

T. BEHRAKIS

**Vers une analyse automatique des textes : le traitement de 42 observations contenues dans les livres *Épidémies I et III* d'Hippocrate**

*Les cahiers de l'analyse des données*, tome 8, n° 4 (1983), p. 475-489

[http://www.numdam.org/item?id=CAD\\_1983\\_\\_8\\_4\\_475\\_0](http://www.numdam.org/item?id=CAD_1983__8_4_475_0)

© Les cahiers de l'analyse des données, Dunod, 1983, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

## VERS UNE ANALYSE AUTOMATIQUE DES TEXTES : LE TRAITEMENT DE 42 OBSERVATIONS CONTENUES DANS LES LIVRES EPIDÉMIES I ET III D'HIPPOCRATE

[HIPPOCRATE]

### 1 Introduction

Nous avons suivi une procédure automatique qui a abouti à l'analyse sémantique d'un texte d'Hippocrate.

1.1 Les données : les 42 observations : Dans les livres Epidémies I et III (2) ont trouvé entre autres choses trois séries séparées d'observations particulières, quarante-deux au total.

Une observation est la description de l'évolution de l'état d'un malade.

Hippocrate examine successivement l'altération du visage, la position dans le lit, les mouvements des mains, la respiration, les sueurs, l'état des hypocondres, le sommeil, les selles, les urines, le vomissement et l'expectoration. En guise d'illustration, nous reproduisons la traduction de Littré :

#### 13. ἈΡΡΩΣΤΟΙ ΤΕΣΣΑΡΕΣΚΑΙΑΕΚΑ.

<sup>6</sup> Ἀρρώστος πρῶτος.

Φιλίσκος ἔφασι παρὰ τὸ τεῖχος, ἡ κατεκλιθῆ · ἡ ἤ  
<sup>10</sup> πρῶτη πυρετός ἐξύς · ἴδρωσεν · ἡ ἐς νύκτα, ἐπιτόνωσ.  
<sup>11</sup> Δευτέρῃ, πάντα ἡ παρυξύνθη · ἀψὲ ἡ δὲ ἀπὸ ἡ κλυ-  
σματίου καλῶς διήλλε νύκτα δι' ἡσυχίας. ἡ Τρίτῃ, πρωί, καί  
μέχρι ἡ μέσου ἡ ἡμέρης, εἶδος γενέσθαι ἡ ἀπύρετος· πρὸς δι-  
λην δὲ πυρετός ἐξύς, ἡ μετὰ ἰδρώτος· διψήσας γλώσσα ἡ ἐπι-  
ξηραίνετο. ἡ μέλαινα ἡ οὐρῆσεν · ἡ νύκτα ἡ δυσάφρων οὐκ ἐπι-  
μήθη ἡ πάντα παρέκρουσε. ἡ Τετάρτῃ, πάντα ἡ παρυξύνθη ·  
οὐρα μέλαινα · ἡ νύκτα ἡ εὐφορωτέρην · ἡ οὐρα ἡ εὐχρόστια.  
<sup>21</sup> Πέμπτῃ, περὶ ἡ μέσον ἡ ἡμέρης, ἡ σμικρὸν ἀπὸ βίνου ἡ ἔστα-  
ξεν ἡ ἀκρήτων οὐρα δὲ ποικίλα, ἔχοντα ἡ ἐναυρήματα ἡ στραγ-  
γύλα, γλυκιδίς, ἡ διεσπαρμένα, ἡ οὐχ ἡ ἴδρωτο ἡ πρὸς ἐμῆν  
<sup>5</sup> δὲ βάλανου, ἡ ζουδία σμικρὰ ἡ διήλλεν νύκτα ἡ ἐπιτόνωσ·  
ἡ ὕπνοι ἡ σμικροὶ λόγιοι. ἡ λήρος ἡ ἀκρα ἡ πάντοθεν ψυχρὰ,  
καὶ ἡ οὐκ ἔτι ἡ ἐπαναθερμαινώμενα οὐρῆσε ἡ μέλαινα ἐκαι-  
μήθη σμικρὰ πρὸς ἡμέρην ἡ ἀφρων ἡ ἴδρωσε ἡ ψυχρὸν · ἡ ἀκρα  
<sup>10</sup> πελιόνα· περὶ ἡ δὲ μέσον ἡ ἡμέρης, ἡ ἐκταίος ἡ ἀπέθωνεν.  
<sup>14</sup> Τοῦτέῳ ἡ πνεῦμα ἡ διὰ τήλωσ, ὡσπερ ἡ ἀνακαλωμένη ἡ ἀραῖον,  
μέγα ἡ σπλην ἐπύρεθ, ἡ περιερεῖ κυρτώματι ἡ ἰδρώτος ψυχροί  
<sup>21</sup> διὰ τελευσῆ ἡ οὐ παρξυμοί ἡ ἐν ἀρτίσων.

#### 13. QUATORZE MALADES.

Premier malade.

Philiscus demeurait auprès de la Muraille; il prit le lit.  
*Premier jour*, fièvre aiguë, sueur, nuit pénible. *Deuxième jour*, exacerbation générale; le soir un petit lavement procura des évacuations favorables; la nuit fut tranquille. *Troisième jour*, le matin et jusqu'au milieu de la journée, la fièvre parut avoir cessé. Mais vers le soir, fièvre aiguë, avec sueur; soif; la langue commença à se sécher; le malade rendit une urine noire; il passa une nuit pénible, ne dormit pas, et eut des hallucinations sur toute chose. *Quatrième jour*, aggravation générale, urines noires; la nuit fut plus supportable, et les urines d'une couleur plus favorable. *Cinquième jour*, vers le milieu de la journée, il eut une petite épistaxis d'un sang très noir. Les urines furent variées, on y voyait nager des nuages arrondis, semblables à du sperme, dispersés; elles ne déposèrent pas. Après un suppositoire, Philiscus rendit peu de matières avec des vents. La nuit fut laborieuse; courts sommeils; il parla beaucoup, et divagua; les extrémités étaient partout froides, et on ne pouvait les réchauffer; il rendit des urines noires; il dormit un peu à l'approche du jour; perte de la parole, sueur froide, extrémités livides, mort vers le milieu du sixième jour. Chez ce malade, la respiration fut, jusqu'à la fin, grande, rare, comme si le malade se souvenait de respirer; la rate se gonfla et forma une tumeur ardoisie; les sueurs froides durèrent jusqu'au bout; les accès furent aux jours pairs.

(1) Docteur 3° cycle.

(2) Les 1-er et 3-ème livres des Epidémies sont considérés comme authentiques d'après le témoignage de tous les commentateurs depuis Galien.

(3) E. Littré. Hippocrate : Oeuvres complètes. 1839-1861. Vol II, p. 682.

Sur les 42 observations, il y a 25 cas de morts. Cette proportion considérable de cas mortels est frappante ; dans l'antiquité, Asclepiade appelait les épidémies une méditation de la mort. Mais on ne peut pas voir, dans ce rapport de nombres, un bilan statistique de la pratique d'Hippocrate. Il est évident que les 42 observations, sont choisies parmi beaucoup d'autres. Quel est le motif qui a décidé ce choix ? Si l'on se rappelle combien ces observations répondent fidèlement aux règles dogmatiques, établies dans le Pronostic (1), on sera convaincu, qu'Hippocrate a été mû par le désir d'éclaircir ce livre, et a fait usage des observations, qui lui paraissaient les plus propres à cette fin.

En effet, le Pronostic contient les règles qui expliquent au médecin, ce qu'il a à examiner, une fois rendu auprès du malade. Ces règles correspondent exactement au travail qu'il a fait dans les 42 observations (cf. § 3.1).

La finalité de ces règles n'est pas d'établir le diagnostic d'une maladie particulière, mais le diagnostic de l'état général du malade. C'est ce qu'Hippocrate appelle faire une prognose (προγνῶσις).

On doit ajouter, que dans ces observations l'auteur mentionne exclusivement trois maladies particulières. On y trouve six cas d'une maladie appelée phrenitis, quatre cas d'une maladie appelée causus, et un cas de phthisie.

1.2 *Le problème posé* : E. Littré, qui reste le principal commentateur de l'oeuvre hippocratique, a posé les problèmes des observations comme suit : "A quelles maladies rapporter en général, les histoires des malades qu'Hippocrate a consignées dans les épidémies ?

Que doit-on entendre par le causus et la phrénétis" ? (2)

Il a répondu sur cette question que : "Elles se rapportent aux fièvres rémittentes et continues des pays chauds" (3)

Dans une réédition (Javal et Bourdeaux ; Paris, 1932), de la traduction de Littré, le professeur Roger (doyen honoraire de la faculté de médecine de Paris) commente ainsi les livres I et III des épidémies.

"Nous retrouvons la description des types cliniques, qui nous sont déjà connus : des infections à prédominance nerveuse, causus, phrénite, léthargie et toute une série de maladies fébriles, qu'on a essayé d'identifier avec les infections actuelles. La tentative est inutile, car à cette époque, où la nosographie était à peine ébauchée, presque toutes les infections étaient confondues".

Le professeur Roger n'en cite pas moins quelques diagnostics précis : paludisme, typhoïde, érysipèle... ; ne prétendant reconnaître avec certitude que cette dernière affection.

Sans prétendre, quant à nous, dépasser les hypothèses formulées par Littré, nous avons décidé de concentrer notre attention sur une question différente, c'est-à-dire sur la finalité de la pathologie d'Hippocrate, d'après le texte de 42 observations.

(1) *Le Pronostic est un autre livre authentique, qui a été écrit juste avant les "Epidémies I et III".*

(2) *E. Littré. Hippocrate : Oeuvres complètes. 1839-1861. Vol II p.538.*

(3) *Même oeuvre : Vol II, p. 539.*

On part de l'hypothèse qu'une maladie est déterminée à partir du moment où un vocabulaire la signifie, et qu'un mot est caractéristique à partir du moment où il a une répartition non uniforme dans le corpus traité.

Le problème des observations se pose pour l'analyse des données d'une façon différente de celle de Littré : on examine le texte pour dégager l'information réduite à ses éléments les plus significatifs et suggestifs à la fois.

De cette façon on pourra répondre sur le point de savoir si Hippocrate a essayé d'établir le diagnostic des maladies particulières ou le pronostic de l'état général du malade.

1.3 La méthode : On a utilisé trois programmes : le programme ANA. LEX. (1), un programme d'analyse factorielle des correspondances et un programme de classification hiérarchique avec comme critère d'agrégation la maximisation de la variance interclasse.

On a écrit le texte grec sur support informatique, en transcrivant l'alphabet grec, en format libre, sans altération ni codage préalable.

A partir de ce texte, ANA. LEX. calcule la table de contingence croisant les observations (textes) et les formes (2) graphiques les plus fréquentes.

On a éliminé de toutes les analyses les mots formés d'une, deux ou trois lettres (le paramètre MOMAX = 3). De cette façon on a éliminé la plupart des mots fonctionnels (articles, prépositions, conjonctions, etc.), parce qu'on considère que ces mots ne jouent pas un premier rôle dans une analyse qui vise un but sémantique.

On a effectué une suite d'analyses portant sur des ensembles de plus en plus étendus de formes, choisies sur le seul critère de la fréquence (on change le paramètre NSEU d'ANA. LEX.).

Après chaque A.F.C. on tente une expérience de C.A.H. sur les observations, en considérant leurs coordonnées, les coordonnées factorielles sur les six premiers axes d'inertie. Ensuite on dessine, sur le plan (1.2), les principales classes d'observations. De cette façon se produisent des polygones de classes, dans lesquels on peut observer aussi les mots, qui y correspondent.

Grace au deuxième tableau extrait du programme ANA. LEX., qui représente les formes classées par fréquence totale, on connaît les mots qui sont responsables d'un changement possible de structure, structure représentée par les graphiques de l'A.F.C. ou les classes de la C.A.H. .

Finalement (cf. § 3.2) après avoir fixé la structure du texte des 42 observations, considérées comme authentiques, on projette sur le plan (1.2) les profils de quelques observations qui appartiennent aux autres livres des Epidémies (II, IV, V, VII), pour traiter le problème d'authenticité d'auteur.

---

(1) Une procédure d'analyse lexicale écrite en langage FORTRAN (ANA. LEX.) par L. Lebart. Cahiers de l'Analyse des Données. 1981, n° 2, pp 229-243.

(2) On appelle forme, une suite de lettres comprise entre deux blancs.

2 Les résultats des analyses des observations des livres I et III

On a fait dix analyses. On est parti d'un seuil égal à 18 et on a terminé au seuil 2. On présente ici trois analyses aux seuils 10, 8, 6 ; puis un synopsis de l'ensemble.

2.1 Analyse au seuil 10 : Le nombre des mots retenus ici pour avoir au moins 10 occurrences de chacun d'eux dans le corpus est égal à 97. Le tableau analysé est de dimension 97 x 42.

Les valeurs propres

$$\lambda_1 = 0.18 \quad \tau_1 = 9.9 ; \lambda_2 = 0.15 \quad \tau_2 = 8.4 ; \lambda_3 = 0.11 \quad \tau_3 = 6.4$$

Plan (1.2) et C.A.H.

Sur le plan (1.2) une cloison oblique sépare les observations des morts et celles des guéris. Il n'y a que trois points qui se situent mal : 1012, 3102, 3311 (1)

La C.A.H. nous a donné trois classes principales :

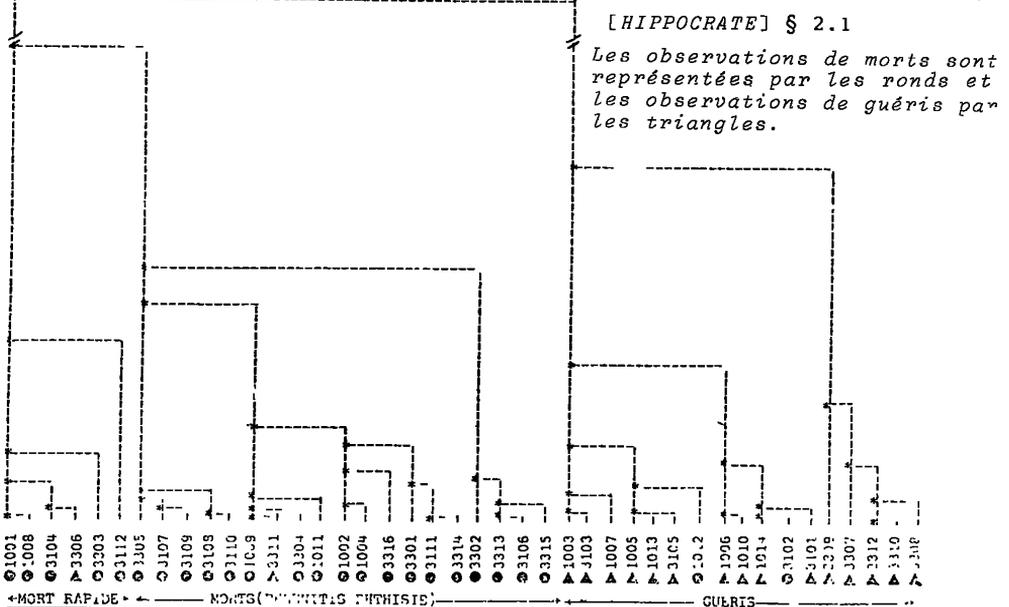
a) La classe de "Mort rapide"

Elle contient des observations de malades qui sont morts après une maladie de durée limitée. On y trouve six malades, cinq d'entre eux sont morts après une durée de 4-5 jours.

Les mots, avec une CTR forte qui y correspondent

ἡμέρας	(la journée)	CTR 66	νύκτα	(la nuit)	CTR 40
πέσον	(le milieu)	CTR 54	ἡσυχίης	(calme)	CTR 35
ἠρωί	(le matin)	CTR 42	ἠαρωζύνη	(aggrav.)	CTR 25

REPRESENTATION DE LA CLASSIFICATION HIERARCHIQUE DES 42 OBSERVATIONS ;



(1) Le sigle de chaque observation se traduit comme suit : Le premier chiffre est le livre (1-er ou 3-ème). Le deuxième est la partie du livre. Le troisième et quatrième chiffres donnent le rang de l'observation.



Il s'agit des mots qui désignent le temps de la maladie dans la journée et ils expriment une évolution violente et rapide.

b) La classe des "Mort (phrénitis-phtisie".

Elle contient les six cas de phrénitis et le seul cas de phtisie. Mais il s'agit en général d'une classe qui comprend des maladies longues. Elle contient dix neuf malades, dont dix-huit sont morts et on trouve des cas de maladies d'une durée qui va jusqu'à cent-vingt jours.

Les mots qui y correspondent.

πάλιν	(de nouveau)	CTR	100	οὐδέν	(nullement)	CTR	33
πολλά	(beaucoup)		56	ἀφωνος	(sans voix)		33
ταχύ	(vite)		50	ψυχρά	(froides)		33
ἄκρα	(extrémités)		43	τέλος	(fin)		30
κοιλίη	(ventre)		36	ἀπέθανεν	(est mort)		28
σβικρά	(peu)		35				

c) La classe des "Guéris"

Elle contient 17 observations de malades, dont 15 guériront. La durée est entre cinq et cent vingt jours.

Les mots qui y correspondent :

εἰκοστή	(vingtième jour)	CTR	62	ἄδύνη	(douleur)	CTR	33
κωφωσις	(surdité)		56	ἐπιήνας	(les jours suivants)		32
ἐκρίθη	(il a guéri)		38	ἄπυρος	(sans fièvre)		25

*La répartition des jours de la maladie*

Au plan (1.2) : premier, deuxième et troisième jours se placent dans le polygone "Mort rapide", tandis que le vingtième jour se trouve en bas entre la classe de "Guéris" et la classe "Mort (phrénitis-phtisie)". On pourrait donc dire, que le 1-er axe est un indice de la durée de la maladie.

*Causus.* On doit remarquer encore que les quatre cas de causus sont dispersés sur le plan (1.2).

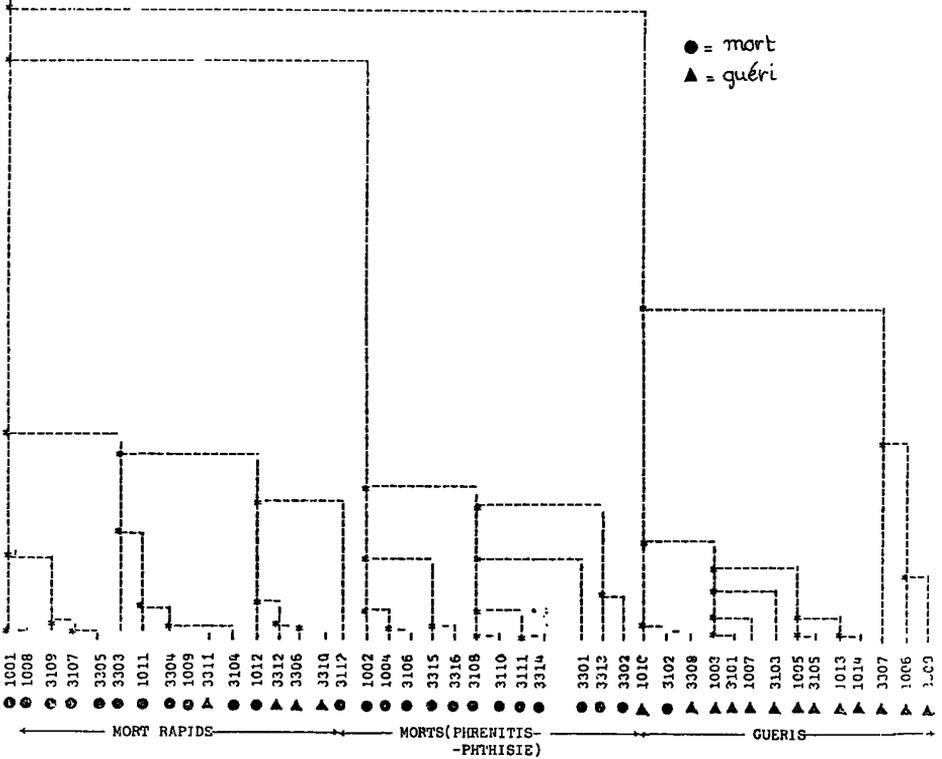
*Un essai d'explication des points "mal situés".*

3102 : On remarque que dans la description de l'évolution de la maladie, certains mots utilisés (comme : ἐκρίθη ≈ il a guéri, ἄπυρος ≈ sans fièvre etc.) - lors d'une période de rétablissement - attirent cette observation du côté des guéris bien que la maladie finisse par la mort.

3311 : De la même façon dans la description d'une maladie, dont l'issue est la guérison, mais courte et violente, des mots usités (comme παροργεῖν ≈ délirait, λόχου ≈ parlait beaucoup) donnent à penser que le malade est atteint de phrénitis.



REPRESENTATION DE LA CLASSIFICATION HIERARCHIQUE DES 42 OBSERVATIONS.  
[HIPPOCRATE] § 2.2 - SEUIL = 8



*Les observations de morts sont représentées par les ronds et les observations de guéris par les triangles.*

1012 : Dans cette observation, on note que malgré la mort du malade, il y a des mots associés à la guérison comme :  $\pi\theta\nu\sigma$   $\approx$  douleur,  $\alpha\pi\upsilon\rho\sigma$   $\approx$  sans fièvre, etc.

2.2 Analyse au seuil 8 : Nombre des mots retenus (pour avoir au moins 8 occurrences) : 124. Tableau analysé :  $124 \times 42$ . Vingt-sept mots ont été ajoutés.

*Valeurs propres*

$$\lambda_1 = 0.18, \tau_1 = 8.9 ; \lambda_2 = 0.16, \tau_2 = 7.7 ; \lambda_3 = 0.12, \tau_3 = 5.7$$

*Plan (1.2) et C.A.H.*

Au plan (1.2) on retrouve la ligne qui sépare les morts et les guéris.

La C.A.H. nous donne les trois classes de l'analyse précédente. La seule différence est qu'un groupe de six observations de "Mort (phre.-phti.) est déplacé à la classe "mort rapide".

2.3 *Analyse au seuil  $\geq 6$*  : Nombre de mots retenus (pour avoir au moins 6 occurrences) : 145. Tableau analysé :  $145 \times 42$ . Vingt-et-un mots ont été ajoutés.

*Valeurs propres*

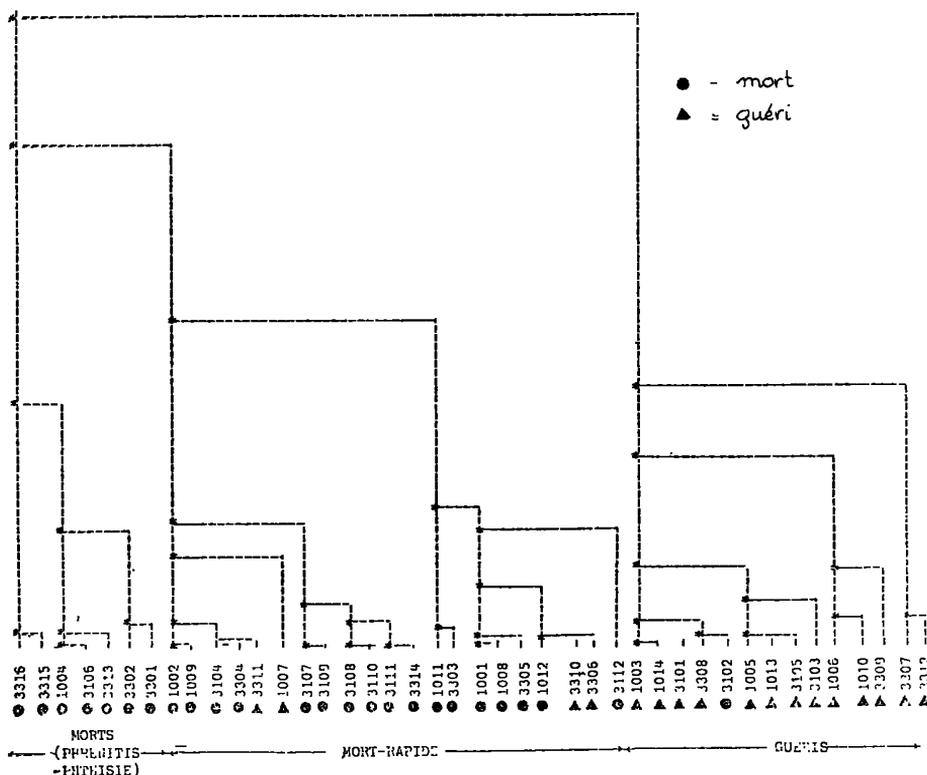
$\lambda_1 = 0.20, \tau_1 = 8.5$  ;  $\lambda_2 = 0.17, \tau_2 = 7.4$  ;  $\lambda_3 = 0.13, \tau_3 = 5.5$

*Plan (1.2) et C.A.H. .*

La situation ne change pas au plan (1.2). Morts et guéris se distinguent par une ligne qui passe par l'origine.

La C.A.H. nous donne les trois classes principales. La classe de "mort (phré - phti)" diminue encore de cinq points. Ces points sont déplacés vers la classe de "mort rapide".

REPRESENTATION DE LA CLASSIFICATION HIERARCHIQUE DE 42 OBSERVATIONS  
[HIPPOCRATE] § 2.3

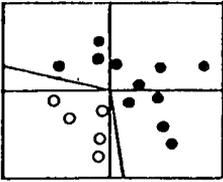


*Les observations de morts sont représentées par les ronds et celles de guéris par les triangles.*

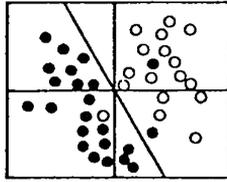


2.4 *Synopsis des analyses* : Dans toutes les analyses quel que soit le seuil, il y a un noyau de points qui caractérise chacune des classes reconnues ci-dessus et aussi un nombre de points qui se déplacent entre elles. Nous avons schématisé sur deux pages les résultats des analyses factorielles et des classifications.

AFCORR [HIPPOCRATE] § 2.4



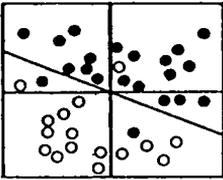
S=2



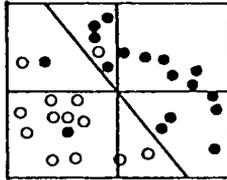
S=3



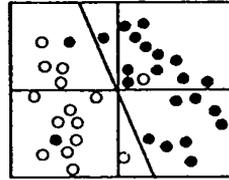
S=4



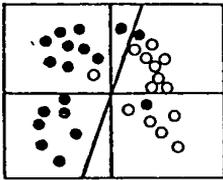
S=6



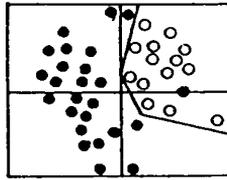
S=8



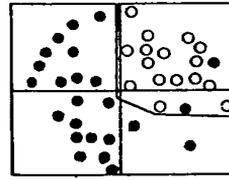
S=10



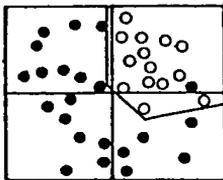
S=12



S=14



S=16



S=18

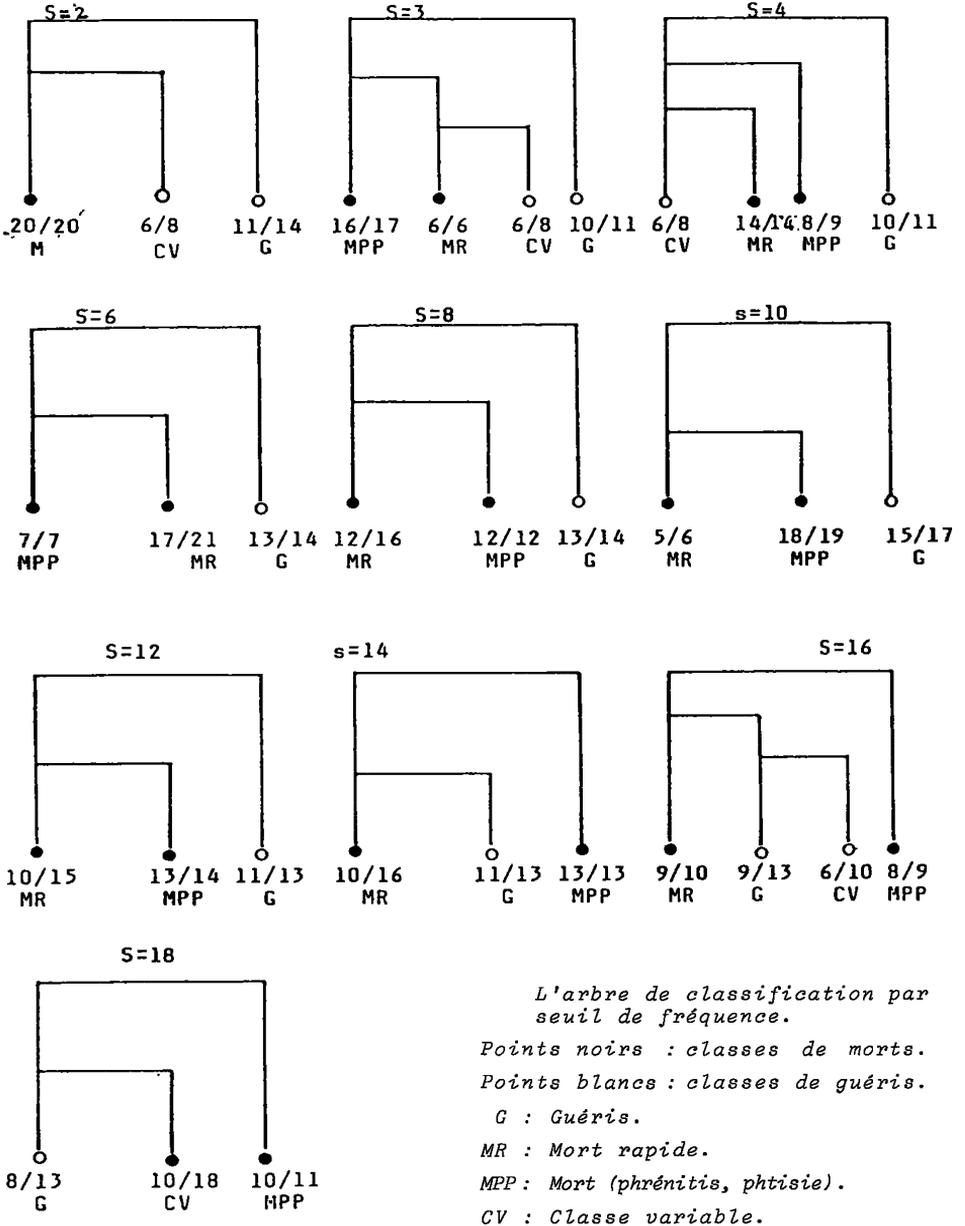
Les plans 1-2 pour les différents seuils de fréquence.

Axe horizontal (2). Axe vertical (1).

Points noirs : morts.

Points blancs : guéris.

Cah : [HIPPOCRATE] § 2.4



### 3 Conclusion sur la structure des textes aux livres 1 et 3 des Epidémies

3.1 Sémantique du pronostic : On a donc (aux livres 1 et 3 des épidémies) trois groupes principaux d'observations, qui concernent la guérison, la mort, et la durée de la maladie. Pour chacun de ces groupes, on a un vocabulaire qui y correspond (cf. tableau).

[HIPPOCRATE] § 3.1

*Points négatifs pour l'issue de la maladie*      *Points positifs pour l'issue de la maladie*

fièvre aiguë	πυρετός δεύς	ρινῶν αἱμορραγῆσαι	épistaxis
frisson	ἐπερρίγσθον	(οὐρα πουλά και ποσά	(urines abondantes
sans dormir	ἀγρυπνοί	την ὑπόστασιν ἔχοντα)	et ayant sédiment)
agités	ἀδῆμονες	(κοιλίην παραχῶδες	(évacuations
ayant soif	διψῶδες	και κοιλῶδες)	bilieuses)
qui avait des nausées	αἰσῶδες	(δυσεντερικοί	dysenterie
sueurs	ἐφίδρουν	γενναίωτο)	
déliraient	παρελεχον		
terreurs	φράβοι		
désespoir	δυσθυμίαι		
les extrémités froides	ἀμρεα περίψυχρα		
urines noires tenues	οὐρα βόγανα ὀχῆα		

*Points pronostiques de la pathologie hippocratique (Epidémie I, 2-ème partie § 9).*

Ce résultat sémantique justifie l'hypothèse, que dans les 42 observations, Hippocrate s'intéresse au pronostic de l'état général du malade et non au diagnostic des maladies particulières.

Dans le livre Epidémie I, § 9 Hippocrate fait des pronostics. Il cite des symptômes qui mènent à la mort et d'autres, qui mènent à la guérison.

Au plan (1.2) de l'analyse au seuil = 6, on peut voir la coïncidence des résultats de notre analyse avec la plupart des points-pronostiques d'Hippocrate.

Pour mieux soutenir ce résultat on va citer le dernier § (25-ème) d'Hippocrate dans le Pronostic. Hippocrate écrit :

"Celui qui veut apprendre à présager convenablement quels malades guériront et lesquels mourront, chez lesquels la maladie durera plus de jours et chez lesquels la maladie durera moins, doit juger toutes choses, par l'étude des signes et pas la comparaison de leur valeur réciproque, tels qu'ils ont été décrits. Il faut avoir une connaissance approfondie des signes et des autres symptômes et ne pas ignorer que, les mauvais signes annoncent du mal et les bons du bien".

On voit donc que notre résultat correspond aux propositions de l'auteur.

3.2 Comparaison sommaire avec les autres livres des épidémies : On a voulu comparer les 42 observations d'Hippocrate avec des observations analogues contenues dans les livres Epidémies II, IV, V, VII en matière d'homogénéité d'auteur (1).

(1) Les 2-ème, 4-ème et 6-ème livres des Epidémies sont considérés comme notes écrites par Hippocrate pour son usage personnel et recueillies par son fils Thessalus, notes qui avaient été augmentées par Thessalus lui-même.

Les 5-ème et 7-ème livres sont considérés comme appartenant à Hippocrate 4, fils de Dracon.

On a donc choisi deux observations par livre, l'une ayant comme issue de la maladie, la guérison et l'autre la mort : soit huit nouvelles observations au total. On a mis ces observations en éléments supplémentaires dans les analyses successives correspondant aux divers seuils.

Au seuil de fréquence minimal égal à 10, on analyse un tableau 100 x 50.

### Plan (1.2)

Au plan (1.2) une ligne sépare les morts et les guéris des 42 observations (Epid. I, III). Cette même ligne ne sépare pas les morts et les guéris des livres qu'on a mis en supplémentaires.

En effet on observe qu'un seul point nouveau se projette du côté des guéris. Cela signifie, que dans les cas nouveaux, la mort et la guérison ne s'expriment pas avec le vocabulaire des 42 observations.

Voici quelques exemples de la différence du vocabulaire : le mot ἀπέθανε (est mort) est remplacé par le mot ἐτελεύτησε ; l'expression ἐσχιν ἐπηράθη (rate gonflée) est remplacée par l'expression ἐσχιν μεγάλη (rate grande); le mot ἤβησε (a vomi) est remplacé par ἤβησεν ἐβητόν (a rejeté des vomissements).

La mise en éléments supplémentaires d'un nombre limité d'observations rencontrées dans les autres livres des Epidémies nous a donné une indication sur l'hétérogénéité d'auteur de l'ensemble des Epidémies, hétérogénéité déjà reconnue depuis Littré.

3.3 Vers une analyse automatique du sens : Il faut rappeler qu'entre la perforation du texte et le plan (1.2) ou l'arbre de C.A.H. , aucune intervention manuelle, ne s'interpose, aucune décision de l'analyste.

Ainsi une élaboration automatique où la subjectivité n'a pas de part, peut contribuer à la recherche du sens d'un texte. Mais il faut souligner le rôle essentiel de l'homogénéité du corpus étudié: ici les textes ne diffèrent ni par l'auteur, ni par le thème, ni par le genre.

Il faudrait continuer cette recherche, en appliquant la même méthode à des textes d'auteurs différents ; et d'abord à l'ensemble des livres des Epidémies ; pour lesquels on a déjà vu sur des exemples que la répartition du vocabulaire trouvée dans les livres authentiques (1, 3) ne se reproduit pas dans les autres livres.

### Références

- T. BEHRAKIS : Paris 6, Thèse 3-ème cycle, 1983. Vers une analyse automatique des textes : Le traitement des 42 observations contenues dans les livres Epid. I et Epid. III d'Hippocrate.
- J.P. BENZECCI et coll. : Pratique de l'Analyse des Données. Linguistique et Lexicologie. DUNOD, Paris 1980.
- L. LEBART : Une procédure d'analyse lexicale écrite en langage FORTRAN [ANA. LEX.] Cahiers de l'Analyse des Données 1981, n° 2, pp229-243.
- E. LITTRÉ : HIPPOCRATE : Oeuvres complètes. Paris 1839-1861.

