

## Comptes rendus

### Problèmes de méthode

DOM JACQUES FROGER: *La critique des textes et son automatisation*.  
Coll. «Initiation aux nouveautés de la science» n° 7.  
Paris, Dunod, 1968, xii + 280 p.

L'édition d'un texte présente un double aspect: il y a un travail passionnant qui consiste à étudier les manuscrits, à pénétrer la langue et la pensée de l'auteur et à établir le texte par des conjectures ingénieuses ou par un choix judicieux entre les leçons offertes par les différents témoins. Mais dès que le texte est long et que les manuscrits sont nombreux, il y a aussi un travail fastidieux, de pure routine, quand il faut collationner les versions et démêler, à partir d'un fouillis de variantes, les relations généalogiques des manuscrits. Parfois c'est une vie entière que le savant doit y consacrer. Aussi les résultats sont-ils souvent décevants: éditions dont le texte et l'apparat critique fourmillent de fautes et d'inexactitudes, éditions inachevées parce que l'éditeur est mort avant d'être arrivé au terme ou qu'il s'est épuisé à la tâche, éditions annoncées qui ne voient jamais le jour parce que des philologues, enthousiastes d'abord, ont sous-estimé l'ampleur et les difficultés de l'entreprise.

Afin de remédier à cet état des choses on a envisagé plusieurs solutions: le travail en équipe tel qu'on le pratique, par exemple, avec grand succès dans certains ordres religieux, le recours à des secrétaires (mais, outre qu'ils sont difficiles à obtenir avec les crédits réservés dans la plupart des pays à cette sorte de recherches, ils ne font peut-être pas toujours preuve de l'acribie nécessaire), et enfin, plus récemment, l'utilisation des ordinateurs électroniques qui se sont montrés très efficaces pour la confection d'*indices verborum* et de concordances.

Ces dernières années, on a assisté en effet à deux tentatives fort intéressantes d'appliquer les ordinateurs aux parties les plus routinières de la critique des textes. L'une est due à un Américain, M. Vinton A. Dearing<sup>1</sup>, professeur à l'Université de Californie et éditeur des œuvres de John Dryden; il s'inspire surtout des méthodes de W. W. Greg<sup>2</sup> qu'il a pourtant assouplies et développées de façon à en faire des instruments efficaces et applicables à des cas beaucoup plus compliqués.

---

1: Vinton A. Dearing, *A Manual of Textual Analysis*. Berkeley & Los Angeles, University of California Press, 1959, xi + 108 p.; *Methods of Textual Editing, a Paper delivered ... at a Seminar on Bibliography held at the Clark Library*. Los Angeles 1962, iv + 35 p.; *Some Routines for Textual Criticism: The Library*, Déc. 1966, p. 309-317; *Some Notes on Genealogical Methods in Textual Criticism: Novum Testamentum*, t. 9 (1967), p. 278-297. Deux études n'ont pas encore été publiées: *The Logical Foundations of Textual Genealogy and the Calculus of Variants* (53 pages dactylographiées) et *Abaco-Textual Criticism* (41 pages dactylographiées).

2: *The Calculus of Variants. An Essay on Textual Criticism*. Oxford 1927.

L'autre tentative est due à un bénédictin français de l'Abbaye de Solesmes, Dom Jacques Froger<sup>3</sup>, qui se place plutôt, comme il est naturel, dans la lignée de Dom Henri Quentin tout en prenant ses distances et en nous offrant des méthodes très originales.

Comme la réalisation technique des procédés avancés par M. Dearing n'est décrite que très sommairement<sup>4</sup> et que le philologue américain est en train d'adapter ses programmes, faits d'abord pour un IBM 7090/94, à un ordinateur de la «troisième génération» (il s'agit probablement d'un IBM Système/360) tout en y apportant des améliorations considérables, nous préférons laisser de côté ces recherches jusqu'à plus ample informé et nous attachons surtout aux réalisations de Dom Froger.

Parmi les nombreuses opérations que comporte l'édition d'un texte, il y en a notamment deux qui se prêtent à un traitement automatique: la collation et une partie des recherches généalogiques, à savoir la détermination de l'enchaînement des manuscrits. Il est nécessaire, en effet, comme l'avaient déjà vu Dom Quentin et W. W. Greg, qu'on décompose l'établissement du stemme en deux temps. On cherchera d'abord à déterminer l'enchaînement à partir des différences qu'on relève entre les diverses versions comparées à un texte choisi arbitrairement comme texte de référence, et on ne se demandera pas si ces différences correspondent à des fautes ou à des leçons authentiques. C'est seulement dans un second temps qu'on donnera, après avoir introduit la notion de faute, une orientation à l'enchaînement et qu'on obtiendra, si le cas s'y prête, un stemme.

Les deux programmes, établis par M<sup>me</sup> Renaud et par M. Porré de la Compagnie des Machines Bull, sont conçus comme des unités indépendantes de façon qu'on puisse, par exemple, collationner les textes à la main et faire établir l'enchaînement par la machine à condition toutefois que les variantes soient présentées à l'ordinateur de la même façon que les résultats de la collation automatique.

Pour expliquer ses théories sur la généalogie des manuscrits, Dom Froger s'inspire dans une large mesure de la logique<sup>5</sup> et des mathématiques nouvelles (notamment de la théorie des ensembles), dont il a une connaissance approfondie et auxquelles il a consacré un chapitre remarquable (p. 139-216) illustré constamment par des exemples tirés de la critique des textes. Malheureusement il ne nous est guère possible, dans les limites d'un compte rendu, d'entrer dans les détails de cet aspect important du livre: cela demanderait de nombreux diagrammes et un exposé

3: En dehors du livre dont nous avons à rendre compte on peut renvoyer aux articles suivants du même auteur: *La critique textuelle et la méthode des groupes fautifs*: Cahiers de lexicographie, t. 3 (1961), p. 207-224; *L'emploi de la machine électronique dans les études médiévales*: Bulletin de philosophie médiévale (S.I.E.P.M.) t. 3 (1961), p. 177-188; *La collation des manuscrits à la machine électronique*: Bulletin de l'Institut de Recherche et d'Histoire des Textes, n° 13 (1964-65), p. 135-171; *La machine électronique au service des sciences humaines*: Diogène, n° 52 (1965), p. 110-144.

4: Notamment dans *Methods of Textual Editing*, p. 18-20 et p. 24-25.

5: Par exemple les diagrammes d'Euler-Venn pour lesquels on se sert de cercles ou de formes arrondies pour représenter les concepts et pour délimiter leur extension.

préalable des notions qu'il faut utiliser. La formulation logique des procédés n'est d'ailleurs pas nécessaire comme l'a démontré l'auteur dans le chapitre II, qui contient un exposé traditionnel de ses méthodes de critique textuelle<sup>6</sup>; elle contribue cependant pour beaucoup à la clarté et à la rigueur, et il est hors de doute qu'elle facilite à merveille le dialogue, souvent difficile, entre l'humaniste et le programmeur de formation scientifique et qui est bien étranger aux problèmes philologiques.

#### *La collation*

Avant de procéder à la collation automatique il y a un certain nombre de préparatifs auxquels il faut faire face:

Pour les sigles des manuscrits retenus, Dom Froger préconise l'emploi généralisé des chiffres arabes à la place des désignations traditionnelles et numérote les témoignages de façon continue. C'est là une exigence peu commode: en général les lettres sont plus faciles à retenir et comme elles permettent de caractériser jusqu'à 26 manuscrits par position (contre les dix des chiffres), elles occupent moins de place dans la mémoire de la machine, sur les rubans magnétiques et sur les cartes perforées. Il est possible que l'ordinateur qu'a utilisé Dom Froger n'admette que des chiffres avec cette fonction, mais pour les ordinateurs plus récents le choix des sigles n'est limité que par le répertoire de caractères de l'imprimante, et celle-ci permet maintenant d'utiliser jusqu'à 240 signes différents; même si les lettres grecques et hébraïques doivent être le plus souvent exclues, il reste possible de conserver la grande majorité des sigles existants.

La transcription se fait à peu près comme d'habitude; seule exception importante: le programme de collation n'admet pas les observations «archéologiques» (déchirure, grattage, etc.), qui s'imposent pourtant dans un appareil critique et qui sont importantes pour qu'on puisse déterminer la valeur de certaines leçons. Il faut donc ajouter tous ces renseignements après la collation automatique ou bien à partir d'une liste qu'on a rédigée en même temps que la transcription ou bien en refaisant la collation une deuxième fois. Nous croyons pourtant qu'on pourrait se tirer d'affaire d'une manière très simple: il suffirait de placer, éventuellement sous une forme codifiée, chaque commentaire entre deux signes spéciaux (par exemple «%» et «+») dont la fonction serait définie dans le programme. La machine réserve alors aux champs ainsi délimités un traitement particulier: elle les imprime dans les listes qui donnent les résultats des collations et elle les ignore quand elle procède aux comparaisons et aux calculs qui devraient permettre de découvrir l'enchaînement des manuscrits.

Les lacunes qu'on peut déterminer sans consulter les autres manuscrits sont indiquées par une série d'*X*. Si l'on veut tenir compte des corrections et additions postérieures, la seule solution satisfaisante serait de considérer chaque main comme une version indépendante; il n'est pas nécessaire pourtant de refaire la transcription:

---

6: Il semble même que Dom Froger ait d'abord élaboré ses méthodes sans s'appuyer sur les mathématiques et que ce soit seulement après coup qu'il s'est rendu compte que sa méthode n'était au fond qu'un cas particulier de la théorie des ensembles.

on peut se contenter de copier les cartes perforées automatiquement en remplaçant celles qui contiennent des modifications de la main postérieure.

Comme la transcription se fait le plus souvent à la main, il est prudent de taper la copie à la machine à moins qu'on n'ait une écriture particulièrement lisible. Les textes sont donnés ensuite à une secrétaire qui en translittère le contenu sur cartes perforées d'après des règles à fixer pour chaque cas. Comme les cartes contiennent forcément des fautes, il faut les faire passer par une «vérificatrice», où le texte est perforé une deuxième fois au-dessus des perforations déjà faites; s'il y a désaccord, on contrôle et on refait la carte en cas de faute. Dom Froger, qui n'a pas beaucoup de confiance dans les vérificatrices<sup>7</sup>, préfère vérifier le texte sur des épreuves imprimées à partir des cartes perforées de façon que chaque ligne de l'épreuve corresponde à une carte et porte le même numéro.

Il faut bien avouer que le procédé est assez compliqué et que le travail préparatoire risque d'être au moins aussi long qu'une collation faite à la main. Pour le simplifier les philologues mettent souvent leur espoir dans la lecture optique qui dispenserait les éditeurs de toute préparation des textes. Nous sommes persuadé pourtant qu'on n'arrivera jamais à faire lire directement des textes manuscrits anciens de façon satisfaisante; même des exemplaires bien conservés d'incunables et d'éditions anciennes demanderaient une programmation tellement compliquée de l'ordinateur et un contrôle si serré de la part du philologue pour détecter les mélectures dues à la qualité médiocre du papier et de l'impression que la lecture optique ne se justifierait dans ces cas qu'à titre d'expérience.

En revanche il est pleinement possible, dès maintenant, de simplifier la procédure si l'on se sert, pour les copies dactylographiées, d'une machine à écrire équipée avec des caractères légèrement stylisés («Farrington Selfcheck 12F/12L», «OCR-A», etc.). Les «lectrices optiques», déjà sur le marché, se chargeraient alors de lire les feuilles à des vitesses qui varient entre 400 signes/sec. («Farrington Optical Page Reader») et 1000 signes/sec. («IBM 1288 Optical Page Reader»).

Lorsque l'ordinateur a enregistré les textes à partir des cartes perforées, il attribue à tous les mots du texte de référence des numéros pairs et à tous les espaces des numéros impairs, l'espace devant le premier mot du texte reçoit le numéro un. Tous les numéros désignent des lieux variants virtuels. Comme les additions de deux ou de plusieurs mots exigent plus d'un lieu par espace, la machine garde la possibilité d'ajouter une ou deux décimales («1,1 – 1,2 – 1,3», etc.). Les autres versions sont comparées ensuite, l'une après l'autre, au texte de référence par l'ordinateur qui superpose pour ainsi dire les deux textes en alignant les mots identiques. En cas de désaccord il prend un nombre de mots plus élevé (par exemple deux) et fait toutes les comparaisons possibles dans les deux sens; s'il n'arrive pas encore à une identité, il prend une tranche de cinq mots et ainsi de suite jusqu'à ce qu'il trouve un mot commun sur lequel il aligne les deux textes. S'il y a une variante, elle est considérée comme une addition située dans le lieu qui précède la leçon du texte de référence qui, à son tour, est considérée comme omise. Toutes les variations sont donc exprimées par des additions et des omissions.

La méthode de Dom Froger, qui consiste à numéroter les lieux et à prendre

7: Diogène, n° 52 (1965), p. 127.

comme unité de comparaison le mot, peut paraître un peu compliquée. Pour les textes en vers, M. Dearing a choisi comme unité le vers que l'ordinateur examine à la fois du début et de la fin; chaque fois que la comparaison révèle une différence, il retient le mot pour l'imprimer ensuite au-dessous du mot dont il pourrait être une variante. C'est seulement après avoir examiné une bonne partie du vers que l'ordinateur décide que les deux versions correspondent. S'il y a des différences près du début ou de la fin de la ligne, il examine d'abord par exemple les dix vers précédents et les dix vers suivants pour être sûr qu'il n'est pas tombé sur une omission, une interpolation ou une transposition; s'il n'y en pas, il continue à comparer les deux lignes; s'il y a transposition, la ligne transposée reçoit le numéro de celle du texte de référence et un signe attire l'attention sur le fait. Si un vers, par une faute assez courante, est composé de deux moitiés de vers différents du texte de référence, le vers composé est inventorié deux fois. Avant d'imprimer les résultats, la machine se demande si la même variante se trouve ailleurs dans le texte ou dans un autre texte déjà collationné et note le fait à titre de renseignement. De cette façon M. Dearing arrive, avec l'IBM 7090, à collationner par exemple cinq textes de 56 lignes en douze secondes et 27 textes de 24 lignes en 30 secondes. Pour des textes en prose, M. Dearing envisage de prendre comme unité la phrase (terminée par un point) et de procéder de la même manière; ceci est possible à la rigueur pour des textes modernes, auxquels il s'intéresse surtout: ils ont des chances d'avoir conservé la ponctuation de l'auteur. Dans le cas des textes anciens et médiévaux, qui sont ponctués de façon peu systématique, une laborieuse précollation s'impose afin que le texte soit coupé en morceaux de même longueur dans tous les manuscrits.

On se rend facilement compte des dangers que présente une collation automatique telle que la propose Dom Froger: aussi longtemps que le texte à collationner n'offre que peu de variantes, tout va bien, mais dès qu'il présente une lacune ou une interpolation importante ou une série de variantes de quelque étendue, c'est l'effondrement total, et on risque que l'ordinateur, en désespoir de cause, aligne les deux textes sur des mots banals (tels que *et*, *est*, *à*, *de*, etc.) ou sur des mots pleins qui se trouvent par hasard dans différentes parties du texte.

Dom Froger ne semble pas encore avoir résolu ces difficultés et, en effet, le cas est difficile. On pourrait se livrer à une précollation afin de détecter les passages critiques qui seraient collationnés à la main, on pourrait établir une liste des mots trop fréquents pour être utiles ou exiger que l'alignement soit fait seulement quand deux ou trois mots de suite correspondent, on pourrait établir un dialogue avec la machine, comme on le fait au L.A.S.L.A. à Liège pour la lemmatisation et l'analyse grammaticale des mots latins<sup>9</sup>, de façon qu'elle appelle à la rescousse chaque fois que les complications dépassent ses capacités. Mais ce ne sont là que des pis-aller.

Un autre inconvénient non négligeable est causé par les variantes graphiques: même si la différence ne porte que sur une seule lettre, des formes du même mot

8: *Methods of Textual Editing*, p. 19.

9: Cf. A. Bodson et E. Evrard, *Le programme d'analyse automatique du latin: Revue de l'Organisation internationale pour l'étude des langues anciennes par ordinateur*, 1966, n° 1, p. 1-50.

sont considérées par la machine comme deux mots entièrement différents. Si l'on collationne, par exemple, un texte picard avec un texte anglo-normand, on risque que des phrases entières qui contiennent exactement les mêmes mots lexicaux soient considérées par l'ordinateur comme deux séries de leçons qui n'ont aucun rapport. Si les variantes graphiques sont assez stéréotypées ou faciles à circonscrire, on peut présenter à la machine une liste disant par exemple que si la seule différence entre deux mots latins est un *ae* pour un *e* (ou l'inverse), les deux mots sont identiques, ou bien on peut établir, si on a le courage, une liste de correspondances graphiques précisant par exemple, pour l'ancien français, que *out* = *ot* = *eut* = *ut* = *aut*, etc. Au cas où les différences graphiques sont trop nombreuses et trop capricieuses, la seule solution est probablement qu'on supprime les variantes inutiles une fois la collation terminée.

Le philologue reçoit les résultats de la collation automatique sous forme de trois types de listes:

1°. *Collations séparées* pour chacun des manuscrits qu'on a comparés avec le texte de référence. L'ordinateur donne les numéros des lieux variants et imprime le texte de façon continue aux passages où il n'y a pas de variantes.

2°. *Collations rassemblées* où les variantes de tous les manuscrits sont données avec les sigles. Nous reproduisons quelques lignes de l'exemple donné par Dom Froger:

«N 1.1 R.OR 2 3 4 5 6 N 2 R.CA 2 3 4 5 6 MA N 3 R.CHERE 2 5 N 4 O.FILLE  
2 5 JE N 7 R.DOIS 4 N 8 O.VAIS 4 VOUS N 11 R.ANNONCER 2 N 12 O.DIRE  
4 UNE N 16 O.ETRANGE 2 3 5... , etc.»<sup>10</sup>

Il est possible qu'on puisse facilement s'habituer à travailler sur ces listes, mais à première vue elles n'inspirent guère un grand enthousiasme. Il est vrai que l'imprimante de l'ordinateur utilisé par Dom Froger ne permet qu'un nombre très restreint de signes. Sans doute pourrait-on obtenir des résultats plus satisfaisants si on employait des crochets, des parenthèses, des astérisques, etc. Il ne serait pas difficile non plus de faire placer les variantes après le mot de référence au lieu de les faire précéder. L'idéal, à notre avis, serait d'avoir des listes où les variantes de chaque manuscrit seraient données dans une ligne à part au-dessous du texte de référence. Pour les collations séparées une telle disposition serait facile à réaliser: on n'a qu'à demander à l'ordinateur de rédiger les deux lignes à la fois en calculant les espaces qu'il faudrait introduire dans le texte de référence afin que les deux textes restent alignés en cas d'additions. En ce qui concerne les collations rassemblées les calculs deviennent plus compliqués, mais ils restent dans la mesure du possible. Il semble que ce soit une solution de ce type qu'a adoptée M. Dearing.

3°. La dernière liste se présente sous forme de cartes perforées, nommées *cartes récapitulatives*; elles ne donnent que le numéro du lieu variant et les sigles des manuscrits qui contiennent la variante en question et qui constituent ainsi un «groupe variant».

C'est sur les collations rassemblées et les cartes récapitulatives que le philologue aura à travailler maintenant. Celles-ci, qui devraient fournir les données pour le

10: *N* est suivi du numéro du lieu variant, *R* = «Rajoute», *O* = «Omet»; la variante précède le mot de référence et est suivie des sigles des manuscrits qui la contiennent.

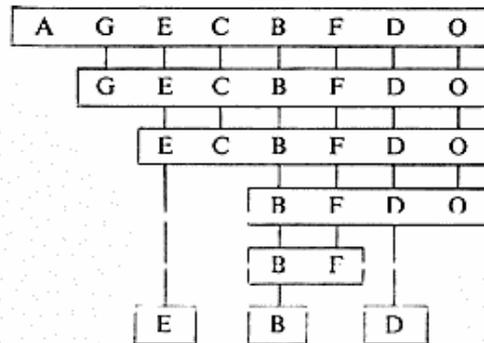
calcul de l'ordinateur, sont presque toujours inutilisables à l'état brut: les variantes graphiques, que la machine a minutieusement inventoriées, sont ainsi inutiles et même nuisibles vu que des manuscrits collatéraux peuvent contenir la même graphie d'un mot contre celle d'un ancêtre; en ancien français deux manuscrits collatéraux peuvent par exemple offrir, indépendamment l'un de l'autre, la leçon *caïr* contre *cheoir* des autres manuscrits («rencontre fortuite»). Aussi faut-il supprimer les cartes dont les numéros renvoient à de telles variantes. De même, dans tous les cas où une variante est décomposée par la machine et exprimée par plusieurs lieux variants ou quand une longue addition se traduit par autant d'additions qu'il y a de mots ajoutés, il est nécessaire, afin de ne pas fausser les fréquences, de ne conserver qu'une seule des cartes qui concernent la variante, par exemple la première.

Le programme qui doit établir l'enchaînement des manuscrits n'admet que des variations de forme binaire (une variante par lieu variant). Dans le cas de variations à variantes multiples il faut donc supprimer les variantes excédentaires en les rattachant ou bien à la leçon du texte de référence ou bien à la variante qu'on considère comme essentielle. Il est regrettable pourtant qu'on soit obligé d'introduire une évaluation aussi subjective dans une opération qui se veut objective. Si l'on s'y refuse ou qu'il soit impossible de décider, on n'a qu'à supprimer purement et simplement le lieu variant.

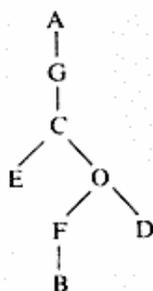
*L'enchaînement des manuscrits*

Si on travaille à la main, on examine ensuite les groupes variants que la collation a fait ressortir et on les dispose par niveaux d'après le nombre de manuscrits que les groupes renferment: en haut le groupe le plus important auquel on joint le manuscrit de référence; ensuite, en descendant, les groupes par nombre dégressif des manuscrits, et à la fin, au niveau 1, sont juxtaposés tous les «groupes» qui ne contiennent qu'un manuscrit. De cette manière on est sûr en tout cas que les pseudo-descendants sont à un niveau inférieur par rapport aux pseudo-ancêtres. Reprenons le texte de Molière (*Le Malade imaginaire*, I,5) que Dom Froger a utilisé comme illustration tout au long du livre. Il est conservé dans huit versions dont sept sont pourvues de différentes fautes introduites par Dom Froger afin de démontrer les mécanismes des opérations de la critique des textes. Il a brouillé les cartes et nous faisons semblant de n'avoir aucune idée des relations véritables. Il a choisi un manuscrit *A* comme texte de référence: la collation montre qu'il y a cinq groupes de 8, 7, 6, 4 et 2 manuscrits respectivement et trois manuscrits isolés.

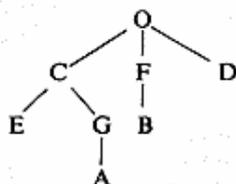
Si l'on dispose les groupes et les manuscrits isolés par niveaux en alignant verticalement les sigles répétés à divers niveaux, on obtient ce résultat:



On n'a alors qu'à biffer les sigles qui se retrouvent à un niveau inférieur, et l'enchaînement est trouvé à partir des manuscrits qui restent :



En réalité c'est *O* qui est l'original, et si on réussit à le trouver en éliminant les manuscrits fautifs, on aboutit au stemme véritable en suspendant l'enchaînement à *O* :



Evidemment l'exemple est privilégié et peu réaliste puisque tous les manuscrits y compris l'original sont conservés. Si la tradition manuscrite est tronquée, comme il arrive presque toujours, nous obtenons à certains niveaux des groupes où tous les manuscrits ont des descendants et où il n'y a pas de pseudo-ancêtre responsable des différences; dans ces cas nous avons repéré des manuscrits perdus qu'on désigne conventionnellement par des lettres minuscules de la fin de l'alphabet (*x, y, z*) ou par des lettres grecques. Nous laissons de côté les cas plus compliqués qui présentent des anomalies dues à la contamination ou aux rencontres fortuites.

Ce sont à peu près les mêmes opérations qu'exécute la machine; en gros on peut distinguer quatre étapes :

1°. Impression d'une liste des groupes variants retenus en fin de compte par le philologue. Ils sont énumérés dans l'ordre du texte.

2°. Tableau où les variantes sont rassemblées par classes et où les classes sont associées aux différents types de groupes correspondants; la machine calcule la fréquence de chaque classe.

3°. Les groupes sont disposés par niveaux d'après le nombre de manuscrits qu'ils contiennent. Comme nous l'avons vu, c'est là une opération tout à fait simple si on la fait à la main, et on pourrait facilement demander à la machine de procéder de la même façon. Cependant Dom Froger a préféré être conséquent avec lui-même en expliquant fort savamment à la machine qu'elle doit avoir recours à l'inclusion logique de façon qu'elle énumère les manuscrits au lieu de les compter.

4°. L'enchaînement des groupes ressort d'un tableau où l'ancêtre immédiat (ou les ancêtres immédiats en cas d'anomalie) est indiqué pour chaque type de groupe

variant. A partir de ces données (niveau et ancêtre immédiat) on peut facilement dessiner le graphe ou le faire dessiner par l'ordinateur s'il est pourvu de l'équipement nécessaire et qu'on dispose d'un programme approprié.

Pour trouver l'enchaînement des manuscrits individuels l'ordinateur ne retient que les sigles des manuscrits qui n'ont pas de descendants et met des sigles conventionnels pour les manuscrits perdus.

Avant d'établir le texte il ne reste donc, si tout c'est bien passé, qu'à trouver l'orientation du pseudo-stemme de façon à en faire un vrai stemme. C'est pour cette opération, faite entièrement à la main, qu'on introduit la notion de faute et qu'on s'appuie également sur les données externes.

Bien qu'il soit parfaitement possible de rédiger l'apparat critique du texte établi à partir des collations rassemblées, il est peut-être plus prudent de procéder à une nouvelle collation automatique en prenant comme version de référence le texte édité; si celui-ci est composite, il est nécessaire de le mettre d'abord dans une forme accessible à la machine. Du même coup, on a la possibilité d'obtenir une concordance ou un *index verborum* qui peut être publié tel quel avec l'édition ou offrir des matériaux précieux pour un glossaire rédigé à la main. Une concordance, de type KWIC par exemple («Key-Word in Context»), faite dès le début à partir d'un des manuscrits qui inspirent le plus de la confiance, rendrait également de très grands services pour l'établissement du texte permettant à l'éditeur de mieux connaître la langue de l'auteur et ses tournures préférées.

Il n'est pas facile, en se fondant sur les procédés élaborés par Dom Froger, de dresser le bilan des progrès qui pourraient être réalisés grâce aux ordinateurs électroniques. En effet, le livre, paru en 1968, fait état d'expériences menées en 1961 avec un petit ordinateur (le «Gamma Tambour» de la Compagnie des Machines Bull), maintenant dépassé et qui impose de nombreuses restrictions techniques (nombre de caractères limité à l'entrée et à la sortie, faible capacité de la mémoire, restrictions dans l'utilisation des 80 colonnes des cartes perforées, etc.); il est vrai que Dom Froger insiste loyalement sur ces inconvénients et qu'il souligne à plusieurs reprises qu'on pourrait obtenir de nettes améliorations si l'on utilisait les machines modernes beaucoup plus puissantes et souples, mais il aurait été plus convaincant de voir réalisées ces améliorations d'une manière plus concrète. Avec les programmes de Dom Froger il faut reconnaître que les résultats pratiques sont assez décevants:

A moins qu'on n'ait affaire à des textes, de préférence récents et imprimés, qui ne présentent que peu de variantes, le gain de temps dont on pourrait se réjouir est plus que problématique; et comme les interventions du philologue restent nombreuses, il n'est même pas sûr qu'on arrive à des résultats plus exacts bien que la machine ait la réputation de travailler d'une manière infaillible.

Si le nombre de manuscrits est suffisamment grand, la restitution de l'enchaînement semble en revanche plus efficace d'autant plus que la méthode de Dom Froger offre d'excellents moyens de détecter à temps les anomalies et de les éliminer à la main si elles ne sont pas trop graves.

Même si les résultats concrets ne sont pas spectaculaires, la fréquentation des ordinateurs présente d'autres avantages incontestables: elle oblige surtout le phi-

logue à se renouveler et à préciser sa doctrine de façon à prévoir les moindres détails et à aboutir à un tout cohérent où toutes les notions sont définies avec rigueur. Ce sont là justement les qualités qui caractérisent le travail de Dom Froger.

B. Munk Olsen

PARIS

B. MUNK OLSEN: *Anvendelsen af elektronisk databehandling ved løsningen af filologiske opgaver: konkordanser, indices verborum.*

(Le Traitement automatique appliqué à la solution de problèmes philologiques: concordances, indices verborum.)

Københavns universitets fond til tilvejebringelse af læremidler. Copenhague, 1968, 128 p.

L'indexation automatique constitue un domaine déjà classique de la mécanolinguistique. C'est peut-être parce que, justement dans ce domaine, les avantages de la mécanisation semblent le plus évidents. Il n'est pas besoin, en effet, de souligner l'intérêt capital que présenterait pour les linguistes et critiques littéraires de toutes sortes (lexicographes, syntacticiens, stylisticiens, historiens de langue, etc.) le fait de posséder des index totaux ou sélectionnés des textes d'une période donnée ou même d'un seul ouvrage. Jusqu'à une époque récente, les entreprises de ce genre semblaient chimériques ou du moins impliquaient un travail énorme et fastidieux. Or, l'adaptation des ordinateurs électroniques au dépouillement des textes a rendu possible l'élaboration automatique d'index et de concordances.

C'est à une tâche de ce genre que s'est attelé B. Munk Olsen en se proposant d'établir des concordances de trois ouvrages du philosophe anglo-français Jean de Salisbury († 1180). Le livre de Munk Olsen rend compte des méthodes qui ont présidé à son travail. Il comporte trois parties: aspects techniques, établissement d'une bande magnétique contenant le texte, utilisation de cette bande. Dans les deux dernières parties, l'auteur donne un excellent aperçu des multiples problèmes que soulève le traitement automatique des textes, et il décrit soigneusement les solutions que lui-même a apportées à ces problèmes. Il s'agit notamment des problèmes de codage, de prédiction, de références, de corrections, de lemmatisation, de présentation des textes de sortie, etc. Particulièrement intéressante me paraît l'utilisation qu'a faite Munk Olsen du lecteur optique.

Je ne m'arrêterai qu'à un seul de ces problèmes, crucial celui-là: comment réaliser les programmes qui dictent à la machine les opérations à effectuer. Munk Olsen esquisse trois procédés: utilisation de programmes déjà existants, programmation par un programmeur professionnel (non linguiste) et programmation par l'utilisateur lui-même (p. 52-54). Munk Olsen a opté pour la deuxième solution tout en reconnaissant qu'elle pose des problèmes délicats de coopération entre le linguiste et le programmeur, ce dernier ne comprenant ordinairement rien aux problèmes spécifiquement linguistiques. Bien que, actuellement, cette solution soit de loin la plus répandue en mécanolinguistique, il me semble qu'il faille absolument préconiser le troisième procédé: le linguiste – comme tout autre scientifique qui désire utiliser l'ordinateur – doit apprendre à programmer. Avec un langage comme l'Algol, c'est chose relativement facile qui fera que le linguiste apprendra à maîtriser la machine