

## DISCUSSIONS SUR D'AUTRES DICTIONNAIRES DE MACHINE

### LE PROJET DU VOCABULAIRE POLITIQUE GREC-LATIN ANCIEN ET SUGGESTIONS DU VOCABULAIRE ITALIEN DE LA CRUSCA

*Pennacini*

Je m'excuse d'attirer l'attention des machinistes sur des sujets qui ne touchent pas l'organisation d'un LEL, mais qui concernent l'emploi de l'ordinateur, quels que soient les caractères du LEL, à l'aide des recherches dans le domaine de la lexicologie.

L'équipe composée de cinq chercheurs dont je suis un membre, au nom de laquelle je fais cette intervention, est occupée, à l'Institut de philologie classique de la Faculté de Lettres et Philosophie de l'Université de Turin, dans une recherche sur la pensée politique ancienne, sous la direction de M. Italo Lana, prof. de Littérature Latine, avec l'aide financière du Conseil National des Recherches. Le but de la recherche est de recueillir tous les passages des auteurs grecs et latins, comprenant mots et locutions qui comportent soit une signification, soit un caractère politique. Pour simplifier le travail on a décidé la répartition de la recherche selon des critères chronologiques; nous avons donc établi de dresser en premier lieu un fichier homérique et d'en tirer un lexique; après on dressera également des fichiers et lexiques des auteurs compris dans des périodes historiques déterminées, par exemple VI et V siècle. la polis, VI - I siècle, l'âge hellénistique, l'âge romain, la latinité chrétienne. Les buts auxquels ces recherches visent sont les suivants:

- 1) Un fichier contenant tous les témoignages que les littératures latine et grecque nous ont transmis à propos de la pensée politique ancienne.
- 2) Des lexiques partiels de chaque auteur ou de chaque période historique.
- 3) Finalement un lexique de la pensée politique ancienne tout entière.

Méthodes: en ce qui concerne le travail sur les textes homériques, nous avons

adopté le procédé suivant : nous avons formé deux équipes ; l'une a lu directement les textes et en a tiré des morceaux qui ont une signification politique ; l'autre équipe a fait un choix des termes politiques parmi les mots réunis dans le lexique homérique de Ebeling. Avant de transférer dans les fiches les matériaux recueillis, une équipe les a vérifiés au moyen des concordances homériques. Afin de constituer le fichier, nous avons à ce moment-là comparé le résultat de deux travaux et ça a été la base de la constitution du fichier, qui est maintenant achevé. Chaque fiche contient un mot, tous les passages dans lesquels nous l'avons trouvé, les glosses, les explications des anciens, et dans une certaine mesure la littérature critique. Ce travail a duré deux ans ; toutes ces opérations ont été accomplies sans l'aide des machines. Nous avons employé seulement des machines à écrire soit en lettres grecques, soit en lettres latines. Le nombre des fiches que nous avons tirées des textes homériques est à peu près 1500 ; nous n'avons pas les éléments suffisants à déterminer l'exact montant de la dépense pour chaque fiche. Mais dans la perspective de continuer ce travail, se pose la question d'automatiser le choix des mots et l'élaboration des fiches pour gagner du temps, pour parvenir à un recueil complet et précis autant qu'il est possible, et rendre aussi la tâche des chercheurs moins lourde. Alors la question que je pose est semblable à la question posée par le Père Valentini. Nous demandons donc aux chercheurs et aux techniciens ici réunis de nous suggérer des solutions possibles de ce problème.

*Busa*

Je pense que parmi nos collègues d'ici, le mieux préparé est le Professeur Duro, qui a une expérience très large et très souple du dépouillement de données lexicographiques.

*Duro*

Pour répondre à la question que nous pose M. Pennacini, il faut dire d'abord que c'est surtout une question de financement et d'organisation. Vous avez en effet une grande masse de textes à examiner pour en extraire un lexique spécial, le lexique de la terminologie politique. Chaque recherche d'un lexique spécial exige la lecture d'un certain nombre de textes pour choisir des mots et des exemples : ce choix peut être facilité si nous disposons de concordances. Mais notre expérience nous conseille de ne pas faire recours à la machine, si nous n'envisageons pas d'extraire d'un texte au moins un mot, c'est à dire un exemple, pour chaque ligne. Or, vous avez dit, si j'ai bien en-

tendu, que de toute l'oeuvre d'Homère vous avez extrait seulement 1500 mots qui concernent le lexique politique.

*Pennacini*

600 mots dans 1500 fiches.

*Duro*

Disons alors 1500 exemples. Et quel est le nombre des mots, ou encore mieux des occurrences, dans les poèmes d'Homère? Si vous faites un pourcentage entre les fiches que vous avez extraites et le nombre des occurrences en Homère, vous verrez qu'il ne vous convient pas d'en faire une analyse électronique. Vous pouvez faire une extraction d'exemples par quelque autre procédé, qui peut être la méthode traditionnelle de la fiche à la main, ou, si vous désirez un procédé plus moderne, à l'aide de la reproduction par xérogaphie. Le recours aux moyens électroniques serait trop cher, vu le rapport entre le peu de fiches que vous pourriez recueillir et la grande quantité de textes que vous devriez lire et analyser.

*Pennacini*

Je ne sais pas si on peut supposer que le pourcentage des mots sur la quantité des lignes sera toujours le même dans les autres oeuvres que nous allons examiner. Le cas d'Homère c'est un cas tout à fait particulier. Par exemple les textes de Platon et d'Aristote seront plus riches.

*Duro*

Nous analysons électroniquement les textes anciens, littéraires ou non, pour extraire de ces textes des exemples qui ont de toute façon un certain intérêt pour l'histoire de la langue, et en moyenne nous n'obtenons presque jamais plus d'une fiche utile par ligne. Cependant, il s'agit d'un lexique général, pas d'un lexique spécial. Si nous avions pour notre but la recherche d'un lexique spécial, nous aurions un mot par 5 lignes, 2 ou 3 mots par page et pas plus. En ce cas, il n'y aurait aucun avantage dans le recours aux moyens électroniques.

*Prosdocimi*

Je n'ai pas bien compris la raison de ce procédé pour l'étude de l'idéologie. Il existe un problème de lexique et un problème de sémantique (je considère cette dernière comme une concentration conceptuelle des idées). Voi-

là donc la première question que je pose: tout ce que vous étudiez ne doit-il être important qu'au point de vue du lexique (et alors c'est une lemmatisation assez simple ayant des caractères d'objectivité) ou doit-il l'être aussi au point de vue de la sémantique (ce qui veut dire que vous y comprenez des idées exprimées par des tours de phrase, et alors c'est une lemmatisation plus complexe et plus subjective)? Car le lexique peut être riche et les idées pauvres et vice versa. Je vous demande aussi si vous faites une recherche en vous fondant sur le lexique ou sur la sémantique? Nous trouvons un grand nombre d'études de stylistique, même du courant spitzerien, affrontant des problèmes du même genre.

*Pennacini*

La recherche se déroule bien sûr dans le champ des concepts, pas seulement dans le champ de la lexicologie, des mots; mais le mot, je crois, est le signe, en quelque mesure, des concepts. Alors nous procédons aussi à isoler des mots et des locutions, des passages dans lesquels les mots se trouvent; et, ainsi faisant, on met aussi les mains sur des concepts plus larges, que les mots seuls n'identifient pas.

*Prosdocimi*

Mais cette identification d'idées, est-elle le "specificum" qui naît de l'étude lexicale, ou bien est-elle une interprétation intuitive -obtenue de toute façon par d'autres moyens- que l'étude des mots devrait confirmer ou faire apparaître d'une façon objective?

*Pennacini*

Nous trouvons des mots dans les auteurs: alors la signification qu'on peut attribuer aux mots, c'est celle qu'on peut extraire du contexte dans lequel le mot se trouve.