

La comparaison de textes



Plusieurs fonctionnalités d'Idéographix permettent de comparer des textes entre eux. En tout premier lieu, les dictionnaires et les occurrences des mots. Mais nous attirons ici l'attention sur une manière rapide pour l'enseignant de procéder à des investigations comparatives entre les textes qui constituent, par exemple, le corpus que la classe a étudié.

Ici, la comparaison ne se fait pas à partir d'un dictionnaire mais d'une liste de mots choisis par l'utilisateur et porte simultanément sur un nombre illimité de textes. Le plus simple est sans doute de prendre un exemple. Une quinzaine de textes ont déjà été rencontrés ; l'enseignant souhaite comparer la fréquence respective d'une dizaine de mots sur lesquels il veut revenir. Quels textes privilégier ?

Il s'agit d'obtenir un tableau où figurent en lignes les différents textes entre lesquels se fait la recherche et en colonnes les différents constituants (mots, parties de mots) recherchés. On peut donc instantanément savoir comment des éléments lexicaux sont représentés dans un ensemble de textes. Les constituants recherchés peuvent être réunis pour cette occasion ou appartenir à des listes, notamment une sélection dans un dictionnaire ou sur les tablettes de l'enseignant ou dans le répertoire préexistant des mots de base, etc. Ainsi l'enseignant peut-il suivre de manière méthodique l'accroissement du lexique que fréquentent ses élèves d'un texte à l'autre.

Les résultats sont donnés au choix sous la forme d'un nombre brut d'occurrences ou d'un pourcentage par rapport au nombre total de mots de chaque texte.

Le calcul d'un pourcentage n'a pas de sens lorsque le constituant recherché n'est pas un mot entier.

La dernière ligne indique pour les mots entiers leur probabilité d'apparition en français écrit.

Le contexte.

Il s'agit d'un travail mené dans une classe de cycle 3 pendant un peu plus d'une période scolaire, sur la lecture en réseau d'un certain nombre de livres de genre fantastique. On n'a pas fait une leçon sur le genre littéraire car on aurait abordé la leçon de lecture non pas par son début : la confrontation d'une écriture, d'un texte et d'une lecture mais par la fin : le résultat savant d'autres leçons de lecture faites par d'autres.

L'intérêt de ces textes est qu'ils sont, le plus souvent, construits pour faire douter le lecteur. Le jeu de l'interprétation est quasi obligatoire. Très vite les enfants se posent la question du « comment l'auteur fait-il cela ? ».

Et pour répondre à ce genre de question, Idéographix fait partie des outils réellement efficaces pour démontrer un mécanisme littéraire bien huilé.

Le premier texte auquel nous nous sommes intéressés est *La bicyclette hantée*.¹

Il s'agit de l'histoire qu'une grand-mère raconte au lecteur (analepse classique en début des histoires fantastiques) d'une bicyclette hantée par l'esprit d'un gentil fantôme. La bicyclette semble se diriger seule et entraîne la petite héroïne vers une curieuse maison où elle rencontre l'ancien propriétaire du vélo, un garçon sympathique aux vêtements démodés et pour cause, il est mort depuis longtemps.

Rapidement les élèves se donnent comme piste de travail de retrouver l'apparition du doute et du fantastique dans le texte. Les moments où l'auteur glisse imperceptiblement du rationnel à l'irrationnel.

Les enfants ont mis à jour trois pistes :

L'auteur exprime le doute par des mots

L'auteur exprime le doute par l'utilisation du conditionnel

L'auteur exprime le doute en faisant agir la bicyclette (anthropomorphisation figée, ici l'auteur joue sur cette ambiguïté).

Les relectures multiples et longues qu'impose un travail de recherche sur un texte entraînent une charge de travail qui

¹ GAIL H./ SIMS B., Folio Benjamin, série Panique

rebute les enfants. On pourrait argumenter que ces recherches sont formatrices, mais ce qui est avant tout formateur, c'est de formuler une piste de recherche, de multiplier ses recherches et d'analyser les résultats pour rebondir encore sur le texte.

On se trouve un peu dans la même situation qu'à l'époque de l'apparition de la calculette. Si l'on parle dans les nouvelles instructions officielles du « calcul instrumenté » on y parlera bientôt de « la lecture savante instrumentée ».

Donc, lors de ce travail sur *La bicyclette hantée*, c'est en groupe de recherche que les enfants ont utilisé Idéographix.

1. L'auteur exprime le doute par des mots :

comme si - incroyable - étonné - l'impression - semblait : on recherche les occurrences grâce à Idéographix et rapidement se pose un problème d'ordre lexical car les enfants sont confrontés à des notions fondamentales de linguistique : la morphologie constructionnelle (affixation, reduplication...) flexionnelle (flexion verbale, les affixes syntagmatiques...)

L'intérêt du logiciel sera de permettre aux enfants de faire des expériences sur un domaine où la démarche par tâtonnements n'avait encore jamais eu sa place : le lexique.

2. L'auteur exprime le doute par l'utilisation du conditionnel :

Là aussi se pose le problème de savoir ce qu'on peut demander à l'ordinateur, quelles suites de lettres. Les enfants proposent **rait** en fin de mot.

Une des propositions ne convient pas, c'est de l'imparfait.

3. L'auteur exprime le doute en faisant agir la bicyclette :

Les enfants proposent bicyclette et elle. Recherche d'occurrences...

Reste à supprimer les **elle** dont l'antécédent n'est pas la bicyclette et à classer les phrases, celles où la bicyclette subit l'action et celles où elle la fait.

Au cours de la leçon de lecture bien d'autres pistes sont exploitées, ne nécessitant pas forcément l'utilisation du logiciel (la recherche des adresses au lecteur du narrateur personnage, les annonces dans la description d'un des personnages de sa qualité de fantôme, le choix de l'explication rationnelle ou irrationnelle dans les lectures en réseau...)

Idéographix peut être aussi utilisé comme outil d'analyse des productions de la classe

Une autre utilisation d'idéographix fut liée à la préparation du travail d'écriture qui accompagne généralement une leçon d'écriture. Quand on a commencé à repérer « les trucs » de cette littérature, on n'a qu'une envie, c'est de les essayer à son tour. Pour préparer ce travail, le dictionnaire du texte fut fourni aux enfants, par groupe, charge à chacun de ces groupes d'établir une liste de mots qui pourrait aider à la

construction d'une histoire fantastique.

Le résultat de chacun des groupes fut saisi sur Idéographix et sauvegardé comme un seul texte. On imprima le dictionnaire, sur un transparent, avec uniquement le nombre d'occurrences comme renseignement.

Le résultat fut l'occasion de débattre des choix, (le nombre d'occurrences signifiant le nombre de groupes ayant sélectionné le mot). Le débat fut très intéressant autour de mots comme fantôme, certains ne l'avaient pas sélectionné, prétextant que si l'on voulait provoquer le doute du lecteur il ne fallait rien lui révéler. On chercha la place de ce mot dans le texte et on le trouva, sans aucune surprise, à la fin du texte, dans la partie « révélation du mystère ». Des mots comme 'seule' provoquèrent une tout autre discussion, les enfants justifiant leur choix en multipliant les exemples de sa possible utilisation dans un texte fantastique. Après cette séance, tous avaient déjà en tête un morceau de leur histoire.

On peut aussi utiliser le texte logiciel avant même la rencontre avec le texte

Si l'on considère que la lecture est un acte référentiel il est certain que tous les enfants n'ont pas le même rapport au texte. Un simple travail préparatoire, de mise en réseau, d'explication du lexique avant et non pendant la lecture, d'explication de l'énonciation... peut rendre l'écoute vraiment rentable pour ces enfants soi-disant inattentifs.

L'utilisation de l'affichage de la **probabilité d'apparition en français écrit** peut ainsi révéler des sources de difficultés inattendues. Lors du travail sur la littérature fantastique, nous eûmes à étudier un texte plus difficile : *La peur* de Maupassant. Un travail préparatoire du texte eut lieu prenant en compte la dimension proactive de la lecture.

Tout d'abord, on demanda aux enfants, de choisir des mots ayant un rapport avec le fantastique qui pourraient se trouver dans les textes qu'ils avaient lus² et aussi dans le prochain texte qu'ils auraient à lire. On insista sur le fait que ce texte était plus ancien et qu'il utilisait un registre plus soutenu.

Ce qu'on cherche, dans un tel travail c'est avant tout l'ouverture d'un horizon d'attente plus précis, tous les résultats ne seront pas exploités et tant pis si en cherchant **sembl** (comme indicateur du doute) on repère **ensemble** qui n'a rien à voir.

Une élève vient de proposer de comparer la version de *La peur* qu'ils ont en classe avec une version qu'elle a apportée de sa bibliothèque personnelle...

Pierre CHOLET

¹ Notamment chez Folio Benjamin, Série Panique : *Trois histoires pour frémir*, O'CONNOR J., / KARAS B., - *L'étrange maison de poupée*, DUSSLING J. / LAMUT S.