

En guise d'éditorial

Ce premier numéro de notre Revue se situe en dehors de la série normale qui commencera avec l'année 1966.

Il s'agit ici d'une prise de contact destinée à répondre immédiatement aux buts que s'est assignés notre Organisation. Il est en effet important que chaque membre fasse connaissance le plus tôt possible avec tous les autres et ait une idée des travaux que chacun a entrepris ou est sur le point d'entreprendre.

On trouvera donc à la suite de ce texte une liste des membres de l'Organisation. Dans la mesure où les renseignements nous sont parvenus, chaque nom est accompagné d'un curriculum vitae et d'une description sommaire des travaux entrepris et des projets.

Nous redoutions, au moment où nous avons lancé l'Organisation, que des travaux identiques ne soient entrepris simultanément sur ordinateur par des chercheurs qui s'ignorent. Un premier examen des réponses reçues jusqu'à présent écarte provisoirement cette crainte.

Ce qui se produit et se produira sans doute encore, ce sont des rencontres entre des travaux réalisés d'une manière artisanale et en quelque sorte manuelle et des travaux faits sur machines. C'est ainsi que viennent de paraître coup sur coup deux Concordances du Corpus Ti-

bullianum, celle de A. DELLA CASA parue en Italie en 1964 et celle de E. O'NEIL parue en Amérique vers la même époque. Ces deux Concordances font double emploi, non seulement l'une avec l'autre, mais encore avec un Index du même Corpus qui vient d'être achevé par un des chercheurs de mon Laboratoire et qui va paraître incessamment chez Mouton à La Haye.

Je citerai encore les cas suivants :

- 1°) Index de Cassiodore auquel travaille depuis longtemps Dom Cappuyns de l'Université de Louvain, à l'aide de fiches manuscrites, et le même Index constitué sur machines par le Professeur Halporn (Université de l'Indiana, U. S. A.).
- 2°) Index des Lettres de Pline le Jeune entrepris par le R. P. Van Ooteghem (Namur - Belgique) sur fiches manuscrites et aux U. S. A. par deux membres de notre Organisation à l'aide des machines. Il s'agit de M. Charles W. Dunmore (New-York University) et de Miss Jane Barlow (Susquehanna University, Pennsylvania). Pour autant que mes renseignements soient exacts à l'heure actuelle, Miss Jane Barlow paraît avoir renoncé à son projet.
- 3°) Enfin, tout récemment, notre Index de la Consolation à Marcia vient d'être, en quelque sorte, doublé par une concordance de la même œuvre faite par un membre de l'équipe du Professeur Grimal (Sorbonne Paris). Je signale dans le prochain numéro de L'Antiquité Classique les graves erreurs commises par l'auteur du travail dont M. Grimal a pris la responsabilité.

La plupart des erreurs qu'on découvre dans les Indices faits à l'aide de fiches manuscrites sont inévitables et imputables à l'humaine nature incapable de copier correctement un texte prolongé (les manuscrits sont là pour nous en convaincre ?), incapable aussi de compter sans se tromper. C'est la raison pour laquelle nous croyons qu'à l'exception de quelques travaux remarquables, tous les Indices existants devront être recommencés sur ordinateur.

Dans le domaine de la recherche mécanographique il semble bien qu'il ne se soit produit qu'une seule duplication de travail - encore est-elle de faible importance - entre le laboratoire de M. Charles W. Dunmore et mon propre Laboratoire : il s'agit d'un Index de l'Apocoloquintose qu'une de mes élèves a présenté comme travail de fin d'études.

Dans des cas semblables, il me paraîtrait intéressant que celui des deux Laboratoires qui désire publier l'Index reçoive en communication le fichier des cartes perforées de l'autre Laboratoire. Un tel échange assurerait un contrôle efficace des données et permettrait par suite une publication où ne subsisterait qu'un minimum d'erreurs.

Naturellement, cela postule un dessin de carte, un système de perforation et une codification symbolique identiques.

L'examen d'une telle standardisation des techniques est d'ailleurs un des buts de notre Organisation. Nous nous proposons dans le prochain numéro de cette Revue de décrire les méthodes et les techniques employées au Laboratoire de Liège et je souhaiterais que d'autres numéros soient consacrés à des descriptions semblables d'autres Laboratoires. La documentation ainsi réunie conduirait à des discussions fructueuses et permettrait sans doute, à la longue, d'adopter une langue unique pour le traitement de l'information dans le domaine philologique. C'est mon plus cher souhait.

✘ ✘ ✘

Tout ceci m'amène à dire un mot, à cœur ouvert, des méthodes qu'il conviendrait de discuter dans cette Revue et des buts plus lointains que nous devrions atteindre. Les idées exprimées ici sont celles du Laboratoire de Liège : il est bien évident que nous serons heureux d'avoir l'avis de tous les membres de l'Organisation.

Je commencerai par quelques observations sur le travail de dépouillement des textes.

- 1°) Compte tenu des grandes facilités qu'offrent les machines, il serait regrettable de constituer des Indices sans y ajouter des listes de fréquence où tous les mots du texte étudié seraient repris en ordre de fréquence décroissante.

De telles listes sont et seront au moins aussi utiles que les Indices proprement dits. J'en donne un exemple dans les pages qui suivent.

- 2°) Dans la méthode mise au point au Laboratoire de Liège, une carte est consacrée à chaque mot du texte étudié, parce que nous avons décidé de faire une analyse morphologique et syntaxique de chaque mot, ainsi qu'un début d'analyse stylistique qui consiste à indiquer la place de certains mots importants de la phrase (sujet, verbe, complément d'objet direct, épithète) par rapport au contexte restreint dont ils font partie.

Une telle étude est lente, mais elle est efficace, car ayant réuni en moyenne une dizaine de renseignements différents à propos de chaque mot, nous pourrons dans la suite explorer la langue de l'auteur sous de nombreux points de vue. En particulier, nous espérons que cette technique nous aidera à résoudre, mieux qu'on ne l'a fait jusqu'aujourd'hui, les grands problèmes d'authenticité et de chronologie relative que posent certains textes anciens.

- 3°) Nous ne voyons pas l'intérêt d'établir des Concordances, c'est-à-dire des indices où chaque occurrence du mot est accompagnée d'un contexte plus ou moins étendu. La raison principale de cette option est qu'il est impossible de déterminer à priori quel est le contexte qui permettra de comprendre le sens du mot. De sorte que, quel que soit le nombre de mots que l'on demandera à la machine d'imprimer avant et après le mot envisagé, la concordance sera ou bien trop explicite, là où le contexte est sans intérêt, ou bien insuffisante, dès que le mot recouvre une notion un peu com-

plexe. J'ajouterai que les concordances sont dangereuses dans tous les cas où le lecteur peut avoir l'illusion de comprendre le mot à l'aide du seul contexte qu'on lui fournit.

Malgré cette prévention contre les concordances, nous avons cependant constitué un programme pour notre ordinateur (IBM 1620) qui permet de réaliser n'importe quel type de concordance : la machine imprime le nombre de mots qu'on lui demande 1, 2, 10, 50 ou 100 mots avant et après le mot intéressant elle peut réaliser aussi une concordance par vers ou par distique, s'il s'agit de poésie, par chapitre, par paragraphe ou par phrase s'il s'agit de prose etc. En effet, il peut être intéressant pour une étude particulière, celle d'un concept par exemple, de retirer d'un fichier tous les passages où l'on trouve ce mot, avec un contexte aussi étendu qu'on le désire; il peut aussi être intéressant pour une analyse purement formelle - au niveau structural - de demander un contexte très restreint : un mot avant, un mot après tel substantif ou tel verbe, etc. Ce programme, comme tous ceux que nous avons réalisés, est écrit en SPS et est évidemment à la disposition des membres de l'Organisation.

Par ailleurs, la constitution d'Indices, même accompagnés de relevés statistiques, ne me paraît être qu'une étape : la réunion d'une documentation en vue de recherches approfondies dans les domaines de la philologie, de la psychologie, de la stylistique et de l'enseignement des langues.

Dans le domaine de la philologie classique, beaucoup de chercheurs de ma génération - et c'est plus vrai encore des jeunes générations - pensent que les études traditionnelles ont atteint un point de saturation au-delà duquel on ne fait plus que tourner en rond, reprendre indéfiniment les mêmes problèmes, sans arriver à une solution, sans même l'espérer, puisqu'aucun élément extérieur (sauf dans des cas très rares) ne peut venir étayer indiscutablement une hypothèse.

Les études que nous entreprenons, qu'on les appelle linguistique appliquée, analyse quantitative, statistique littéraire, etc. apportent, elles, des éléments nouveaux.

Mon père, Armand Delatte, que d'admirables recherches ont conduit du pythagorisme ancien à la magie byzantine, du commentaire de texte à la description des gemmes antiques et qui possédait une merveilleuse connaissance du grec et du latin, expliquait volontiers à ses élèves quelques Lettres à Lucilius de Sénèque. Dans les dernières années de sa vie, il aurait souhaité entreprendre une édition des Lettres et lui, qui connaissait cependant admirablement le philosophe stoïcien ne cessait de regretter l'absence d'un index complet de Sénèque dont l'utilité, dans ce cas, lui paraissait évidente, aussi bien sur le plan des idées pour mieux cerner un concept que sur le plan du vocabulaire employé par l'auteur, pour corriger une leçon fautive.

C'est probablement le rayonnement profond qui

émanait de sa personne qui nous a conduits à commencer nos recherches par l'étude de Sénèque et qui, dès la création du Laboratoire, nous a amenés à ajouter à nos Indices les relevés statistiques importants auxquels je viens de faire allusion et qui concernent la langue et la tradition manuscrite.

L'utilité des Indices et d'une description statistique d'une œuvre apparaît évidente aussi lorsqu'il s'agit d'étudier l'évolution de la langue.

Un de nos membres a entrepris une étude grammaticale de Cassiodore. Il me semble qu'il trouverait une aide précieuse s'il pouvait comparer la langue de Cassiodore à celle de Sénèque. Or c'est un fait d'expérience personnelle que la présence de certains phénomènes dans une œuvre n'est mise en lumière, assez paradoxalement, que par la comparaison avec des textes où ne se manifestent pas ces phénomènes. C'est donc surtout par une confrontation en ordinateur des deux fichiers de cartes perforées de Sénèque et de Cassiodore que l'on aboutira à des découvertes tant sur le plan du vocabulaire - préférences, exclusives de certains mots - que sur le plan de la syntaxe où les constructions ont naturellement évolué beaucoup au cours des siècles.

Il en ira de même lorsqu'on abordera l'étude des



grands problèmes d'authenticité et d'évolution de style. Les critères appliqués jusqu'ici à la solution de ces problèmes ne me paraissent guère pertinents, en raison de l'influence perturbatrice que peuvent exercer sur le style d'une œuvre des éléments extérieurs à l'individualité de l'auteur, par exemple la personnalité de l'interlocuteur : Sénèque n'écrira pas de la même façon s'il s'adresse à sa mère ou s'il s'adresse à l'empereur.

Au surplus, notre expérience dans l'analyse statistique d'une œuvre littéraire n'est pas suffisante pour décider si l'appréciation de la valeur des écarts par rapport à une norme ou par rapport à une autre œuvre doit reposer sur les mêmes critères statistiques que ceux qui sont employés dans le contrôle d'une fabrication. Avant d'arriver à une méthode sûre, il faudra procéder à de nombreux essais et à de nombreuses vérifications que seule permettra l'utilisation massive des Indices et des relevés statistiques.

Dans le domaine des recherches psychologiques consacrées aux auteurs anciens - et ce domaine est encore, en vérité, de la philologie bien comprise -, l'intérêt des Indices et des listes de fréquence est considérable. J'ai eu l'occasion de l'observer dans une étude des mots-clés chez les élégiaques Tibulle et Propertius. Cette étude, fondée sur une méthode suggérée par M. P. Guiraud pour les auteurs français, consiste essentiellement à comparer la liste de fréquence d'un auteur à celle du vocabulaire de la langue. La comparaison permet de dégager une liste

de mots qui ont une fréquence anormale chez l'auteur par rapport à l'usage de la langue. Ces mots sont donc propres à l'auteur considéré et sont révélateurs sur le plan psychologique. J'ai complété la méthode en confrontant les mots de Tibulle au vocabulaire de Propertius : comparaison fructueuse puisque tous deux appartiennent au même genre littéraire. Grâce à cette technique, on est conduit tout naturellement à dresser le profil caractérologique d'un auteur, à mettre en évidence certains thèmes poétiques, à mieux comprendre le sens de certains mots souvent employés et, à ce titre, chargés d'affectivité.

Les études stylistiques bénéficieront aussi de l'abondance des relevés statistiques. Si l'on adopte la définition du style comme étant une déviation, un écart par rapport à une norme, il faudra définir sur le plan formel quelle est la déviation personnelle de chaque auteur, dans quel domaine préférentiel elle s'exerce; pour y arriver, une fois encore, il faudra comparer, à différents niveaux, de nombreuses œuvres les unes aux autres.

Enfin, en ce qui concerne l'apprentissage des langues anciennes, nous nous sommes aperçus qu'on pouvait en rationaliser l'étude en concevant d'une part un vocabulaire de base général scientifiquement établi sur les listes de fréquence des principaux auteurs de la latinité et des microlexiques pour chaque œuvre que l'on souhaite lire dans les

classes, en établissant d'autre part une grammaire fréquentielle où les phénomènes sont répartis selon la fréquence avec laquelle ils apparaissent dans les textes.

✕ ✕ ✕

On aura noté que j'ai insisté souvent dans les pages qui précèdent sur la nécessité d'obtenir rapidement un grand nombre d'Indices d'œuvres antiques accompagnés de listes de fréquence et de relevés statistiques sur la morphologie et la syntaxe. C'est que la plupart des études que l'on peut imaginer à partir de ces relevés supposent des comparaisons.

Or, notre expérience nous a montré que, quelle que soit la rapidité des machines, lorsqu'on veut codifier l'analyse morphologique et syntaxique et dans une certaine mesure, stylistique de chaque mot, il faut y consacrer un temps considérable. En réalité, un philologue ne peut guère analyser et codifier plus de 250 mots par jour; au-delà de ce chiffre, apparaissent les erreurs de distraction dues à la fatigue. Cette lenteur nous a paru archaïque en un temps où les ordinateurs sont doués de logique et dotés de prodigieuses mémoires. Nous avons donc décidé de faire un programme d'analyse automatique de la morphologie et de la syntaxe des textes latins : ce programme est opérationnel à Liège

depuis quelques jours. Il permet d'analyser, avec l'aide du philologue pour certains mots ambigus, 3.000 mots par jour. On trouvera dans les pages qui suivent une description sommaire du programme auquel nous consacrerons un numéro spécial de cette revue.

Les études qui sont ici suggérées ne sont qu'une faible partie des immenses possibilités que nous ouvre l'utilisation des ordinateurs. Il me paraît cependant indispensable de noter, pour qu'on ne nous accuse pas de croire au miracle, que les machines ne font que nous apporter une description correcte, précise et rapide de tous les phénomènes linguistiques possibles. Mais elles se limitent à cela, ce qui est déjà remarquable. Elles sont incapables de répondre à des questions pour lesquelles on ne leur aurait pas fourni les éléments de la réponse.

Je le dis parce que je suis très frappé par la foi incroyablement infantile que certains hommes de science - dans le domaine des sciences humaines - mettent dans les machines : ils s'imaginent que les machines font tout le travail, ont réponse à tout, qu'il suffit de poser une question : la réponse arrive, rapide et infaillible. Les bandes dessinées ont contribué beaucoup à répandre la croyance que les machines sont intelligentes !

Peut-être est-il bon de rappeler que c'est l'homme qui est intelligent, pas la machine; que le travail de ceux qui ont un

ordinateur à leur disposition ne consiste pas seulement à pousser de temps à autre sur un bouton; qu'enfin, lorsque sortent les résultats d'une imprimante ou d'une tabulatrice, alors seulement commence le vrai travail philologique.

x x x

Les méthodes et les techniques qui sont décrites ici sont aussi d'application pour les langues autres que le latin; je pense au grec et aux langues modernes. Ce m'est l'occasion de souhaiter la bienvenue aux membres de cette Organisation qui ne sont pas des philologues classiques et qui nous ont cependant fait l'honneur de s'inscrire.

En ce qui nous concerne, dans le même temps où nous continuons nos travaux sur Sénèque - analyse des Lettres à Lucilius - nous entreprenons un programme d'analyse morphologique et syntaxique pour le français et un pour le grec.

x x x

Je voudrais dire en terminant un mot plus personnel concernant mon Laboratoire.

Depuis la création du Laboratoire d'Analyse statistique des Langues anciennes de l'Université de Liège, je n'ai cessé d'écrire que ces nouvelles recherches doivent être le fait d'une équipe, que les chercheurs isolés peuvent difficilement soutenir la cadence qu'imposent les machines, qu'une étroite collaboration au niveau de l'équipe est indispensable avec ce que cela comporte de renoncement, mais aussi avec ce que cela comporte d'exaltant.

Je suis heureux d'avoir pu réunir un telle équipe: elle est composée de mon chef de travaux, M. Et. Evrard, de deux assistants, MM. A. Bodson et J.M. Moitroux, d'un aspirant au Fonds National de la Recherche Scientifique, Melle S. Govaert, et de moi-même. Sans doute, est-ce une équipe restreinte mais - on me pardonnera de le dire - elle possède un tel enthousiasme, elle a une telle foi dans ce qu'elle fait qu'elle compense largement son infériorité numérique par son ardeur à un travail qui, au niveau où nous sommes parvenus, est devenu un plaisir.

Toutes les idées qui sont exprimées ici sont le fruit d'entretiens quotidiens et de discussions parfois passionnées, jamais stériles. En quatre ans - la création du Laboratoire date de septembre 1961 - l'équipe a fait beaucoup de chemin, depuis l'unique perforatrice qu'elle possédait. Ce qu'elle a créé de plus beau est cet esprit qu'autour de nous, on appelle Lasla.

L. DELATTE