

1

*Études*

# Enseignement mathématique et linguistique

*ou comment utiliser les faits de langue à des fins d'enseignement  
de la mathématique : quelques exemples de problèmes\**

**Yves Gentilhomme**

*Faculté des Lettres et Sciences Humaines,  
(Besançon).*

## 1. Introduction pédagogique.

Dans un processus pédagogique quel qu'il soit il importe, dès le départ, de ne pas perdre de vue un certain nombre de questions :

- 1) dans quel but on enseigne?
- 2) qu'est-ce qu'on enseigne?
- 3) à qui l'on enseigne?
- 4) dans quelles conditions l'on enseigne?

---

(\*) N.D.L.R. Nous sommes heureux de publier ici un texte mis au point par l'Auteur à partir de sa conférence si vivante prononcée à Besançon lors des journées A.P.M.E.P. de juin 1969. M. Gentilhomme nous a donné également des exercices avec corrigés qui paraîtront dans un Bulletin suivant.

Ces diverses questions, d'ailleurs, n'étant pas indépendantes les unes des autres.

Ce n'est qu'après avoir précisé les données que l'on peut se prononcer sur les méthodes praticables, sur le « rendement » de chacune d'elles, sur une optimisation souhaitable...

1.1. La question (1) est susceptible de multiples interprétations. Nous la schématisons (\*) comme suit :

Soit un ensemble H d'êtres humains (par ex. les élèves d'une classe de lycée) soumis à notre influence (enseignement mathématique); ces élèves possèdent certaines propriétés P, en ce sens que chaque élève possède certaines propriétés dont les unes sont spécifiques, d'autres générales, plus restrictivement les éléments de H sont aptes à se conduire de certaine façon dans un complexe situationnel donné (passer un examen, exercer un métier...). Il nous appartient d'effectuer une transformation T sur les éléments de H de façon à obtenir un nouvel ensemble  $T(H) = H'$  doué de nouvelles propriétés P'. Ainsi, pour fixer les idées, après transformation, les éléments de H' devront être capables de suivre un cours de niveau plus élevé.

S'agit-il d'un bachotage, c'est-à-dire de l'aptitude à triompher des arcanes d'un examen ou, au contraire, d'une formation de la réflexion à des fins diverses : spéculation abstraite, pragmatique... ou d'un compromis entre divers impératifs? Il est clair que selon l'objectif à atteindre le contenu du symbole T ne sera pas le même.

Les exercices proposés, ci-dessous, n'ont pour objet ni de former des mathématiciens par excellence, ni des linguistes de profession, mais d'apprendre à se servir du puissant outil qu'est la mathématique dans un domaine qui, jadis, paraissait lui échapper. Ce faisant, ipso facto, on contribue à fixer les notions et les procédures mathématiques dans l'esprit des élèves; on les incite à réfléchir sur le sens et l'adéquation à la réalité expérimentale des concepts mis en œuvre, à mettre en doute les poncifs et les clichés acceptés par tradition, enfin on oblige à préciser, autant que faire se peut, les notions tirées de l'expérience linguistique de façon à ce qu'elles se plient à un raisonnement déductif présentant un commencement de rigueur.

1.2. Le point (2) est déterminé, en principe, par les programmes scolaires en vigueur. Toutefois, si l'on se place du point de vue de la linguistique, ce programme doit être remis en question. Ainsi, par exemple, les propriétés des angles inscrits, le signe du trinôme, ... trouvent rarement une application naturelle en langue, par contre une initiation naïve à la théorie des réseaux... semble éminemment souhaitable.

(\*) Que le lecteur veuille bien ne pas rechercher une « rigueur mathématique » dans le symbolisme proposé ci-dessous; ce symbolisme n'est qu'une sténographie et une figure de style, bien que les symboles mentionnés obéissent, grosso modo, à une certaine « algèbre » que l'on pourra s'amuser à esquisser, par exemple :

$$|T(H_1 \cup H_2)| < |H_1 \cup H_2|$$

Ceci posé, le professeur de mathématique n'ayant pas pour fonction de se substituer à son homologue littéraire, le matériau linguistique utilisé ici sera « ce que tout élève est censé connaître ». Ce matériau, à première vue pourra paraître naïf, désuet, voire sans intérêt du point de vue du linguiste moderne mais c'est là une controverse où il ne nous appartient pas de nous engager dans le cadre de cet exposé (pas plus que nous ne discuterions par ex. de l'intérêt pour la chimie de tel ou tel mélange que l'on fait analyser à l'étudiant chimiste. C'est un exercice. L'intérêt réside dans le processus et non dans l'objet même).

1. 3. Selon que les éléments de H sont des adultes ou des enfants, selon qu'ils adorent la mathématique ou l'ont en horreur, selon les connaissances acquises, selon les façons dont celles-ci ont été acquises, selon les aptitudes, les goûts, les préoccupations, les centres d'intérêt de tout un chacun..., la nature de la transformation T est susceptible de varier du tout au tout. Nous laissons délibérément en suspens ce problème pédagogique fondamental et ce n'est qu'occasionnellement que nous nous risquerons à de modestes suggestions. La seule hypothèse qui nous tient à cœur est que l'étude de la mathématique n'est pas toujours aisée pour tout le monde, qu'il existe des esprits « obèses » qui ne se plient pas facilement aux notions abstraites sans le support d'exemples concrets, être nul en mathématique n'est pas équivalent à être nul tout court, un peu d'humanité de la part des esprits « subtils » à l'égard de leurs semblables moins chanceux, aboutirait à rendre ces derniers un peu moins « obèses » et leur redonnerait, peut-être, le goût des quintessences (\*). Notre ambition « La mathématique à travers les faits de langue » est, pour une part, une tentative de « récupération ».

1. 4. C'est là toute la question des matériels didactiques, des horaires, du cardinal de H, de l'état d'esprit de la direction et des parents d'élèves, etc. Dans cette étude nous essaierons de ne pas enfreindre les contraintes existantes dans les conditions usuelles de travail.

N.D.L.R. Nos lecteurs voudront bien excuser plusieurs fautes que nous avons commises dans les figures; il nous était difficile de les corriger sans retarder encore le *bulletin*; d'où un certain nombre d'errata : le lecteur est invité à corriger lui-même les figures fautives.

---

(\*) D'autant plus que les « obèses » en mathématique s'avèrent parfois « supra-subtils » en d'autres matières.

## 2. Manipulations de listes.

### 2.1. Le corpus — objet d'investigation.

Soit la poésie de J. Prévert :

#### *Le Cancre (\*)*

- 1 *Il dit non avec la tête*
- 2 *mais il dit oui avec le cœur*
- 3 *il dit oui à ce qu'il aime*
- 4 *il dit non au professeur*
- 5 *il est debout*
- 6 *on le questionne*
- 7 *et tous les problèmes sont posés*
- 8 *soudain le fou rire le prend*
- 9 *et il efface tout*
- 10 *les chiffres et les mots*
- 11 *les dates et les noms*
- 12 *les phrases et les pièges*
- 13 *et malgré les menaces du maître*
- 14 *sous les huées des enfants prodiges*
- 15 *avec des craies de toutes les couleurs*
- 16 *sur le tableau noir du malheur*
- 17 *il dessine le visage du bonheur.*

J. PRÉVERT.

Quels objets de langue et quelles structures peut-on chercher à mettre en évidence?

### 2.2. Les observables.

Disons, en gros : il y a le texte imprimé formé de mots graphiques, de caractères typographiques, de vers disposés sur une feuille blanche... et derrière cette représentation conventionnelle sur un support rigide, il y a la langue parlée, avec les sons, les phonèmes, les groupes phonétiques, l'intonation... Entre la parole et l'écriture on observe certaines correspondances, tout n'est pas fortuit. Enfin entre les divers objets rencontrés il existe des liens de nature variée, plus ou moins rigides ou lâches, on observe des contraintes portant

---

(\*) Le choix du texte est un atout capital dont dispose l'enseignant. Ce choix crée dès le départ une ambiance, un état d'esprit dans lequel se poursuit le travail. C'est un peu comme un parfum qui agit de façon plus ou moins inconsciente sur l'entourage en provoquant un climat de détente, de confiance, de gaieté... ou au contraire de méfiance, d'ennui, de dégoût... Attention, les goûts du maître ne coïncident pas nécessairement avec ceux des patients qui lui sont soumis.

sur les objets appartenant aussi bien au domaine de la graphie, de la phonétique, de la grammaire, de la versification et qui sont fonction de l'information à communiquer, des sous-entendus à évoquer, de la joie esthétique que l'on désire faire partager...

Tous les points de vue ne sont pas accessibles à nos élèves. Notamment le point de vue linguistique fondamental, je veux dire celui de la langue sous sa forme parlée, exige des connaissances qui ne figurent pas dans nos programmes. C'est pourquoi, paradoxalement, nous nous appuyerons principalement sur les données écrites, que l'on « voit » avec ses yeux aussi longtemps qu'on le désire, et qui, de ce fait, paraissent, en quelque sorte, plus « matérielles ».

### 2.3. *Le mot graphique.*

Supposons que les objets que nous avons l'intention de manipuler soient des mots graphiques. Par mot graphique on entend habituellement une suite ininterrompue de lettres, limitée par des espacements (des « blancs »), des signes de ponctuation ou des apostrophes.

Dans la poésie de J. Prévert cette définition ne paraît susciter aucun commentaire et pourtant il nous faut nous décider pour savoir si nous distinguons les majuscules des minuscules : *H* et *h* forment-ils le même mot graphique ou des mots différents? Adoptons, pour simplifier, la convention arbitraire, mais raisonnable, de confondre majuscules et minuscules. D'une façon générale, il faudra se fixer sur ce que l'on entend par lettres identiques et par lettres différentes. Confond-on les divers corps de caractère (italique, bas de casse, romains, ...)? Distingue-t-on les lettres accentuées des non-accentuées? On sait que l'accent ne figure pas obligatoirement sur les majuscules, faut-il le rétablir?

Comment appliquer notre définition à l'extrait tiré des « Soliloques du pauvre » de Jehan Rictus (problème de l'apostrophe)?

*Merd'! V'là l'Hiver et ses dur'tés,  
V'là l'moment de n'pus s'mettra à poil:  
V'là qu'ceuss' qui tiennent la queue d'la poêle  
Dans l'Midi vont s'carapater.*

Une définition claire et universelle n'est pas aisée à formuler, et la recherche d'une définition applicable à un texte, comme celui de Jehan Rictus, constitue, en soi, un problème souvent fort délicat et susceptible de plusieurs solutions plus ou moins conformes aux usages et à notre bon sens.

Autre problème : distingue-t-on les mots graphiques *le* dans *avec le cœur* et dans *on le questionne* ; autrement dit, la valeur grammaticale et le sens sont-ils pris en considération? Plaçons-nous dans les conditions réalistes d'une machine électronique qui ne « comprend » pas le texte, qui ne l'« entend » pas et à

laquelle on n'a fourni aucune instruction spéciale à ce propos. Nous confondrons donc les homographes tels que *vers* (animaux) et *vers* (dans la direction de), *parent* (être humain) et *parent* (du verbe *parer*). La convention inverse eut été tout aussi rationnelle bien que plus difficile à formuler, c'est une affaire de technologie et les deux points de vue trouvent leur utilisation dans la pratique.

Laissons également de côté la question épineuse des mots composés (*chemin de fer, chemin de ronde, chemin de terre, chemin de rencontre...*) et celui des tirets indiquant tantôt l'union (pare-brise) tantôt la coupure dans un passage à la ligne, tantôt les deux à la fois. Quand a-t-on affaire à « un mot », quand à plusieurs? Limitons-nous à notre poésie (titre et signature exclus).

#### 2.4. *La théorie naïve des ensembles, ou l'art de s'en tirer avec des listes longues et embrouillées.*

Mini-problème (1) sur les opérations :

Établir les vocabulaires des mots graphiques suivants :

A : Les mots (2) des 4 premiers vers,

B : Les mots (2) des 4 vers suivants,

C : Les mots (2) des vers 8 à 11 compris,

R : Les mots de toute la poésie, vocabulaire que nous appellerons « référentiel », en ce sens que nous n'envisagerons pas d'autres mots que ceux-là au cours de notre étude.

En quoi les vocabulaires A, B et C sont-ils semblables, en quoi sont-ils différents? Quelle est la part originale de chaque vocabulaire? Quels sont les mots indispensables pour écrire les 8 premiers vers, les 11 premiers vers, les vers de 5 à 11? Quels sont les mots qui n'ont pas été utilisés dans les 4 premiers vers, dans les 4 vers suivants? Les mots non utilisés dans les 8 premiers vers peuvent-ils être obtenus simplement à partir des deux listes précédentes?

Faire des schémas « visuels ».

Comparer les vocabulaires partiels obtenus entre eux, ainsi que leur étendue (nombre de mots).

---

#### Commentaires :

(1) Dans la pratique des laboratoires les mêmes calculs sont effectués sur des textes pouvant comporter des milliers de mots, ex. : le vocabulaire des œuvres de Proust, ce qui est inadmissible dans les conditions scolaires normales — faim de temps et surtout d'intérêt pour des manipulations longues et fastidieuses (à moins de disposer de machines). Dans un mini-problème on essaie de faire à une petite échelle ce qui en réalité peut se passer à grande échelle.

(2) Entendez : « mots graphiques ». Les abus de langage « par abréviation » sont inévitables parce que nécessaires pour qu'un texte soit plus aisément compris à condition de ne pas entraîner d'équivoque.

2 5. *Corrigé* (partiel)

Données : texte considéré comme séquence de mots graphiques.

Traitement (1) : formation de listes de mots du texte répondant à certaines conditions.

Vocabulaires (2) :

A = { à, aime, au, avec, ce, cœur, dit, il, la, le, mais, non, oui, professeur, qu (3), tête (4)},

B = { debout, est, et, fou, il, le, les, on, posés, prend, problèmes, questionne, rire, sont, soudain, tous },

C = { chiffres, dates, efface, et, fou, il, le, les, mots, noms, prend, rire, soudain, tout },

R = { à, aime, au, avec, bonheur, ce, chiffres, cœur, couleurs, craies, dates, de, debout, des, dessine, dit, du, efface, enfants, et, est, fou, huées, il, la, le, les, mais, maître, malgré, malheur, menaces, mots, noir, noms, non, on, oui, phrases, pièges, posés, prend, problèmes, professeur, prodiges, qu, questionne, rire, sont, soudain, sous, sur, tableau, tête, tous, tout, toutes, visage }.

Étendue des divers vocabulaires (cardinaux des ensembles) :

$$|A| = 16, \quad |B| = 16, \quad |C| = 14, \quad |R| = 58$$

Mots communs aux divers vocabulaires (intersection)

$$A \cap B = \{ il, le \}, \quad |A \cap B| = 2,$$

$$B \cap C = \{ et, fou, il, le, les, prend, rire, soudain \}, \quad |B \cap C| = 8,$$

$$A \cap C = \{ il, le \}, \quad |A \cap C| = 2$$

*Commentaire du corrigé :*

(1) La pratique de l'enseignement montre que, trop souvent, les élèves « se lancent » dans le calcul sans trop savoir ce qu'ils cherchent. C'est pourquoi, à l'instar de la tradition pédagogique mathématique qui conseille de mettre en évidence les « hypothèses » et les « conclusions », il est bon de faire précéder la solution des « données » sur lesquelles porte le travail et du « traitement » que l'on compte faire subir aux données.

(2) Par vocabulaire on entend généralement un ensemble d'éléments-mots, les vocables, autrement dit l'ordre de présentation n'est pas pertinent, contrairement à la séquence formée d'occurrences-mots. Là au contraire, l'ordre est essentiel, un même mot peut être répété plusieurs fois, chaque occurrence doit être considérée comme un objet complexe, formé d'un élément-mot avec la place qu'il occupe, celle-ci pouvant être repérée, par exemple, par un numéro d'ordre.

Bien que l'ordre d'écriture des éléments soit sans importance théorique, il est bon dans la pratique, pour mieux s'y retrouver, d'inscrire les mots dans l'ordre alphabétique usuel.

(3) En accord avec notre convention de ne pas considérer l'apostrophe comme une lettre, nous écrivons *qu* et non *qu'*. La convention inverse serait tout aussi défendable. L'esprit de tolérance est capital en matière de sciences. La description d'un système n'admet pas obligatoirement une forme unique. Elle est souvent tributaire du but qu'on se pose.

(4) Il importe de bien faire distinguer, dans une étude linguistique, la langue-objet, correspondant ici au texte traité, de la métalangue, c'est-à-dire la langue dont on se sert pour décrire l'objet donné, pour parler du texte considéré. Par convention nous réservons les caractères italiques pour la langue-objet, les caractères normaux pour la métalangue. En classe il est commode d'utiliser des couleurs différentes.

On constate que l'on a ici  $A \cap B = A \cap C \subset B \cap C$ , relations non vérifiées pour des vocabulaires quelconques. Il s'ensuit un certain nombre de simplifications.

Mots spécifiques aux vocabulaires comparés deux à deux (différence non symétrique) :

$A \setminus B = A \setminus (A \cap B) = \{ \text{à, aime, au, avec, ce, cœur, dit, la, mais, non, oui, professeur, qu, tête} \}$ ,

$A \setminus C = A \setminus B$ ,

$B \setminus C = \{ \text{debout, est, on, posés, problèmes, questionne, sont, tous} \}$ , etc...

Mots spécifiques à chaque vocabulaire opposé aux deux autres :

$(A \setminus B) \setminus C = A \setminus B \approx A \setminus C$  (dans ce cas particulier)...

Diagramme d'Euler-Venn.

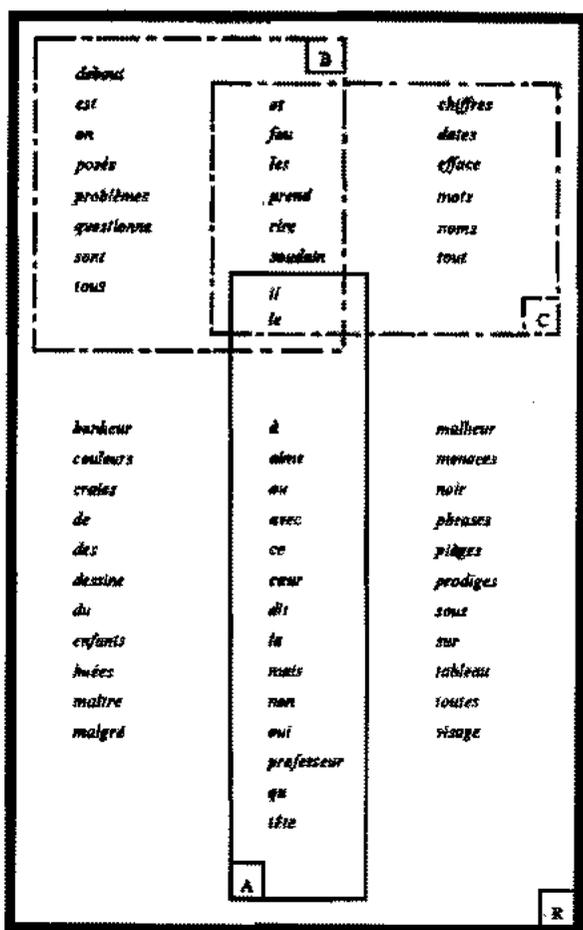


FIG. 1.

Ce par quoi diffèrent deux vocabulaires (différence symétrique) :

$$A \Delta B = (A \setminus B) \cup (B \setminus A) =$$

{à, aime, au, avec; ce, cœur, debout, dit, est, la, mais, non, on, oui, posés, problèmes, professeur, qu, questionne, sont, tête, tous }...

Mots indispensables pour écrire les 8 premiers vers (réunion) :

$A \cup B =$  {à, aime, au, avec, chiffres, cœur, debout, dit, et, est, fou, il, la, le, mais, non, on, oui, posés, prend, problèmes, professeur, qu, questionne, rire, sont, soudain, tête, tous }...

Mots faisant partie du vocabulaire général restés inutilisés dans les 4 premiers vers (complément) :

$\bar{A} = \bar{B} = R \setminus A =$  {bonheur, chiffres, couleurs, craies, dates, de, debout, des, dessin, du, efface, enfant, et, est, fou, hutes, les, maître, malgré, malheur, menaces, mots, noir, noms, on, phrases, plèges, posés, prend, problèmes, prodiges, questionne, rire, sont, soudain, sous, sur, tableau, tous, tout, toutes, visage }...

Tous ces résultats se lisent commodément sur le schéma ci-contre (diagramme d'Euler-Venn). Chaque rectangle y représente une feuille de papier sur laquelle est inscrite une liste de mots. Les listes se chevauchent partiellement.

On constate notamment que :

$A \cup B = \bar{A} \cap \bar{B}$  (formule de Morgan), en d'autres termes, pour établir la liste des mots restés inutilisés dans les 8 premiers vers, il suffit de copier les mots communs aux listes complémentaires des 4 premiers vers et des 4 vers suivants.

Il existe une troisième façon d'obtenir toutes les listes par la méthode dite de pointage des présents et des absents.

## 2.6. Méthode par pointage.

Dans le tableau à double entrée, ci-après, la présence d'un mot écrit à gauche d'une ligne, dans l'ensemble nommé en tête d'une colonne est enregistrée par 1, son absence par 0 (pour alléger les notations les zéros peuvent être sous-entendus lorsqu'aucune équivoque n'est à craindre).

Ainsi tout sous-ensemble de R, dont les éléments sont inscrits dans un ordre donné, ici l'ordre alphabétique usuel, est caractérisé par un nombre binaire de  $|R| = 58$  chiffres figurant dans la colonne correspondante. Ex., pour  $A \cap B$

$$\begin{array}{ccccccc} 000 & \dots & 00 & 101 & 00 & \dots & 000 \\ \hline & & 24 & & 31 & & \end{array} \text{ zéros}$$

Cette notation, peu maniable pour un être humain est parfaitement adaptée aux mémoires des machines électroniques et à certaines considérations théoriques sur les ensembles.

R	A B C	A∩B A∩C B∩C	A∪B A∪C B∪C	A B A C B A B C C A C B	A∩B A∩C B∩C	A∪B∩C A∪B∩C (A B) C (B A) C (C A) B	$\bar{A} \bar{B} \bar{C}$
à	1		1 1	1 1	1 1	1 1	1 1
aitrac	1		1 1	1 1	1 1	1 1	1 1
au	1		1 1	1 1	1 1	1 1	1 1
avec	1		1 1	1 1	1 1	1 1	1 1
bonheur							1 1 1
ce	1		1	1 1	1 1	1 1	1 1 1
chiffres	1	1	1 1 1		1 1	1 1 1	1 1 1
cœur	1		1 1	1 1	1 1	1 1	1 1 1
couleurs							1 1 1
crâtes							1 1 1
dates		1	1 1		1 1	1 1	1 1 1
de							1 1 1
debout	1		1 1		1 1	1 1	1 1 1
des							1 1 1
dessine							1 1 1
dû	1		1 1	1 1	1 1	1 1	1 1 1
du							1 1 1
efface		1	1 1		1 1	1 1	1 1 1
enfants							1 1 1
et	1 1	1	1 1 1	1 1	1 1	1 1	1 1 1
est	1		1 1	1 1	1 1	1 1	1 1 1
fou	1 1	1	1 1 1	1 1	1 1	1 1	1 1 1
huées							1 1 1
il	1 1 1	1 1 1	1 1 1			1 1	1 1 1
la	1		1 1	1 1	1 1	1 1	1 1 1
le	1 1 1	1 1 1	1 1 1			1 1	1 1 1
les	1 1	1	1 1	1	1	1	1 1 1
tout	1		1 1		1 1	1 1	1 1 1
toutes							1 1 1
visage							1 1 1
Total 58	16 16 14	2 2 8	30 28 22	14 14 14 8 12 6	28 26 14	2 36 14 8 6	42 42 44

On découvre aisément les règles d'écriture qui permettent de remplir une colonne en fonction des colonnes précédentes et, partant de là, les opérations associées sur les nombres binaires.

**2.7. Classement des vocabulaires.**

Dans l'énoncé, on nous demande de classer les vocabulaires étudiés, listes que dans un but de concision nous désignerons par les symboles ensemblistes correspondants :

$$\begin{aligned}
 &R, A, B, C; \\
 &A \cap B = A \cap C = A \cap B \cap C, \quad B \cap C; \\
 &A|B = A|C = (A|B)|C, \quad B|A, \quad B|C = (B|A)|C, \\
 &\quad\quad\quad C|A, \quad C|B = (C|A)|B;
 \end{aligned}$$

$$\begin{aligned}
 &A \Delta B, A \Delta C, B \Delta C; \\
 &A \cup B, B \cup C, A \cup B \cup C; \\
 &\bar{A}, \bar{B}, \bar{C}, \quad \overline{A \cup B} = \bar{A} \cap \bar{B}.
 \end{aligned}$$

Remarquons qu'en partant des listes A, B, C et R, on le voit sur le schéma, il est possible d'obtenir 6 listes élémentaires non empiétant les unes sur les autres (au lieu de 8 dans le cas général\*) :

$$A \cap B \cap C, (A \setminus B) \setminus C, (B \setminus A) \setminus C, (C \setminus A) \setminus B, (B \cap C) \setminus A \text{ et } R \setminus (A \cup B \cup C).$$

Toute liste, simple ou composée, déduite de A, B, C et R comprendra ou ne comprendra pas chacune des listes élémentaires, ce qui donne au total  $2^6 = 64$  possibilités. Heureusement on ne nous invite pas à les comparer toutes les unes aux autres, mais seulement 26 d'entre elles qui, compte tenu des égalités, se réduisent à 21 listes.

Un tel classement ne s'obtient pas du premier coup. Il est bon d'opérer par étapes, de faire d'abord des classements partiels. A titre indicatif nous donnons le réseau général conçu comme suit. Lorsqu'une liste en contient une autre, on les réunit par une flèche allant de la 1<sup>re</sup> à la 2<sup>e</sup>. Comme le nombre de flèches s'avère très important, il est judicieux de faire l'économie de celles qui vont de soi (transitivité de l'inclusion). Ainsi  $A \cup B \cup C$  contient  $A \cup B$ , laquelle contient A; il est clair que  $A \cup B \cup C$  contient également A. Nous sous-entendons la flèche correspondante. D'où deux problèmes accessoires de dénombrement : combien de flèches sont nécessaires et combien font double emploi.

Il appartient à chaque élève de rendre son réseau le plus lisible et le plus harmonieux possible en évitant notamment certains chevauchements, en plaçant judicieusement les sommets, en donnant aux lignes une sinuosité agréable.

Sous-réseaux :

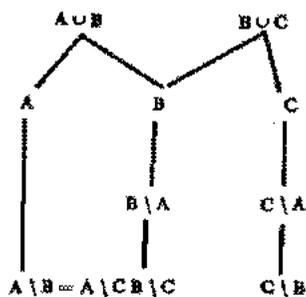
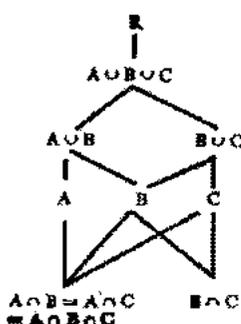


Fig. 2\*\*.

(\*) A savoir, dans le même ordre :  $A \cap B \cap C, A \cap \bar{B} \cap \bar{C}, \bar{A} \cap B \cap \bar{C}, \bar{A} \cap \bar{B} \cap C, \bar{A} \cap B \cap C, \bar{A} \cap \bar{B} \cap \bar{C}$  et  $A \cap \bar{B} \cap C, A \cap B \cap \bar{C}$ . Ici les deux dernières classes sont vides.

(\*\*) Dans les fig. 2 et 4, effacer l'arc (C, A ∩ B) et le remplacer par l'arc (B ∩ C, A ∩ B).

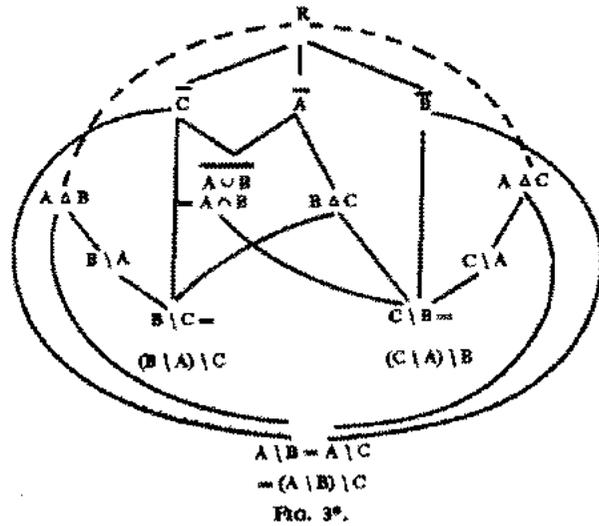


FIG. 3°.

Et par raccordement des 3 sous-réseaux :

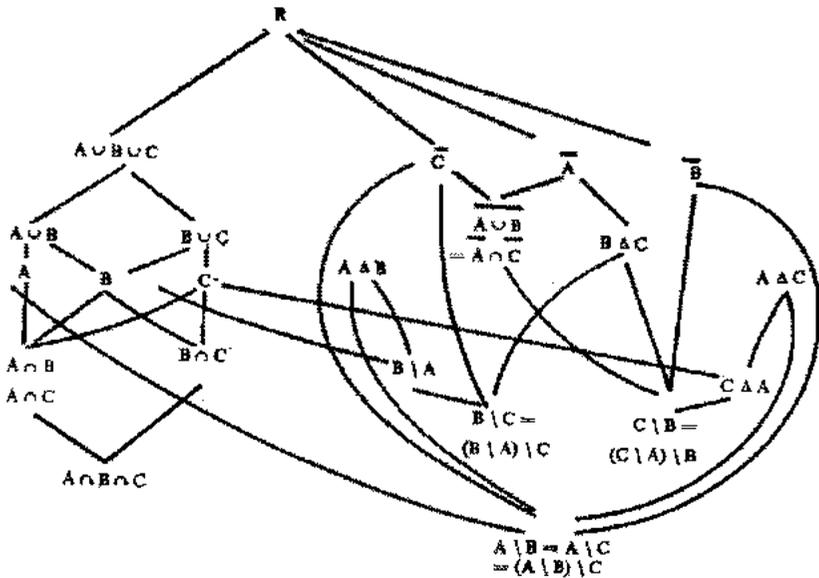


FIG. 4°.

Un tel réseau présente l'avantage de permettre de répondre immédiatement à toute question posée sur la hiérarchie par inclusion des vocabulaires étudiés.

(\*) Dans les fig. 3 et 4, déplacer  $A \cup B$  (qui est égal à  $\overline{A \cap B}$ ) sous  $\overline{A}$  et  $\overline{B}$ . Ajouter dans la fig. 4 les arcs  $(A \cup B \cup C, A \Delta C)$ ,  $(B \cup C, B \Delta C)$ ,  $(A \cup B, A \Delta B)$ ,  $(B \cap C, A \cap B)$ . Effacer  $(A \cap C, A \cap B \cap C)$ ,  $(B \cap C, A \cap B \cap C)$ ,  $(C, A \cap B)$ . Sous  $A \Delta C$ , lire  $C \setminus A$ .

### 2.8. Quelques problèmes d'adressage.

Dans la poésie de Prévert chaque occurrence-vers (\*) porte un numéro  $v$ ; chaque occurrence-mot un numéro  $m_p$  dans la poésie et  $m_v$  dans le vers; chaque occurrence-lettre un numéro dans la poésie  $l_p$ , un dans le vers  $l_v$ , un dans le mot  $l_m$ . Étudier les relations entre ces diverses adresses.

Exemple : pour la lettre  $e$  de *cœur*,

$$v = 2, m_p = 13, m_v = 7,$$

$$l_p = 37, l_v = 19, l_m = 1.$$

Bien entendu il convient de préciser, au préalable, ce qu'on entend par lettre, si on compte ou non les blancs, les signes de ponctuation, les apostrophes, si on dédouble les groupes tels que  $\alpha$ .

On obtient ainsi une série d'injections de l'ensemble des occurrences dans l'ensemble des entiers, des couples et des triplets d'entiers qu'il est loisible de représenter à l'aide de tableaux. On en déduit des applications diverses entre les éléments envisagés.

Ainsi, par exemple, une occurrence-mot admet comme adresse soit  $(v, m_v)$  soit  $m_p$ .

$v \backslash m_v$	1	2	3	4	5	6	7	8	...
1	1	2	3	4	5	6			
2	7	8	9	10	11	12	13		
3	14	15	16	17	18	12	20	21	
4	22	23	24	25	26				
5	27	28	29						
6	30	31	32						
7	33	34	35	36	37	38			
.									
.									

Tableau qui permet de résoudre divers problèmes : domaine de définition de la fonction  $m_p \mapsto (v, m_v)$ , détermination de  $v$  et  $m_v$  connaissant  $m_p$  et inversement, l'étude de  $\max m_p$  en fonction de  $v$ , etc...

### 2.9. Quelques mini-problèmes de comptage.

Supposons qu'il vous faille enseigner le vocabulaire nécessaire à la compréhension de la poésie de Prévert. Existe-t-il un ordre de priorité pour un tel enseignement?

(\*) Pour la notion d'occurrence, cf. § 2.5., note (2).

L'idée qui vient spontanément à l'esprit est que les mots les plus souvent rencontrés sont les plus utiles, d'où l'intérêt d'un classement par ordre d'effectifs (fréquences) décroissants.

Comme, en outre, il est souhaitable de savoir apprécier le rendement d'une telle méthode, on calculera les effectifs cumulés (fréquences cumulées) qui déterminent la proportion du texte couverte par les mots les plus fréquents retenus et, en un certain sens, le pourcentage de compréhension de ce texte.

Il est évident que le problème de la compréhension est, en fait, bien plus complexe et que la connaissance du mot n'est pas le seul facteur à prendre en considération. Néanmoins, à titre d'essai préliminaire une démarche fondée sur des hypothèses simplifiées au maximum n'est pas totalement déraisonnable. La seule attitude scientifique correcte est de s'en remettre à l'expérience, quitte à critiquer en deuxième instance certaines « balvetés ».

La partition de l'ensemble des occurrences-mots de la poésie en fonction des éléments-mots associés conduit au tableau des effectifs suivant (les mots étant classés par « paquets » et non pas individuellement, on parlera plutôt de préordre).

N°		eff.	eff. cum.	N°		eff.	eff. cum.	N°		eff. cum.
1	<i>les</i> .....	10	10	20	<i>dates</i> .....	1	39	39	<i>phrases</i> ..	1
2	<i>il</i> .....	7	17	21	<i>de</i> .....	1	40	40	<i>pièges</i> ....	1
3	<i>et</i> .....	6	23	22	<i>debout</i> ..	1	41	41	<i>posés</i> ....	1
4	<i>le</i> .....	6	29	23	<i>dessine</i> ..	1	42	42	<i>prend</i> ....	1
5	<i>avec</i> .....	4	33	24	<i>efface</i> ...	1	43	43	<i>problèmes</i>	1
6	<i>dit</i> .....	4	37	25	<i>enfants</i> ..	1	44	44	<i>prodiges</i> ..	1
7	<i>du</i> .....	3	40	26	<i>est</i> .....	1	45	45	<i>professeur</i>	1
8	<i>des</i> .....	2	42	27	<i>fou</i> .....	1	46	46	<i>qu</i> .....	1
9	<i>non</i> .....	2	44	28	<i>huées</i> ....	1	47	47	<i>questionne</i>	1
10	<i>oui</i> .....	2	46	29	<i>la</i> .....	1	48	48	<i>rbr</i> .....	1
11	<i>à</i> .....	1	47	30	<i>mais</i> .....	1	49	49	<i>sont</i> .....	1
12	<i>aime</i> ....	1	48	31	<i>matre</i> ...	1	50	50	<i>soudain</i> ..	1
13	<i>au</i> .....	1	49	32	<i>malgré</i> ..	1	51	51	<i>sous</i> .....	1
14	<i>bonheur</i>	1	50	33	<i>malheur</i> ..	1	52	52	<i>sur</i> .....	1
15	<i>ce</i> .....	1	51	34	<i>menaces</i> ..	1	53	53	<i>tableau</i> ..	1
16	<i>chiffres</i> ..	1	52	35	<i>mots</i> .....	1	54	54	<i>tête</i> .....	1
17	<i>cœur</i> .....	1	53	36	<i>noir</i> .....	1	55	55	<i>tous</i> .....	1
18	<i>couleurs</i>	1	54	37	<i>nour</i> .....	1	56	56	<i>tout</i> .....	1
19	<i>crates</i> ...	1	55	38	<i>on</i> .....	1	57	57	<i>toutes</i> ....	1
							58	58	<i>visage</i> ....	1

On remarque le nombre considérable d'hapax (mots rencontrés une fois) seulement et que c'est parmi ces hapax que figurent les mots les plus chargés de sens tels que : *questionne, soudain, bonheur*... De sorte que les 10 premiers mots, recouvrant près de 49 p. 100 du texte, en font ignorer l'essentiel. Concluons que l'expérience d'un tel enseignement ne paraît guère probante d'après ce

mini-comptage. Cependant il serait faux de vouloir inconsidérément étendre ces conclusions, à de plus larges perspectives.

Un mini-comptage analogue peut-être repris avec des lettres (\*) au lieu des mots et on peut se demander dans quelle mesure il est possible de deviner les lettres les plus rares si par un malencontreux hasard elles étaient effacées et que seules étaient restées les lettres les plus fréquentes.

Voici un autre exercice. Imaginons qu'un « espion » veuille transmettre secrètement la poésie en l'encodant comme suit. Chaque élément-mot est remplacé par son numéro dans une certaine liste. Comment devra-t-il fabriquer la liste de façon à ce que le message soit le plus bref possible?

D'une façon générale le thème-message secret -- est inépuisable en matière d'exercices de combinatoire et de statistique descriptive.

## 2.10. *Enoncés divers.*

### 2.10.1. *MM. Dupont et Durand se disputent.*

Xavier, Yvonne et Zoé publient chacun un dictionnaire.

M. Dupont remarque que les mots vedettes qui figurent à la fois dans le Xavier et le Zoé, figurent également dans le Yvonne et le Zoé. Il en déduit que le Yvonne est plus complet que le Xavier et même, plus précisément, que le Yvonne contient tous les mots vedettes du Xavier. Son ami Durand conteste le fait. On vous demande d'être juge et de mettre d'accord MM. Dupont et Durand.

Quelque temps après, les maisons d'édition Alesia, Babylone et Carnot sortent encore trois dictionnaires. Cette fois-ci c'est M. Durand qui s'aperçoit que si l'on fondait ensemble d'une part l'Alesia et le Carnot, d'autre part le Babylone et le Carnot, le premier recueil contiendrait tous les mots vedettes du second. Il en conclut que le Babylone n'apporte rien de nouveau par rapport à l'Alesia. M. Dupont l'attaque et la dispute recommence. A vous d'intervenir.

### 2.10.2. *Les peines de cœur de Nicole.*

Nicole travaille dans un centre de lexicologie. Son patron lui confie la tâche suivante : trouver les entrées communes au Larousse et au Quillet et ensuite comparer cette liste au dictionnaire de l'Académie, plus précisément, relever les entrées originales, c'est-à-dire celles qui figurent dans le dictionnaire de l'Académie, ou dans la liste mais pas dans les deux à la fois.

Sachant Nicole très distraite, le patron confie la même tâche à Monique afin de confronter les résultats.

Or, il se trouve que Monique est la meilleure amie de Nicole, elle connaît les peines de cœur de cette dernière et désire la décharger un peu des soucis du laboratoire, aussi lui propose-t-elle l'arrangement qui suit :

« Ma chérie, pour cette fois-ci tu te contenteras de rechercher ce que le

---

(\*) Les comptages de mots nécessitent souvent des textes fort longs et il est pratique de traiter les mêmes questions en comptant des lettres, tout en rappelant les divergences qui en résultent.

Larousse et le Dictionnaire de l'Académie ont d'original. Moi à la place du Larousse je prendrai le Quillet. Ensuite je recopierai les mots communs aux deux listes. Une autre fois c'est toi qui feras le gros du travail ».

Que pensez-vous de cet arrangement?

Pouvez-vous comparer la durée des travaux demandés et effectués, en essayant de « chiffrer les temps » de façon très approximative? (\*)

### 2.10.3. *Y-a pas d'justicel* (\*\*)

Mon papa y m'a dit comme ça : « Zizi, j'veux qu'tu soyes une fille bien, qu'tu saches t'faire valoriser dans l'beau monde. T'as encore chopé un zéro en autograffe. Alors, écoute-moi bien, pour te punir et pour t'apprendre à aimer la belle littérature, dimanche après-midi, tu m'copieras tous les noms qui sont à la fois dans l'acte 1, scène 2 d'Athalie, dans l'acte 2, scène 3 de Bérénice et dans l'acte 3, scène 1 de Clitandre. T'ajouteras à cette liste tous les noms de l'acte 1, scène 2 d'Athalie, après(s) en zavoir enlevé ceux de l'acte 3, scène 1 de Clitandre et ceux de l'acte 2, scène 3 de Bérénice. Tu m'suis? Puis tu compléteras ce qu't'auras fait avec tous les noms qui sont à la fois dans l'acte 1, scène 2 d'Athalie et dans l'acte 3, scène 1 de Clitandre dont t'auras biffé auparavant tous les mots qui sont dans l'acte 2, scène 3 de Bérénice. Puis, n'oublie pas, t'ajouteras encore à tout ça la même chose mais en remplaçant Clitandre par Bérénice et Bérénice par Clitandre.

Enfin, fais attention, dans c'que t'auras écrit, t'effaceras tous les mots qui sont dans Athalie, même acte et même scène que j'ai dit tout à l'heure, Et tâche de m'remettre une copie propre sinon j'te passe une torgniole. Quand tu seras grande, tu m'remercieras pour tout c'que j'ai fait pour toi ».

— Je lui ai rendu une copie blanche, toute propre et j'ai ramassé quand même ma torgniole. C'est pas juste!

Qu'en pensez-vous?

### 2.10.4. *Kséduresésplike*

Queiqu'un vous demande d'effectuer une série de manipulations que nous résumons par la formule :

$$\overline{A \cap \overline{B}} \cap (\overline{A \cap B}) \cup (A \cap B) \cap (A \cap B)$$

où A et B sont les ensembles de mots graphiques déterminés dans « Le cancre ».

Imaginez une conversation en langue naturelle entre ce quelqu'un et vous-même où il vous présente sa demande et où vous lui suggérez des manipulations beaucoup plus simples permettant d'obtenir les mêmes résultats.

(\*) On ne saurait accorder trop d'importance à ce petit calcul de grossière estimation, fondé sur le bon sens et des hypothèses simplistes, que tout directeur de laboratoire est constamment obligé de pratiquer afin de planifier le travail.

(\*\*) La « traduction » de l'énoncé en langage ensembliste fournit aisément la solution. Attention/Cet énoncé est, à certains endroits, volontairement et malicieusement ambigu. La critique de l'énoncé fait partie du problème, c'est un moyen de forcer l'élève à garder son esprit en éveil, à ne pas accepter passivement ce qu'on lui raconte.

### 2.10.5. *La dactylo est distraite.*

Un centre de recherches a mis sur cartes perforées tous les mots graphiques d'« Andromaque », une carte par occurrence-mot, d'où un ensemble A de cartes.

Les mots graphiques ont été ensuite classés par ordre alphabétique de façon purement mécanique et on n'a conservé qu'une seule carte par mot graphique distinct. (Dans ce classement, des formes comme *chantons*, *chantez* sont considérées comme distinctes tandis que *vers* (préposition) et *vers* (subst. pluriel) comme identiques).

La même opération est faite pour « Iphigénie » et « Phèdre », il en résulte 3 jeux de cartes A, I et P.

Cependant à la suite de quelques distractions de la perforatrice, l'orthographe ayant été parfois légèrement altérée, des mots graphiques identiques peuvent apparaître comme distincts et vice-versa.

Prévoir toutes les anomalies qui en résultent pour A, et ensuite lorsqu'on veut fusionner les jeux A et B, puis A, B et C, pour les jeux X et Y qui en résultent.

Pouvez-vous donner un ordre de grandeur de l'importance des diverses anomalies sachant que, par exemple, « Phèdre » contient 14 217 occurrences dont 2 961 formes différentes et en admettant que la perforatrice distraite se trompe environ une fois sur 100 cartes?

### 2.10.6. *Polysémie grammaticale.*

Effectuer une partition de l'ensemble E, ci-dessous, en fonction de la nature de la polysémie grammaticale des éléments-mots qui le constituent.

E = { *bois, cache, lit, part, partie, juste, pris, croissant, bleu, cardinal, levant, bave, avant, jaune, coucher, lave, cuit, vue, ris, caille, dîner, mousse, rôti, pas, dément, coulant, chauffe, tris, lancer, meuble, devant, voile, porte, brosse, chasse, érudit, noirs, robe, pu, causes, dérouté, violette, griffe, vieux* }.

Exemple de polysémie grammaticale :

*bois* peut être : substantif, verbe 1<sup>re</sup> ou 2<sup>e</sup> personne de l'indicatif.

*voile* peut être : substantif masculin, substantif féminin, verbe 1<sup>re</sup> ou 3<sup>e</sup> personne de l'indicatif ou du subjonctif.

## 3. Structure syntaxique.

3.1. *Réseaux.* Sans désirer innover quoi que ce soit en matière de syntaxe (\*), on peut simplement se poser comme objectif de rendre d'une certaine façon « visibles », à l'aide de schémas appropriés, des relations grammaticales

(\*) Le but de cet article n'est pas d'apporter des solutions nouvelles dans le domaine de la linguistique mathématique, mais de montrer que des résultats élaborés par d'éminents théoriciens peuvent être adaptés à des fins pédagogiques modestes.

bien connues telles que : la relation sujet-verbe, verbe-complément, substantif-adjectif, article-substantif, préposition-substantif... Il n'est pas absurde, dans une première approche, de faire abstraction de toute hiérarchie (gouverneur, gouverné) en adoptant comme unités de travail les occurrences-mots. Figurons celles-ci par des points, placés n'importe où dans le plan, et les relations syntaxiques par des lignes joignant ces points.

En procédant ainsi avec la phrase : *il dit non avec la tête*, certains aboutiront au réseau (1) d'autres au réseau (2) mais personne ne proposera le réseau (3).

Pour fixer les idées, adoptons en accord avec certains linguistes, le premier point de vue.

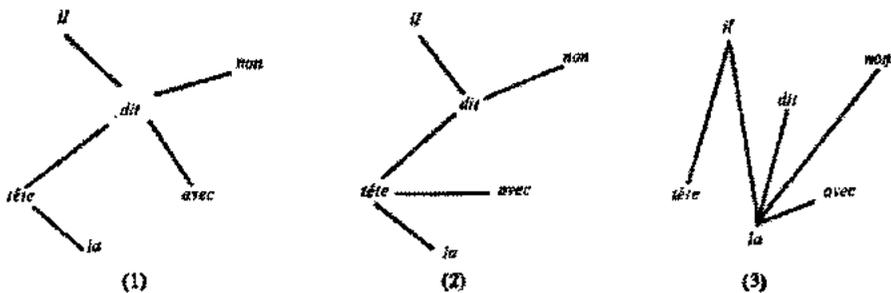


FIG. 5. — Effacer la liaison *dit-tête*, la remplacer par *tête-avec*.

Remarquons que ce réseau se présente sous l'aspect formel d'un « arbre » avec une « cime » et des « branches ». Imaginons ce réseau construit de tiges articulées, pour faire ressortir la structure d'arbre il suffit de le soulever par un quelconque de ses sommets, d'où 6 possibilités. Voir schéma ci-dessous.

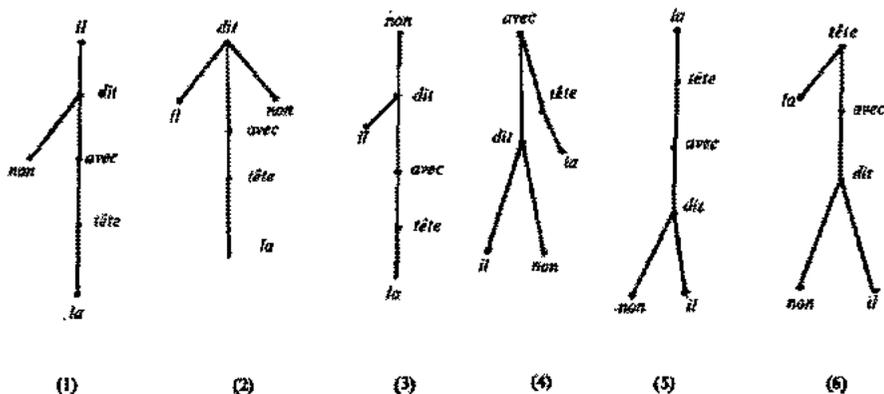


FIG. 6.

Ceci étant on peut tenter de poursuivre notre ébauche par une description plus riche en information en faisant successivement diverses conventions.

1. Relation d'ordre. Admettons (ce qui parfois est contestable) que toute paire reliée du réseau est orientable et devient un couple constitué d'un élément-mot gouverneur et d'un gouverné. Les solutions proposées par diverses écoles linguistiques ne seront peut être pas rigoureusement identiques, adoptons en une, par exemple la 2<sup>e</sup>.

Nous obtenons une arborescence (arbre orienté à partir de la cime vers les extrémités des branches),

Si de plus on tient à rappeler la structure d'ordre total de la chaîne parlée elle-même, on obtient une arborescence avec sa projection sur l'axe des temps.

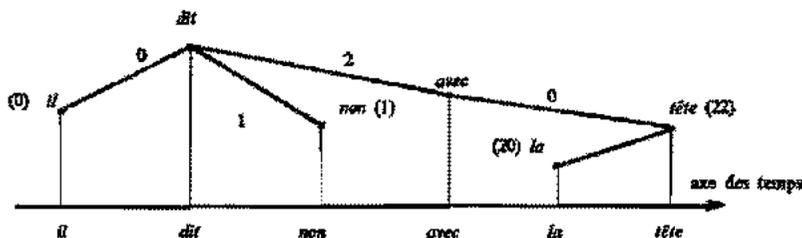


FIG. 7.

Un problème d'adressage se pose. Comment désigner un sommet déterminé sans citer le mot qui l'accompagne. Il existe une solution classique de repérage, à chaque arc on attribue un numéro d'ordre à partir de la gauche, chaque mot (cime exclus) se trouve doté d'une adresse non équivoque. On peut en imaginer d'autres et établir des correspondances.

Le réseau ci-dessus possède la propriété importante : les projectantes verticales ne coupent pas les arcs de liaison. On dira que la phrase est projective. La projectivité n'est d'ailleurs pas une propriété universelle. Exemple : Voir schéma ci-dessous.

L'inversion du complément de nom de *humour* en est responsable.

Remarquons également qu'une même suite de mots peut parfois être analysée de différentes façons, selon le sens qu'on attribue à la phrase. *Le petit*

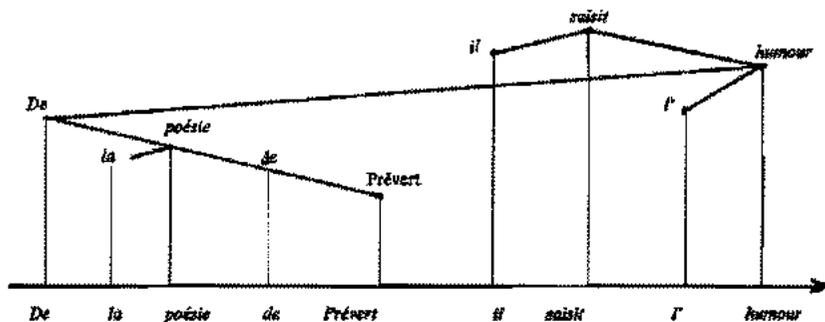


FIG. 8.

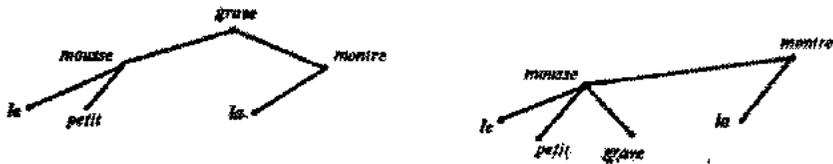


FIG. 9.

*mousse grave la montre.* S'agit-il d'une montre qu'on est en train de graver ou de l'attitude pleine de gravité d'un mousse qui est en train de montrer quelque chose de connu à son interlocuteur.

Il en résulte 2 schémas structurels distincts : voir schémas ci-dessus.

Il ne semble pas que de telles ambiguïtés d'interprétation se présentent dans le texte de Prévert,

D'une façon générale si une machine avait à déterminer la nature des parties du discours (adjectif, substantif, verbe, article, pronom...) des divers mots de la phrase, elle hésiterait entre de nombreuses possibilités qu'on obtient en suivant les trajets de gauche à droite du réseau : voir schéma ci-dessous.

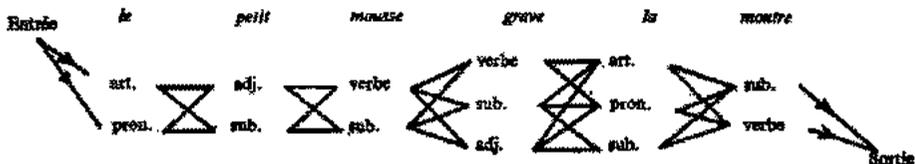


FIG. 10.

Soit  $2^4 \times 3^3 = 144$  solutions a priori, dont 2 apparemment s'avèrent correctes.

Des conditions supplémentaires peuvent réduire ce nombre, par exemple, on peut exiger que la phrase possède 1 verbe, qu'un article ou un adjectif accompagnent nécessairement un substantif (sauf dans des cas spéciaux), que le pronom *la* précède le verbe... ce qui conduit à des problèmes de dénombrement parfois laborieux. Voir schéma ci-dessus.

### 3.2. « Bourgeoisement » et « élagage » des arborescences.

Partons de la phrase simple :

- 1) *Le cancre dit non au professeur.*

On peut procéder aux expansions suivantes :

- 2) *Le cancre dit toujours non au professeur.*
- 3) *Le cancre dit toujours non au professeur inhumain.*

- 4) *Le cancre paresseux dit toujours non au professeur inhumain.*  
 5) *Le cancre très paresseux dit toujours non au professeur inhumain.*

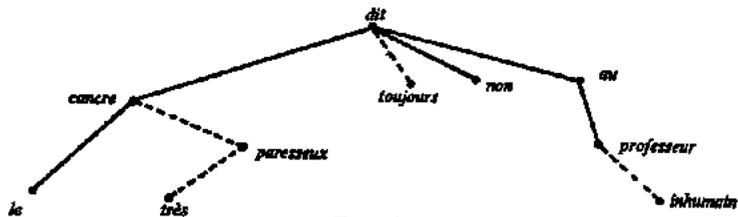


FIG. 11.

De nouveaux rameaux bourgeonnent sur les arborescences. Inversement on peut chercher à élaguer une arborescence touffue de façon à la réduire au strict minimum, linguistiquement acceptable. Voir schéma ci-dessus.

Il n'y a pas qu'une seule façon de faire bourgeonner ou d'élaguer des arborescences pour passer d'une extrême à l'autre. Cependant on ne peut pas opérer n'importe comment. *Très* doit obligatoirement être effacé avant *paresseux*, certaines séquences de mots constituent des phrases, d'autres non.

On peut étudier les suites de phrases intermédiaires possibles, d'où des problèmes de dénombrement. Ainsi, dans cet exemple, on peut commencer par effacer ou *très* ou *toujours* ou *inhumain*.

Deux cas se présentent ensuite : *très* a été effacé, il reste *paresseux*, *toujours*, *inhumain*; *très* n'a pas été effacé il reste *très*, *toujours* ou *très*, *inhumain*, etc.

Si on désigne par *a*, *b*, *c*, *d* les 4 mots *paresseux*, *très*, *toujours*, *inhumain*, la succession des opérations se laisse figurer à l'aide d'un réseau que l'on « traverse » soit dans un sens soit dans l'autre. Voir schéma ci-dessous.

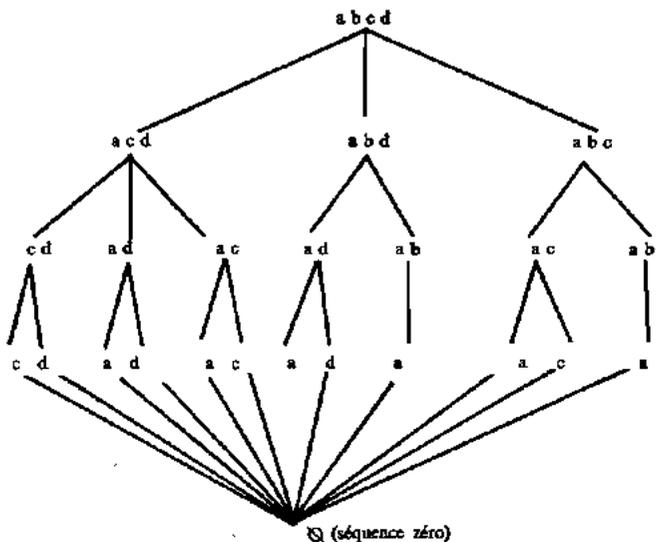


FIG. 12.

Soit 12 solutions.

On peut chercher à réduire le schéma en confondant les sommets de même nom, on parvient, ici, à un réseau planaire :

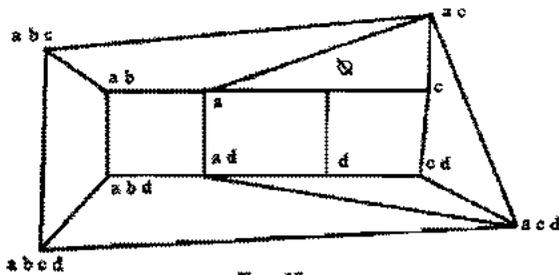


FIG. 13.

Voici une recherche légèrement plus compliquée. Soit :

$\varphi_1$  = notre cancre pas particulièrement méchant oppose cependant un non catégorique au professeur qui l'interroge de façon inhumaine.

$\varphi_2$  = ce cancre, la honte de la classe, dit non, sous la huée des autres élèves, au professeur pourtant pas très exigeant.

Comment passer par réduction-expansion de  $\varphi_1$  à  $\varphi_2$  (sans introduire de mots nouveaux)? Donner au moins une solution concrète. Combien existe-t-il de solutions linguistiquement acceptables, c'est-à-dire toutes les étapes intermédiaires étant des phrases correctes?

### 3.3. Relations et matrices.

Examinons la situation du point de vue de la notion ensembliste de relation. L'ensemble {il, dit, non, avec, la, tête} est structuré par une relation binaire, d'ordre strict (on ne s'occupe pas des boucles sur le réseau) que l'on peut présenter comme une relation de gouvernement direct, notons-la par «  $\succ$  », soit direct ou indirect, notée «  $>$  ».

dit $\succ$ il	avec $\succ$ tête	dit $>$ la
dit $\succ$ non	tête $\succ$ la	dit $>$ tête
dit $\succ$ avec		avec $>$ la

Ces relations admettent une représentation matricielle (lire les tableaux de gauche à droite).

$\succ$	il	dit	non	avec	la	tête
il	0	0	0	0	0	0
dit	1	0	1	1	0	0
non	0	0	0	0	0	0
avec	0	0	0	0	0	1
la	0	0	0	0	0	0
tête	0	0	0	0	1	0

$>$	il	dit	non	avec	la	tête
il	0	0	0	0	0	0
dit	1	0	1	1	1	1
non	0	0	0	0	0	0
avec	0	0	0	0	1	1
la	0	0	0	0	0	0
tête	0	0	0	0	1	0

On peut en déduire d'autres relations : les relations réciproques (est gouverné par, dépend de), complémentaires (ne domine pas, ne dépend pas de), domine ou est dominé... conduisant à d'autres matrices puis on peut rechercher les règles formelles permettant de passer d'une matrice à une autre.

dépend directement de	dépend de	ne domine pas directement	ne dépend pas directement de
$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}$

Il est facile d'imaginer des sous-relations telles que « il existe un accord grammatical » : *il-dit, la-tête*, « il y a relation sans accord » *dit-non, avec-tête*, etc.

### 3.4. Structure parenthétique.

Il existe une autre façon de voir la structure d'une phrase en la considérant comme faite de séquences bien formées (syntagmes) qui s'emboîtent les unes dans les autres. Là encore les points de vue divergent, certains préféreront le 1<sup>er</sup> réseau, d'autres peut-être le 2<sup>e</sup>, mais personne n'aura idée d'insister sur le 3<sup>e</sup>.

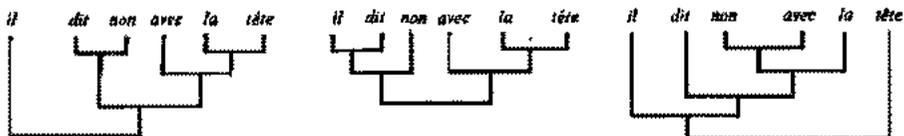


FIG. 14.

Adoptons, par exemple, la 1<sup>re</sup>. Les mêmes choses peuvent être dites à l'aide de parenthèses :

(il ((dit non) (avec (la tête))))

et comme chaque couple de parenthèses associées ne contient que deux objets, on peut faire l'économie de la 2<sup>e</sup> et écrire en notations polonaises (\*) :

il|dit non|avec|la tête

(\*) Dans la notation polonaise le symbole opératoire précède les objets sur lesquels on opère. Ici l'opération binaire consiste à considérer les deux mots (ou suites de mots) successifs comme formant un tout. Une seule parenthèse suffit.

ou au contraire, en faisant l'économie de la 1<sup>re</sup> parenthèse :

*il dit non|avec la tête|||*

Là encore divers problèmes formels peuvent être proposés. Retrouver les parenthèses fermantes. Tout découpage conduit-il à une arborescence? Combien peut-on en imaginer de différents? Ou au contraire des problèmes d'imagination : on donne un découpage possible, inventer une phrase qui lui convienne.

Faire une partition d'un ensemble de phrases d'après le découpage associé, etc.

Rien n'oblige à se restreindre au binarisme et d'aucuns préfèrent ne pas prendre position sur certains parenthésages.

*(il dit non (avec (la tête)))*

D'où une série d'autres problèmes.

On peut rechercher des procédures purement formelles permettant de passer d'un genre d'arborescence à un autre et préciser quelles sont les informations nécessaires pour y parvenir.

### 3.5. Raccordement de réseaux.

Deux phrases sont susceptibles parfois d'être combinées en une seule.

et  $\left\{ \begin{array}{l} \text{il dit oui avec le cœur} \\ \text{il dit non avec la tête} \end{array} \right. \rightarrow \text{il dit oui avec le cœur et non avec la tête}$   
 puis  $\left\{ \begin{array}{l} \text{il écrit les chiffres} \\ \text{il efface les chiffres} \end{array} \right. \rightarrow \text{il écrit puis il efface les chiffres}$

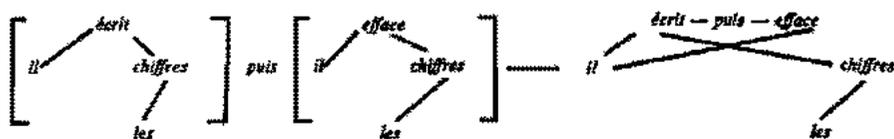


FIG. 15.

Ajouter la liaison (*efface, chiffres*) : mettre une flèche au milieu, entre les deux parties du schéma.

Cette fois-ci le résultat n'est plus une arborescence.

Quelles sont les opérations correspondantes sur les matrices?

#### 4. Déclinaisons. Réseaux, algèbre, information numérique.

##### 4.1. Problème.

Présenter sous forme de réseaux le paradigme de la déclinaison latine de *rosa*.

a) On suppose connues les variables grammaticales :  
 — singulier S, pluriel P, nominatif N, vocatif V,  
 — accusatif  $A_s$ , génitif G, datif D, ablatif  $A_p$ , et on veut découvrir la désinence correspondante (grammaire générative de formes graphiques ou point de vue de l'élève qui fait un thème latin).

b) On suppose connues les formes graphiques et on cherche à déterminer les variables correspondantes (grammaire de reconnaissance ou grammaire de l'élève qui fait une version latine).

##### Solution partielle.

###### a) Thème

On peut concevoir de nombreux réseaux selon l'ordre dans lequel on présente les données. Nous retenons :

Réseau : (1) on choisit d'abord le nombre (S ∨ P, singulier ou pluriel) et ensuite le cas (N ∨ V ∨  $A_s$  ∨ G ∨ D ∨  $A_p$ ).

Réseau : (2) on choisit d'abord le nombre et ensuite le cas.

La solution est chaque fois unique, il s'agit d'une surjection d'un ensemble de paires {nombre, cas} sur un ensemble de 6 formes.

{*rosa, rosae, rosam, rosarum, rosas, rosis*}

###### b) Version

Les formes étant classées, par exemple par ordre alphabétique, quelles sont les paires susceptibles de leur être associées. Nous emploierons une notation sténographique, inspirée de la logique formelle :

##### Réseaux de déclinaison.

Grammaire de celui qui écrit :



En outre, certaines associations donnent soit 0 soit 1.

$S \wedge P = N \wedge V = N \wedge A_c = \dots = 0$  (on admet qu'aucun mot ne peut être à la fois singulier et pluriel dans un texte donné).

$S \vee P = N \vee V \vee A_c \vee G \vee D \vee A_p = 1$  (aucune information n'est fournie par un tel message dans le cadre que nous nous sommes posé).

$V \wedge 0 = 0, \quad V \wedge 1 = V.$

#### 4.2. Calcul sur l'information.

Les six cas du latin, au singulier ou au pluriel, conduisent à 12 valeurs grammaticales possibles. Se donner une terminaison diminue l'indétermination I évaluée comme le logarithme de base 2 du nombre de cas possibles :

$$I = \log_2 12 = 3,58$$

*ae* réduit l'indétermination à 4 valeurs

$$I_{ae} = \log_2 4 = 2$$

*arum* supprime toute indétermination

$$I_{arum} = \log_2 1 = 0.$$

L'information J apportée par une terminaison est calculée comme une variation d'indétermination.

D'où les mesures de l'information :

$$J_{ae} = \log 12 - \log 4 = 1,58$$

$$J_{arum} = \log 12 - \log 1 = 3,58.$$

Autres exercices : quelle est l'information apportée par la pénultième, quelle est l'information apportée par la dernière lettre connaissant déjà la pénultième? etc.

On parle beaucoup de la synthèse des enseignements divers; ceci constitue un exemple profitable et au professeur de latin et au professeur de mathématique.

## 5. Opérations.

Voici un exemple classique de composition des personnes grammaticales (1<sup>re</sup>, 2<sup>e</sup>, 3<sup>e</sup> pers.).

Considérons la phrase :

*toi et moi, nous sommes des calfs*

$$2 * 1 = 1$$

On voit que la 2<sup>e</sup> et la 1<sup>re</sup> pers. conduisent à une 1<sup>re</sup> pers.

De même :

<i>vous et moi, nous sommes des kraks</i>	$2 * 1 = 1$
<i>toi et toi, vous ferez les pluches</i>	$2 * 2 = 2$
<i>toi et lui, vous êtes des dégonflés</i>	$2 * 3 = 2$
<i>elle et lui, ils se disputent toujours</i>	$3 * 3 = 3$
<i>lui et moi, nous ne sommes jamais d'accord</i>	$3 * 1 = 1$
<i>moi et moi, nous ne nous ennuyons jamais</i>	$1 * 1 = 1$

(dira Sacha Guitry)

etc.

Ce que l'on résume dans la table de Pythagore

*	1	2	3
1	1	1	1
2	1	2	2
3	1	2	3

L'opération « \* » jouit de la commutativité ( $a * b = b * a$ ), de l'associativité  $(a * b) * c = a * (b * c)$ .

Voici un autre exemple :

Trois personnages sont en scène : Aglaé, Brigitte, Camomille. C'est Aglaé qui parle à Brigitte de ce qu'elle a dit ou de ce que Brigitte ou Camomille ont pu lui dire :

I. Discours direct :

*Je dis: « tu es une sottie »*  
*Tu dis: « je suis une sottie »*  
*Elle dit: « je suis une sottie ».*

II. Discours indirect :

*Je dis que tu es une sottie*  
*Tu dis que je suis une sottie*  
*Elle dit que je suis une sottie.*

Qui est sottie dans l'histoire? Est-ce 1 (Aglaé), est-ce 2 (Brigitte qui écoute ou qui a parlé auparavant), est-ce 3 (Camomille qui n'est pas là mais qui a pu parler).

	1	2	3		1	2	3
1	1	2	3	1	1	2	3
2	2	1	3	2	1	2	3
3	3	1	2	3	1	2	3

Les deux tables ne sont pas pareilles.

Étudier la loi de composition de  $m$  (masculin),  $f$  (féminin),  $s$  (singulier),  $p$  (pluriel), dans :

*Un chat et un chat ça fait deux chats*  
*Un chat et une chatte ça fait deux chats*  
*Une chatte et une chatte ça fait deux chattes*  
*Deux chats et trois chattes ça fait cinq chats, etc. (\*)*

ou celle de :

*fautueil f, chaise c, tabouret t, siège s, meuble m, armoire a.*

On a par exemple les égalités symboliques

$f+f = f$  (*des fauteuils et des fauteuils donnent des fauteuils*)

$f+c = s = f+t = c+t$

$a+c = m = a+s = a+m$

	a	c	m	s	t
a	a	m	m	m	m
c	m	c	m	s	s
m	m	m	m	m	m
s	m	s	m	s	s
t	m	s	m	s	t

## 6. Modèles « à états finis ».

Soit à noter de la façon la plus « économique » possible les mots de la liste :

$L = \{\text{détermination, interminable, déterminant, prédéterminé, déterminatif, déterminisme, terminé, déterministe, déterministique}\}.$

Par « économique » entendons : éviter d'écrire plusieurs fois les mêmes segments de mots. On peut hésiter entre beaucoup de solutions. Le procédé

(\*) Ceci ne marche pas avec *monsieur* et *dame* car on ne peut pas dire : *un monsieur et une dame ça fait deux monsieurs*, c'est donc bien un fait de langue.

traditionnellement utilisé est de remplacer par des guillemets, les morceaux qu'on se dispense de récrire, par exemple :

	<i>termin-é</i>	
dé	'' ''	
pré-	'' ''	
''	'' -ant	
''	'' a-tif	
''	'' n-tion	
''	'' -is-me	
''	'' -te	
''	'' -tique	
in	'' -able	

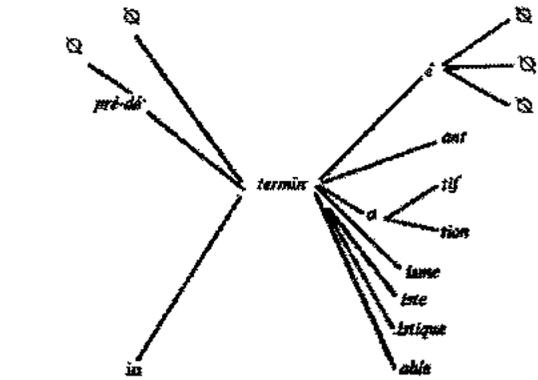


FIG. 17.

Réseau (1)

La même économie (75 lettres) peut être réalisée à l'aide d'un réseau. Il est possible d'imaginer beaucoup d'autres réseaux analogues, toutefois en économisant les lettres on dépense des arcs.

Inversement le réseau le plus économique en arcs est constitué uniquement de points isolés figurant la liste donnée elle-même, il ne fait économie d'aucune lettre.

Remarquons que le réseau (1) peut être parcouru de gauche à droite sans que le départ et l'arrivée aient lieu sur la même ligne horizontale. A ce moment le nombre de mots obtenus augmente, de 10 on passe à  $4 \times 8 = 32$  mots, dont certains sont consignés dans les dictionnaires courants (*indétermination, déterminable...*), d'autres ne figurent que dans des ouvrages spécialisés (*prédéterminatif, termination*) d'autres peut-être, sans qu'il soit aisé de dire lesquels, n'ont jamais été employés dans aucun texte sérieux. Notons que certains néologismes pris hors contexte nous paraissent choquants mais passent inaperçus dans des contextes bien choisis :

« Cette tâche colossale ne peut être considérée comme terminable dans des délais aussi stricts ».

*Déterministique* est usité en philosophie et dans la théorie des algorithmes, pourquoi pas *indéterministique* ou *prédéterministique*.

Aussi à côté d'un vocabulaire consacré par l'Académie ou par d'autres autorités accréditées, on devine l'existence d'un glossaire de termes qui n'ont qu'une existence en quelque sorte « potentielle » — qui par exemple, n'ont eu qu'une vie éphémère ou qui sont susceptibles d'apparaître un jour, lorsque le besoin s'en fera sentir, qui seront compris par un locuteur français et qui peut-être seront retenus pour l'usage.

Voici une phrase recueillie dans une conversation entre jeunes étudiants :  
*Il se coupe, se beurre et se confiture une tartine. Le verbe confiturer fait partie*

de ce vocabulaire potentiel. On conçoit, dans ces conditions, que parler de « l'ensemble des mots de la langue française » constitue un abus, dans un ensemble on doit pouvoir décider, en principe, si un élément fait ou ne fait pas partie de cet ensemble. Or, il s'avère que les « mots », apparemment les plus nombreux, ne possèdent justement pas cette propriété. Tout ce que l'on est en droit de dire, si l'on veut être rigoureux (et encore!) c'est « un ensemble des mots de tel dictionnaire, de tel ouvrage ».

Le réseau (1) peut-être modifié comme suit :

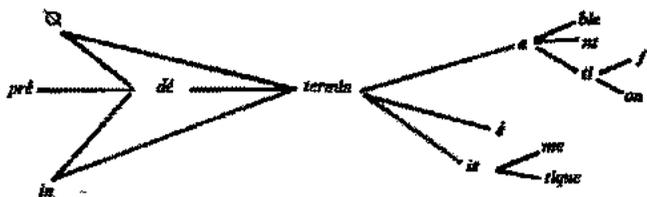
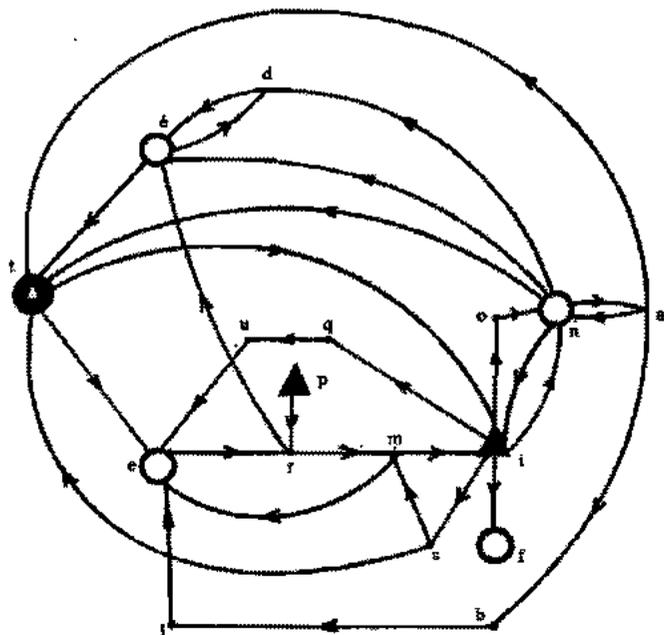


FIG. 18.

Réseau (2)



- Stop
- ▲ S'engager avec prudence
- Stop-départ

FIG. 19. Ajouter un triangle noir en d.

Réseau (3)

Il offre  $5 \times 8 = 40$  trajets de gauche à droite, cependant certains découpages de mots paraissent tout à fait contestables du point de vue de la grammaire traditionnelle (termin-a-nt).

On peut aller plus loin et faire en sorte qu'aucune lettre ne soit écrite deux fois. Sur le schéma, les triangles indiquent des points de départ possibles et les ronds des stops possibles pour les trajets considérés (on aurait pu aussi bien réunir tous les triangles à un point dit d'entrée du réseau et tous les ronds à un point dit de sortie).

Parmi les produits de ce genre de mécanisme on retrouve les vocabulaires, traditionnel et potentiel, formés sur le radical *-term-* mais en plus des mots attestés qui n'ont rien à voir avec le radical *-term-* exemple : *ion, if, dédédé, tique, ...* des suites de lettres que l'on ne sait rattacher à rien : *ismion, ininable, déttiste...* mais qui conservent encore une « petite allure française », puis des monstres à peine prononçables pour un locuteur français : *prmistermintinint* et même des suites infinies : *istermistermistermi...* *stermin* correspondant à des cycles du réseau.

À propos de mots, ayant comme une petite allure française et qui, à la rigueur, peuvent être compris on peut citer cet extrait d'un poète contemporain qui cherche à s'exprimer justement dans un langage qui n'est pas du français mais qui néanmoins suggère quelque chose à un français.

#### BABELLOSTE

A MONZAMI A. TOM. Ce premier jodi du Florimois aux Vinchènes en Panamie, à la Pacificavia, en l'An 0 du Mirivis.

Potami A. Tom,

Des zozidissent queuté un mochant. Cépapossib.

Cétipatoi quiè dans les planures à l'infond des vervefouts alpagés?

Cétipatoi qui colormignote dans les blus des bluets, chandèle dans les candilya, lyrodore dans les séduisances des lyrias en Vallonarève? Alorquoi!

Zon surpose queuté dun cractère rudoyant. Cépavrai. Tèzum caresseur aux charmuses des trillagorges zozibosques. Tu serpentines en aquavive. Cédutoi qui cièle dans les fiéroments. Finquoi tocequè danla Réals du Terréciel, cédutoi, toutautant aux articulets dune pattemouche quala queudunchinge.

...

E pouah! porieu chocotétatom!

Surça, Caripanpote, récepte les sentimieus calorés d'admire. Moitatoi cardialement.

André MARTEL,  
dit le Martélandre  
Papapafol du Paralloïdre.

Ainsi l'étude des productions d'un réseau rejoint dans une certaine mesure certaines préoccupations d'ordre littéraire.

Le classement de toutes les productions d'un réseau n'est pas toujours chose aisée et on peut bâtir de nombreux exercices sur ce thème.

Voici quelques énoncés :

Appelons « rendement d'un réseau relativement à un dictionnaire » le rapport entre le nombre de mots produits par le réseau et figurant dans ce dictionnaire au nombre total de mots produits. Étudiez les rendements des

réseaux (1) et (2) relativement au nouveau petit Larousse, au Quillet, au dictionnaire de l'Académie, au dictionnaire encyclopédique Larousse, etc.

Peut-on obtenir le réseau (1) à partir d'une sous-liste de L? En d'autres termes, la liste L est-elle redondante?

Peut-on obtenir le réseau (1) avec une liste différente de L. Pour un référentiel fixé de mots formés des mêmes préfixes, radical et suffixes, combien peut-on imaginer de listes, analogues à L, conduisant au réseau (1)?

Existe-t-il une ou plusieurs listes minimums?

De la liste L on peut extraire 2 sous-listes de façon que le raccordement des réseaux associés redonne le réseau (1). Étudiez ce problème.

Il est clair que ces questions sont susceptibles d'être exprimées en termes de la théorie des automates et conduisent tantôt à des résultats classiques tantôt à des problèmes dont la solution générale est loin d'être évidente. Il ne peut être question au niveau du lycée de rentrer dans le détail de ces théories, cependant maints problèmes peuvent être posés en termes naïfs, sur des ensembles suffisamment réduits pour qu'on puisse mettre en œuvre des raisonnements simples ne nécessitant aucune érudition mathématique.

## 7. Mots et contextes (phrases à trou).

7.1. Soit l'ensemble des pronoms dits personnels :

$P = \{ \text{je, tu, il, elle, nous, vous, ils, elles, me, te, se, le, la, les, leur, moi, toi, soi, lui, eux} \}$

et l'ensemble des contextes C :

- a = Azor... mordit cruellement.
- b = Bernard... raconte des histoires drôles.
- c = Casimir... défendra contre Azor.
- d = Denise, regarde... bien en face, pas nous.
- e = Ernestine, ..., n'aime pas les maths.
- f = François... bat à la pétanque.
- g = Gabriel, ..., c'est un ange.
- h = Hector ferait n'importe quoi pour ....
- i = Isidore, que faites ...?
- j = ... tire ma révérence.
- k = Kate ... dit une sottise.
- l = Lucie ... griffa au visage.
- m = Monique ... donna une gifte.
- n = Nicolas, ..., n'aime que le vin.
- o = Odile, ... je vous dis tout ce que je pense.
- p = Pierre ... nettoie tous les jours.

Effectuer la partition des pronoms P, d'après les divers contextes où ils peuvent se placer, et une partition des contextes C, d'après les divers pronoms qu'ils peuvent recevoir, sans que les phrases complètes qui en résultent chagrinent le bon usage (tenir compte de la ponctuation!, négliger les traits d'union).

Formons le tableau à double entrée : pronoms personnels, contextes. A la croisée d'une ligne et d'une colonne marquons 1 si le contexte accepte le pronom, 0 dans le cas contraire.

pr. pers.	cont.															
	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
— je .....										1						
— tu .....																
— il .....																
— elle .....					1			1								
— nous .....	1	1	1			1		1			1	1	1			1
— vous .....	1	1	1			1		1	1		1	1	1			1
— ils .....																
— elles .....								1								
— me .....	1	1	1			1					1	1	1			1
— te .....	1	1	1			1					1	1	1			1
— se .....	1	1	1			1					1	1	1			1
— le .....	1		1	1		1						1				1
— la .....	1		1	1		1						1				1
— les .....	1		1	1		1						1				1
— leur .....		1									1		1			
— moi .....				1				1								1
— toi .....				1				1								
— soi .....								1								
— lui .....		1						1	1		1		1	1		
— eux .....								1								

D'où la partition de l'ensemble C des contextes (colonnes identiques), {a, c, f, l, p}, {b, k, m}, {g, n}, {d}, {e}, {h}, {i}, {j}, {o}, et la partition  $\Pi$  de l'ensemble P des pronoms (lignes identiques).

{nous, me, te, se}, {le, la, les}, {tu, il, ils}, {elles, eux}, {vous}, {elle}, {leur}, {moi}, {toi}, {soi}, {lui}, {je}.

Il est clair qu'avec d'autres contextes la partition des pronoms eut été différente et vice-versa. Bien plus, avec un choix de contexte ad hoc chaque classe pourrait être réduite à un seul élément. Mais il n'est pas certain que n'importe quelle partition puisse être réalisée en choisissant de façon astucieuse les contextes.

Partant de là on peut construire un exercice d'imagination : étant donné une partition que l'on sait être réalisable, inventer les contextes appropriés et vice-versa.

## 7.2. Partitions dérivées.

Considérons la partition  $\Pi'$  des pronoms personnels, que l'on rencontre dans certaines grammaires, d'après la personne et le nombre :

{je, me, moi}, {tu, te, toi}, {il, elle, le, la, lui, se, soi}  
 {nous}, {vous}, {ils, elles, leur, eux, les}.

A partir des deux partitions  $\Pi$  et  $\Pi'$  on peut en déduire deux autres :

La première, dite partition croisée,  $\Pi \wedge \Pi'$ , est plus fine que chacune des composantes  $\Pi$  et  $\Pi'$ , elle retient comme principes de classement tous les arguments qui ont servi de critère pour les deux composantes (distribution dans divers contextes et valeurs grammaticales). Les nouvelles classes de pronoms seront contenues (sans chevauchement) et dans les classes de  $\Pi$  et dans celles de  $\Pi'$ . Plus précisément, de toutes les partitions plus fines que  $\Pi$  et  $\Pi'$  c'est celle qui l'est le moins.

Pour l'obtenir formons le tableau à double entrée. Les lignes en seront définies par les classes de l'une des partitions, les colonnes par celles de l'autre. A la croisée d'une ligne et d'une colonne on retiendra le pronom qui figure à la fois et dans la ligne et dans la colonne. Il est clair que si  $\Pi$  détermine  $p$  classes et  $\Pi'$   $p'$  classes,  $\Pi \wedge \Pi'$  en déterminera un nombre au plus égal à  $p \times p'$ . Voir tableau page suivante.

La deuxième partition notée  $\Pi \vee \Pi'$  est moins fine que chacune des composantes  $\Pi$  et  $\Pi'$ . Les nouvelles classes de pronoms devront être suffisamment étendues pour contenir (sans chevauchement) les classes des deux composantes. Plus précisément de toutes les classes moins fines que  $\Pi$  et  $\Pi'$  c'est celle qui est la plus fine.

	je me moi	tu te toi	il elle le la lui se soi	nous	vous	ils elles leur eux les		
nous .....	me	te	se	nous				
me .....								
te .....								
se .....								
le .....			le			les		
la .....			la					
les .....								
tu .....		tu	il			ils		
il .....								
ils .....								
elles .....					vous	elles		
eux .....						eux		
vers .....								
elle .....			elle					
leur .....						leur		
moi .....	moi							
toi .....		toi						
soi .....			soi					
lui .....			lui					
je .....	je							

Sur les 18 classes non vides, 16 ne contiennent qu'un seul élément, les 2 autres, 2 éléments chacune : {le, la} et {elles, eux}.

Le tableau précédent permet de l'obtenir. Du moment qu'une ligne et une colonne se croisent en une case non vide, tous les pronoms de la ligne et de la colonne doivent être groupés dans une même classe.

Ici on obtient seulement les 2 classes :

{il, elle, le, la, lui, se, soi, nous, me, te, les, ils, moi, je, tu, toi, elles, leur, eux} et {vous}.

Les pronoms sont rangés dans la même classe soit parce qu'ils peuvent appartenir aux mêmes contextes soit parce qu'ils possèdent des mêmes valeurs grammaticales de personne et de nombre.

### 7.3. Substituabilité.

Reprenons la partition II et le tableau qui nous a permis de l'obtenir. On remarque que tout contexte qui peut contenir *nous*, peut contenir également *me*, *nous* et *me* seront dits « intercommutables ». Il n'en est pas de même pour *nous* et *le*, on ne peut remplacer *nous* par *le* et vice-versa que dans certains cas, pas dans tous. La situation de *nous* et *vous* est encore différente. Partout où figure *vous* on peut mettre *nous* mais non l'inverse, nous dirons que *vous* « domine » sur *nous*.

Ces remarques établissent dans l'ensemble P des pronoms une relation de « préordre ».

Il y a des pronoms hiérarchisables :  $nous > vous$ .

Il y a des pronoms non comparables :  $nous \parallel le$ .

Il y a des pronoms équivalents :  $nous \sim me$ .

Si à chaque pronom on attache un nombre binaire, suite de 0 et de 1 de la ligne correspondante

$$\begin{aligned} [nous] &= 1110010100111001 \\ [vous] &= 1110010110111001 \\ [le] &= 1011010000010001 \\ [me] &= 11100100000111001 \end{aligned}$$

On peut convenir d'une sorte de « soustraction » des nombres binaires : un « 0 » peut se retrancher de « 1 » mais non l'inverse; d'où l'automatisation du processus de comparaison :

$[nous] \underline{*} [me] = 0000000000000000$  : équivalence

$[vous] \underline{*} [nous] = 0000000010000000$  : dominance

$[nous] \underline{*} [le]$  et  $[le] \underline{*} [nous]$  sont impossibles : incomparabilité

Pour obtenir la hiérarchisation de façon systématique il est commode d'abord de réduire le tableau en ne gardant de chaque classe qu'un seul élément représentatif dit « élément distingué ».

	a	b	d	e	g	h	i	j	o
<i>je</i> .....								1	
<i>tu</i> .....				1		1			
<i>elle</i> .....	1	1				1			
<i>nous</i> .....	1	1				1	1		
<i>vous</i> .....			1			1			
<i>elles</i> .....				1		1			
<i>le</i> .....						1			
<i>leur</i> .....						1			
<i>moi</i> .....						1			1
<i>toi</i> .....						1			
<i>soi</i> .....						1			
<i>lui</i> .....		1			1	1			

On en déduit le tableau de comparaison 2 à 2 des éléments distingués. Nous marquerons « 1 » chaque fois qu'un pronom en domine un autre, rien dans le cas contraire. Pour simplifier, négligeons la réflexivité ainsi que la soustraction avec la classe *tu* car, concrètement, cela manque d'intérêt.

domine →	<i>je</i>	<i>tu</i>	<i>elle</i>	<i>nous</i>	<i>vous</i>	<i>elles</i>	<i>le</i>	<i>leur</i>	<i>moi</i>	<i>toi</i>	<i>soi</i>	<i>lui</i>	(degré)
<i>je</i> .....	—	—											0
<i>tu</i> .....		—											0
<i>elle</i> .....			—			1					1		2
<i>nous</i> .....			1	—		1		1			1		4
<i>vous</i> .....			1	1	—	1		1			1		5
<i>elles</i> .....						—					1		1
<i>le</i> .....							—						0
<i>leur</i> .....								—					0
<i>moi</i> .....							1		—	1	1		3
<i>toi</i> .....										—	1		1
<i>soi</i> .....											—	1	0
<i>lui</i> .....						1		1				—	3
Total ... (valence)	0	0	2	1	0	4	1	3	0	1	7	0	19

D'où le réseau de domination, figurant de façon visuelle, le préordre (les pronoms sont classés par paquets et non individuellement). (Voir schéma page suivante.)

Chaque paquet est accompagné de son degré, ou nombre d'arcs qui en sont issus, et de sa valence, ou nombre d'arcs dont il est l'origine.

Le réseau comporte au total 19 arcs. Cependant certaines liaisons sont redondantes en ce sens qu'elles peuvent être déduites par transitivité. Elles ont été marquées en pointillé. Les liaisons directes (10) figurent en trait plein.

Le réseau contient 2 points isolés : {*je*} et {*tu, il, ils*}, ce dernier, rigoureusement parlant, devrait être « dominé » par tous les autres.



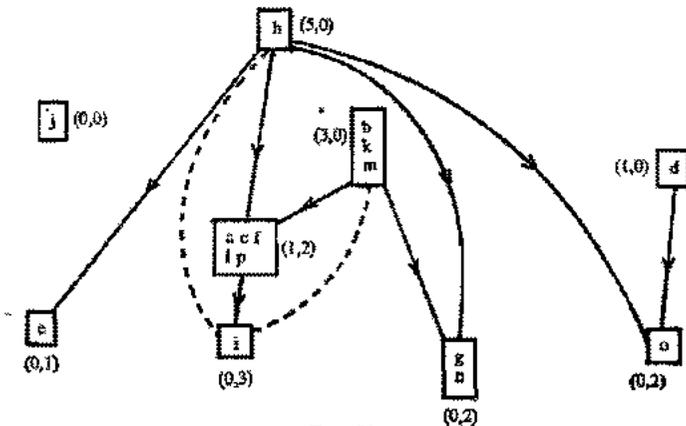


FIG. 21.

## 8. Algorithmes.

L'application effective de règles de grammaire suppose que, plus ou moins consciemment, l'élève met en œuvre un algorithme. Qu'il s'agisse d'accorder un participe passé, de mettre un substantif au pluriel ou de placer une virgule, l'élève doit examiner les données et selon ce qu'il constate il agira de telle ou telle façon.

Voici un exemple.

On trouve dans la grammaire latine de Ch. Georgin et H. Berthaut l'ensemble des règles suivantes, permettant de déterminer la position de l'accent tonique.

« Chaque mot latin possède une syllabe sur laquelle la voix s'élève ou appuie. On dit de cette syllabe qu'elle porte l'*accent tonique* et on l'appelle *syllabe accentuée*.

La place de l'accent tonique est déterminée par les règles suivantes :

1° Les monosyllabes sont accentués, sauf les prépositions, les conjonctions et les enclitiques :

*Nōx, la nuit; per nōctem, pendant la nuit.*

2° Les mots de deux syllabes portent l'accent tonique sur la première syllabe : *cōgo, dōmus.*

3° Dans les mots de plus de deux syllabes, la place de l'accent dépend de la quantité de l'avant-dernière syllabe (*pénultième*).

a) Quand l'avant-dernière syllabe est longue, l'accent porte sur elle : *natūra, Romāni.*

b) Quand l'avant-dernière syllabe est brève, l'accent est reporté sur la syllabe précédente (*antépénultième*): *dōminus, pueritia.*

4° L'enclitique attire l'accent sur la syllabe qui le précède : *omniāque.*

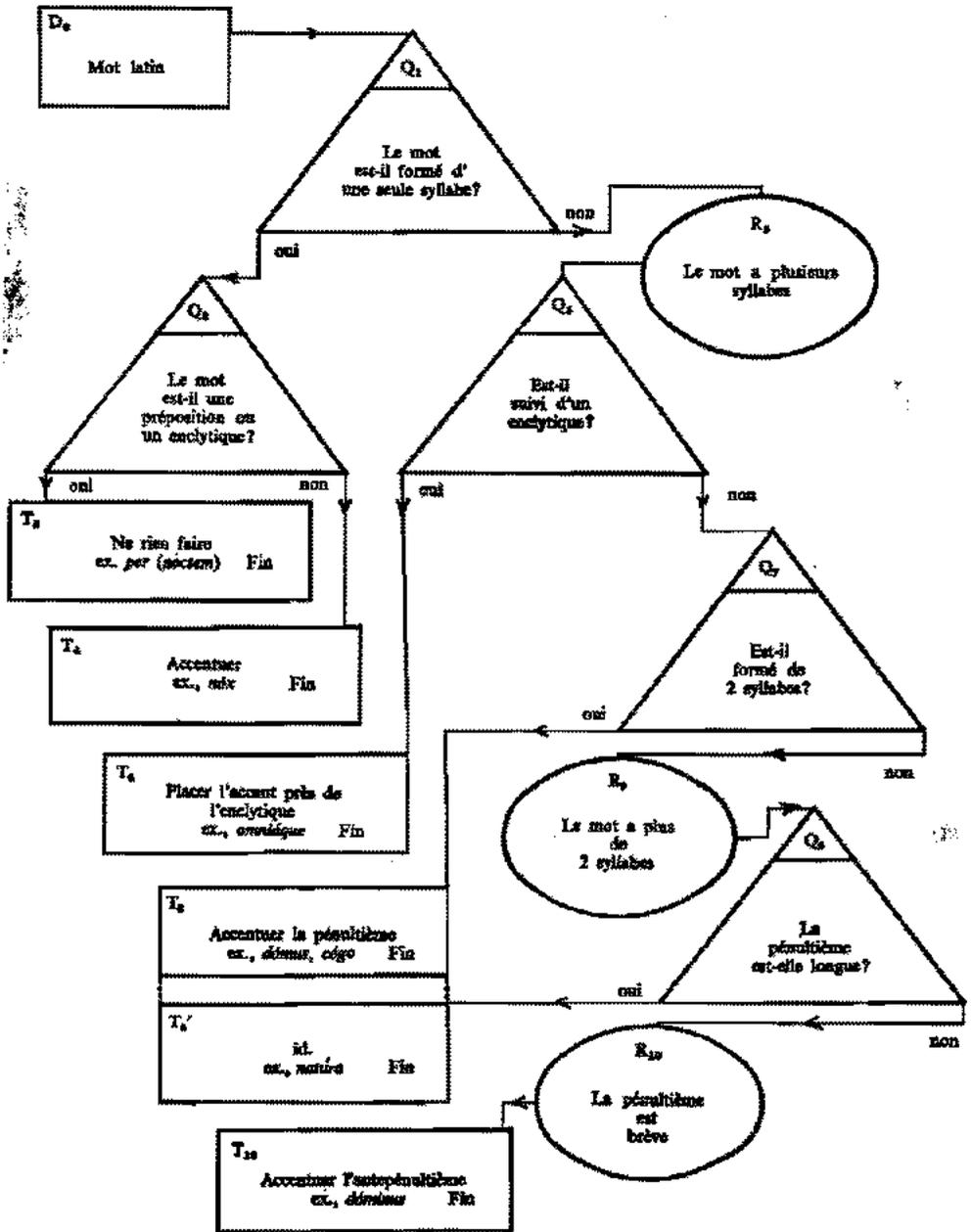


FIG. 22.

L'ensemble de règles ainsi constitué peut-il donner lieu à une procédure du type *algorithme*? Pourquoi? Dessinez l'*organigramme* correspondant. Quelle information apportent les variations typographiques du corps des caractères dans l'énoncé proposé par les auteurs.

Appelons *complexité d'application de l'algorithme au traitement d'un mot donné*  $\mu$ , le nombre  $n(\mu)$  de démarches élémentaires que cet algorithme exige pour obtenir la position de l'accent sur  $\mu$ . Calculez  $n(\mu)$  pour chacun des exemples donnés, puis la *complexité moyenne* caractéristique de l'algorithme.

Nous ne donnons ici qu'un organigramme fournissant une solution possible.

Il est bon de faire réfléchir sur la notion même d'algorithme de façon à éviter les utilisations abusives de ce terme.

Extrait de « De la politesse au pensionnat »  
par M<sup>me</sup> la Comtesse Drohojowska (Paris 1872)

*« Pour demander à boire ou pour boire, il faut tenir son verre de la main droite, avec le pouce et les deux premiers doigts. On ne doit ni tenir son verre à deux mains, ni tousser dedans, ni le porter à sa bouche quand elle est pleine, ni le laisser presque plein sur la table. Il faut donc éviter de se faire verser à boire plus qu'on n'en peut prendre chaque fois. On doit boire lentement, ne pas faire de bruit en buvant, s'essuyer la bouche après avoir bu. Gardez-vous d'avancer le verre de votre voisine ou de poser le vôtre pendant qu'on vous sert; ne relevez pas vivement et de manière à frapper le goulot de la bouteille votre verre, pour témoigner que vous êtes suffisamment servie, mais faites un léger mouvement, presque imperceptible, accompagné du simple mot merci. »*

S'agit-il d'un algorithme? Précisez.

## 9. Conclusion.

Le présent compte rendu doit être considéré comme un premier survol, permettant d'entrevoir quelques possibilités qu'offre la matière linguistique pour un raisonnement à caractère mathématisant. De là il y a loin au panorama exhaustif de toutes les questions que l'on peut soulever à ce propos (\*). Notamment l'auteur s'est volontairement tenu en marge des théories générales développées dans de nombreux ouvrages techniques supposant une préparation mathématique particulière. Il a tenu à insister sur le fait qu'en se maintenant à un niveau volontairement très élémentaire, ne nécessitant pratique-

(\*) Par ex. : aucun exemple explicite n'a été donné pour les probabilités, la notion de distance, les espaces vectoriels, la fonction exponentielle, les combinaisons, la notion d'infini, la continuité, etc... alors que de tels exercices sont aisés à fabriquer.

ment aucune érudition linguistique préalable il était possible de « modeler » un esprit d'observation, d'expérimentation, de formalisation, de déduction... sans nullement avoir en vue une spécialisation linguistico-mathématique ultérieure.

Il est clair que l'on peut imaginer des exercices analogues plus spécifiques (l'auteur en a fabriqué de nombreux, par ex., à l'usage de ses étudiants en linguistique), mais c'est là une tout « autre histoire ».

Y. G.