

Autres Programmes du L.A.S.L.A. à la disposition
des Membres de l'Organisation

- Programme de comptages grammaticaux.

Ce programme permet de déterminer le nombre d'occurrences de chaque catégorie grammaticale et de leurs subdivisions, par exemple, les nombres de substantifs, le nombre de substantifs de la première déclinaison, le nombre de substantifs de la première déclinaison au singulier, le nombre de substantifs de la première déclinaison au nominatif singulier, et ainsi de suite, quel que soit l'ordre dans lequel se trouve le fichier : ordre du texte, ordre alphabétique, ordre grammatical.

- Programme permettant la Constitution des listes des mots d'un texte en ordre de fréquence décroissante, et partant de l'Index mis sur cartes perforées dans l'ordre du texte. Ce programme donne la fréquence des lemmes, soit dans la totalité du texte, soit dans des parties de longueur déterminée. En outre, il fournit la distribution du vocabulaire pour l'ensemble des mots puis pour chaque catégorie grammaticale séparément.

- Etablissement automatique de Concordances : j'en ai parlé dans l'article précédent. L'ordinateur qui possède en mémoire le mot ou les mots dont on désire le contexte, recherche lui-même toutes les phrases où figurent le mot, soit en explorant un fichier de cartes, soit en lisant ses disques magnétiques où a été transcrit le texte.

- Programme de recherche des séquences d'ordre 1 dans la langue d'un auteur
Ce problème consiste à rechercher avec quelle fréquence une catégorie

grammaticale donnée suit une autre catégorie grammaticale, dans l'ordre naturel du discours. Par exemple, combien de fois trouve-t-on un substantif suivi d'un substantif, un substantif suivi d'un adjectif, un substantif d'un verbe, etc. ? L'ordinateur calcule en même temps les fréquences théoriques de ces suites, compare les fréquences observées aux fréquences théoriques et calcule les χ^2 pour déterminer les écarts significatifs. L'ordinateur rechercherait avec la même facilité les séquences d'ordre 2, c'est-à-dire les suites de trois catégories grammaticales du type substantif - substantif - adjectif.

De telles explorations dans le domaine de l'inconscience de la création littéraire constituent une approche pour résoudre avec des critères purement formels, les problèmes d'authenticité ou de chronologie relative des œuvres d'un auteur.

- Programme de recherche sur le rythme d'enrichissement du vocabulaire d'une œuvre.

L'ordinateur lit le texte dans l'ordre naturel du discours. D'une part il compte chaque mot du texte et d'autre part, chaque fois qu'il rencontre un mot nouveau, il compte une unité supplémentaire; il cumule les chiffres des deux séries. On obtient ainsi des listes qui permettent de tracer des courbes de l'enrichissement ou plus exactement du vieillissement de l'œuvre littéraire. Ces courbes doivent être différentes d'auteur à auteur et intéressent la solution des mêmes problèmes que le programme précédent.

- Programme permettant d'étudier la corrélation qui existe entre le nombre de syllabes par mot (chaque carte-mot dans notre codification porte, en effet, en colonne 1, un chiffre indiquant le nombre de syllabes du mot) et le nombre de mots par phrase. Il semble que chez certains auteurs, plus la phrase est longue et plus elle contient de mots longs. Cette constatation curieuse demande à être vérifiée.
- Par ailleurs, nous envisageons la constitution d'un programme d'analyse automatique morphologique et syntaxique du français et un programme semblable pour le grec.

Tous ces programmes sont écrits en SPS pour un ordinateur IBM 1620 équipé de deux armoires à disques magnétiques type 1311 avec I/O à cartes perforées.

Ces programmes sont disponibles au siège du L. A. S. L. A. Il n'en coûte que la duplication du jeu de cartes-programmes et l'envoi.

L. D.