

DE L'UTILISATION DE LA LEXICOMÉTRIE EN COLLÈGE

Chantal BERTAGNA

La lexicométrie consiste, pour simplifier, en un comptage du nombre d'occurrences (emplois) d'un mot dans un texte ou un groupe de textes et en un traitement statistique de ces chiffres qui permet de comparer l'emploi d'un mot dans les différentes parties d'un texte ou dans plusieurs textes. Cet instrument, employé depuis longtemps par les sciences humaines, trouve toute son utilité grâce à la puissance de la machine, pour des textes longs.

Or, ce que nous allons évoquer ici serait plutôt de la lexicométrie détournée puisqu'elle s'applique à des textes courts. Nous ne prétendons donc pas à la rigueur scientifique qui est celle de l'équipe de l'INRP (cf bibliographie) mais nous proposons des exercices qui, plus modestes dans leur approche, permettent néanmoins de renouveler et surtout d'améliorer la qualité de lecture et d'analyse des textes en collège et ont été pratiqués en classe de 3^{ème}, 4^{ème} et même 5^{ème}.

Après avoir connu bien des problèmes de diffusion, la lexicométrie est désormais aisément accessible grâce à la parution de deux logiciels : "Lexicométrie" coproduit par Cédic-Nathan et le CARFI de Versailles et "Pistes" diffusé par le CNDP. Ces deux programmes permettent au professeur d'entrer ses propres textes. Mais si le novice veut se familiariser avec cet outil de travail, il peut s'appuyer sur des textes déjà "prêts à consommer" : "Lexicométrie" propose *Les Précieuses ridicules*, des extraits de Corneille et de Mérimée. L'INRP édite dans la série "Logitexte" des analyses du *Tartuffe*, des *Fleurs du Mal* et des Déclarations des droits de l'homme. Tous ces programmes fonctionnent sur compatibles PC.

Avant d'envisager des exemples précis de travaux, il convient de préciser quel intérêt pédagogique il peut y avoir à utiliser un tel outil. Le logiciel fournit un index hiérarchique et un index alphabétique des mots d'un texte. Le lecteur a alors son attention attirée par la fréquence d'emploi d'un vocable, par un champ lexical dominant (par exemple les

mots "amour, aimer, amante(e), amoureux"). Il va alors demander à la machine de lui fournir les contextes du mot retenu ce qui lui permettra d'affiner son analyse. S'il travaille sur plusieurs textes (extraits d'un même auteur à des époques différentes, textes traitant d'un même thème, etc.), par le module statistique, il saura si tel mot est en sur-emploi ou en sous-emploi ce qui peut trahir une intention de l'auteur ou une caractéristique essentielle du texte et mérite en tous cas une étude attentive. C'est précisément cette sollicitation de l'attention à accorder aux mots qui rend l'outil très efficace. Plus de risque de paraphrase mais une observation de faits linguistiques, la formulation d'hypothèses à partir des données numériques, hypothèses que l'on va ensuite confronter au texte lui-même ; tout ceci constitue bien une démarche de chercheur qui est rarement mise en œuvre dans le premier cycle en cours de français et qui, l'expérience le prouve, passionne les élèves même si elle peut leur paraître ardue au début.

Concrètement, que peut-on faire ? Si l'enseignant désire faire travailler sa classe sur une œuvre entière, il pourra ainsi attirer l'attention de ses élèves sur la spécificité de langage de tel ou tel personnage (par exemple le langage populaire plein d'interjections de Mascarille dans *Les Précieuses*), sur l'importance et le sens de quelques mots-clés (ex : "honneur, gloire, feu ou flamme" dans une tragédie classique ou "Dieu" dans le *Tartuffe*). Après étude de ces mots dans leur contexte immédiat, le travail habituel d'explication de texte se trouvera singulièrement facilité.

Mais nous avons parlé plus haut de textes courts. En effet, pour faire comprendre ce qu'est une typologie de textes (différences d'énonciation suivant le genre littéraire envisagé), pour amener un travail sur un thème ou un champ lexical que ce soit la publicité ou le bonheur, pour faire étudier l'évolution d'un auteur à travers ses œuvres ou encore pour faire découvrir la fonction spécifique de l'introduction d'un roman ou de l'exposition d'une tragédie, il peut être très formateur de choisir plusieurs extraits que la machine va traiter et pour lesquels on va faire étudier la fréquence des mots, leurs contextes, dans chaque texte et comparativement.

Quelle que soit la démarche adoptée, la lexicométrie offre toujours le même intérêt : elle fait pour un temps éclater le texte, met en valeur les mots, en fait des objets d'étude en eux-mêmes, éclaire de façon bien nette les mots importants et amène ainsi les élèves à s'interroger sur le lexique et le poids des mots dans les textes étudiés.

Si elle met l'élève en position de chercheur, elle constitue aussi un précieux instrument de préparation de cours pour le professeur. D'ailleurs, selon l'environnement informatique dont il dispose, celui-ci utilisera seul la machine et fournira à ses élèves des tirages sur imprimante des index, des statistiques, ou bien il fera travailler les apprenants sur un ou plusieurs postes. Les obstacles matériels ne sont donc pas insurmontables et il serait regrettable de ne pas profiter de performances informatiques qui favorisent chez le jeune une lecture plus rigoureuse, plus minutieuse, bref plus intelligente.

Chantal BERTAGNA - Professeur agrégé
Collège Le Prieuré - 78600 Maisons-Laffitte
Groupe de Pilotage Informatique et Lettres
CARFI de Versailles

BIBLIOGRAPHIE

- * Des textes avec ou sans ordinateur, INRP, *Rencontres pédagogiques* 1984, n° 3 (50 F).
- * *L'ordinateur en classe de français, Fascicule 1 : Lectures*, CARFI, 3 bd de Lesseps, 78000 Versailles, 1987, (35 F).
- * *Informatique et étude de textes*, P. Muller (sous la direction de), INRP, Collection rapports de recherches, 1989 n° 9 (80 F).
- * à paraître en 90-91 une autre brochure du CARFI contenant des exemples d'utilisation en collège.