

P.A.E. STYLISTIQUE INFORMATIQUE

Pierre COURANJOU, Bernard LACHAMBRE

L'étude stylistique est une entreprise toujours délicate à mener en classe : chaque professeur de français déplore en effet, que la paraphrase soit l'écueil où, bien souvent, viennent s'échouer les tentatives d'analyse des élèves. Par nature, l'analyse stylistique est une démarche rigoureuse, en cela très voisine du "raisonnement géométrique", puisqu'aussi bien il s'agit, à partir de notions grammaticales précises de formuler des hypothèses sur lesquelles l'explication du texte se construit.

Sans doute les élèves de Lycées Technologiques sont-ils mieux armés pour affronter cette analyse méthodique de la littérature, en raison du caractère "pluriscientifique" de leur formation. Nous avons songé que nous pourrions mettre à profit une recherche inter-disciplinaire en mathématiques et en français, grâce à l'outil informatique. Ainsi, pratiquement, les élèves pourraient-ils percevoir qu'un style se définit objectivement, par la mise en évidence de constantes d'écriture.

Pour ce qui est des mathématiques proprement dites, ce genre d'étude permet un travail sur des statistiques en grandeur réelle, et dans d'autres domaines que ceux qui sont dévolus à l'INSEE : l'économie, la population, etc. De plus, il implique une analyse des types de représentation graphique des données, et le choix de la meilleure représentation pour une statistique donnée.

Nous avons choisi de faire porter l'étude sur le premier des trois derniers romans de CELINE : "D'un château l'autre", parce que l'écriture romanesque est l'une des plus modernes que nous possédons, et qu'elle allait permettre, pour cela, des étalonnages et des comparaisons avec d'autres oeuvres et d'autres auteurs. Nous avons limité notre recherche à l'étude du rythme de CELINE : à notre connaissance, en effet, aucune étude de ce type n'a jamais été menée ; H. GODARD pressent indéniablement les questions que soulèvent le rythme célinien, mais rien ne prouve ses intuitions (1). Comme le rythme ne saurait s'analyser en dehors d'une

(1) "Poétique de Céline" de Henri GODARD chez Gallimard 1985.

description précise des éléments de la phrase, nous avons décidé d'étudier systématiquement chacun de ces éléments. Nous espérions ainsi parvenir à mieux cerner la notion de phrase chez CELINE.

Ce travail a été réalisé dans le cadre d'un P.A.E. avec des élèves des classes de Première E et Première F1 au Lycée Technologique VIETTE à Montbéliard en 1987-88, deux enseignants de mathématiques et français étaient concernés.

I - LA DÉMARCHE

Il a fallu tout d'abord, trouver un outil informatique de traitement d'un texte, pouvant faire apparaître deux éléments : la structure grammaticale et le rythme.

Les premiers essais ont été réalisés avec un traitement de texte, faisant apparaître, grâce à une mise en page appropriée, les éléments importants (Annexe 1). Cette outil, cependant s'est montré trop limité dans ses possibilités. L'idée est donc venue de considérer le texte sous forme d'un "fichier".

A l'aide du logiciel de gestion de données DBASE, un petit fichier test a été réalisé ; il prenait comme unité de base la phrase, une fiche équivalant à une phrase. Les différentes rubriques étant constituées des différents mots de la phrase (Annexe 2). Mais il fallait opérer un choix : ou bien conserver le texte complet, et donc prévoir tout les cas de figure - nombre des substantifs, des verbes, etc. dans leur plus grande occurrence -, ou ne prendre en compte que les rubriques grammaticales à priori importantes, en en éliminant certaines qui, par la suite, pouvait s'avérer intéressantes.

Un dernier essai a été réalisé, toujours avec DBASE, mais en prenant comme unité non la phrase, mais le mot, ou l'élément de ponctuation. Cette structure de fichier (Annexe 3) était simple puisqu'une fiche ne contient qu'un élément, et quelques-unes de ses caractéristiques : sa nature grammaticale, en particulier, et son rythme - le nombre de ses syllabes . De plus cette structure pouvait aisément être modifiée, si besoin était. Sans doute les fichiers deviendraient-ils vite volumineux, mais quelques essais ont permis de constater que les temps de traitement ne seraient pas trop longs. Enfin, cette méthode de travail conservait le texte dans son intégralité.

II - LES LOGICIELS UTILISÉS

D'autres logiciels de traitement de bases de données auraient pu être utilisés. Signalons simplement que, par programmation, DBASE offrait de multiples possibilités. Enfin, DBASE est un des logiciels proposés à l'achat en licence mixte.

Pour la représentation graphique, nous nous sommes servis de GRAPH IN THE BOX qui est, lui aussi, un logiciel en licence mixte.

III - LA MÉTHODE DE TRAVAIL

Un premier fichier grandeur nature a été saisi. Il comprenait les deux premières pages de l'édition de la pléiade de "D'un château l'autre". Chaque groupe de deux élèves a pour cela saisi huit à dix lignes de texte, puis, les différents fichiers ainsi constitués ont été regroupés, pour obtenir le texte complet. Outre cet intérêt pratique, le travail sur de petits fichiers est bien plus rapide, et évite les fausses manoeuvres lourdes de conséquences.

Des traitements ont été réalisés, soit directement, soit par l'intermédiaire de petits programmes, écrits avec les élèves, ou par les enseignants dans les cas plus complexes (le cahier des charges étant cependant fait avec les élèves).

Ces programmes permettent :

- de réécrire le texte à partir du fichier (ce qui permet des corrections rapides),
- de reproduire le texte, avec saut de ligne après chaque ponctuation, en indiquant le nombre de syllabes de chaque groupe,
- de faire des calculs statistiques relatifs au rythme de la phrase et à la nature des mots.

Un second extrait du roman, situé quelques pages plus loin, a été saisi ; les mêmes opérations ont été effectuées. Les résultats ont permis d'émettre plusieurs hypothèses concernant le rythme de l'élément minimal de texte et d'avoir quelques éléments de caractérisation de l'écriture de CELINE.

Des représentations graphiques ont mis en évidence les ressemblances et les différences des deux textes saisis ; d'ailleurs, l'étude des différences et des résultats aberrants a permis de proposer des

hypothèses de lecture pour le comptage des "e" caducs, et le respect de certaines pauses non marquées par une ponctuation, mais qui doivent être faites à la lecture. On trouvera en annexe (Annexe 4), sur un court extrait, la comparaison des rythmes avant et après le travail effectué.

Pour l'instant, si le travail réalisé (Annexe 5) permet de nuancer les hypothèses d'H. GODARD, il les confirme globalement. La mise en fiche systématique a conduit les élèves à constater qu'un texte - fût-il romanesque - est avant tout un texte à lire, et à bien lire. Mais cette lecture, notamment pour des raisons de cohérences grammaticales, de constantes stylistiques, ne peut être faite qu'au terme d'un travail précis d'explication. Dans ce sens, et avec toute l'humilité nécessaire à de semblables entreprises, il nous semble qu'un travail interdisciplinaire se trouve fondé en raison. De plus, nous espérons bien, dans la suite, établir un fichier beaucoup plus important : il nous permettra de vérifier les hypothèses d'analyse du rythme, et d'analyser de près la syntaxe de CELINE.

Bernard LACHAMBRE
 Lycée Tehnologique VIETTE
 25200 MONTBELIARD
 C.U.F.I.A.P.
 Ecole Normale Fort Griffon
 25000 BESANCON

ANNEXE 1

Utilisation du traitement de texte pour mettre en valeur la structure de la phrase.

Un exemple (les verbes sont tous les uns sous les autres) :

Pour parler franc,
là entre nous,
je finis encore plus mal
que j'ai commencé ...
Oh, j'ai pas très bien commencé ...
je suis né,
je le répète,
à Courbevoie,
Seine,
je le répète pour la millième fois ...
après bien des aller et retour je termine vraiment au plus mal ...

ANNEXE 2

Premier fichier réalisé, ici : un enregistrement (une fiche) = une phrase, tous les mots ne figurent pas dans le fichier, l'ordre des mots du texte n'est pas respecté.

Un exemple :

	Phrase 1	Phrase 2	Phrase 3	Phrase 4	...
VERBE1	parler		finis	ai commencé	
VERBE2			ai commencé		
ADVERBE1	franc	là	encore	pas	
ADVERBE2			plus	très	
ADVERBE3			mal	bien	
ADVERBE4					
INTERJECTION				Oh	
SUBSTANTIF1					
SUBSTANTIF2					
SUBSTANTIF3					
ADJECTIF1					
ADJECTIF2					
ADJECTIF3					
PREPOSITION1	Pour	entre			
PREPOSITION2					
RYTHME	4	4	13	8	

Le texte était le suivant :

Pour parler franc, là entre nous, je finis encore plus mal que j'ai commencé... Oh, j'ai pas très bien commencé

ANNEXE 3**Structure des fichiers**

Champ	Nom du champ	Type	Longueur
1	MOT	Caractère	24
2	NATURE	Caractère	3
3	NPERS	Numerique	1
4	RYTH	Numerique	2
5	FIN	Logique	1
Total			32

Nota : Deux champs n'ont pas été utilisés dans le travail réalisé

1 - NPERS pour indiquer le numéro de la personne dans le cas des verbes et des pronoms.

2 - FIN (champ logique, qui ne peut être que vrai, noté .T. ou faux, noté .F.) pour mettre en valeur la notion de phrase chez Céline. Ce travail a été esquissé mais n'a pas donné de résultats intéressants actuellement.

Extrait du fichier

Enregistrement numéro	MOT	NATURE	NPERS	RYTH	FIN
1	Pour	PRE	0	1	.F.
2	parler	V	0	2	.F.
3	franc	ADJ	0	1	.F.
4	,	PON	0	0	.F.
5	là	ADV	0	1	.F.
6	entre	PRE	0	2	.F.
7	nous	PRO	4	1	.F.
8	,	PON	0	0	.F.
9	je	PRO	1	1	.F.
10	finis	V	1	2	.F.
11	encore	ADV	0	2	.F.
12	plus	ADV	0	1	.F.
13	mal	ADV	0	1	.F.
14	que	C	0	1	.F.
15	j'	PRO	1	0	.F.
16	ai commencé	V	1	4	.F.
17	...	PON	0	0	.T.

La nature des mots

Liste des codifications utilisées

ADJ	adjectif	PON	punctuation
ADV	adverbe	PRE	préposition
ART	article	PRO	pronom
C	conjonction	PRS	présentatif (2)
I	interjection	S	substantif
N	négation	V	verbe
PAU	pause (1)		

(1) Ce code a été introduit après une première étude des résultats. Il permet d'introduire dans le fichier des mots "vides" qui correspondent à une rupture dans le rythme de la phrases, non marqué par une ponctuation. Ces enregistrements ne sont pris en compte que dans le traitement qui met en évidence les éléments de phrase, il sont alors considérés comme des ponctuations (voir Annexe 4).

(2) Sont regroupées sous cette nature des locutions comme "y a", "c'est que"...

A noter aussi qu'en cas de mots composés, locutions, verbes conjugués avec un auxiliaire, le champ nature du premier mot est rempli normalement, alors que le champ nature des suivants contient un nombre qui permet de retrouver le mot avec lequel il est associé (ex : 4 indique que le premier mot de la locution est situé 4 enregistrements avant) Ceci a permis d'affiner les statistiques sur la nature des mots.

Enfin, dans quelques cas, deux mots consécutifs peuvent tous les deux jouer le rôle d'adjectif ou de substantif (ex : le chenu pauvre). Pour les repérer et les analyser, ils sont suivis d'une astérisque * .

ANNEXE 4**Travail sur le rythme du texte de Céline**

Extrait du fichier

Numéro	MOT	RYTHME (après élision)	RYTHME (1ère saisie)
1	tout le	2	2
2	chaume	1	2
3	s'	0	0
4	envolait	3	3
5	...	0	0
6	la	1	1
7	neige	1	1
8	,	0	0
9	le	1	1
10	vent	1	1
11	dansaient	2	2
12	la dedans	2	3

quelques lignes plus loin ...

137	c'	0	0
138	était	2	2
139	presque	1	2
140	...	0	0
141	deux	1	1
142	ans	1	1
143	en	1	1
144	fosse	1	2
145	,	0	0
146	trois	1	1
147	mètres	1	2
148	sur	1	1
149	trois	1	1
150	!...	0	0

Voir ci-dessous l'évolution de l'analyse du rythme sur un extrait du texte de Céline (mise en page faisant apparaître les unités de rythme).

1 – Après la première saisie

7 tout le chaume s'envolait ...
 2 la neige ,
 7 le vent dansaient la dedans !...
 2 cinq ans ,
 5 cinq mois à la glace !...
 2 Lili ,
 5 malade opérée ...
 15 et allez pas croire que cette glacière était gratuite !
 3 pas du tout !...
 4 confondez rien !...
 4 j' ai tout payé
 5 les notes sont là ,
 8 et signées par mon avocat ...
 8 certifiées par le Consulat ...
 11 ce qui vous explique que je suis si raide !...
 16 pas seulement du fait des pirates de la Butte Montmartre...
 9 les pirates de Baltique aussi !...
 28 les pirates de la Butte Montmartre voulaient me saigner que mes
 tripes dégoulinent la rue Lepic ...

2 – Après élision

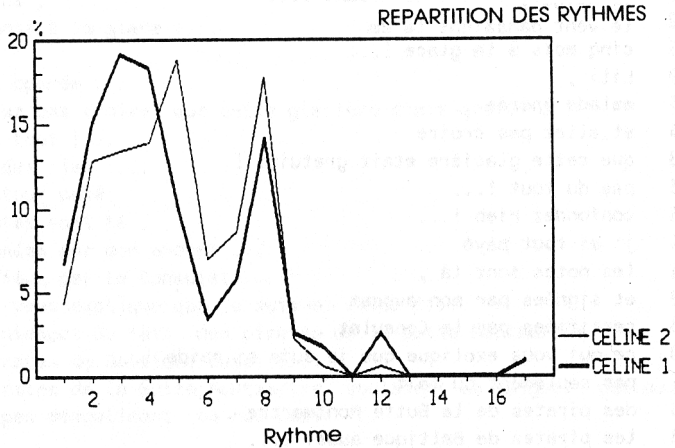
6 tout le chaume s'envolait ...
 2 la neige ,
 6 le vent dansaient la dedans !...
 2 cinq ans ,
 5 cinq mois à la glace !...
 2 Lili ,
 5 malade opérée ...
 13 et allez pas croire que cette glacière était gratuite !
 3 pas du tout !...
 4 confondez rien !...
 4 j' ai tout payé
 4 les notes sont là ,
 8 et signées par mon avocat ...
 8 certifiées par le Consulat ...
 10 ce qui vous explique que je suis si raide !...
 13 pas seulement du fait des pirates de la Butte Montmartre ...
 8 les pirates de Baltique aussi !...
 23 les pirates de la Butte Montmartre voulaient me saigner que mes
 tripes dégoulinent la rue Lepic ...

3 – Après prise en compte des pauses non indiquées par une ponctuation

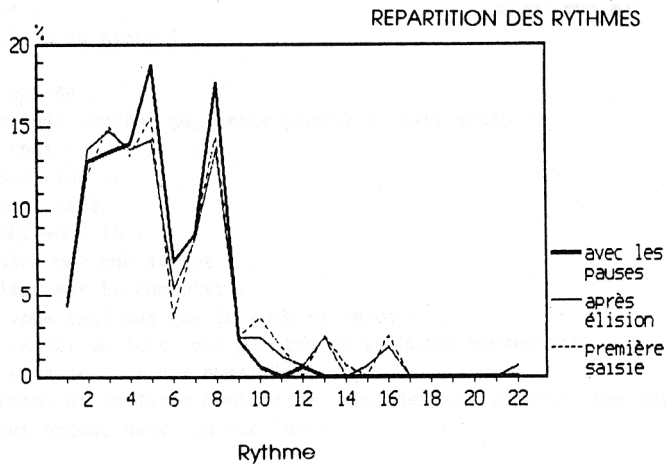
6 tout le chaume s'envolait ...
 2 la neige ,
 6 le vent dansaient la dedans !...
 2 cinq ans ,
 5 cinq mois à la glace !...
 2 Lili ,
 5 malade opérée ...
 5 et allez pas croire
 8 que cette glacière était gratuite !
 3 pas du tout !...
 4 confondez rien !...
 4 j' ai tout payé
 4 les notes sont là ,
 8 et signées par mon avocat ...
 8 certifiées par le Consulat ...
 10 ce qui vous explique que je suis si raide !...
 5 pas seulement du fait
 8 des pirates de la Butte Montmartre ...
 8 les pirates de Baltique aussi !...
 8 les pirates de la Butte Montmartre
 5 voulaient me saigner
 6 que mes tripes dégoulinent
 4 la rue Lepic ...

ÉTUDE STATISTIQUE DU RYTHME DU TEXTE (répartition du rythme des éléments de phrase)

1 - Comparaison de deux extraits de texte



2 - Évolution des résultats



ANNEXE 5

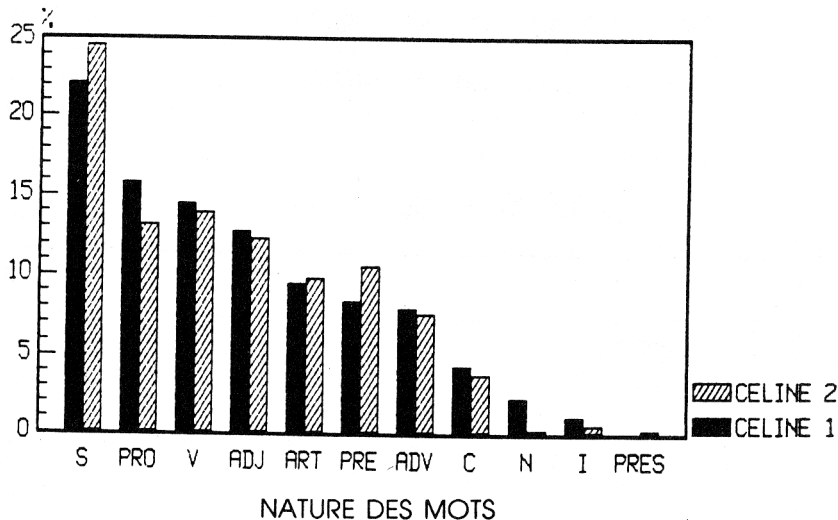
Les résultats obtenus

Le travail a porté principalement sur la nature des mots et sur le rythme du texte. Il a été fait à partir de deux extraits du livre de Céline "D'un château l'autre", ces deux extraits sont notés dans les graphiques "CELINE 1" et "CELINE 2".

En ce qui concerne le rythme du texte, on pourra se reporter à l'annexe 4 ci-dessus. On remarque une bonne corrélation des résultats obtenus sur les deux extraits.

On notera l'importance des groupes de 8 syllabes mais également de 3, 4 et 5 syllabes. Dans ce dernier cas, ils se trouvent généralement juxtaposés en 3 + 5 ou 4 + 4.

Pour la nature des mots, une bonne corrélation existe également, il suffit d'observer le graphique ci-dessous.



Les deux graphiques suivants indiquent les résultats obtenus sur l'ensemble des deux extraits ayant servi à l'étude.

Une étude plus complète devra être faite pour savoir si ces résultats sont représentatifs de ce roman, et même des trois derniers romans de Céline. Une étude comparative faite sur d'autres auteurs permettrait peut-être une caractérisation de l'écriture de Céline sur la fin de sa vie.

