

PREMIERS PAS DANS L'APPLICATION DE L'INFORMATIQUE A L'ÉTUDE DES TEXTES PHILOSOPHIQUES

par ANDRÉ ROBINET

Je voudrais d'abord très rapidement vous rappeler les diverses étapes qui nous ont amené, au CENTRE D'HISTOIRE DES SCIENCES ET DES DOCTRINES du CNRS, à préciser les points que j'ai mis à l'ordre du jour pour ce matin. Ce sont des publications que vous connaissez peut-être sur les applications de l'informatique aux textes philosophiques¹. En ce qui concerne les publications, il y a d'abord des résultats d'occurrences avec les index alphabétiques classés de toutes les références qui permettent de retrouver les formes dans les textes; il y a ensuite des concordances qui sont effectuées à partir de ces occurrences; et il y a enfin différents types de tableaux: ce sont des tableaux de co-occurrences où l'on prend comme repère, le mot que l'on veut étudier dans un texte et on se demande comment les autres items se disposent aux alentours. Cette présentation a un avantage pour les philosophes, c'est qu'elle est claire et distincte, car il ne faut pas montrer des choses trop géométriques ni des choses trop compliquées dans la présentation externe. Ici, on a à la fois l'éloignement des formes par rapport au mot-pôle, leur éloignement dans le sens droit ou dans le sens gauche jusqu'à

¹ *Documentation et discipline philosophique*, Editions du CNRS, 1969; *Les Applications de l'informatique aux textes philosophiques*, Editions du CNRS, 1972; *Les Etudes philosophiques et l'ordinateur*, in « *Revue Internationale de Philosophie* », n° 103, 1973; « *Cirpho* » (*Revue de Cercle International de Recherches Philosophiques par Ordinateur*).

plus 10 et moins 10 items, et dans chaque colonne le nombre le plus grand dans lequel on rencontre la forme; par exemple ici, pour le mot 'substance' dans le *Discours de Métaphysique* on s'aperçoit que le mot 'individuel' est 5 fois à la position + 1 et une fois à la position — 1. Par conséquent on peut toujours commencer à émettre l'hypothèse qu'entre 'substance' et 'individuel', dans ce texte, il y a des relations préférentielles et qu'il y a constitution d'une association binaire. Je ne voudrais pas m'arrêter à cela; mais on vous a donné hier plusieurs documents et je voudrais seulement vous soumettre quelques indications sur l'usage qu'on en peut faire. Il est bien évident qu'au sens que l'on vient de définir, ces éléments-là nous permettent d'armer les éditions d'une manière tout à fait neuve; il faut quand même reconnaître que ces possibilités ne nous étaient pas offertes auparavant (voir *Philogramme I*).

Le *Philogramme II*, tiré de PIM 71 analysait 10 textes de Malebranche, sur le thème 'Théorie de la connaissance chez Malebranche', c'est-à-dire la vision en Dieu. Dix textes différents de T¹ jusqu'à T¹⁰ qui étaient datés de 1674 jusqu'à 1712. La machine a évidemment dépouillé chacun des textes sur les différentes années et nous a donné les listes pour leurs diverses occurrences et leurs diverses fréquences. Cela peut attirer notre attention sur les reliefs lexicographiques; ces reliefs lexicaux ont-ils une signification? C'est ce que nous pouvons nous demander purement et simplement; par conséquent je n'émetts pas d'hypothèse philosophique au départ; j'observe purement et simplement ce que nous donne cette manière d'investigation du texte. En même temps, j'ai sollicité la machine de totaliser l'ensemble des textes T¹-T¹⁰ et de constituer un grand ensemble de données qui est tout le vocabulaire, lequel par une réduction au coefficient d'une fréquence absolue, n'a évidemment aucune signification: il faut qu'elle soit référée au nombre total des items qui se présentent dans ce texte pour avoir un sens; à partir de cela, quand nous avons les coefficients de fréquence

pour chacun des textes, ils deviennent comparables entre eux. Par ce procédé, j'ai donc une moyenne d'utilisation des items dans l'ensemble du lot. Dans chaque sous-ensemble ce sont les contextes qui vont se présenter par rapport à cette moyenne de façon différente; je vais donc avoir des textes qui vont être situés au-dessus. Voici un exemple précis. Le mot qui est utilisé à plus haute fréquence dans ces textes est le mot 'Dieu'. D'ailleurs c'est assez curieux, car dans les dépouillements de PIM 71 qui traitent de Malebranche, MONADO 72 qui traitent de Leibniz, PENSÉES 74 qui traitent de Pascal, dans certains textes de Descartes dont on vous parlait tout à l'heure, « Dieu » intervient toujours en très haute fréquence, ce qui n'est pas le cas dans d'autres ensembles philosophiques, et il y a là manifestement l'indication d'une sorte de concentration à l'intérieur de la pensée de l'auteur, en fonction de quoi la pensée de l'auteur se déroule et se développe, s'organise. Ce grand relief apparaît avec des termes de très haute fréquence et la disposition alors sur mon tableau est la suivante: les chiffres sur la première ligne signifient pour Dieu: 6 dans les textes 6, 7, 9, 5 c'est-à-dire dans des textes assez tardifs et 4, 8, 10, 3, 1, 2. 10, 3, 1, 2 ce sont les premiers des textes; je vous donne tout de suite une clef: T¹, T³, et T¹⁰ c'est en fait la *Recherche de la Vérité*, c'est le même texte fondamental de 1674, revu en 1678 et corrigé en 1712. Il est donc assez normal qu'on retrouve T¹⁰, bien qu'il soit le dernier chronologiquement, dans un lot lexical qui porte toujours la trace de ses origines. L'association 10, 3, 1 représente donc le même texte; mais comme vous allez le voir, avec des nuances lexicographiques importantes pour les reliefs intérieurs à ce groupe. Cela veut dire que dans les textes de 10, 3, 1, et 2, qui sont les *Conversations chrétiennes*, dans les premiers textes de Malebranche où il réfléchit à la vision en Dieu, le mot 'Dieu' apparaît avec une fréquence beaucoup plus élevée que lorsqu'il en parle dans les textes plus tardifs. Prenons d'autres termes significatifs qui nous sont signalés par

les fréquences, précisément. Vous avez ensuite le mot 'idée', le mot 'esprit', le mot 'chose'. J'avais parlé à Montréal de parallélisme lexico-philosophique: la vision en Dieu par les idées qui est pour un esprit la manière de connaître les choses entraîne en quelque sorte sous la plume de l'auteur une redondance que l'analyse fréquentielle nous signale tout à fait normalement. Or cette redondance fait que nous nous apercevons que les textes 1, 3, 10 de la *Recherche de la Vérité* et le texte 2 également, se trouvent toujours en très bonne position par rapport aux autres textes. C'est-à-dire que le quatuor Dieu-idée-esprit-chose dans ces textes de Malebranche sur la vision en Dieu, viendraient préférentiellement dans les premiers écrits et seraient beaucoup moins utilisé par la suite. C'est tout ce que je peux dire.

Deuxièmement, je fais le même sondage pour d'autres termes et je prends des termes qui ont des indices moyens importants, qui ne sont pas de très hautes fréquences, et justement il est bon comme on le disait tout à l'heure de fouiller à tous les niveaux des fréquences. Nous avons 'étendue', 'perception' et 'intelligible'. Ce sont des termes intéressants, bien mis en relief par là, et on s'aperçoit que ceux-ci par contre, dans les premiers textes sont en somme très peu employés au début et vont être plus employés dans les textes qui viennent plus tardivement. Constatons ceci purement et simplement; il y a une sorte de balancement qui s'effectuerait entre d'une part un groupe lexical comme celui des quatre premiers mots de ce schéma, où il semblerait que l'auteur les utilise beaucoup dans les premières de ses rédactions, puis qu'ils soient un peu délaissés et que l'accent soit mis ensuite sur étendue-perception-intelligible qui n'apparaissent pas beaucoup au départ.

A ce moment-là, et je réponds ici à Mr. Garin: l'historien de la philosophie, sûr des données extrêmement précises de ce genre, peut précisément confirmer quelque chose que l'on connaît fort bien: c'est la transition, de la théorie de connaissance de Malebranche qui, au départ, s'interroge sur l'origine de l'idée:

quelle est l'origine de l'idée? — elle est en dieu. Et qui, deuxièmement, dans une phase postérieure, s'interroge sur la nature de l'idée. Donc, quand il aborde le problème de la connaissance, il change d'orientation lexicale, en quelque sorte, ce que nous philosophes désignons par un changement, non pas d'optique sur le problème de la connaissance, mais un changement d'attention apporté par un aspect du problème de la connaissance. Je vais maintenant en bas du tableau avec un mot comme 'efficace'. Regardez ce qui se produit à ce niveau-là. Je ne trouve 'efficace' ni dans 1, ni dans 2, ni dans 3, 4, ou 5. Il n'est pas dans les 5 premiers textes où Malebranche s'interroge sur la vision en Dieu. Il apparaît dans 6, 9; il est dans 10, dans 8 et dans 7. Que se passe-t-il ici? Vous vous apercevrez que 10 est subitement séparé des autres versions par l'intervention du mot 'efficace' dans cette rédaction. Il s'agit en fait de la *Recherche de la Vérité*, III, II, VI. C'est ce texte qui a été analysé trois fois, sous sa forme 1674, sous sa forme 1678, sous sa forme 1712, il nous met là, bien en relief ce qui n'est pas tout à fait un hapax, mais deux occurrences du mot 'efficace' qui apparaissent dans cette rédaction de 1712. Il y a effectivement, si l'on va voir le texte à la suite de cette révélation lexicale, un nouveau paragraphe que Malebranche a composé et où il apporte la preuve de l'existence de l'idée par son efficace sur l'esprit. On a donc là une dissociation qui intervient et qui est mise en relief par cette analyse sémiotique uniquement. J'ai pris ce premier exemple pour insister sur le rôle de ces reliefs lexicaux sur la génétique de la pensée philosophique.

C'est également très important et j'insiste sur le deuxième exemple, sur l'étude des structures. Le *Philogramme* III, c'est ce petit soleil qui se rapporte à l'opération MONADO 72 et où j'ai ensuite essayé de repérer par rapport à un mot pris comme centre, comme pôle, le mot 'substance'. Les deux textes de Leibniz analysés sont le *Discours de Métaphysique* (DM) qui est indiqué en pointillé; je vous rappelle que le *Discours de Méta-*

physique est de 86 et que la *Monadologie* est de 1714. Si j'observe les co-occurrences et les fréquences qui me sont données, je m'aperçois d'une chose: c'est que (regardez s'il vous plaît, la ligne du mot 'simple' sur la ligne continue, c'est-à-dire le *Discours de Métaphysique* de Leibniz), je n'ai aucune position où 'simple' vient à droite du mot 'substance', par exemple, ou s'associe d'une manière binaire. Par contre dans le texte de la *Monadologie*, à la position + 1, j'ai 24 fois le mot 'simple'; je l'ai une fois à la position + 3, je l'ai une fois à + 7, + 9 ... Par contre, si je me demande ce qui jouait le rôle de 'simple' dans le *Discours de Métaphysique*, je m'aperçois que le mot 'particulière' y est très important ainsi que le mot 'individuelle'. Ils jouent un peu le rôle que 'simple' jouait dans la *Monadologie*. Voyez, sur la ligne 'individuelle', vous avez 8 fois 'individuelle' à + 1, vous avez 4 fois 'particulière' à + 1, une fois à + 2, une fois à + 4. C'est un constat purement lexical, mais je suis en droit comme historien de la philosophie de me demander: est-ce que Leibniz n'avait pas le mot 'simple' à ce moment-là? Non, il ne l'avait pas avec 'substance'. Il avait encore, et comme Arnauld précisément qui raisonnait en raison d'une conception plus aristotélicienne de la 'substance' où intervenait le mot 'individuelle' et le mot 'particulière' mais la 'substance individuelle ou particulière' ce n'est pas du tout la 'substance simple'.

Ces deux premiers schémas sont des schémas manuels, tirés par moi-même de l'étude de ces listes, par un historien de la philosophie. Pouvons-nous automatiser ce genre de sortie? C'est le problème que vous avez sur cette étude de dispersion. Ce problème se rapporte au programme MONADO 74 qui est en cours d'exécution à Bruxelles. On y a abouti aux mêmes résultats, par élaboration automatique, et j'aime assez cette procédure car elle est obtenue par l'automatisation intégrale des données. Car je crois que l'historien de la philosophie peut aussi mettre à son service la mathématisation et l'intelligibilisation de ces grands ensembles que sont les textes qu'ils manipulent.

CLASSEMENT DES CO-OCCURRENTS EN FONCTION DE LEUR DISTANCE ET DE LEUR CO-FRÉQUENCE LOCALE

[illegible]

Philogramme I

Tiré de PIM 71; Coeff. F. R. comparés sur T_1-T_{10}





