

LES TEXTES CONTAMINES ET L'ORDINATEUR

Dans son livre *La critique des textes et son automatisation* (Paris, 1968), dom Jacques Froger a offert aux éditeurs une base théorique pour la collation des manuscrits et pour l'établissement du stemme à l'aide de l'ordinateur. De sa première application, faite par lui-même dès avant la parution du livre, il s'ensuit que cette méthode ne convient qu'à une tradition textuelle homogène. Mais il n'y a pas beaucoup de cas de ce genre dans la littérature latine médiévale. Les oeuvres latines du XIIe - XVe siècle, surtout celles qui étaient destinées à l'usage de l'Ecole et de l'Eglise, sont conservées dans de nombreux manuscrits, qui sont presque contemporains, mais dont la généalogie est tout à fait obscure. Dans ce cas, on ne peut appliquer la méthode proposée par dom Froger parce que, pour l'employer, il faut d'abord éliminer l'anomalie, tandis que dans notre cas l'anomalie est impliquée intégralement dans la tradition textuelle et nous ne sommes point capables de l'éliminer quand nous commençons le travail d'édition. Souvent on croit que, dans ce cas, il n'est guère possible d'établir le texte critique, et c'est aussi l'avis de dom Froger. Il suggère aux éditeurs, quand il y a trop d'anomalies dans la transmission textuelle, de renoncer à l'édition critique et d'éditer le texte d'après un seul manuscrit. C'est une solution acceptable, mais seulement à la condition qu'il y ait moins de cinq manuscrits et que l'anomalie soit si étendue et si profonde qu'il ne soit pas possible, dans la plupart des lieux variants, de conjecturer la base commune. Cela concerne les oeuvres écrites en langues vulgaires, où le nombre de manuscrits est d'ordinaire moindre et où le copiste est en même temps le rédacteur du texte copié. Mais, avec un nombre plus grand de

manuscrits, et surtout s'il s'agit d'oeuvres latines, où le scribe reste, presque toujours, un simple copiste, il y a une possibilité de restituer le texte de l'original; plus grand est le nombre des manuscrits, plus s'accroît la probabilité que le texte critique soit authentique, au moins en grande partie; en tout cas, il sera plus proche de l'original que le texte du meilleur manuscrit, à l'exception naturellement de l'orthographe, qui se refuse à toute restitution. Je crois que c'est précisément dans ce cas - c'est-à-dire quand on a plusieurs manuscrits dont la généalogie n'est point claire - que l'ordinateur peut nous aider substantiellement dans l'établissement du texte critique. Seulement, nous ne l'interrogerons pas de la même manière que dom Froger.

Dans ce cas, on doit principalement renoncer à la collation par ordinateur. Celle-ci ne convient - à vrai dire - que s'il n'y a, dans les manuscrits collationnés, que des fautes de copiste. Mais - et cette remarque concerne, autant que la méthode de dom Froger, celle, toute différente, de MM. Zarri et Maretti de Milan - la collation automatique ne convient pas aux manuscrits où il y a un assez grand nombre de variantes d'interprétation telles que synonymes, anomalies grammaticales, rectifications de scribe, émendations stylistiques, etc. A l'égard de telles anomalies, la collation traditionnelle, faite à la main par le philologue lui-même est plus rapide, même quand il collationne chaque manuscrit deux fois et, surtout, elle est plus exacte. Conformément à la tradition, cette collation faite à la main est négative, c'est-à-dire que le philologue transcrit le texte du manuscrit de base qu'il a choisi arbitrairement, avec toute la précision possible, et que, des autres

manuscrits, il ne note que les variantes. Il n'enregistre pas les variantes purement orthographiques telles que confusion entre majuscules et minuscules ou entre *i* et *y*, addition et omission de *h*, surtout au commencement des mots, *n* prothétique avant le groupe *-gn-*, *m* prothétique avant le groupe *-pn-*, *n* omis devant le groupe *-gw-*, etc. Il n'enregistre pas non plus les variantes résultant de la nationalité des copistes, telles que la confusion *s-sc-x-z* et *u-o* dans les manuscrits d'origine italienne, la confusion des consonnes sonores et sourdes dans les manuscrits dont le scribe était allemand, la confusion de *g* et *j* devant les voyelles antérieures dans tous les manuscrits d'origine quelconque, etc. En fait, il note ces variantes émanant de la nationalité du scribe, parce qu'elles apportent un témoignage important sur le niveau de civilisation du scribe et nous aident à caractériser le manuscrit en question, mais il les met à part. On élimine de l'apparat critique les variantes orthographiques et celles qui sont causées par les différences de prononciation, parce qu'elles l'inonderaient d'une masse dépourvue de signification. Au surplus, chaque nationalité et chaque période ayant des conventions d'orthographe et de prononciation propres, nous obtiendrions - en ne les éliminant pas - des variantes communes pour des manuscrits qui n'ont en commun que la nationalité des scribes, mais qui, généalogiquement, sont loin l'un de l'autre.

Sur la base de cette collation qu'il a faite à la main, le philologue établit de façon préliminaire le texte et l'apparat critiques, en se servant de toutes les méthodes critiques convenables, surtout de la règle *lectio difficilior verior*

et de la psychologie du scribe. On peut objecter que cette méthode laisse le champ trop libre à la conjecture, qui ne peut qu'être arbitraire. Mais, comme pour un lieu variant, on ne rencontre qu'exceptionnellement deux leçons variantes qui conviennent toutes les deux également aux principes de la critique textuelle, au sens du texte et à l'usage de l'auteur - c'est-à-dire à son spectre, comme le dit Froger, - le choix de la leçon restituée est en règle générale facile. D'ailleurs, plus souvent que deux leçons, on rencontre, pour un même lieu, plusieurs leçons variantes qu'on peut expliquer l'une par l'autre, de sorte que, même ici, le choix de la leçon originale ne fait pas difficulté. Le texte critique préliminaire qu'on obtient de cette façon fournit la lecture correcte pour 90 % des lieux variants. Ce sont les 10 % restants qui nous obligent à continuer le travail.

A ce moment, on répète la collation sur la base, non plus d'un manuscrit, mais sur le texte critique préliminaire. La tâche principale de cette deuxième collation est de suppléer les lacunes et de corriger les fautes qui se sont introduites lors de l'établissement préliminaire du texte et de l'apparat, soit que les sigles des manuscrits aient été confondus, soit qu'on ait commis quelque erreur de copie. Alors seulement vient l'étape dans laquelle, à mon avis, l'ordinateur pourrait rendre au philologue des services importants.

Déjà dans la première étape de son travail, le philologue remarque que certains manuscrits présentent les mêmes leçons dans un grand nombre de

lieux. Comme les relations entre les manuscrits sont données surtout par les fautes communes, le philologue a besoin de la statistique : combien de fautes tel ou tel scribe a-t-il commises, combien de fautes a-t-il en commun avec un ou plusieurs autres scribes. Si nous faisons cette statistique à la main - en enregistrant sur des fiches et en comptant, pour chaque leçon variante, les sigles des manuscrits - c'est un travail lourd et ennuyeux où l'on n'évite pas les fautes de distraction. Le travail va plus vite si l'on emploie, au lieu des fiches manuscrites, les fiches tringlées (cartes à perforations marginales) sur lesquelles on enregistre les groupes des manuscrits (pour chaque leçon variante séparément) en les perforant selon un code. Dans ce cas, on peut aussi noter - de nouveau selon un code - des indications sur le type de la leçon variante (faute accidentelle, variante explicable généalogiquement, lacune, addition, variante au sens propre, permutation de mots). Ensuite, nous rangeons les fiches à l'aide des tringles ou d'un petit rangeur mécanique. Ce procédé est à peu près trois fois plus rapide que le précédent et les résultats sont plus précis. Mais il reste, malgré tout, très lent. Au surplus, le philologue ne peut traiter ainsi qu'une toute petite partie de l'oeuvre éditée; il lui serait tout à fait impossible d'établir à la main les relevés statistiques de l'oeuvre entière, surtout parce qu'avec le nombre croissant des manuscrits croît très vite le nombre de groupes dans lesquels les manuscrits se rencontrent. Je crois que c'est l'ordinateur qui pourrait préparer ces relevés statistiques qui forment la base du travail du philologue dans la deuxième étape.

Les données de ces relevés sont tirées de l'apparat critique préliminaire. On transcrit sur cartes perforées le numéro du lieu variant (dans un texte de prose, c'est le numéro sous lequel le lieu variant figure dans l'apparat critique dactylographié et qui ensuite, dans le livre imprimé, est remplacé par le numéro de la ligne; s'il s'agit d'un poème, ce numéro est donné par le numéro du vers et par le numéro d'ordre du mot dans le vers). Cette donnée, nécessaire pour permettre un recours à l'apparat critique, demande de trois à huit colonnes de la carte perforée, selon la longueur du texte en question. Les deux colonnes suivantes distinguent rencontre fortuite et rencontre généalogique. Les quatre colonnes suivantes sont réservées à l'indication du type de variante : omission, addition, variante simple et permutation des mots. Si le texte est tellement compliqué qu'il y a des niveaux divers, on peut noter dans les colonnes suivantes, le niveau dont il s'agit (par exemple le niveau du texte propre, des gloses, des citations). Tout le reste de la carte est libre pour les sigles des manuscrits.

Pour le travail ultérieur, le philologue a besoin des statistiques suivantes :

- 1) Le nombre de variantes isolées de chaque manuscrit. Ce nombre permet de dresser la liste des manuscrits en ordre croissant de la quantité de variantes isolées qu'ils présentent.
- 2) Le nombre de variantes qu'un manuscrit donné a en commun avec un ou plusieurs autres manuscrits.

- 3) Le nombre de variantes communes à deux manuscrits quelconques, avec distinction des cas où ces combinaisons sont isolées et de ceux où elles se rencontrent avec d'autres manuscrits.
- 4) Le nombre de variantes de trois manuscrits.
- 5) Le nombre de variantes communes à quatre manuscrits.
- 6) Selon le nombre de manuscrits, on fait ensuite le dénombrement des variantes communes à cinq, six, etc. manuscrits. On calcule ensuite toutes les combinaisons possibles selon le nombre des manuscrits. Pour les relevés 4, 5 et 6, on fera la même distinction que pour le relevé 3.

Enfin, il faut dresser la liste des groupements de manuscrits qui se rencontrent dans les listes 3 à 6. Il est souhaitable de compter à part les cas de rencontre fortuite (notés par la colonne correspondante), et ceux de rencontres généalogiques. Si c'est nécessaire, on peut aussi faire les relevés pour chaque genre de faute (de nouveau désigné par la colonne correspondante).

Voici comment le philologue emploie ces statistiques, qu'elles soient faites à la main, ou sur des fiches à perforations marginales, ou à l'aide de l'ordinateur. C'est le poème latin *Rudium doctrina* qui nous servira d'exemple.

Ce poème fut écrit à l'usage de l'école vers la fin du XIIe ou au commencement du XIIIe siècle par un auteur inconnu. Pour l'éditer, je disposais de 33 manuscrits dont le plus ancien est d'un siècle au moins postérieur à l'original. Comme la lecture de ce poème fut "obligatoire" dans les écoles de l'Europe entière pendant trois siècles au moins, on comprend bien que les manuscrits que nous avons à notre disposition sont très différents, qu'ils sont contaminés, interpolés, etc. Malgré cela, il a été possible d'atteindre un assez haut degré d'exactitude dans l'établissement du texte critique. (L'édition a paru sous le titre *Quinque claves sapientiae : Incerti auctoris Rudium doctrina, Bonvicini de Ripa Vita scolastica*, Bibliotheca Teubneriana, 1969). L'apparat critique des cent premiers vers a servi de base à l'exposé suivant qui décrit une expérience faite à l'aide de fiches tringlées, tandis que, pour établir le stemme dans l'édition, je n'ai employé que les fiches traditionnelles. Naturellement, je désigne les manuscrits par les mêmes sigles que dans l'édition où l'on trouvera aussi la description des manuscrits.

Le nombre des combinaisons possibles des manuscrits est donné par l'analyse combinatoire. Puisque nous trouvons le *Rodium doctrina* dans 33 manuscrits, nous pouvons obtenir :

33 33 33 33 33 33 33 33 33 33 33 33 33 33
1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 13 + 14 +

33 33 33 33 33 33 33 33 33 33 33 33 33 33
15 + 16 + 17 + 18 + 19 + 20 + 21 + 22 + 23 + 24 + 25 + 26 + 27 + 28 +

33 33 33 33 33
29 + 30 + 31 + 32 + 33 = 33.442.494,591 combinaisons au total.

Naturellement, toutes ces combinaisons ne sont pas réalisées et ne pourraient pas l'être même si le texte était assez long et ne comptait pas seulement 376 vers comme le *Rudium doctrina*. Cela ne serait possible que si, dans la transmission, régnait non pas une anomalie mais une anarchie absolue. Dans ce cas, il faudrait vraiment renoncer à l'établissement du texte critique. Mais la réalité n'est pas si triste et les relations généalogiques qui lient les manuscrits laissent leurs traces dans les combinaisons réalisées. Plus grande est la partie de l'oeuvre que l'on étudie, plus évidentes deviennent les traces de la généalogie dans les combinaisons réalisées. Et c'est pourquoi nous voudrions consulter l'ordinateur pour que nous puissions traiter une partie vraiment représentative, ce qui n'est jamais possible sans l'ordinateur.

Dans les 100 premiers vers du poème, il y a au total 646 mots. De ce total, 294 mots sont conservés dans tous les manuscrits sans aucune variante. Pour les 352 mots restants, qui représentent 352 lieux variants, on rencontre 603 variantes, dans lesquelles sont réalisées 220 combinaisons.

- 1) De ces 603 leçons variantes, plus de la moitié (389) est due aux scribes individuels, parce que chacune de ces 389 leçons variantes se trouve dans un seul manuscrit. Ces leçons variantes isolées sont distribuées entre les manuscrits de la façon suivante :

2 - K,Δ	13 - M,S
3 - Π,Θ	15 - R
4 - I	16 - Ψ
5 - L,U,Λ	17 - P
6 - W	19 - V,Z
7 - F,Ω	20 - N
8 - Y	22 - X
9 - B,E,T	23 - A
10 - H,Γ	24 - C,Q
11 - D,Σ	26 - G
12 - J	

Ce tableau des variantes isolées nous montre d'une part la confiance que

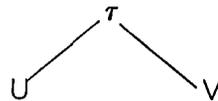
mérite chaque scribe en tant que copiste, d'autre part, la distance approximative qui sépare un manuscrit de son prédécesseur. Aucun copiste, quelque négligent qu'il soit, n'augmente le nombre de fautes de son prédécesseur de plus de dix fautes en 100 vers. C'est pourquoi nous supposons la perte du prédécesseur immédiat (éventuellement de l'avant-dernier), si le nombre des leçons variantes isolées dans la partie traitée est plus grand que dix.

- 2) Une autre donnée dont nous avons besoin est le nombre des variantes où un manuscrit coïncide avec un ou plusieurs autres manuscrits. Pour le poème que nous étudions, nous avons obtenu la liste suivante :

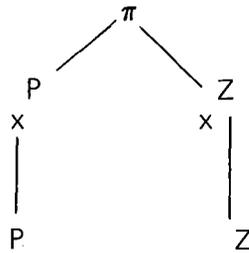
	21 - A,Λ	31 - M		
12 - I		32 - E,R		
	23 - S	33 - N		
	24 - X			
		35 - D	45 - B,T	55 - P
16 - Σ	26 - L,Q,V,Y,			56 - U
	27 - H			
18 - Δ,Π	28 - K,Γ	38 - C		58 - Z
19 - Ψ	29 - J		49 - G	
20 - W,Θ,Ω	30 - F			

- 3) Les combinaisons les plus importantes sont celles où deux manuscrits au moins ont une variante en commun. On compte pour chaque variante

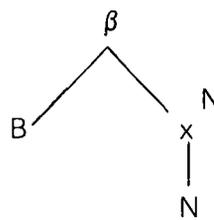
toutes les combinaisons possibles. Des 603 variantes, il y en a 98 qui ne se rencontrent que dans deux manuscrits formant 73 combinaisons parmi les 528 possibles (c'est-à-dire $\binom{33}{2}$: AC (1 leçon isolée, 14 leçons où AC se rencontre dans un groupe des manuscrits ultérieurs) AF (1,14), AP (1,4), AQ (1,8), AA (1,11), BG (1,19), BN (3,18), BP (1,12), BV (1,12), BZ (1,16), CE (1,9), CG (1,16), CP (2,14), CZ (1,19), DK (1,15), DQ (1,9), DΠ (1,7), EF (1,15), EI (1,4), EN (1,9), EZ (1,5), EΨ (1,4), FG (1,11); FN (1,9), FR (1,9), FA (1,9), GK (1,6), GM (1,10), GN (2,19), GO (1,6), GW (1,6), GZ (1,18), HP (1,5), HQ (1,5), HZ (1,6), HΠ (2,5), JL (1,11), JW (1,6), JΓ (1,12), JA (1,10), KN (1,9), LM (1,7), LP (1,5), LV (1,9), MP (2,2), MQ (1,4), MU (1,10), MY (1,9), MZ (1,5), NP (1,17), NU (1,15), NZ (2,20), NΓ (1,10), PQ (1,5), PZ (7,26), PΓ (1,6), PΠ (1,3), PΣ (1,4), QS (1,4), QZ (1,5), QΨ (1,2), RΨ (1,8), SΔ (1,7), TU (12,31), VW (1,5), VZ (1,9), WΓ (1,4), XZ (1,8), XΘ (1,9), ΓΩ (1,4), ΨΨ (1,6), ΠΨ (1,7). C'est le groupe TU qui se rencontre le plus souvent (12 fois isolément, 31 fois dans des combinaisons plus larges), ce qui révèle que les manuscrits T et U dépendent d'un ancêtre commun :



La combinaison PZ apparaît sept fois isolément et 26 fois dans des groupes plus larges. Il faut donc supposer un ancêtre plus lointain Π . Entre celui-ci et nos manuscrits P et Z, on doit insérer (vu le nombre de leçons isolées de P et Z) des maillons intermédiaires :



Les manuscrits B et N se rencontrent trois fois en groupe isolé et 18 fois dans des groupes plus larges. Ils ont, eux aussi, un ancêtre commun β duquel - à ce qu'il semble - B a été copié directement, tandis que N l'a été par l'intermédiaire d'un intermédiaire perdu x^N :



Ce qui conduit aux rapports généalogiques, ce n'est pas le nombre des leçons où deux manuscrits se rencontrent isolément, mais bien la relation entre la somme des leçons où apparaissent des groupes isolés de deux manuscrits et celle des leçons où les deux manuscrits se trouvent dans un groupe plus vaste. Par exemple, la combinaison MP apparaît deux fois isolément. Cependant, on ne peut supposer aucune liaison généalogique entre ces manuscrits, parce qu'ils ne se trouvent que deux fois dans un groupe plus vaste et que le type de ces deux variantes révèle qu'il s'agit d'une rencontre fortuite.

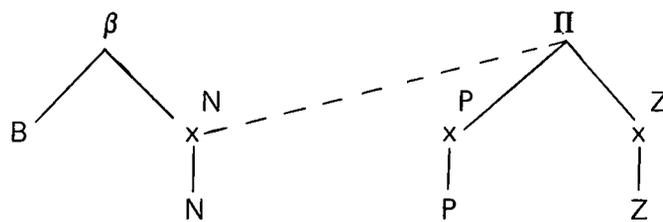
A partir de la liste des combinaisons de deux manuscrits isolés, il faut dresser celle des paires dans lesquelles entre un manuscrit donné :

avec	1	manuscrit	-	I, T, Δ, Σ, Ω
avec	2	manuscrits	-	R, S, X, Y
avec	3	manuscrits	-	D, K, U, Λ
avec	4	manuscrits	-	H, J, L, V, W, Γ, Π
avec	5	manuscrits	-	A, B, C, Θ, Ψ
avec	6	manuscrits	-	E, F
avec	7	manuscrits	-	M
avec	9	manuscrits	-	G
avec	10	manuscrits	-	Q, Z
avec	12	manuscrits	-	P
avec	19	manuscrits	-	N

Cette liste permet de comprendre le degré de contamination de chaque manuscrit. Les manuscrits qui s'apparient avec le plus grand nombre de manuscrits, apparaissent comme produits par une transmission mêlée. Dans notre cas, cela vaut sans doute pour 8 manuscrits (depuis E, qui se rencontre avec six manuscrits, jusqu'à N, qui se combine avec 19 manuscrits). Pour les manuscrits qui se combinent avec deux, trois, quatre ou cinq autres, on doit décider si la rencontre est fortuite ou généalogique.

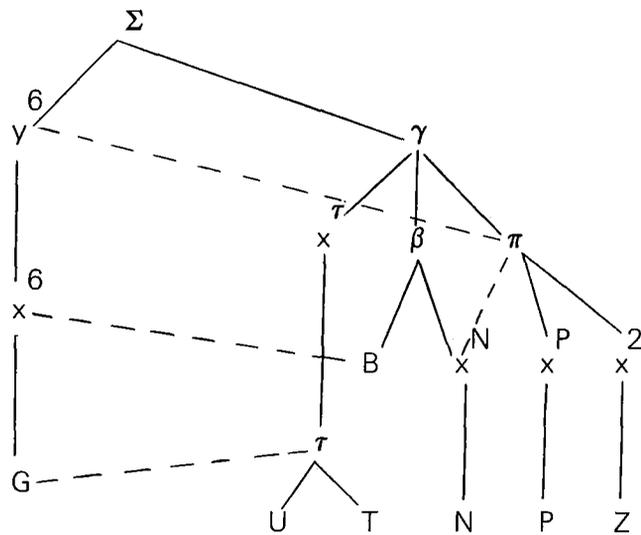
En effet, dans la première étape du travail, toutes les fautes simples sont notées comme résultant d'une rencontre fortuite, alors que certaines d'entre elles résultent d'une dépendance généalogique. Dans ces conditions, T est un descendant direct du manuscrit perdu τ , sans aucun autre intermédiaire, alors que les combinaisons des manuscrits I, Δ , Σ , Ω sont fortuites.

- 4) Des combinaisons possibles de trois manuscrits, qui pourraient être au nombre de 5456 c'est-à-dire $\frac{33}{3}$, il ne s'en réalise que 28. De plus, seul le groupe NPZ est représenté deux fois isolément et 9 fois dans des groupes plus vastes. Comme nous savons déjà que P et Z proviennent du même ancêtre Π , seule la relation entre N et Π est à traiter. Cette relation a laissé des traces jusque dans les combinaisons NP et NZ, soit isolées, soit comprises dans des groupes plus vastes (voir plus haut). On en déduit le fragment de stemme que voici :



- 5-6) La combinaison de quatre manuscrits apparaît 17 fois, celle de cinq manuscrits 12 fois, toujours pour une seule variante qui, au surplus, représente une rencontre fortuite.
- 7) Importante est la combinaison de six manuscrits. Des 1.107.568 types possibles, dix seulement se réalisent, dont un seul, BGPTUZ, se répète deux fois : de plus, ce groupe n'apparaît jamais combiné avec un autre manuscrit. Il s'ensuit que ce groupe n'est pas généalogique et qu'il est causé par l'origine mixte du manuscrit G. Mais, il reste une question à résoudre : y a-t-il quelque relation généalogique entre les ancêtres perdus β , Π , τ que nous connaissons par les traces qu'ils ont laissées dans les manuscrits conservés B et N, P et Z, T et U. La combinaison $\beta\tau$ n'apparaît pas isolément; groupée avec d'autres manuscrits, nous la trouvons sept fois. Il s'ensuit que β et τ appartiennent à la même transmission textuelle, mais que leur ancêtre commun est éloigné. La combinaison $\beta\Pi$ n'apparaît pas non plus dans une leçon isolée; groupée

avec d'autres manuscrits, elle se rencontre cinq fois. Tel est aussi le cas du groupe $\Pi\tau$. Les ancêtres perdus β , τ , Π sont donc collatéraux (il faut supposer un intermédiaire x^T en raison du nombre de variantes qui apparaissent seulement dans la combinaison TU). Voici la partie du stemme qui en résulte :



On continue de la même manière en traitant les combinaisons plus larges. Plus grande est la partie de l'oeuvre qui peut être statistiquement étudiée, plus claires apparaissent les relations généalogiques, même si l'anomalie est

assez profonde. Pour la dominer entièrement, il faut, si possible, étudier l'oeuvre tout entière. Si, à l'aide de l'ordinateur, l'on obtenait la statistique complète, on pourrait espérer établir aussi la partie du stemme la plus proche à l'archétype, pour laquelle la statistique partielle ne nous offre pas des données suffisamment représentatives.

Après l'établissement du stemme, on traite de nouveau les variantes de l'apparat critique pour voir si quelque leçon décisive n'est pas en contradiction avec lui (certaines leçons résultant d'une rencontre fortuite sont naturellement en contradiction). Si c'est le cas, on corrige soit le texte au moyen de l'apparat critique, soit le stemme.

Telles sont quelques-unes des idées d'un philologue qui, jusqu'ici, n'a pas travaillé directement avec l'ordinateur. Il ne peut évidemment juger si l'ordinateur est actuellement capable de faire un nombre si élevé d'opérations ou si l'on peut trouver le moyen d'épargner à l'ordinateur les nombreuses opérations inutiles (par exemple, si deux manuscrits ne sont pas liés dans une paire, il est inutile de chercher leurs combinaisons plus vastes). Cette tâche attend son programmeur.

PRAGUE

Anežka VIDMANOVÁ