mots écrits semble cependant capable d'en déterminer certains traits sémantiques. Par exemple, si les items "enfants" et "citron" lui sont présentés par écrit, il leur associera, sans difficulté, les adjectifs "petit", pour le premier ou "jaune" pour le second, choisis parmi une liste présentée oralement. Cet exemple de comportement pathologique (détaillé, comme le précédent dans E. Andreewsky, 1975) ne peut s'expliquer sans postuler la réalité psychologique de la mémorisation des concepts en traits sémantiques.

IV - DESCRIPTION DE LA MEMOIRE (ou "BANQUE DE DONNEES") SEMANTIQUE

Les données expérimentales de Psycho et de Neuro-linguistique nous conduisent à l'élaboration d'un modèle de mémoire, associant à chaque item lexical un descripteur en traits sémantiques.

Les activités de lecture et de dénomination correspondent à certains traitements de ces descripteurs. Nous avons simulé ces traitements au moyen de procédures : il s'agit de différents types de comparaison entre descripteurs.

IV - 1 Les systèmes de descripteurs.

Un descripteur est associé à tout item lexical (pour l'instant, substantif ou verbe).

Ce descripteur est un vecteur d'un espace correspondant à la classe syntaxico-sémantique de l'item (Ex : substantif-objet). Pour les substantifs-objets, par exemple, un descripteur sera caractérisé par les axes (ou types) suivants :

On donne en exemple le descripteur de l'item "fourchette" :
Catégorie(couvert), dimension (2), forme (allongée), couleur (gris) matière (acier, argent), composants (manche, dents), items associés (cuiller), verbe d'usage (piquer), compléments du verbe (viande, etc.), propriétés (), synonymes (), divers ().

A un item lexical comportant n homographes correspondent n entrées lexicales (et leurs descripteurs), numérotées de 1 à n.

Les classes syntaxico-sémantiques peuvent être en nombre quelconque : pour en ajouter une, il suffit d'en définir l'étiquette, et les rubriques permettant de ranger les traits sémantiques des descripteurs des items correspondants.

La mémoire sémantique comporte, actuellement, différentes classes syntaxico-sémantiques, utiles aux expériences effectuées. Il s'agit des classes suivantes :

- substantif-objet (Ex : orange)
- " animal (Ex : tigre)
- " être humain (Ex : cuisinier)
- " action (Ex petit déjeuner)
- " générique (Ex : fruit)
- verbes d'action (divisés en 11 sous groupes).
(Ex : manger).

La mémoire se présente sous la forme d'un tableau où :

- la première colonne est réservée à l'entrée lexicale
- la deuxième à l'étiquette de la classe syntaxico-sémantique
- chacune des colonnes suivantes réfère à une rubrique définie au niveau de la classe syntaxico-sémantique; elle peut comporter de 0 à 4 propriétés sémantiques.

Il est à noter que les propriétés (ou "traits") qui constituent ainsi les descripteurs sont aussi des entrées lexicales; c'est à dire que ces propriétés sémantiques figurent aussi en première colonne du tableau correspondant à la mémoire sémantique. Comme nous le verrons dans ce qui suit (cf § IV - 3 et IV - 4), cette organisation de la mémoire sémantique favorise les calculs récursifs permettant de comparer deux descripteurs.

IV - 2 Les programmes de gestion de la mémoire.

La mémoire -ou "banque de données sémantiques"- contient les entrées lexicales et leurs descripteurs. Ces derniers sont (pour l'instant) définis d'une manière pragmatique. (Une définition automatique des traits sémantiques est à l'étude).

Les programmes suivants permettent la gestion dynamique de cette banque de données :

- Lecture d'une entrée lexicale; contrôles et mise en forme (entrées par cartes ou par clavier TSO)
- Insertion d'une entrée lexicale nouvelle et du descripteur correspondant.
- Tris alphabétiques.
- Cartes à disque et contrôles.
- Lecture du disque
- Edition de tout ou parties de la mémoire.

On peut ainsi ajouter une nouvelle entrée lexicale, modifier ou supprimer les descripteurs et les entrées lexicales existants, ajouter ou supprimer une classe syntaxico-sémantique. Tous ces programmes sont accessibles soit en traitement par lots, soit en "TSO".

IV - 3 Comparaison des entrées lexicales similaires.

Deux entrées lexicales ne peuvent être comparées, au moyen d'un calcul basé sur l'intersection de leurs descripteurs, que si elles appartiennent à la même catégorie syntaxico-sémantique. On pourra ainsi, par exemple, comparer "orange" à "pomme" ou à "oreiller" (substantifs-objets) et évaluer leurs similarités; mais on ne pourra pas, en s'appuyant sur l'intersection, comparer cet item à un item appartenant à la classe des verbes (comme "manger")

ni au substantif générique "fruit". Les comparaisons d'entrées lexicales telles que "orange" avec "manger" ou "fruit" font l'objet du paragraphe suivant (il est intuitivement évident que ces items ne sont pas étrangers l'un à l'autre).

Soit alors deux entrées lexicales a_0 et b_0 appartenant à la même catégorie syntaxico-sémantique. Leurs descripteurs sont représentés dans le même espace :

$$\begin{array}{l} \left| \begin{array}{cccc} a_0 & a_1 & a_2 & a_j & a_n \\ b_0 & b_1 & b_2 & b_j & b_n \end{array} \right. \end{array}$$

Chaque couple de propriétés sémantiques (a_j, b_j) constitue, d'autres part, un couple d'entrées lexicales, qui appartiennent à un même espace -pouvant être identique ou différent de celui de a_0, b_0 . On aura, par exemple :

$$\begin{array}{l} \left| \begin{array}{cccc} a_j & a_{j1} & a_{j2} & a_{jm} \\ b_j & b_{j1} & b_{j2} & b_{jm} \end{array} \right. \end{array}$$

Similarité directe de a_0 et b_0 .

Nous définirons la "similarité directe" de a_0 et b_0 , au moyen d'une distance de Hamming entre leurs descripteurs :

$$Sd(a_0, b_0) = \frac{1}{n+1} \sum_{i=0}^n \theta(a_i, b_i), \text{ avec :}$$
$$\left| \begin{array}{l} \theta(a_i, b_i) = 0 \quad \text{ssi} \quad a_i = b_i; \\ \theta(a_i, b_i) = 1 \quad \text{si} \quad a_i \neq b_i. \end{array} \right.$$

La similarité directe ne tiens compte que de l'identité des éléments correspondants des deux descripteurs.

Similarité de a_0 et b_0 .

Pour tenir compte, non seulement de l'identité des éléments correspondants, a_j et b_j , mais encore de la mesure, S_d , de leur similarité directe, nous définirons la procédure suivante, S_1 , que nous appellerons "Similarité" de a_0 et b_0 . :

$$S_1 (a_0, b_0) = \frac{1}{n+1} \sum_{j=0}^n S_d (a_j, b_j).$$

On aura $S_1 = S_d = 0$ ssi $a_0 = b_0$.

S_1 , similarité de a_0 et b_0 a les propriétés d'une distance, dans l'espace des traits sémantiques de la classe syntaxico-sémantique commune à a_0 et b_0 .

(On peut, de même, calculer $S_2 (a_0, b_0) = \frac{1}{n+1} \sum_{j=0}^n S_1 (a_j, b_j)$ et ainsi de suite récursivement, S_d , représentant en fait S_0 .)

IV - 4 Comparaisons "prédicatives".

Si les deux entrées lexicales à comparer, a_0 et b_0 , n'appartiennent pas à la même classe syntaxico-sémantique, le calcul de leur similarité n'a pas de sens, ainsi que nous l'avons indiqué plus haut.

Nous pouvons, cependant, trouver des relations entre leurs descripteurs.

Une première évaluation consiste à tester si l'un des items fait partie des propriétés sémantiques de l'autre. Nous appellerons ce type de rapport "relation prédictive directe"; le terme de "prédictatif" vient du fait que les relations entre une entrée lexicale et les traits caractéristiques de son descripteur sont déterminées par un apprentissage qui fait apparaître ces deux items dans un même prédicat; par exemple, les rapports entre "orange" et "manger" apparaissent dans des prédicats tels que "je mange une orange".

Si a_0 appartient à un espace à n traits sémantiques, et b_0 , à un espace à m traits, on aura :

Relations prédicatives directes entre a_0 et b_0 .

$$R_p(a_0, b_0) = \prod_{i=0}^n \theta(a_0, b_i) \cdot \prod_{j=0}^m \theta(b_0, a_j)$$

(avec $\theta(x, y) = 0$ ssi $x = y$.)

Pour que R_p soit nul, il faut et il suffit que a_0 soit égal à b_0 ou à l'une de ses propriétés sémantiques (ou que b_0 soit l'une des propriétés sémantiques de l'entrée lexicale a_0). Quand R_p est nul, il est possible d'obtenir le type de relation unissant a_0 et b_0 ; par exemple, si on compare "orange" et "fruit" ce type sera indiqué par la rubrique correspondant, dans le descripteur de "orange", à la propriété sémantique "fruit" : à savoir "catégorie" (ou "terme générique").

Pour tenir compte, non seulement de l'identité entre les entrées lexicales a_0 ou b_0 et l'une des propriétés sémantiques de b_0 ou a_0 , mais encore des similarités correspondantes, nous définirons, en utilisant la similarité S_1 définie plus haut, la relation suivante, que nous appellerons "distance prédicative" :

Distance prédicative entre a_0 et b_0 .

$$R_q(a_0, b_0) = \prod_{i=1}^n S_1(a_0, b_i) \cdot \prod_{j=1}^m S_1(b_0, a_j)$$

R_q a les propriétés d'une distance; cette relation entre a_0 et b_0 donne une mesure, comprise entre 0 et 1, des relations prédicatives implicites que peuvent déterminer les similarités entre a_0 et les propriétés sémantiques de b_0 (ou l'inverse).

V - SIMULATION D'EXPERIENCES PSYCHOLINGUISTIQUES.

Nous disposons d'une "mémoire sémantique" et de procédures permettant de calculer les relations d'ordre sémantique entre les divers items représentés dans cette mémoire, que nous avons présentés au paragraphe précédent.

Nous pouvons donc simuler les expériences psycholinguistiques qui font intervenir les relations entre items lexicaux; ces simulations peuvent avoir deux objectifs :

- la vérification de la plausibilité du modèle que nous avons proposé;

- la prédiction que tel résultat expérimental devrait être obtenu dans telles conditions; ce dernier objectif devant, bien entendu, être contrôlé au moyen d'une expérience psycholinguistique.

V - 1 Simulation de l'expérience des catégories associatives.

Reprenons l'expérience de H. Goodglass et E. Baker, 1976, qui utilise une technique classique en Psychologie expérimentale : le relevé des erreurs et des temps de latence faites par les sujets, dans diverses conditions. Ici, la question posée aux sujets est de déterminer s'il existe, ou non, une relation sémantique entre deux entrées lexicales : la première est perçue visuellement : c'est un objet, et la seconde, auditivement : ce sont les items lexicaux qui sont (ou non) en relation sémantique avec cet objet.

La figure 1 présente les erreurs dans les réponses des sujets normaux (c'est à dire ne présentant pas de troubles aphasiques). Les points A, B, C, E, et F, représentent respectivement ces réponses pour les relations de types suivant : (les exemples étant donnés relativement à l'objet-cible "orange") identité ("orange"), classe générique ("fruit"), attribut ("juteux"), contexte fonctionnel ("petit déjeuner") et verbe associé ("manger"). Le point D représente les items lexicaux similaires ("pomme"). Il est tout à fait remarquable que ce point D correspond à un

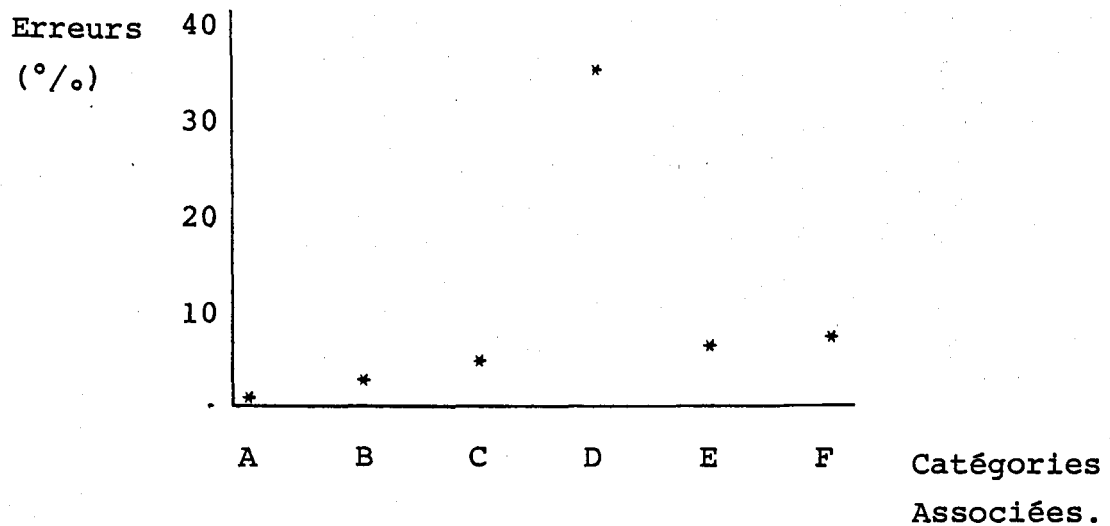


Fig 1.

un taux d'erreurs plus de quatre fois supérieur à celui des autres catégories.

On peut, avec notre modèle, expliquer, et simuler ce résultat de la manière suivante :

- pour 5 (sur 6) items lexicaux présentés, la recherche de l'existence de relations sémantiques avec l'objet-cible ne peut aboutir qu'en utilisant les procédures dites "prédicatives" une similarité n'apparaissant que dans le cas des seules comparaisons du type "orange-pomme". Si (par un effet d'apprentissage) les sujets n'utilisent que des procédures, du type modélisé par les relations R_p ou R_q , ils n'arriveront pas à déterminer de relations entre les items tels que "orange" et "pomme".

- Les distances prédicatives, R_q , calculées pour les points A, C et F, reproduisent la configuration de la figure 1 :

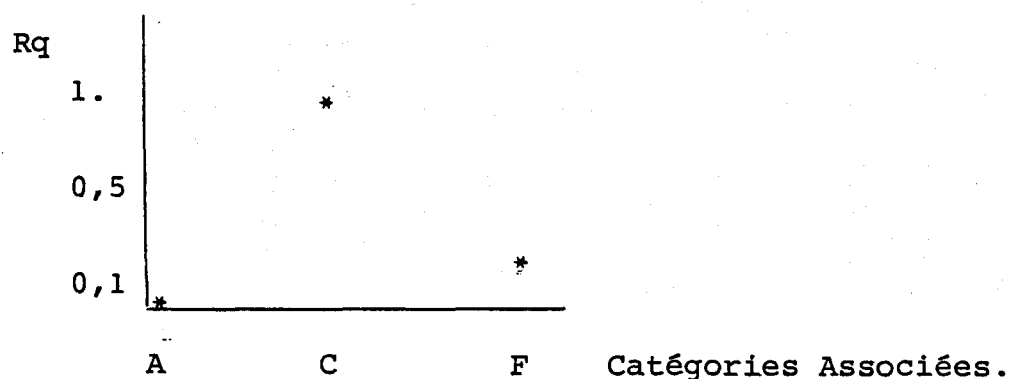


Fig. 2.

V - 2 Simulation d'une expérience centrée sur les similarités.

On peut, avec notre modèle, obtenir une configuration de résultats exactement opposée à celles que présentent la figure 1 et sa simulation, la figure 2. En effet, si nous n'utilisons, pour chercher les relations entre entrées lexicales, que les procédures de similarité, S_1 , les distances relatives aux points C (items similaires) et F (verbe associé) seront en rapport inverse à celui de la figure 2. Seul le point A (même item) demeurera inchangé. Les calculs donnent la configuration de la figure 3 :

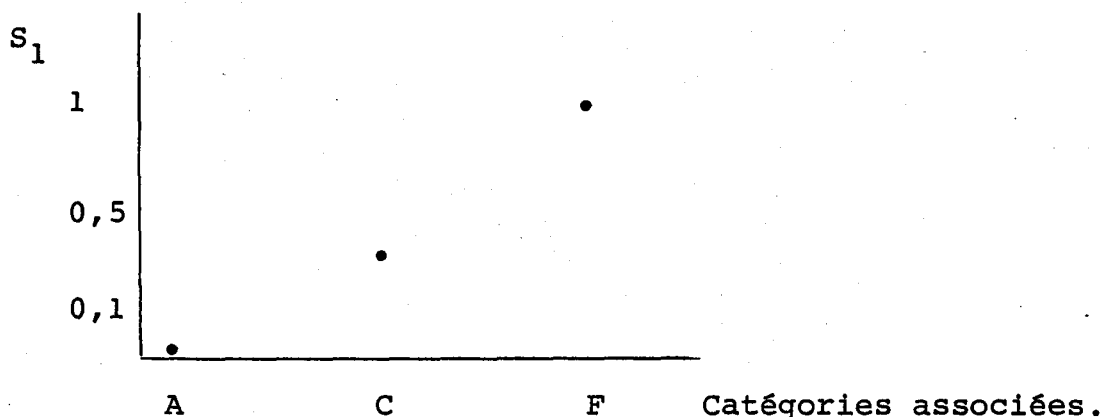


Fig. 3.

On peut alors postuler que si, par un biais expérimental, nous induisons les sujets à utiliser essentiellement des procédures d'évaluation de similarité entre items, nous obtiendrons une configuration de résultats (pourcentages d'erreurs) analogue à celle de la figure 3.

Un biais expérimental possible est la présentation d'un nombre beaucoup plus grand d'items lexicaux similaires à l'objet cible que d'items qui lui soient reliés par des relations de type prédicatif.

V - 3 Contrôle expérimental des résultats postulés.

Une expérience psycholinguistique, analogue à celle ayant donné les résultats de la figure 1, mais ne comportant que peu d'items associés à l'objet-cible, autres que des items lexicaux similaires, a été entreprise, pour vérifier s'il est ainsi possible d'obtenir des résultats analogues à ceux prévus par la figure 3.

Les objets (en fait, des images) cibles étaient au nombre de 5 : orange, banc, scie, accordéon, et pelican. Une liste de mots étaient présentés oralement, et les sujets devaient répondre "oui" ou "non" selon qu'ils voyaient ou non une relation entre le mot entendu et l'image présentée. 50 % des items présentés oralement étaient similaires à l'image présentée; les autres items étaient le verbe associé ("manger", pour "orange", ou "jouer" pour "accordéon") ou des items totalement étrangers.

Les résultats (nombre d'erreurs) obtenus, figure 4, concordent avec ceux qui ont été postulés au moyen de la figure 3.

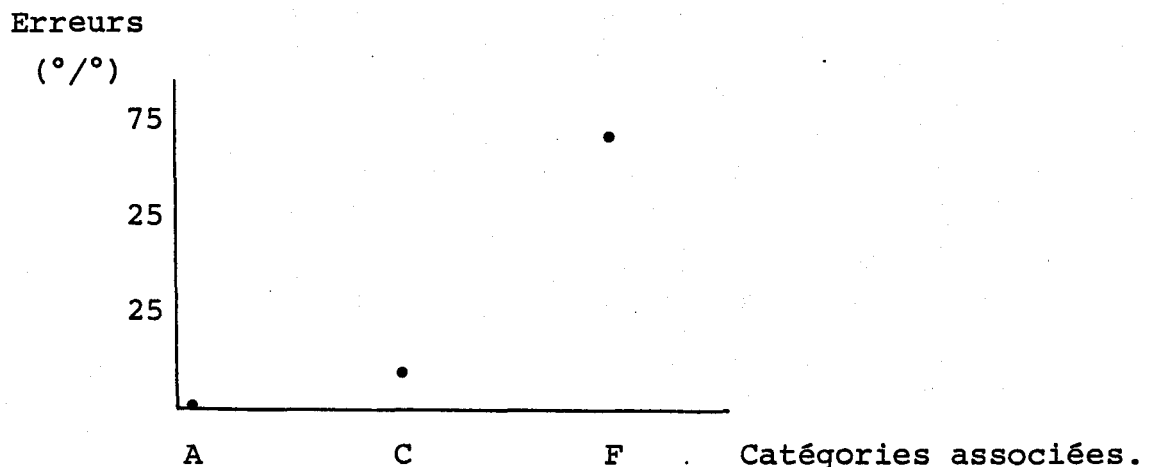


Fig 4.

L'inversion dans la configuration des erreurs relatives à l'estimation de la similarité des items, telle que donnée ci-dessus, par rapport aux résultats de la figure 1, confirme ainsi la validité du modèle de relations entre items lexicaux que nous avons proposé.

VI - CONCLUSIONS.

L'organisation que nous avons proposé, concernant le stockage en mémoire des propriétés sémantiques, d'une part, et leur utilisation dans l'évaluation des relations entre items lexicaux, d'autre part, nous paraît être un modèle fonctionnel. En effet :

- on peut rendre compte, par simulation, de certains comportements linguistiques impliquants une évaluation des rapports sémantiques entre items;

- On peut, en raisonnant sur les propriétés de la mémoire et des procédures qui l'utilisent, prévoir les comportements linguistiques des sujets dans de nouvelles conditions expérimentales (ainsi que nous l'avons montré au paragraphe précédent.)

Les simulations sur ordinateur de processus psychologiques ne sont pas en général destinées à une réflexion sur l'expérimentation en Psychologie, comme dans le travail que nous avons présenté. En effet, ces simulations sont le plus souvent du domaine de l'Intelligence Artificielle, et se proposent de résoudre un problème déterminé, en dehors de considérations sur des conditions psychologiques expérimentales.

Nous pensons, avec le présent travail, avoir montré qu'il est possible et bénéfique de lier les recherches en Psychologie à celles d'Intelligence Artificielle.

VII - BIBLIOGRAPHIE.

- (1) Albert M.L., Yamadoris A., Gardner H. & Howes D. "Comprehension in Alexia", Brain, 1973, Vol. 96, part 2, 317-328.
- (2) Andreewsky A., Debili F., Fluhr C. "Computational learning of semantic lexical relations for the generation and automatic analysis of content" (soumis pour présentation au congrès IFIP, 1977)
- (3) Andreewsky E. "The mechanisms of comprehension; an interpretation of Language Pathology" Proceeding of the workshop on Computational Semantic, Lugano, 1975

- (4) Andreewsky E. "Combination and Selection in Comprehension; a procedural approach to aphasia" (sous presse).
 - (5) Goodgass H., Baker E. "Semantic field, Naming and Auditory comprehension in Aphasia"; Brain and Language 1976, N°3, 359-374.
 - (6) Le Ny J.F. "Problèmes de sémantique psychologique" Langage, 40, 1975.
 - (7) Miller G. "Semantic relations among words" (Communication présentée au M.I.T., 1976).
-

(Liste des items lexicaux

utilisés, et de leurs descripteurs)

1 *** ARBRE
TYP VEGETAL
DIM 4
COL VERT
MAT BOIS
CMP RACINE
TRONC
FEUILLE
ASS FORET
PRP BRULER

2 *** CLOCHE
TYP USTENSILE
DIM 2
FRM CONIQUE
COL GRIS
MAT METAL
CMP BATTANT
ASS EGLISE
VRB SONNER
TINTER
SUB TOCSIN

3 *** GATEAU
TYP ALIMENT
DIM 2
FRM DIVERS
COL DIVERS
MAT FARINE
SUCRE
OEUF
ASS FAIM
VRB MANGER
PRP DOUX

4 *** GUITARE
TYP INSTRUMENT DE MUSIQUE
DIM 3
FRM ALLONGE
MAT BOIS
CMP CORDE
MANCHE
ASS ORCHESTRE
VRB JOUER
GRATTER

5 *** INSTRUMENT
TYP OBJET
-DIVERS
FRM QUELCONQUE

MAT METAL
BOIS
PLASTIQUE
VRB EXECUTER
SUB OPERATION
PRP UTILE

6 *** ORANGE
TYP FRUIT
AGRUME
DIM 2
FRM SPHERIQUE
COL ORANGE
MAT COMESTIBLE
CMP QUARTIER
ECORCE
ASS JUS
CITRON
VRB PRESSER
MANGER
SUB PETIT-DEJEUNER
PRP JUTEUX
COMESTIBLE

7 *** ORCHESTRE
TYP SOCIAL
MUSIQUE
CMP INSTRUMENT DE MUSIQUE
ASS MUSICIEN
VRB JOUER
INTERPRETER
SUB PARTITION
MORCEAU
MUSIQUE
PRP BRUYANT
BON

8 *** PAIN
TYP ALIMENT
DIM 2
FRM ALLONGE
CYLINDRIQUE
COL BEIGE
MAT FARINE
ASS FAIM
VRB MANGER

9 *** PETIT-DEJEUNER
TYP REPAS
CMP CAFE
FRUIT
PAIN
VRB MANGER
PRENDRE
SUB MATIN

10 *** PLANTE
TYP VEGETAL
DIM 2
COL VERT
CMP RACINE
FEUILLE
CHLOROPHYLLE
POT
ASS

11 *** POMME
TYP FRUIT
DIM 2
FRM SPHERIQUE
COL VERT
JAUNE
ROUGE
CMP PEPIN
PEAU
ASS JUS
CICRE
VRB MANGER
PRP COMESTIBLE

12 *** TAMBOUR
TYP INSTRUMENT DE MUSIQUE
DIM 2
FRM ROND
CYLINDRIQUE
MAT PEAU
CMP CAISSE
HAGUETTE
ASS ORCHESTRE
VRB BATTRE
PRP BRUYANT

13 *** TROMPETTE
TYP INSTRUMENT DE MUSIQUE
DIM 2
MAT METAL
ASS ORCHESTRE
VRB JOUER
PRP BRUYANT

TABLEAU DES RELATIONS INTER - LEXICALES

S _d	R _p	x1	x2													
N° ***	1er item															
N° ***	2ème item															
x3	x4	S ₁	x5 etc.													
0.0	0.0	C	C													
1	***	ARBRE														
1	***	ARBRE														
0.0	7.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1.0000	1.0	0														
2	***	ARBRE														
1	***	CLOCHE														
6.000	6.000	1.000	1.000	0.0	1.000	1.000	1.000	0.0	1.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1.0000	1.0	0														
2	***	ARBRE														
1	***	GATEAU														
6.000	6.000	1.000	1.000	0.0	1.000	1.000	0.0	0.0	1.000	0.0	0.0	1.000	0.0	0.0	0.0	0.0
0.0000	1.0	0														
4	***	ARBRE														
1	***	GUIRE														
4.000	5.000	0.000	1.000	1.000	0.0	0.0	0.0	1.000	0.0	1.000	0.0	0.0	0.0	0.0	0.0	0.0
0.0000	1.0	0														
5	***	ARBRE														
1	***	INSTRUMENT														
2.000	3.000	0.000	1.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000	0.0	0.0	0.0	0.0
1.0000	1.0	0														
6	***	ARBRE														
1	***	ORANGE														
7.000	7.000	1.000	1.000	0.0	1.000	1.000	1.000	0.0	1.000	0.0	0.0	1.000	0.0	0.0	0.0	0.0
1.0000	1.0	0														
7	***	ARBRE														
1	***	ORCHESTRE														
4.000	4.000	1.000	1.000	0.0	0.0	0.0	1.000	0.0	1.000	0.0	0.0	1.000	0.0	0.0	0.0	0.0
1.0000	1.0	0														
8	***	ARBRE														
1	***	PAIN														
5.000	5.000	1.000	1.000	0.0	1.000	1.000	0.0	0.0	1.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1.0000	1.0	0														
9	***	ARBRE														
1	***	PETIT-DEJEUNER														
2.000	2.000	1.000	1.000	0.0	0.0	0.0	0.0	1.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.4000	1.0	0														
10	***	ARBRE														
1	***	PLANTE														
2.000	5.000	0.400	0.0	1.000	0.0	0.0	0.0	0.0	1.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0333	1.0	0														
11	***	ARBRE														
1	***	POMME														
5.000	6.000	0.033	1.000	1.000	0.0	0.0	0.0	1.000	0.0	1.000	0.0	1.000	0.0	0.0	0.0	0.0
1.0000	1.0	0														
12	***	ARBRE														
1	***	TAMBOUR														
6.000	6.000	1.000	1.000	0.0	0.0	1.000	1.000	0.0	1.000	0.0	0.0	1.000	0.0	0.0	0.0	0.0
1.0000	1.0	0														
13	***	ARBRE														
1	***	TROMPETTE														
6.000	6.000	1.000	1.000	0.0	0.0	1.000	0.0	0.0	1.000	0.0	0.0	1.000	0.0	0.0	0.0	0.0
0.0	0.0	C														
2	***	CLOCHE														
2	***	CLOCHE														
0.0	6.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0571	1.0	0														
3	***	CLOCHE														
3	***	GATEAU														
6.000	7.000	0.057	1.000	0.0	1.000	1.000	1.000	0.0	0.0	1.000	1.000	0.0	0.0	0.0	0.0	0.0
1.0000	1.0	0														

