

## EPIGRAPHIE ET INFORMATIQUE

Consacrée aux "*Applications à l'Épigraphie des Méthodes de l'Informatique*" (1), une *table-ronde* du *Centre National de la Recherche Scientifique* a été organisée à Marseille le 8 décembre 1972 par l'*Institut d'Archéologie Méditerranéenne* et l'*Unité de Recherche Analyse Documentaire et Calcul en archéologie*, avec la collaboration de l'Université de Provence. A l'origine, le but de cette réunion était triple. Il s'agissait d'une part de recenser les problèmes rencontrés par les épigraphistes dans l'étude et l'exploitation de leur matériel, d'autre part de faire apparaître un appel à l'informatique comme un recours naturel, enfin d'examiner les solutions proposées au terme de travaux entrepris depuis 1969 par un groupe de chercheurs de Marseille et d'Aix-en-Provence, en vue d'une exploitation automatique d'un Corpus d'Inscriptions latines. L'annonce de ce colloque souleva un certain intérêt, des participations nouvelles furent proposées et l'objectif initial en fut quelque peu modifié. En fait, quatre méthodes d'exploration de corpus épigraphiques furent proposées à une cinquantaine de participants, et, dépassant quelque peu les simples problèmes documentaires, quatre exposés évoquaient les possibilités offertes par les moyens de calcul d'atteindre à des résultats que les méthodes traditionnelles pouvaient difficilement espérer.

La densité du programme ainsi fixé n'a malheureusement pas permis, sauf sur certains points précis, d'aboutir à des conclusions assurées. A la confrontation entre épigraphistes et informaticiens s'est juxtaposée une confrontation entre différents procédés et systèmes. Les débats en souffrirent.

Néanmoins quelques lignes de force se sont dessinées au cours des discussions. Il est probable que, malgré ces difficultés, les principes d'une utilisation rationnelle des méthodes informatiques ont été mieux compris par les participants. Il est apparu très clairement en particulier que la technique offrait la possibilité d'atteindre à des objectifs, et que ces objectifs devaient être définis par les épigraphistes. Reste à savoir si les moyens matériels seront donnés pour mettre en oeuvre la méthode qui apparaîtra la mieux adaptée à des travaux particuliers ou à une refonte globale de la documentation épigraphique. Avant de résumer les diverses communications, il faut encore préciser qu'il a surtout été question à Marseille d'épigraphie latine, et à un seul moment d'épigraphie grecque. Cela ne veut pas dire d'ailleurs que certaines des méthodes utilisées ne puissent être transposées dans d'autres domaines culturels.

#### *PROBLEM AND PROSPECTS FOR THE PRODUCTION OF COMPUTER COMPILED INDICES TO EPIGRAPHIC WORKS (2)*

Le Dr. E. J. JORY n'ayant pu assister à la *table-ronde*, le texte anglais de sa communication fut distribué et un résumé en fut fait en français. L'auteur a rendu compte des raisons qui l'ont incité à fabriquer les *Indices* du Tome VI du *C.I.L.*, des difficultés qu'il a rencontrées dans ce travail, et a signalé les extensions possibles qu'il voit à l'utilisation de l'informatique.

Il est certain que les *indices* de ce volume du *C.I.L.* n'ayant jamais été publiés, un outil essentiel puisqu'il devrait recenser les inscriptions trouvées à

Rome fait cruellement défaut aux épigraphistes. Cela aussi bien en raison du grand nombre des textes (près du quart de l'ensemble de ceux réunis dans la totalité des volumes du *C.I.L.*) que de leur importance qualitative. Evidemment, une expérience réussie sur une collection aussi variée a toute chance de pouvoir être étendue et transposée.

L'ordinateur utilisé par le Dr. JORY est un PDP 6 (Digital Equipment Corporation). Un code a été mis au point pour traduire les caractéristiques physiques des inscriptions. D'abord enregistrée sur bande perforée, l'information est transférée sur bande magnétique et grâce au programme KWIC (Key Word In Context) on obtient un total imposant de 7.315 pages de 90 lignes, chaque ligne correspondant à un "mot-clef" et à son contexte. Dans le procédé KWIC, les mots-clefs, rangés par ordre alphabétique, apparaissent au milieu de la page, flanqués d'environ 90 caractères de contexte, avec la référence de l'inscription et le numéro de ligne. L'index est divisé en quatre rubriques :

*Vocabula*  
*Cognomina*  
*Vocabula acephala*  
*Graeca*

Les problèmes rencontrés ont eu pour cause d'abord le manque de régularité dans l'emploi des signes diacritiques par les éditeurs du *C.I.L.* Un très long travail d'uniformisation a donc été rendu nécessaire. Ensuite, le matériel utilisé, malgré des avantages (fonctionnement en time-sharing avec accès multiples), présentait des inconvénients dûs à la déficience des périphériques. Le support de l'information utilisé à l'avenir sera sans doute la carte perforée,

assurée d'une meilleure conservation. Enfin, l'édition étant assurée par des procédés de reproduction photographique des *outputs*, une bonne impression était nécessaire et imposait de fréquents changements des rubans de machines, un papier et un format spéciaux.

Pour terminer, le Dr. JORY, en tenant compte des impératifs scientifiques et économiques en cause, indique un certain nombre de directions de recherche possibles : fabrication d'*indices* du vocabulaire avec ou sans contexte, pour tous les catalogues existants; extension du type de service offert par l'Année Epigraphique (l'auteur souligne que ce projet serait le plus rentable et le plus utile); création enfin d'une "Banque de Données Epigraphiques" dans la mesure où une collaboration internationale est acquise. Les moyens techniques dont le Dr. JORY dispose à Nedlands rendent théoriquement possibles l'établissement et la gestion de cette banque de données.

### **UNE EXPERIENCE DE TRAITEMENT AUTOMATIQUE DES INSCRIPTIONS LATINES A L'UNIVERSITE DE LIEGE (3)**

Sous ce titre, le Professeur Et. EVRARD rend compte des travaux entrepris au *Laboratoire d'Analyse statistique des Langues anciennes* de l'Université de Liège, créé et dirigé par le Professeur L. DELATTE. L'équipe liégeoise bénéficie d'une déjà longue expérience pour ce qui concerne les textes littéraires anciens. Elle pense pouvoir l'utiliser moyennant certains aménagements aux textes documentaires (Papyrus et inscriptions). M. EVRARD

insiste sur le fait que les méthodes utilisées limitent la mise en forme préalable des textes épigraphiques au strict nécessaire. Il souligne que l'intrusion des codages alourdit les opérations de traitement et que leur présence ne se justifie que par un rendement supérieur.

La première expérience "épigraphique" menée par le laboratoire liégeois concerne un groupe relativement restreint d'inscriptions (à peu près 170). Le fichier est organisé par mot, chaque carte perforée portant de surcroît une série de descripteurs (références, lexicographie, morphologie, etc.). L'ordinateur assure automatiquement une série d'opérations de référencement. Il propose pour chaque mot, par comparaison avec un lexique et une table des désinences préalablement enregistrés, un ou plusieurs *lemmes* et analyses morphologiques possibles. Le choix revient au philologue qui doit assurer aussi un certain nombre d'opérations de correction et d'ajustement. Le résultat de ce travail est "réinjecté" en machine et l'ordinateur produit :

- un index des mots
- une liste des lemmes rangés en ordre de fréquence
- des relevés statistiques concernant la morphologie
- une liste des verbes.

Pour les documents épigraphiques, des conventions spéciales ont été adoptées pour indiquer soit la "mise en page de l'inscription", soit les "abréviations et restitutions". De plus, l'introduction sur les cartes perforées de nouveaux descripteurs autorise la sélection des mots en 8 classes. (Divinité, *Praenomen*, *Cognomen*, gentilice, magistrature, et fonctions, tribu, Nom d'empereur, *geographica*). Différents programmes peuvent être utilisés sur

l'ensemble informationnel ainsi constitué et en particulier produire une liste des mots accompagnés de leur contexte. La méthode est un peu différente de celle du Dr. JORY, dans la mesure où le format du contexte n'est pas défini par un nombre de caractères fixé à l'avance, mais peut être modulé en utilisant comme "bornes" n'importe lequel des descripteurs portés sur la carte-mot de façon à produire un fragment du contexte dont on pense qu'il révélera l'important, c'est-à-dire la relation, implicite ou explicite, entre deux termes.

En terminant, M. EVRARD insiste sur les facilités qu'a l'automate de tenir à jour un corpus d'inscriptions latines. L'ordre d'entrée des textes n'influe pas sur les réponses. Il souligne d'autre part que l'édition, de ce point de vue, ne présente plus qu'un intérêt limité. Ainsi conçue, l'opération est beaucoup moins coûteuse qu'on ne pourrait le croire.

*NOTES POUR UN PROGRAMME DE REEDITION ET DE MISE A JOUR DU  
VOLUME V DU CORPUS INSCRIPTIONUM LATINARUM COMPORTANT  
UNE REDACTION AUTOMATIQUE DES INDICES (4)*

Déjà par son titre, la communication de nos collègues italiens se démarque quelque peu de celles qui l'ont précédée. La mécanisation intervient pour fabriquer des indices et sur un matériel entièrement révisé, l'objectif étant la publication des inscriptions traditionnellement recensées dans le volume V du *C.I.L.* Cette révision doit aussi assurer l'adéquation aux visées actuelles

de la science épigraphique. Ainsi, outre l'édition, indispensable intermédiaire qu'ils nous promettent abondamment illustrée, les auteurs proposent de mettre à la disposition de ceux qui en auront l'utilisation, les "archives" enregistrées sur bande magnétique. L'intérêt de l'opération sera singulièrement multiplié avec l'extension du domaine géographique pris en compte. L'idée d'une "banque de donnée" est là aussi présente. Cette banque pourrait être gérée et contrôlée par une association internationale d'Epigraphie. Les indices fabriqués à Milan adoptent la structure de ceux des *Inscriptiones Latinae Selectae*. On pourrait sans peine y adjoindre un *index verborum*, première ébauche d'un dictionnaire du vocabulaire épigraphique. Toute classification figée des textes est abandonnée; rares sont en effet les documents susceptibles d'une seule caractérisation (exclusivement funéraire, militaire ou administratif). Les auteurs exposent ensuite les principes du système POLEMON I.

1. L'input se fait à l'aide d'un lecteur magnétique I.B.M. Chaque document se compose de trois parties :
  - le *contexte* (5) comprend une description sommaire des monuments épigraphiques et des informations utiles à leur organisation ultérieure (localisation, date, références bibliographiques, etc.);
  - le *texte* est entrecoupé d'une série de *codes* désignant le début et la fin d'un segment et indexant ce segment par référence aux chapitres de l'index. Un code différent indique l'appartenance à une sous-rubrique et le format du segment devant apparaître dans l'index;
  - le *système de transcription* abandonne bien sûr toute tentative de "représentation figurée" et se fonde sur une division très nette entre les traits attestés ou sûrement restituables et tout ce qui est interprétation, grâce à l'utilisation des capitales ou des bas de casse. Des signes diacritiques

indiquent les particularités graphiques, les restitutions, etc.; les italiques sont réservées au commentaire.

### *Le langage de programmation*

C'est le langage APL, particulièrement performant dans le domaine de la linguistique qui a été adopté.

Au moment de la mise en mémoire des données, trois problèmes se sont posés. Celui de la transcription des textes et plus particulièrement de la présence des signes diacritiques traduisant des représentations graphiques, un état de conservation, une interprétation : grâce aux méthodes utilisées, il a été possible d'utiliser des signes assez proches de ceux utilisés normalement en épigraphie et de répondre aux besoins du système de transcription. L'introduction des données se fait sous le contrôle d'un programme destiné à relever les erreurs formelles.

L'élaboration des données comprend essentiellement une mise en ordre, selon différents critères, de provenance par exemple, ou d'autres choisies dans la partie "référence" des documents, et une préparation du texte en vue de l'édition. L'impression se fait directement en offset grâce à une photocomposeuse. L'introduction de nouveaux documents ne pose aucun problème. Les résultats présentés par l'équipe de Milan sont très élaborés sur le plan de la présentation et tout à fait acceptable par l'épigraphiste le plus traditionaliste. La publication d'un premier fascicule contenant les 1.500 inscriptions de Milan et de la région ne saurait tarder à paraître.

## *UN SYSTEME D'EXPLOITATION AUTOMATIQUE DU CORPUS INSCRIPTIONUM LATINARUM : LE SYCIL (6)*

En raison des principes définis au début de ces travaux, les objectifs de l'équipe d'Aix-Marseille ont pu paraître très ambitieux aux participants de la réunion. Il a semblé en effet peu raisonnable de faire "moins bien" que le *C.I.L.* Au contraire, les nécessités actuelles de la recherche imposent un accroissement considérable du champ de l'analyse. C'est ainsi que le SYCIL réserve une large place à l'archéologie du monument épigraphique, intégrant le plus grand nombre possible de ces caractéristiques morphologiques et des renseignements sur la situation et l'histoire de l'inscription. Malheureusement, cette démarche est limitée par l'état présent de la documentation. Comment peut-on, par exemple, fixer avec précision le cadre réservé à une analyse paléographique ? Cette insuffisance de la documentation, ainsi que le souci de rendre possible des recherches ultérieures relativement imprévisibles, a conduit à la conception d'un système très ouvert à toute sorte de perfectionnement dans toutes les directions.

Un lexique documentaire a donc été étudié. Il comprend outre le chapitre "Contexte" et un chapitre "Thématique", destiné à recenser les caractéristiques générales des textes de façon non exclusive, deux chapitres destinés à l'analyse, l'un du "contenu", l'autre de la "stylistique". Ces chapitres, organisés de façon hiérarchique en rubriques et sous-rubriques, rassemblent une série de "descripteurs" ou "codes" de quatre signes au maximum. Il va de soi que les termes de ce lexique ne sont pas destinés à une traduction mot-à-mot du vocabulaire épigraphique, mais qu'ils dénotent l'appartenance à une catégorie lexicale (chapitre, rubrique, etc.) d'un "segment" de

texte, comptant un ou plusieurs mots. Ces descripteurs vont donc essentiellement servir à la fabrication des indices prévus. Cependant, la compréhension d'un segment peut imposer sa mise en relation avec un ou plusieurs autres segments qui peuvent ne pas lui être accolés, mais se trouver assez loin dans le texte. Le système SYCIL assure la mise en relation des segments par un jeu de liaisons syntactiques. Ainsi, non seulement les segments et les mots, mais aussi les relations elles-mêmes (entre personnes, entre personnes et fonctions, entre fonction et lieux d'exercice, etc.) peuvent-elles faire l'objet d'une recherche explicite.

Au moment de l'entrée des données, des "marques formelles" traduisant l'état des lettres, indiquant les lacunes et les restitutions, sont placées dans le texte. Selon le programme d'"édition" choisi, ces marques et le texte qu'elles encadrent sont traités de façon à obtenir différents états du texte. Deux de ces états correspondent l'un à la "lecture", l'autre à l'"interprétation" traditionnelles. Un autre état, ramenant tous les mots à une forme régulière (nominatif singulier pour les noms, infinitif pour les verbes, etc.), permet de répondre aux préoccupations d'ordre lexicologique. D'autres états non prévus initialement pourront être fournis à la demande, pour des recherches très particulières, sur les abréviations par exemple.

Les *indices* et les tables, tels que ceux fournis par *C.I.L.*, apparaissent dans le système de consultation comme un produit possible parmi une infinité d'autres, intégrant comme éléments de recherche non seulement les caractéristiques textuelles, mais aussi les données "archéologiques" des inscriptions. Le format unique de consultation prend en compte ces données choisies et organisées, définit l'état du texte souhaité et donne les directives

d'édition et de mise en page.

\*

\*

\*

Ces exposés, et les discussions qui les ont suivis, ont fait ressortir deux conceptions assez éloignées, pour ne pas dire opposées, de la recherche documentaire en épigraphie. Selon les uns, c'est, en gros, à l'épigraphiste à faire son métier sur des textes préalablement enregistrés; on s'en remet, pour éclairer le sens des mots, à un *contexte* dont le format est fixé arbitrairement ou modulable selon certaines règles. Selon les autres, qui ont pris comme référence les *indices* traditionnels, les mots, les groupes de mots ou les concepts doivent être notés comme appartenant à un ou plusieurs ensembles homogènes, lors même de l'introduction des documents dans le système documentaire. Cette attitude conduit nécessairement à l'utilisation de *codes*, ou *descripteurs*, dont l'ensemble constitue un lexique documentaire qui pourra être, le cas échéant, hiérarchisé pour permettre une recherche à différents niveaux. On comprend mal le point de vue de M. EVRARD qui considère que cette démarche n'est que le résultat d'une "fascination" produite par le codage. Tout dépend encore une fois des objectifs que l'on assigne à un système documentaire. Il ne s'agit même plus d'automatisation, mais de poser une question plus générale, et essentielle, sur la finalité de la documentation épigraphique et de son organisation.

Néanmoins, et puisqu'il était question d'informatique à Marseille, il faut savoir ce que l'on est en droit d'attendre de l'utilisation des moyens

modernes de traitement de l'information. Pourquoi donc une automatisation de la recherche documentaire en épigraphie ? Deux raisons très évidentes apparaissent immédiatement. Ainsi qu'il a été souligné dans toutes les communications, ces systèmes peuvent être sans problème tenus à jour. Ils pourraient aisément être mis en oeuvre sur de vastes collections, et être ainsi ce manoeuvre intellectuel dont le travail prolonge et multiplie la démarche intellectuelle. A ces deux motifs, on peut en ajouter un autre, moins sensible : une nouvelle mise en forme de la documentation épigraphique pourrait en rendre l'accès plus facile aux non spécialistes. Etre épigraphiste ne consisterait plus à "savoir chercher dans les indices du corpus".

Néanmoins, puisque l'épigraphie bénéficie d'une longue et solide tradition documentaire, il paraît assez vain de mettre sur pied aujourd'hui des outils qui ne seraient en mesure de rendre au moins les services offerts par le C.I.L. dans ses indices et ses tables. Bien sûr, l'utilisation constante, à domicile, d'un ordinateur permet de lire tout un corpus inorganisé à la recherche d'un mot, de rechercher le contexte de ce mot pour savoir s'il appartient bien à la catégorie recherchée et de reprendre ce travail sur chacun des mots de cette catégorie. Pensons, par exemple, au nombre de "mini-recherches" qu'il faudrait poursuivre pour venir à bout d'un problème simple : production d'un index des noms de militaires d'une région donnée. Il faudrait d'abord faire l'inventaire des termes désignant un militaire; ensuite obtenir la production de tous les "contextes" de ces termes, vérifier s'il s'agit bien de soldats (un *décurion*, un *princeps* ne sont pas toujours des militaires) et trouver les noms des personnages ainsi désignés. Il est beaucoup plus simple et efficace d'accoler à tous les termes désignant sûrement un militaire, un descripteur dénotant, entre autres, cette qualité et d'explicitier, grâce à des liaisons syntactiques, ou, si c'est possible

par le jeu d'imbrications, la relation entre ce mot et le nom du militaire. C'est là, sans doute, faire au préalable le travail de l'épigraphiste, ou du moins une partie de son travail, la mise en forme des données documentaires et l'élucidation des rapports qui les unissent. Le lexique documentaire est donc l'outil indispensable pour arriver à produire des *indices* comparables à ceux du *C.I.L.* Il est tout à fait clair que les *indices* du *C.I.L.* ne sont même qu'une partie de ceux qu'il est possible et légitime d'escompter en combinant les termes du langage documentaire d'une façon différente de la façon traditionnelle, en vue d'obtenir des produits mieux adaptés à telle ou telle recherche. C'est à ce point qu'il convient de s'interroger sur les éléments qui doivent être analysés en fonction des besoins actuels de la recherche, et de leur extension, plus ou moins prévisible. De plus en plus, on s'attache aux aspects "archéologiques" des inscriptions et à leurs caractères graphiques et de mise en page. L'ennui est que les démarches préalables à une formalisation de ces données font cruellement défaut à un niveau suffisamment général. Par contre de nombreuses études plus particulières montrent tout l'avantage qu'on peut attendre de la prise en compte des traits contextuels. D'ores et déjà, il est évident que ces données ne prendront tout leur sens que dans la mesure où elles seront tabulées de la même manière et en même temps que les données textuelles. On peut prévoir, par exemple, que l'on recherchera la co-occurrence d'une forme de lettre, d'une origine géographique et de caractéristiques thématiques, si l'on essaie de mieux définir les frontières de "provinces épigraphiques" et si l'on considère que cela peut révéler des phénomènes culturels. Le document épigraphique est beaucoup trop riche de renseignements pour qu'on en réduise l'analyse à son texte seul, rejetant les traits contextuels dans un "commentaire" imprécis et irrégulier. Il faut donc d'abord faire appel à tout le savoir de ceux qui sont concernés par la documentation épigraphique, et en premier lieu aux

“épigraphistes” eux-mêmes pour qu’ils expriment clairement le genre de produits qu’ils désirent. Cela ne peut se faire que par une collaboration très étendue, dont les conditions ne sont sans doute pas encore réunies. En attendant, il est raisonnable de réserver au moins la possibilité d’introduire dans le système documentaire les éléments d’analyse définis ultérieurement.

Dans la deuxième partie de la réunion de Marseille, la mise en oeuvre de moyens de calcul semble offrir la possibilité de choisir des critères “utiles” pour cette analyse et la nécessaire réduction qui en est le corollaire. Il s’agit là, dans l’ensemble, d’un domaine très nouveau pour ce qui concerne l’épigraphie, et les communications qui furent présentées à la Table-ronde ne sauraient être considérées que comme des expériences préliminaires, situées très en amont de la démarche aboutissant à une connaissance proprement historique.

#### *PROBLEMES DE LA STATISTIQUE DE L’AGE D’APRES LES INSCRIPTIONS FUNERAIRES ROMAINES (7)*

M. M. CLAUSS, dans un exposé très charpenté et qui, selon le mot de M. M. LEGLAY, “vient à son heure”, met en cause la validité d’une démarche tendant à s’appuyer sur les seuls documents épigraphiques pour élucider les problèmes démographiques de l’Antiquité. Il énumère six phénomènes susceptibles de perturber une étude statistique globale de l’âge moyen.

- Les indications d'âge sont souvent ramenées à une dizaine ou au chiffre 5.
- Selon les régions, la représentation relative des couches sociales varie considérablement. Certaines classes (militaires, esclaves) sont démesurément représentées, d'autres absentes ou du moins indiscernables.
- De même la représentation des différents groupes d'âges est très variable, particulièrement quand il s'agit de vieillards et d'enfants.
- Les femmes sont très rarement attestées.
- Les témoignages sont trop peu nombreux par rapport à la totalité de la population pour pouvoir être significatifs.
- La pratique épigraphique varie selon les régions.

M. CLAUSS conclut donc à l'inutilité d'une recherche démographique régionale ou étendue à l'ensemble de l'Empire. Une discussion animée s'engage entre MM. PELAUM, LEGLAY et ETIENNE. Il en ressort que, si une étude globale est effectivement impossible, on pourrait faire une "statistique différenciée" selon les groupes sociaux. Encore faudrait-il avoir la possibilité de déterminer des groupes sociaux suffisants selon les régions. D'autre part, les données anthropologiques que l'on peut obtenir lors des familles de nécropoles devraient compléter et vérifier les informations épigraphiques.

Le problème de la représentativité des inscriptions par rapport à la société dont elles émanent est l'objet de la communication suivante, présentée hors

des séances et qui offre, semble-t-il, une possibilité d'atteindre à cet égard des résultats suffisamment sûrs.

#### **APPLICATION DE LA STATISTIQUE INFERENCELLE A LA DETERMINATION DE LA REPRESENTATIVITE D'INSCRIPTIONS LATINES (8)**

Le travail de l'équipe de Valence expose précisément les difficultés rencontrées pour assurer avec une "suffisante" certitude que les inscriptions provenant d'une région ou d'une cité sont bien un échantillon représentatif, en termes de classes sociales, de la totalité des inscriptions qui y furent gravées, et dont une partie nous est inconnue. Cela ne tend pas à dire que cet échantillon reflète la réalité sociale de la région ou de la ville considérée.

L'équipe de Valence isole, par site étudié, deux groupes d'inscriptions, un groupe correspondant aux textes publiés dans le *C.I.L.*, l'autre les inscriptions "*post-C.I.L.*". Dans le cas où ces deux groupes peuvent être considérés comme des échantillons aléatoires, dûs au hasard des découvertes, leur comparaison doit révéler une certaine homogénéité, sous laquelle ils ne sauraient refléter le total originel.

Les résultats des tests statistiques donnent des résultats très différents selon les villes. Le plus souvent, l'hypothèse se trouve vérifiée, avec

plusieurs degrés de certitude. Parfois elle est refusée. Il convient de s'interroger sur les motifs de cette irrégularité : la rigueur du procédé mathématique n'est-elle pas de mise dans une discipline où l'on se contente souvent d'approximations très vagues ? Mais alors, ainsi que le souligne M. MENAUT, on peut, en épigraphie, démontrer à peu près n'importe quoi. Peut-être aussi, dans les cas où l'hypothèse a été rejetée, faut-il tout simplement mettre en cause les conditions de découverte des inscriptions. Nos collègues de Valence postulent en effet que les deux échantillons, "pré-" et "post-C.I.L.", sont aléatoires. Cela n'est pas assuré. On peut imaginer - et les exemples ne manquent pas de situations semblables - que les textes proviennent de différentes nécropoles utilisées par des classes sociales ou à des époques différentes, et que la découverte de ces cimetières et des inscriptions qu'ils contenaient n'a pas été simultanée, mais échelonnée dans le temps, sujette à des contingences comme l'extension de l'urbanisation moderne, l'acquisition et la fouille systématique, etc. Quoiqu'il en soit la méthode utilisée par M. MENAUT et ses collègues est le premier pas d'une démarche qui doit permettre d'assurer une base solide à un raisonnement ultérieur sur la représentation épigraphique de la société.

La statistique peut utiliser d'autres données que les données textuelles pour différents travaux tendant soit à proposer une chronologie, soit à définir des groupes sur le sens desquels l'historien aura à se prononcer. Notons que pour ces travaux, même un échec relatif est productif, dans la mesure où il incite à la prudence. Le calcul se fonde en effet sur l'explicitation des données fournies par l'épigraphiste. Il n'y a aucune raison pour qu'une démarche fondée sur des données implicites atteigne un niveau de crédibilité plus élevé.

**PALEOGRAPHIE GRECQUE ET SERIATION AUTOMATIQUE DES  
INSCRIPTIONS (9)**

Un exposé des travaux de Mme A. STEFAN, dont l'absence de Marseille a été regrettée, sera remplacé dans la publication des Actes par un article où elle résume ses études sur les inscriptions grecques d'Istrie. Considérant que, comme tout produit de l'activité humaine, la graphie des monuments épigraphiques est sujette à des influences géographiques et sociales, et à une évolution dans le temps, Mme STEFAN a utilisé pour proposer une séquence chronologique, la méthode mise au point par F. HOLE et M. SHAW pour classer des ensembles de données archéologiques. Les variables, dans le cas présent, sont les différentes formes que peuvent prendre les lettres dans les inscriptions sur pierre. La procédure est simple : à partir d'une *matrice d'incidence* où est signalée la présence ou l'absence de toutes les formes des lettres, l'ordinateur dresse une *matrice de similitude* et l'ordonne en permutant les lignes jusqu'à obtenir l'écart le moins important entre les lignes et dans l'ensemble de la matrice. La succession ainsi proposée traduit le degré de ressemblance des inscriptions et une évolution chronologique. De la même façon, on peut obtenir un tableau reflétant l'évolution d'un alphabet.

Mme STEFAN indique qu'il est nécessaire d'utiliser des inscriptions provenant d'un même site : les variations locales pourraient perturber les calculs. Par la suite, la comparaison des résultats provenant de différentes villes pourrait être intéressante. Il faudrait aussi déterminer des groupes selon le genre de l'inscription (publiques, privées, etc.). Il est nécessaire que les textes soient assez longs pour comprendre la totalité des lettres

de l'alphabet et il est préférable qu'ils proviennent assez régulièrement de périodes historiques longues. Un hiatus important dans la séquence ou une distribution irrégulière est cause d'erreurs.

L'auteur a considéré la lettre, dans la forme où elle apparaît, comme une variable en soi et n'a pas tenté une analyse au niveau du trait. Pour l'étude de la paléographie latine, il semble que cette deuxième solution doive être adoptée. En tout cas, le calcul permet une comparaison objective et sûre entre les différentes graphies des lettres et permet de prendre en compte toutes les variables aussi nombreuses soient-elles. On obtient une proposition de chronologie relative qu'il faut ensuite "accrocher" à une chronologie absolue fournie par d'autres moyens à propos de certaines des inscriptions ainsi analysées. On trouve là d'ailleurs un moyen de valider les propositions faites par le calculateur et, aussi à certaines occasions, une occasion de remettre en cause des datations effectuées par les épigraphistes selon des critères qu'il faudrait discuter.

C'est à peu près un résultat semblable qu'atteint une recherche menée à Marseille sur un corpus de textes latins, et en utilisant comme variables non plus les données paléographiques, insaisissables en l'occurrence, mais des caractéristiques du formulaire, de la morphologie (au sens grammatical), voire de la mise en page des inscriptions.

**DOSSIER SUR LES PROBLEMES SOULEVES PAR L'INTRODUCTION DES  
METHODES FORMELLES DANS LE RAISONNEMENT EN HISTOIRE (10)**

Il s'agit d'un dossier présentant un travail en cours, donc sans prétention d'aboutir à des résultats historiques. Un corpus de 60 inscriptions funéraires de vétérans d'Afrique du Nord a d'abord été analysé selon 51 des caractéristiques graphiques et sémantiques, dont il faut souligner qu'elles peuvent être produites directement par le système SYCIL. Le choix des caractéristiques a été fait de façon très empirique. Un premier travail a consisté à passer d'une grille descriptive I à 51 caractérisations à une grille descriptive II à 15 caractérisations. Il s'est agit d'éliminer certains traits, trop ou pas assez souvent attestés, de réduire des groupes de caractérisations à une seule donnée mieux utilisable. Un certain nombre des textes analysés étant datés par ailleurs, la suite du travail a consisté à tenter d'établir une séquence où ils auraient trouvé une place logique. Les calculs ont révélé une incompatibilité entre la chronologie absolue proposée par les épigraphistes et la chronologie relative fournie par le calcul matriciel. Par contre, ils firent apparaître un groupement compact des inscriptions en fonction de leur origine géographique.

L'hypothèse selon laquelle les textes en question se distinguaient nettement les uns des autres a été vérifiée au moyen d'un programme fondé sur l'analyse discriminante. Les résultats avec différents systèmes descriptifs sont illustrés par des graphiques montrant que les "centres" de chacun des groupes géographiques sont très éloignés les uns des autres. D'autre part, certaines inscriptions sont situées sur ces graphiques relativement près du "centre" d'un groupe auquel elles n'appartiennent pas.

Ces textes se trouvent appartenir pour la plupart à une zone géographique déterminée, mais non définissable en termes administratifs. On a là un indice permettant de proposer l'hypothèse d'une différenciation régionale superposée aux divisions provinciales. Sur un de ces groupes géographiques relativement homogènes, une nouvelle procédure de sériation a donné une séquence acceptable, à une exception près. Un des textes est placé avant une inscription de la fin du second siècle, alors que l'épigraphiste propose une date plus tardive. Il convient avant tout de s'interroger sur la valeur des critères qui ont conduit à cette datation tardive. Dans ce cas précis, il semble que le raisonnement soit appuyé au moins en partie sur des conjectures. De surcroît, on peut déterminer, grâce au calcul, les caractères les plus discriminants. Ce résultat est particulièrement important. Il devrait permettre lors de l'exploration de domaines mal connus de choisir les traits "utiles" qu'il convient de retenir pour analyser un certain matériel en fonction de certaines visées, et de ne pas retenir systématiquement la totalité de la documentation dans un souci assez vain par ailleurs d'exhaustivité.

\*

\*

\*

L'historien de l'Antiquité est sans doute plus traumatisé que ses collègues préhistoriens ou modernistes par cette intrusion du calcul sous toutes ses formes dans une discipline à la fois si traditionnaliste et si littéraire. "Il ne peut se défendre d'une sensation de vertige devant les perspectives qui s'ouvrent", a même dit l'un des assistants de la *Table-ronde*. Pourtant, il ne s'agit pas de perdre pied, même si certains des points d'appui dont il a

l'habitude lui semblent soudain chanceler. "Les fils de la Tradition ont besoin de beaucoup d'énergie pour ne pas laisser se scléroser l'organisme qu'ils ont à entretenir, pour le renouveler perpétuellement et l'adapter" (11). Aussi bien dans un projet strictement documentaire que pour des travaux plus sophistiqués, dont le but est en fin de compte de lui fournir des bases solides, l'historien reste à la fois le maître-d'oeuvre et l'arbitre, et sera le principal intéressé. A cet égard, les positions furent nettement définies à Marseille : c'est aux historiens d'écrire l'histoire.

C.N.R.S.  
Institut d'Archéologie méditerranéenne  
Université de Provence

Michel JANON

## NOTES

- (1) Les "Actes" de cette réunion seront publiés dans le Tome 9 de la Revue *Antiquités africaines*.
- (2) Dr. E.J. JORY, Senior Lecturer - Department of Classics and Ancient History. The University of Western Australia, NEDLANDS, Australie.
- (3) Prof. Et. EVRARD, Laboratoire d'Analyse statistique des Langues anciennes, Boulevard de la Sauvenière, 110, B-4000 LIEGE (Belgique).
- (4) Ce travail est le fruit de la collaboration de P. BALDACCI, O. IANOVITZ, E. MARETTI, C.-P. ZARRI, Istituto di storia antiqua, Università degli Studi di Milano, via Festa del Perdono, 7, I-20122 MILAN (Italie).
- (5) Le terme *contexte* prend un sens différent selon le cas. Ou bien il s'agit d'une partie du texte, contexte d'un mot ou d'une phrase, ou bien il s'agit de toute la documentation qui accompagne une inscription, contexte archéologique ou historique.
- (6) Ont participé à ces travaux MM. E. CHOURAQUI, J. VIRBEL (U.R.A.D.C.A. - CNRI, 31, Rue J. Aiguier, F-13274 MARSEILLE (France) Cedex 2), P. CORBIER et P.-A. FEVRIER (Université de Provence, 29, avenue R. Schuman, F-13621 AIX-en-PROVENCE (France)), M. JANON (I.A.M. - C.N.R.S., Université de

Provence, 29, avenue R. Schuman, F-13621 AIX-en-PROVENCE  
(France)).

- (7) Il s'agit d'un résumé d'un article paru dans la Revue *CHIRON*, Tome 3, 1973, pp. 395-417, sous le titre : *Probleme der Lebensalterstatistiken aufgrund römischer Grabinschriften*, et dû à M. M. CLAUSS, Ruhr-Universität, Buchenstrasse GA 6/157, D-463 BOCHUM QUERENSBURG (R.F.A.).
- (8) Traduction française de Joaquin AGUILALLA ALMER, Marco A. LOPEZ CERDA, Francisco MONTES SUAY (Departamento de Estadística Matematicas - Facultad de Ciencias de Valencia), Gerardo PEREIRA MENAUT (Departamento de Arqueologia, Epigrafia y Numismatica - Facultad de Filosofia y Letras de Valencia).
- (9) Mme A. STEFAN, Institut d'Archéologie, 11, rue I.C. Frimu, BUCAREST (Roumanie).
- (10) Dû à la collaboration à des titres divers de M. BORILLO, J. GUENOCHÉ, J. VIRBEL, M. JANON.
- (11) L. ROBERT, *Conférence inaugurale*, Actes du deuxième Congrès international d'Épigraphie grecque et latine (Paris 1952), Paris, 1953, p. 17.