

le fragment comme échantillon

Il est possible, moyennant l'application de méthodes statistiques, de déterminer certaines caractéristiques du langage et du style d'une oeuvre. Mais de nombreuses oeuvres anciennes, grecques ou latines, ne nous sont pas connues dans leur intégrité; nous ne possédons d'elles que des fragments. Par conséquent, le problème se pose d'apprécier la valeur du fragment comme échantillon, de fixer des critères aptes à mesurer la légitimité de l'élévation du fragment à la dignité d'échantillon d'un univers.

Ce problème présente différents aspects.

On se demandera d'abord si les fragments sont aléatoires ou non. En effet, ils peuvent avoir été conservés par tradition directe, dans un papyrus ou un manuscrit, où seuls quelques passages sont lisibles; ils peuvent aussi nous être parvenus par voie indirecte, c'est-à-dire par l'intermédiaire d'ouvrages d'autres auteurs. Dans ce cas, il s'agit de citations allant d'un seul mot à quelques dizaines de vers ou de lignes.

Dans le premier cas, les fragments sont à considérer, sans aucun doute, comme des échantillons aléatoires et, en tant que tels, ils peuvent être pris, sans restriction aucune, comme représentatifs des caractères de l'ensemble. Naturellement, les statisticiens préfèrent opérer eux-mêmes

les prélèvements, dans les conditions qu'ils ont choisies, de manière que les échantillons répondent, quant à leur extension et à leur composition, à des critères fixés par les chercheurs, en rapport avec l'hypothèse de travail.

Mais cela n'est pas possible au philologue. Il devra se contenter du hasard de l'échantillonnage, qui est, dans ce cas, hors de doute. Il devra rassembler un nombre satisfaisant de fragments suffisamment étendus pour que l'on puisse raisonnablement supposer que les principales caractéristiques lexicales et stylistiques du texte y sont représentées. Les constatations se feront fragment par fragment et, de chacun d'eux, on extraira des données quantitatives sous forme de moyennes, de pourcentages, de fréquences. Réunies, ces données en fourniront d'autres, plus générales, relatives à l'ensemble des fragments. Ces résultats, pris isolément ou en corrélation, serviront de base à l'indication des probabilités de certaines caractéristiques de l'oeuvre tout entière.

Quoique les échantillons - toujours dans le cas de fragments aléatoires - n'aient pas été prélevés selon des critères répondant à l'hypothèse de travail, ils pourront toutefois être analysés selon des critères suggérés par cette même hypothèse, la fortuité de l'échantillonnage n'étant pas dérivée de l'application de normes préétablies mais étant, pour ainsi dire, historique.

Le chercheur établira à quelle caractéristique de l'oeuvre il va s'intéresser sur la base des résultats de l'analyse des échantillons; on pourrait dire : sur la base de l'inventaire des matériaux repérables dans les échantillons. Il pourra, par exemple, laissant de côté les contenus et ne s'occupant que des paroles, relever, à l'intérieur de chaque membre du discours (période

ou proposition), la succession des vocables, classés au point de vue morphologique ou au point de vue syntaxique. Cette recherche pourra se faire si le matériel est suffisant. Souvent, hélas! le fragment ne permet que des constatations d'ordre lexical, car il ne possède pas l'extension nécessaire pour rassembler des éléments significatifs sur les modules expressifs employés.

En substance, tant dans le cas de fragment à transmission directe que dans le cas de fragment à transmission indirecte, l'hypothèse de travail doit être déterminée après l'inventaire du matériel présent dans les fragments. Bien que nous soyons tous d'accord pour dire que chaque partie d'un texte est reliée au reste (en d'autres termes, que l'emploi des méthodes statistiques suppose l'admission de l'hypothèse structurale), il est possible, à mon avis, de différer l'utilisation du schéma structural ou d'en réduire l'ampleur, dans une mesure correspondant à la richesse et à la signification du matériel repérable dans les fragments. Dans la plupart des cas, l'extension insuffisante du fragment n'autorise aucune constatation concernant, par exemple, la marche syntaxique : type ou durée des propositions ou des périodes, distribution de la coordination et de la subordination, correspondance de certaines sphères lexicales et de certaines constructions etc. Il conviendra donc de déterminer des éléments simples, dont la signification et l'appréciation ne dépendent pas ou dépendent très peu du contexte. Je dirais même qu'il faut choisir des éléments qui peuvent être considérés comme la base des choix syntaxiques et stylistiques. En cette matière, c'est l'élément lexical qui est le premier et peut-être l'élément fondamental.

Une fois définies, dans les limites du possible, toutes les données qualitatives et quantitatives se rapportant à chaque élément, on pourra en induire, par voie statistique, l'univers probable, je veux dire l'ensemble des choix lexicaux, ou même, si l'on veut, l'univers des choix concernant l'emploi, par exemple, des adjectifs qualificatifs, des noms abstraits, des noms concrets, etc. Par là, je ne veux pas nier, chose évidente, que l'oeuvre est (ou mieux, fut, parce qu'elle a disparu) un ensemble, dans lequel chaque partie était solidement liée par ses fonctions avec toutes les autres, mais je veux suggérer de renvoyer à plus tard l'analyse des connexions, qui font exister la structure de l'oeuvre. Je veux suggérer aussi que le matériel fourni par les fragments offre rarement la possibilité de construire une image complète, bien que probable, de l'ensemble de la structure de l'oeuvre perdue : nature, fonction et signification des rapports et des connexions, qui couraient entre les grandes parties de l'oeuvre, par exemple, entre les chapitres ou même entre les livres d'une oeuvre historiographique perdue; ou bien entre les compositions, qu'un poète rassembla en un même livre; ou encore entre les actes d'une oeuvre dramatique (dont, bien entendu, le canevas ne serait pas connu à partir d'une autre source) : en fait, la nature, la fonction et la signification des rapports et des connexions entre ces parties ne seront pas apparentes; elles seront au-delà de l'horizon du chercheur, qui ne dispose que de fragments. Bref, la vision d'ensemble, depuis les microstructures jusqu'aux mégastructures, reconstruite selon des probabilités sur la base des échantillons, n'est jamais ou n'est que rarement possible dans le cas considéré.

Une fois envisagés les ensembles de tous ces éléments, que le matériel fourni par la constatation a permis de définir de manière claire et, autant que possible, non équivoque, ce sera le moment d'établir des cor-

rélations entre les ensembles pour arriver à une définition statistique du style. Naturellement, de ce style, le lecteur a, bien qu'à travers des fragments, une impression; et souvent il est même en mesure d'indiquer les situations de langage, qui donnent naissance aux effets phoniques et sémantiques, qui se nomment style; en plus d'être qualitatif, le style est, comme tous les phénomènes qui ont lieu dans l'emploi de la langue, quantitatif; c'est pourquoi il est intéressant de quantifier les faits, pour le dire comme de Saussure, syntagmatiques et les faits associatifs, et d'établir des corrélations (ce sera au chercheur à décider quelles corrélations établir, sur la base de sondages qualitatifs opérés sur les échantillons, même d'extension limitée); ou bien, pour employer le vocabulaire traditionnel, il sera intéressant d'établir des corrélations entre le lexique (c'est-à-dire les ensembles se rapportant au lexique) et la syntaxe et la morphologie (c'est-à-dire les ensembles se rapportant à l'une et à l'autre).

Tout ce qui précède tente de répondre à la question que je posais en commençant : est-il possible, sur la base du matériel repérable dans des fragments de courte ou de très courte extension, d'avancer des hypothèses raisonnables, ou mieux, d'avancer des hypothèses selon la méthode scientifique sur l'univers des choix de langage et le style d'une oeuvre ? Je ne dis pas d'un auteur, je ne dis pas de toute l'oeuvre d'un auteur, pour des raisons évidentes : mais, en se basant sur l'hypothèse structurale, il est sans aucun doute légitime de supposer que d'un nombre suffisant d'échantillons l'on puisse dériver un ensemble d'éléments tels qu'ils définissent la signification et la fonction des choix de langue et de style opérés par un auteur dans la composition d'une oeuvre.

Mais, comme j'ai tenté de l'expliquer, la réalisation de ce résultat, son importance, son ampleur et son étendue, dépendent de la quantité et de la

qualité du matériel repérable dans les échantillons. A ce sujet, l'extension moyenne de l'échantillon est particulièrement importante.

*

* * *

Ce que je viens d'exposer concernait le fragment parvenu par voie directe. Il peut être traité comme un échantillon fortuit et offre toutes les garanties, quant aux constatations faites sur lui et aux déductions qui peuvent en être tirées en vue de la reconstruction selon les probabilités de certaines caractéristiques de l'oeuvre disparue, toutes les garanties, dis-je, que les statisticiens attribuent à l'échantillon fortuit.

Toutefois, il y a aussi dans la littérature grecque et latine des fragments qui nous sont parvenus par voie indirecte, c'est-à-dire par d'autres écrivains qui les ont cités, et qui ont donc été conservés dans le corps du texte d'une autre oeuvre. Ces fragments posent un problème difficile : ils ne sont pas fortuits, car l'auteur qui les a insérés comme citation dans son ouvrage, en a considéré le plus souvent le côté lexical, mais aussi parfois l'aspect morphologique ou syntaxique. J'ai défini les trois points selon la théorie traditionnelle de la langue; toutefois, cela ne cache pas au lecteur leur importance par rapport à la formation d'un style.

Le problème consiste en ceci : est-il légitime de considérer l'élément, à cause duquel le fragment, ou mieux, le passage a été extrait de l'oeuvre à laquelle il appartenait (il s'agit précisément d'*excerpta*) comme correctement calculable dans les constatations des données ? La reconstruction

des caractéristiques de l'oeuvre disparue ne sera-t-elle pas faussée par des éléments dont la présence ne dépend pas du choix fortuit d'un échantillon ?

Que l'on considère le cas de fragments-*excerpta* conservés à cause de la présence en eux de formes vulgaires et archaïques par des grammairiens et des érudits de la langue latine. Une fois identifié l'auteur de l'oeuvre ou des oeuvres, d'où les *excerpta* ont été extraits, une fois vérifié quels fragments-*excerpta* proviennent de chaque oeuvre, une fois effectuées les constatations, faits les comptes et constaté que, dans le vocabulaire constitué sur la base des *excerpta*, sur cent vocables, douze présentent des formes archaïques et sept des formes vulgaires, il n'est pas possible d'en déduire que l'oeuvre d'où les *excerpta* ont été extraits, présentait, sur le total des vocables utilisés, 12 % de formes archaïques et 7 % de formes vulgaires. Et cela pour une raison très simple : il n'est absolument pas certain que, dans les parties disparues, la distribution des vocables respecte les mêmes valeurs; on pourrait même supposer, à vouloir être candide ou pessimiste, que dans les parties disparues il n'y avait aucune forme vulgaire ou archaïque, et que celles qui ont été conservées étaient les seules. Ce sont là, disons-le, les cas limites : dans les parties disparues, ou bien pourcentages égaux à ceux des *excerpta* ou bien absence totale. Il est donc clair qu'il n'est pas légitime, vu l'intentionnalité de l'*excerptum* (en ce sens que l'*excerptum* a été fait pour conserver, en tant qu'exemple, un élément linguistique jugé intéressant par l'auteur), de calculer la probabilité de l'élément à cause duquel l'*excerptum* nous est parvenu. Faut-il pour cela refuser de prendre en considération l'*excerptum* en entier ? Cette solution d'exclusion radicale libérerait le chercheur de toute préoccupation; mais dans la plupart des cas, elle le libérerait aussi de la possibilité de mener des recherches sur des

oeuvres fragmentaires; en effet, la majeure partie des oeuvres fragmentaires, qui nous sont parvenues de l'antiquité grecque et latine, sont précisément constituées par des *excerpta* et des citations.

Mais d'autres solutions sont possibles : en premier lieu, procéder à une classification des fragments sur la base de l'élément qui en a provoqué la conservation; éliminer de la recherche les groupes de fragments, distingués par un élément révélateur des choix de langue et de style. Mais le résultat serait d'affaiblir précisément la reconstruction de ces choix; ou même d'enlever tous les fondements ou la majeure partie d'entre eux au procédé statistique de reconstruction.

Reste alors le procédé le plus efficace et en même temps le plus économique, en ce sens qu'il permet d'utiliser au maximum les échantillons dans les limites des critères de correction de la méthode statistique : partant de la répartition des fragments en classes, comme nous l'avons déjà mentionné, par rapport à l'élément qui en a provoqué la conservation et procédant à la constatation des données se rapportant aux caractéristiques de la langue et du style, but de la recherche, l'élément "intentionnel" ne sera pas considéré au cours de la constatation concernant cet élément, tandis que tous les autres éléments repérables dans le fragment seront relevés et auront leur juste poids dans la reconstruction des choix linguistiques et stylistiques.

Procédant par exemple à la constatation des éléments aptes à déterminer les caractéristiques de la langue employée dans l'oeuvre perdue, ni l'archaïsme, ni le vulgarisme, ni l'idiotisme etc. ne seront pris en considération, au cas où ces éléments ont été la cause de la transmission du fragment, alors que l'on se servira d'archaïsmes, de vulgarismes, etc. figurant

dans le texte de fragments transmis pour d'autres raisons. Mais cet archaïsme sera considéré non pas en tant qu'archaïsme, mais en tant, par exemple, que technicisme, au cas où, transmis comme archaïsme, il porterait aussi une signification technique et où, pour déterminer le modèle de langue employé, il conviendrait de rechercher aussi la présence de vocabulaire technique dans l'oeuvre perdue.

En procédant ainsi, il est naturel que l'on ne tienne pas compte de manière rigoureuse des exigences de l'hypothèse structurale; parce que, à vouloir être des structuralistes rigoureux, il semble naturel de supposer que la présence d'un vulgarisme dans un passage n'ait pas été sans conséquences tant sur le choix des autres vocables que sur le choix des structures syntaxiques, des figures rhétoriques etc. Toutefois, si cette affirmation est raisonnable du point de vue structuraliste, bien qu'elle soit fondée plus sur un raisonnement intuitif que sur une preuve scientifique ou une démonstration empirique, il reste le fait qu'il ne semble pas y avoir d'autre voie de sortie entre celle que je propose et l'exclusion du fragment tout entier. Du reste, l'exigence structuraliste, mise de côté pour le moment, resurgira lorsque les résultats des recherches et des constatations sectorielles seront reliés. On constituera, par exemple, des tables, où la donnée lexicologique rencontrera la donnée syntaxique, cette dernière la donnée métrique, celle-ci la donnée rhétorique etc. Il est naturel que le vide créé par l'exclusion d'un élément ne sera pas rempli par cette opération structurale successive; mieux vaut cependant pécher par omission que par excès.

A. PENNACINI,
Université de Turin.