

PROJET D'ETUDE STATISTIQUE DE L'ELECTRE D'EURIPIDE

Le travail que l'on présente ici a pu être entrepris au Laboratoire d'Analyse statistique des Langues anciennes, grâce à l'équipement spécial dont il dispose (perfo IBM 836 modifiée permettant de reporter sur cartes un texte grec avec tous les signes spéciaux), et surtout en raison de la constante bienveillance de M. le Professeur L. DELATTE, Directeur du Laboratoire, et de M. E. EVRARD, Professeur à la Section de Philologie Classique, qui a bien voulu m'initier à l'analyse statistique appliquée à la philologie.

Le choix d'ELECTRE, tragédie d'Euripide, a été déterminé par diverses considérations. Au départ, la lecture de ce texte, en vue d'en faire l'objet d'un cours dans l'Enseignement moyen, avait attiré mon attention sur les nombreux emplois de synonymes exprimant la douleur du personnage principal. On sait qu'Euripide passe pour avoir été l'élève des Sophistes; il me paraissait intéressant d'étudier, grâce aux techniques du traitement de l'information, l'influence que leurs recherches dans le domaine du vocabulaire avaient éventuellement exercée sur l'auteur d'Electre.

Or, il se fait que le même thème légendaire a été traité par les deux grands

Tragiques, et que nous avons heureusement conservé les trois oeuvres presque complètes (les CHOEPHORES d'Eschyle comportent plusieurs lacunes, mais on peut estimer que les quelque 1.000 vers qui nous sont parvenus représentent les deux tiers de la tragédie complète). Une étude sur le texte d'Euripide présentait donc un intérêt supplémentaire, puisqu'elle pouvait servir de base à une recherche comparative des procédés et de la langue des trois grands de la Tragédie grecque : en comparant des pièces de théâtre traitant le même sujet, on peut espérer cerner avec plus de certitude les différences essentielles entre les conceptions personnelles de chacun des trois auteurs. On sait par ailleurs que la tragédie d'Euripide comporte des passages qui parodient le texte d'Eschyle, sinon celui de Sophocle; ceci montre bien que l'intention de l'auteur était d'opposer ses idées à celles de ses prédécesseurs.

La première partie du travail a porté sur la constitution du fichier complet sur cartes perforées. Le texte suivi est celui de l'édition des Belles-Lettres; il compte 7.778 mots. Chaque carte (qui ne concerne qu'un seul mot) reprend les indications ci-après, suivant le dessein de carte et les codes proposés par M. J. DE BIE (dans la *Revue*, 1966, n° 4, pp. 31 et suivantes).

- | | | |
|----|------------|--|
| 1. | Col. 1 | Code carte-mot |
| 2. | Col. 2-21 | Lemme (d'après le dictionnaire de Liddell et Scott) |
| 3. | Col. 22-43 | Forme lue dans le texte (les formes comportant une crase sont perforées sur deux cartes successives, la première est lemmatisée sur le premier élément - <i>καί</i> pour <i>κοῦ</i> , par exemple - la |

- seconde sur le deuxième élément - oŭ pour
koŭ en reprenant le même exemple).
4. Col. 44 Ponctuation (fin de vers, de phrase, ou de vers et de phrase).
 5. Col. 45-59 Référence automatique (codes d'auteur et d'oeuvre, n° de vers, n° d'ordre dans le vers, n° d'ordre dans la phrase).
 6. Col. 60 Code personnage (avec distinction par le "zoning" entre les parties iambiques, lyriques et mélodramatiques).
 7. Col. 61-66 Analyse morphologique (cf. J. DE BIE, article cité).
 8. Col. 71-75 N° d'ordre dans l'oeuvre.

Les colonnes 67 à 70 et 76 à 80 restent libres et permettent donc d'introduire d'autres indications dans le fichier (par exemple, une analyse syntaxique).

Un premier comptage sur l'ordinateur du L.A.S.L.A. a été réalisé pour l'ensemble du texte, sur la base des analyses morphologiques, en distinguant successivement les catégories grammaticales, les cas ou les personnes, les nombres, les modes et les temps. A partir des données numériques ainsi obtenues et réparties en quatre tableaux (voir annexes), j'ai réalisé deux séries de tests de χ^2 ; la première série porte sur les formes déclinées (substantifs, adjectifs, numéraux, pronoms, article et participes); l'autre porte sur les formes conjuguées de l'indicatif (ce mode étant le seul à comporter tous les temps et toutes les personnes de la conjugaison).

Les tableaux en annexe représentent donc :

Tableau 1^a : la répartition de chaque catégorie grammaticale suivant les cas de la déclinaison.

Tableau 1^b : la répartition des formes aux trois nombres (Sg, Pl, Duel), suivant les cas de la déclinaison.

Tableau 2^a : les formes de chaque temps (le plus-que-parfait étant regroupé avec le parfait, le futur perfectif avec le futur simple), distribuées suivant les personnes de la conjugaison.

Tableau 2^b : les formes employées aux trois nombres, distribuées suivant les personnes de la conjugaison.

Dans chaque tableau, la distribution des effectifs réels a été comparée au moyen du test de χ^2 à une distribution théorique aléatoire calculée à partir de la distribution des cas ou des personnes sur l'ensemble du texte (dernière colonne de chaque tableau). On pose ainsi l'hypothèse que, dans chaque catégorie grammaticale (chaque temps pour les formes verbales), chaque cas sera représenté par un effectif proportionnel à celui que l'on observe, pour ce cas, sur l'ensemble des catégories. Les différences constatées entre l'effectif réel et l'effectif théorique manifesteront donc les particularités du style (recherche des substantifs en fonction de sujet, emploi de la première personne plus fréquemment au futur ... par exemple). L'examen des résultats a permis quelques constatations intéressantes.

1. *Formes déclinées.* Le tableau 1^a montre la distribution des catégories grammaticales, par cas de déclinaison, sans distinguer entre les nombres.

La colonne R indique l'effectif réel, la colonne T l'effectif théorique, la colonne χ^2 l'expression décimale de $\frac{(R - T)^2}{T}$.

T

Le groupe des Numéraux n'est repris que pour mémoire, les effectifs étant trop faibles pour permettre des calculs significatifs.

On remarque d'abord que la valeur de χ^2 est très élevée pour toutes les catégories grammaticales, sauf pour l'article ($\chi^2 = 0,27$ soit une probabilité de 0,99 pour $\nu = 5$). Il n'y a donc aucune intention particulière dans l'emploi de l'article; celui-ci a perdu toute valeur de signification particulière.

Pour les autres catégories grammaticales, il convient d'observer quelles sont les valeurs partielles composant l'indice de χ^2 qui entraînent, à elles seules, une valeur particulièrement élevée. Nous arrivons ainsi aux observations qui suivent.

- a) au Nominatif, on note un important déficit des substantifs (valeur partielle 65,57), et un excès de participes (valeur partielle 323,36); la situation est inverse si on considère le Génitif (41,44 et 36,01 respectivement). Il conviendrait, me semble-t-il, d'examiner si l'on se trouve plus généralement en présence d'une corrélation : le participe, forme verbale à l'origine, serait considéré comme forme nominale principalement lorsqu'il est employé comme sujet d'un verbe conjugué.
- b) au Datif, il y a un excès de pronoms. Cela paraît correspondre à la nature même de ces mots. Comme ils remplacent essentiellement des

substantifs, l'excès constaté à ce cas dans la catégorie substantifs, parallèlement à la catégorie pronoms, serait à rapprocher du déficit des participes au même cas.

- c) le tableau 1b permet de noter un effectif relativement plus important du Datif au pluriel par rapport au singulier; cet excès correspond à un déficit au Nominatif. Une étude plus poussée pourrait peut-être déterminer s'il ne s'agit pas d'une manifestation de l'importance particulière de l'action individuelle (le singulier, au total, est d'ailleurs nettement plus important que le pluriel).

2. *Formes conjuguées.* Comme je l'ai indiqué plus haut, les tableaux 2^a et 2^b ne tiennent compte que des formes du mode indicatif; en effet, seul celui-ci comprend tous les temps et toutes les personnes.

Le tableau 2^a répartit les différents temps selon les personnes, sans distinguer les nombres. On examinera les valeurs de χ^2 suivant le principe exposé plus haut pour les formes nominales.

- a) le présent ne révèle pas de caractéristique particulière ($\chi^2 = 1,49$; $p = 0,48$ pour $\nu = 2$).
- b) l'imparfait est en déficit à la première personne et en excès à la troisième, ce qui paraît l'associer à l'aoriste (pour la première personne seulement) et l'opposer au parfait (première et troisième personnes).
- c) le futur est en excès à la première personne, ce qui l'oppose dans ce cas à l'imparfait et à l'aoriste.
- d) l'aoriste est en déficit à la première personne et en excès à la deuxième; cela paraît l'associer à l'imparfait pour la première personne;

l'excès à la deuxième personne ne correspond pas à un déficit marqué ailleurs.

- e) le parfait est en excès à la première personne et en déficit à la troisième; ce qui l'oppose à l'imparfait.

De ces observations paraît se dégager une répartition des temps de l'indicatif en trois groupes (le présent étant exclu puisqu'il ne révèle pas de particularités marquantes) :

le Parfait
l'Aoriste et l'Imparfait
le Futur

Il y a opposition entre le premier et le deuxième groupe et entre le deuxième et le troisième à la première personne. Une opposition semblable existe entre le premier et le deuxième groupe à la troisième personne. Enfin, l'Aoriste présente un excès isolé à la deuxième personne.

Sans vouloir risquer une interprétation sur cette seule base, une explication possible de ces divergences me semble pouvoir se formuler ainsi :

- L'excès de première personne au futur et au parfait pourrait provenir de l'attitude des personnages face à la situation qui les entoure : ils n'envisagent le passé que dans ses conséquences, et l'avenir dans sa probabilité.
- L'excès de deuxième personne à l'aoriste indiquerait que ces mêmes personnages ne considèrent le passé récent que chez l'interlocuteur.
- L'excès de troisième personne à l'imparfait correspondrait aux actes des

personnages généralement absents de la scène (Egisthe, Clytemnestre). Leurs actes n'ont pas encore connu de conclusion; cela expliquerait en outre l'opposition du parfait et de l'imparfait.

De l'examen du tableau 2^b (répartition des personnes suivant les nombres), résultent les constatations suivantes :

- a) le Singulier l'emporte nettement, ce qui semble indiquer que l'action est le fait d'un seul individu;
- b) au Pluriel, la deuxième personne est en déficit très marqué (valeur partielle 6,05) et la troisième personne est en excès (valeur partielle 2,60 seulement). Cela pourrait révéler que les personnages ne s'adressent pas directement à un groupe et qu'ils évoquent fréquemment des actions collectives auxquelles ils n'ont pas été mêlés.

Il va sans dire que les tentatives d'explication ci-dessus devront faire l'objet de vérifications à divers niveaux, notamment en recourant à l'examen des formes remarquables placées dans leur contexte; il faudra aussi se demander si ces constatations faites sur l'ensemble de la tragédie se répètent dans les diverses parties (épisodes, chœurs, mélodrames). En outre, il sera particulièrement intéressant d'étudier de ce même point de vue le texte des trois ou quatre rôles principaux (Electre, Oreste, Clytemnestre, la Coryphée, par exemple).

Tels sont les premiers résultats de l'exploitation du fichier de l'ELECTRE

d'Euripide. Ils me semblent permettre d'envisager des recherches dans diverses directions, recherches dont le résultat pourrait éclairer de façon nouvelle une tragédie et un auteur dont l'aspect "moderne" a été signalé plusieurs fois, et qui connaît actuellement un regain d'intérêt.

G. RIGO

TABLEAU 1^a

	SUBSTANTIFS			ADJECTIFS			NUMERAUX	PRONOMS			PARTICIPES			ARTICLE		
	R	T	χ^2	R	T	χ^2	R	R	T	χ^2	R	T	χ^2	R	T	χ^2
N	351	539	65,57	228	220	0,29	6	246	269	1,96	306	114	323,36	84	74	} 0
V	107	76	12,64	47	31	8,25	0	10	38	20,63	8	16	4,00	0	10	
A	706	702	0,02	313	286	2,54	8	400	350	7,14	66	148	45,43	95	97	0,04
G	495	371	41,44	129	151	3,20	3	136	185	12,97	25	78	36,01	51	51	0,00
D	310	278	3,68	86	113	} 7,75	1	180	139	} 16,90	11	59	} 40,01	42	38	} 0,23
Autres	4	7	1,28	0	3		0	11	3		0	1		0	1	
TOTAL	1973		124,63	803		22,03	18	983		59,60	416		448,81	272		0,27

TABLEAU 1^b

	SINGULIER			PLURIEL			DUEL			TOTAL		
	R	T	χ^2	R	T	χ^2	R	T	χ^2	Eff	%	
N	976	847	19,64	229	365	50,67	16	9	} 6,40	1221	27,34	<i>ENSEMBLE DE L'OEUVRE</i> <i>Formes nominales</i>
V	141	119	4,06	29	51	9,49	2	1		172	3,85	
A	1074	1101	0,66	513	475	3,04	1	12	10,08	1588	35,56	
G	536	582	3,63	295	251	7,71	8	6	0,66	839	18,79	
D	356	437	15,01	268	188	} 30,88	6	5	0,20	630	14,10	
Autres	14	10	1,60	1	4		0	0	15	0,33		
TOTAL	3097		44,60	1335		101,79	33		17,34	4465		

TABLEAU 2^a

	PRESENT			IMPARFAIT			FUTUR			AORISTE			PARFAIT			TOTAL	
	R	T	χ^2	R	T	χ^2	R	T	χ^2	R	T	χ^2	R	T	χ^2	Eff	%
1	109	105	0,15	16	28	5,14	41	28	6,03	40	58	5,58	31	19	7,57	237	27,78
2	57	66	1,22	15	17	0,23	17	17	0	47	36	3,36	12	12	0	148	17,35
3	212	207	0,12	69	55	3,56	41	54	3,12	120	114	0,31	26	38	3,78	468	54,86
TOTAL	378		1,49	100		8,93	99		9,15	207		9,25	69		11,35	853	

TABLEAU 2^b

	SINGULIER			PLURIEL			DUEL		
	R	T	χ^2	R	T	χ^2	R	T	χ^2
1	206	204	0,01	31	33	0,12			0
2	137	127	0,78	9	20	6,05			2
3	390	402	0,35	78	65	2,60			0
TOTAL	733		1,14	118		8,77			2

ENSEMBLE DE L'OEUVRE

*Formes verbales
Indicatif seul
(Parfait = 5 + 6)
(Futur = 3 + 7)*