::Savoirs LE FIL DES IDÉES

Manier le thésaurus grec

Lieux de savoir, 2. Les mains de l'intellect, Albin Michel, 2011, p. 555-578

Aurélien Berra

1 Une conférencière vient d'analyser les explications antiques d'un épisode célèbre de l' *Iliade ;* un spécialiste de rhétorique l'interroge sur une scholie qu'elle a traduite : selon le commentateur, les Siciliens Corax et Tisias auraient « orné » l'art de la parole. Ne l'ont-ils pas plutôt « ordonné » et institué, comme le veut la tradition ? Le verbe grec qui figure dans le texte, *kosmeô*, peut avoir les deux sens. Pendant la discussion, un auditeur ouvre son ordinateur portable. Il commence par y contrôler l'article du dictionnaire de référence. Mais il s'agit en l'occurrence d'une forme peu habituelle, *ek- kosmeô*, qui n'y est pas répertoriée. Une rapide recherche lui permet de préciser que le verbe ne se rencontre pourvu de ce préfixe que dans quatre textes. Un auteur du iie siècle de notre ère évoque les inscriptions qui « ornent » un temple de Delphes ; quelques siècles plus tard, sa remarque est glosée par deux érudits anonymes. Le quatrième emploi connu figure dans le texte examiné[1].

2 En prétendant qu'Homère connaissait à la fois le nom et l'art de la rhétorique, le scholiaste affirme que les fondateurs n'ont fait que perfectionner une invention antérieure. Si l'on admet qu'il tente de concilier une information attendue avec ce qu'il considère comme le témoignage du poète, les deux interprétations sont envisageables : soit les premiers techniciens ont conféré une belle apparence à ce qui existait longtemps avant eux, soit ils ont organisé un ensemble de pratiques et de savoirs. Ce ne sont pas des passages parallèles ni les autres emplois du terme qui autoriseront à faire la part de la logique singulière du passage et celle de la culture commune aux savants grecs. Cependant, il faut avoir observé les usages du mot dans les textes qui sont parvenus jusqu'à nous pour faire des hypothèses en connaissance de cause, proposer une traduction du commentaire et compléter la description de cette langue éteinte.

3 Cette scène de la vie scientifique ne relève pas de la fiction. Au moyen

d'ordinateurs de plus en plus portatifs et de moins en moins coûteux, les hellénistes accèdent à un gigantesque corpus en cours de constitution depuis les années 1970, dont les millions de mots sont désormais reliés aux entrées du *Greek-English Lexicon* publié par Liddell et Scott en 1843. Ce corpus est le *Thesaurus Linguae Graecae (TLG)*. Présenter cet outil fondamental, l'histoire dans laquelle il s'inscrit et les usages qui en sont faits, ce n'est pas fournir un mode d'emploi ou un compte rendu – du reste, nous n'avons pas affaire à un logiciel, mais avant tout à une base de données, devenue « bibliothèque numérique » à l'ère du réseau mondial. C'est bien plutôt explorer un « lieu » emblématique des transformations de la philologie : en réfléchissant au maniement de la bibliothèque grecque, on discerne les problèmes et les promesses de la science des textes la plus contemporaine.

L'inventaire du trésor : jadis, naguère et aujourd'hui

4 Le fait que les études anciennes aient bénéficié très tôt du développement des technologies informatiques est assurément dû à l'importance des cultures latine et grecque dans le monde occidental et à leur place dans l'éducation, dans l'enseignement universitaire et dans la recherche. Ce sentiment d'une valeur particulière, longtemps entretenu par les institutions, avait en outre favorisé une exploitation aussi intensive qu'extensive de la littérature classique. Non seulement le champ était prêt à fournir aux machines des informations relativement fiables et complètes, mais l'inventaire du « trésor » et le repérage des textes avaient été conduits d'une façon extrêmement précise depuis la Renaissance. Aux XIXe et XXe siècles, les textes canoniques avaient déjà fait l'objet de plusieurs éditions, fondées sur les manuscrits et sur les corrections savantes. Pour les spécialistes des langues, des littératures, des sociétés et des monuments, les conventions de référence étaient si fermement établies qu'elles possédaient l'efficacité d'un cadastre, au sein d'une « science de l'Antiquité » hautement organisée. Éditions des sources, traductions, commentaires, études, lexiques, concordances et encyclopédies : une longue tradition avait accumulé dans les bibliothèques des documents de divers niveaux qui visaient à transmettre les matériaux culturels hérités, mais aussi à fonder leur interprétation. Il fallait les mettre en relation, d'une façon critique, selon des protocoles de citation et de renvoi ouvrant la possibilité d'un contrôle par les lecteurs instruits et par les pairs. En d'autres termes, un vaste ensemble de données et de métadonnées avait été élaboré pour l'étude du monde antique.

5 Pour resituer pleinement le *Thesaurus Linguae Graecae*, ce « trésor de la langue grecque », dans l'évolution des pratiques savantes, il faudrait

construire une perspective dont on cherche encore le lieu et la formule. Cette histoire ne saurait se résumer à une chronologie. Le défi est de comprendre la dynamique des sciences de l'Antiquité modernes au sein du mouvement des « humanités numériques », que l'on peut considérer comme l'application de l'informatique aux sciences de la culture. Si cette discipline en devenir – parfois désignée comme une « interdiscipline » ou une « transdiscipline » – s'institutionnalise à grande vitesse depuis les années 1990, c'est à partir du milieu du xxe siècle qu'elle s'est constituée et il est évident qu'elle plonge ses racines dans une (pré)histoire longue et complexe[2].

6 Il suffira de s'interroger sur le nom anglo-latin du *TLG* pour apercevoir certains problèmes cruciaux. À travers trois textes représentant ici trois âges de la philologie, nous préciserons l'origine et la portée de la métaphore qui fait d'une « langue » un « trésor ».

7 Commençons par lire un article de dictionnaire qui servira notre propos[3] :

> **THESAURUS, THÉSAURUS, subst. masc.**
> **A.** – *LING.* Lexique de philologie ou d'archéologie ; dictionnaire renfermant le vocabulaire aussi complet que possible d'une langue ancienne. Synon. *trésor. Thésaurus de la langue latine.* P. métaph. *Cette institution mal connue* [*l'internat des hôpitaux*] (...) *permet à une élite de jeunes gens* (...) *d'acquérir un incomparable « thesaurus » organique, clinique, thérapeutique et, au bout du compte, philosophique* (L. Daudet, *Dev. douleur*, 1931, p. 6).
>
> **B.** – *DOCUMENTOL.* « Langage documentaire fondé sur une structuration hiérarchisée d'un ou plusieurs domaines de la connaissance et dans lequel les notions sont représentées par des termes d'une ou plusieurs langues naturelles et les relations entre notions par des signes conventionnels » (*Documentation* 1985). *Thesaurus de linguistique, de médecine ; thesaurus documentaire en sociologie ; thesaurus de l'armement. Au moment de l'indexation des documents, les mots-clés sont choisis d'après un thesaurus hiérarchisé* (...) *qui fait l'objet de refontes successives* (Funck, Moureau, dans *B. Bibl. Fr.* , t. 11, 1968, p. 343).
>
> **Prononc. :** [tez ɔ rys]. **Étymol. et Hist. 1.** 1904 « Recueil, lexique de philologie ou d'archéologie ; dictionnaire exhaustif » (*Nouv. Lar. ill.*) ; **2.** 1962 docum. (B. C. Vickery, *Techn. mod. de docum.* [trad. de l'angl.], Paris, p. 22) ; 1964 (Cros-Gardin). 1 empr. au lat. *thesaurus*, gr. θησαυρός (v. *trésor*), empl. comme titre de lex. (R. Estienne, *Dictionarium seu Latinae linguae thesaurus*, 1531 ; H. Estienne, *Thesaurus graecae linguae*, 1572) ; 2 prob. empr. à l'angl. *thesaurus* « liste

de mots ou de concepts classés d'après leur sens » (1852 dans *NED Suppl.2*), puis terme de docum. (1957, *ibid.*), de même orig. que 1.

Trésor de la langue française (1971-1994)[4].

8 Ce *Trésor* est l'aboutissement de la plus grande entreprise lexicographique du siècle dernier pour ce qui concerne le français. La structure de l'entrée nous est familière. Deux sens du mot vedette sont distingués par leur domaine d'origine, la « linguistique » et la « documentologie ». Les définitions sont suivies d'exemples, qui sont soit des syntagmes typiques, soit des emplois contenus dans des citations. Les dates des extraits signalent l'antériorité du sens A, aspect développé dans une rubrique finale consacrée à l'« étymologie » et à l'« histoire » du mot. Quoique souvent abrégées, les références bibliographiques suivent un format standard, où la séquence maximale d'informations comporte auteur, titre, lieu, date, tomaison et numéro de page.

9 Un tel degré de précision ne s'explique pas seulement par les exigences scientifiques des dictionnaires contemporains. Il découle d'un processus de composition, car les rédacteurs du *TLF* se sont fondés sur les 2 600 textes des xixe et xxe siècle que la base informatique Frantext comprenait dans sa première version. Entièrement indexé, ce corpus de textes littéraires et non littéraires assortit en outre de repères statistiques un grand nombre de notices. Peut-être le mot « thésaurus » est-il trop rare pour que les données soient pertinentes ; pour son synonyme « trésor », par exemple, nous disposons d'une fréquence absolue globale et de fréquences relatives par demi-siècle.

10 Quant à l'évolution du terme, retenons que le sens premier de dictionnaire des langues anciennes – ou, pour l'archéologie, de dictionnaire encyclopédique – s'est infléchi dans deux directions, en abandonnant la référence à l'Antiquité : l'exhaustivité est devenue le trait central de « thésaurus », tandis que le sens documentaire plus récent repose sur l'idée d'une organisation conceptuelle des lexèmes au sein d'un répertoire de mots clés[5]. Cependant, la fortune de ce mot technique dans les langues européennes est intimement liée à l'étude des textes anciens : calquant l'usage du français « trésor », employé avec cette valeur depuis le xiiie siècle, l'humaniste Robert Estienne a emprunté le latin *thesaurus* pour nommer son dictionnaire de la langue latine[6].

11 À la génération suivante, son fils Henri est l'auteur du plus célèbre des dictionnaires grec-latin, qui reprend le titre de *Thesaurus*. En voici l'article θησαυρός *(thêsauros)*, « trésor » :

> Θησαυρὸς, οῦ, ὁ, idem quod θήκη, Repositorium, Conditorium : Locus in quo congestæ opes reponuntur & conduntur, ut cellæ granariæ & ærariæ. Poll. libro 9, καὶ θησαυροὶ καὶ ταμιεῖα, ἵνα τὰ χρήματα καὶ οἱ πυροὶ ἔκειντο. Aristot. Œconom. 1, τούς θησαυροὺς τοὺς παρὰ τὰς ὁδοὺς τὰς βασιλικὰς ἀναπληροῦν ἐκέλευε. Et mox, ἐπώλει τὰ ἐκ τῶν θησαυρῶν. Sic Joseph. κατασκευάσας θησαυρὸν καὶ κλείσας παντα- χόθεν ὀπὴν ἐν αὐτῷ μίαν ἤνοιξεν, Ærarium, Pecuniarum repositorium seu conditorium. Et Plutarch. ἐκέλευσεν αὐτῷ τούς θησαυροὺς ἀνοτξαι τῶν χρημάτων. Vbi nota additum genitiuum, omissum in locis præcedentibus. Idem in lib. De solert. anim. γνωρίσασα τὸν ἑαυτῆς ἑκάστη θησαυρόν, ὡς οὐδεὶς χρυσίου θήκην ἄνθρωπος, ἀσμένως ἀνοίγει καὶ προθύμως. [...] Latini quoque thesauros dicunt : ut uidebis apud Gellium lib. 2, c. 10. Sed & ipsæ opes alicubi repositæ seu reconditæ, uocantur θησαυρός. Athen. libro 6, ’Αλεξάνδρου τοὺς ἐκ τῆς ’Ασίας θησαυροὺς ἀνελομένου. [...]
>
> Henri Estienne, *Thesaurus graecae linguae* (1572)[7].

12 Si son lexique est le grand œuvre de cet imprimeur et érudit, c’est parce qu’il a résumé et stimulé une inlassable activité d’éditeur de la littérature grecque. Le *Thesaurus graecae linguae* – à ne pas confondre avec le *TLG* moderne, qui lui doit son nom – a bénéficié des dizaines d’éditions qu’Henri Estienne a préparées avant et pendant sa composition. Une évidente volonté de totalisation lui a ensuite fait poursuivre conjointement la réflexion sur le texte des auteurs antiques et le perfectionnement du dictionnaire issu de leur fréquentation. Ce *work in progress*infini a donc pour substrat la mise en fiches des classiques ou, si l’on veut, une vie entière d’indexation[8]. Du reste, le travail effectué était si considérable que l’ouvrage est demeuré incontournable pendant trois siècles, jusqu’à la première édition du *Lexicon* d’Oxford déjà mentionné. Au xixe siècle encore, on s’est efforcé de l’augmenter et de le purger de ses imprécisions afin d’en rendre la consultation plus aisée et plus sûre[9].

13 L’explication que la notice fournit du mot grec correspond aux renseignements de nos dictionnaires de référence : *thêsauros*désigne un « dépôt », un « magasin où l’on enferme provisions et objets précieux », un « trésor », parfois une « cassette » et, dans les papyrus, un « magasin à grains » ; son emploi dans un sens figuré est attesté dès l’époque archaïque [10].

14 Au fil d’un texte continu, mais qui fait usage de la ponctuation et des majuscules, exemples grecs et gloses latines se succèdent. Comme on le voit aussitôt, les références n’obéissent pas à un format systématique. Les citations sont la source même des définitions. Elles

donnent un contexte au mot *thêsauros*, sans permettre toujours de retrouver le passage, lorsque seul le nom de l'auteur est indiqué. L'ouvrage appartient néanmoins à une époque décisive pour l'organisation du patrimoine littéraire, désormais imprimé, car on met au point une présentation et un découpage des textes en chapitres qui autorisent une consultation efficace et la pratique savante de la référence[11]. C'est ainsi que l'on divise la Bible en versets et que les pages numérotées des classiques portent à intervalles réguliers les premières lettres de l'alphabet, en marge ou entre deux colonnes. Notre système de référence est souvent l'héritier de ces tentatives. Pour citer Platon, Plutarque, Strabon ou Athénée, nous utilisons la pagination des éditions humanistes, en la dissociant des *in-folio* d'origine et en la faisant cohabiter parfois avec un découpage « logique » de l'œuvre.

15 Le troisième texte que nous observerons est d'une nature assez différente. Il s'agit d'un extrait d'Athénée, qui n'est pas lexicographe au premier chef, mais dont la compilation montre de façon exemplaire le mode de référence des érudits anciens les plus précis. Ainsi dans ce passage, où l'on apprend qu'une ritournelle populaire était parodiée dans une comédie intitulée *Le Trésor* [12] :

> Ὑγιαίνειν μὲν ἄριστον ἀνδρὶ θνητῷ,
> δεύτερον δὲ καλὸν φυὰν γενέσθαι,
> τὸ τρίτον δὲ πλουτεῖν ἀδόλως
> καὶ τὸ τέταρτον ἡβᾶν μετὰ τῶν φίλων.
>
> Αἰσθέντος δὲ τούτου καὶ πάντων ἡσθέντων ἐπ' αὐτῷ καὶ μνη- μονευσάντων ὅτι καὶ ὁ καλὸς Πλάτων αὐτοῦ μέμνηται ὡς ἄριστα εἰρημένου, ὁ Μυρτίλος ἔφη Ἀναξανδρίδην αὐτὸ διακεχλευακέναι τὸν κωμῳδιοποιὸν ἐν Θησαυρῷ λέγοντα οὕτως
>
> f Ὃ τὸ σκόλιον εὑρὼν ἐκεῖνος, ὅστις ἦν,
> τὸ μὲν ὑγιαίνειν πρῶτον ὡς ἄριστον ὂν
> ὠνόμασεν ὀρθῶς· δεύτερον δ' εἶναι καλόν,
> τρίτον δὲ πλουτεῖν, τοῦθ', ορᾷς, ἐμαίνετο·
> μετὰ τὴν ὑγίειαν γὰρ τὸ πλουτεῖν διαφέρει·
> καλὸς δὲ πεινῶν ἐστιν αἰσχρὸν θηρίον.
>
> Être en bonne santé, voilà ce qu'il y a de meilleur pour un mortel,
> en second lieu, avoir la chance d'être beau,
> en troisième, s'enrichir sans fraude
> et, en quatrième, être jeune et entouré d'amis.

16 En entendant cette chanson, tout le monde se réjouit et se rappela que le noble Platon la citait comme une excellente formule. Myrtilos dit alors que le poète comique Anaxandride tournait en dérision la

chanson en question dans *Le Trésor*, par ces mots :

> L'homme qui inventa cette chanson, quel qu'il fût,
> en disant qu'être en bonne santé est le premier de tous les biens
> nomma la chose avec exactitude ; mais avoir dit qu'être beau est le second,
> et le troisième s'enrichir, là, vois-tu, c'est de l'égarement :
> après la santé, être riche est ce qu'il y a de meilleur.
> Un bel homme qui a faim, c'est une bête sans foi ni loi !
>
> Athénée, *Les Deipnosophistes*, XV, 694 e-f (vers 200 de notre ère)[13].

17 À la fois construction littéraire et enquête sur les cadres sociaux et le lexique de l'alimentation, *Les Deipnosophistes* sont eux-mêmes une sorte de thésaurus. On perçoit à travers ces quelques lignes de quelle manière l'auteur met en scène la circulation des textes dans un banquet. Le personnage, un « grammairien », introduit ses citations tout comme le fait Athénée, en rapportant au moins l'auteur et le titre du texte ; en l'occurrence, la chanson est complète et l'allusion à Platon est un cas rare de connivence sans référence bibliographique. La compilation s'appuie sur un usage fréquemment explicite de la littérature secondaire antique et sur une pratique personnelle d'annotation et de copie.

18 Nos trois exemples témoignent de méthodes de documentation et de structures de communication différentes. Ce parcours sommaire retrace la constitution progressive d'un mode de référence spécialisé, qui est sans doute l'une des caractéristiques les plus fondamentales de l'activité savante. De ce point de vue, les textes grecs offrent un exemple extrême de réticulation ou de balisage en vue de l'interprétation collective qu'est la recherche. Cet usage social particulier repose sur l'usage concret des livres. En effet, s'il est vrai qu'« un livre est une machine avec laquelle on pense[14] », la formule est plus juste encore au pluriel : à travers le lecteur, ce sont des communautés, avec leurs traditions, qui mettent en œuvre et poursuivent le dialogue des livres dans les bibliothèques. Cela implique que les auteurs, les titres et les maisons d'édition soient des données identifiables et stables, d'une part, que des collections soient accessibles, d'autre part.

19 On sait le rôle crucial qu'a joué le texte de la Bible dans l'histoire de l'herméneutique. D'une façon comparable, l'étude des cultures antiques s'est accompagnée d'une pratique sans cesse plus fine des corpus. Ce champ a eu les moyens, par accumulation, d'approcher l'exhaustivité et la systématicité vers lesquelles tend logiquement l'exploration d'un monde passé. Tout document produit par la culture grecque ancienne est un échantillon qui contribue à la compréhension

des autres documents et n'est lui-même compris que par leur intermédiaire. Le thésaurus semble avoir pour vocation de totaliser ces vestiges.

20 Le premier gain de la conversion numérique est de rendre le système de référence plus efficace, puisque le renvoi d'un texte à un autre se fait d'une manière presque instantanée. Cette opération intertextuelle connaît une véritable métamorphose lorsqu'une bibliothèque entière est contenue dans un disque, puis lorsque la bibliothèque numérique devient un site au sein d'un réseau[15]. Tel est le cas du *Thesaurus Linguae Graecae*, dont nous pouvons à présent mesurer ce qui le sépare d'un thésaurus comme celui d'Estienne, au-delà des hommages que constituent son nom et son adresse sur la toile[16]. La différence entre les deux projets réside-t-elle dans un maniement plus commode, dans un changement d'échelle ou dans un renouvellement de l'usage des textes ?

Le *Thesaurus Linguae Graecae*

21 Dans l'archipel des humanités numériques, son ampleur et sa longévité ont fait du *TLG* un « lieu de mémoire ». Ce statut répond à la haute ambition d'un projet phare, qui a transformé la pratique quotidienne des hellénistes. Il reflète également un niveau exceptionnel d'institutionnalisation et de financement, grâce auquel le *TLG* peut représenter l'accomplissement d'un rêve humaniste, sinon la fin d'une histoire.

22 Pour employer une catégorie typique de notre xxi e siècle commençant, le *TLG* est « né numérique » *(« born digital »)*. Au début des années 1970, le choix de l'informatique est audacieux, mais il n'est pas sans précédent. Il s'inscrit dans l'implantation progressive, au sein des sciences humaines, d'une technologie dont les avantages auraient dû être immédiatement sensibles aux linguistes et aux philologues. L'index des œuvres de Thomas d'Aquin était alors en chantier depuis une vingtaine d'années[17] ; dans le monde anglophone, on voyait apparaître les premières revues et associations professionnelles consacrées à l'usage de l'ordinateur dans le domaine des *humanities*. En revanche, le *TLG* est sans conteste l'un des premiers grands projets collectifs intégrés à une institution de recherche. Son succès repose ainsi sur l'alliance opportune d'un nouveau support et d'une infrastructure[18].

23 L'informatique promettait de résoudre les problèmes documentaires qui avaient condamné à l'échec les tentatives antérieures. Il était clair dès la fin du xixe siècle qu'il ne suffisait plus de réviser le *Thesaurus* d'Estienne pour rendre compte des progrès de l'inventaire – c'est-à-dire du témoignage des documents exhumés ou édités plus

récemment et des résultats de la critique textuelle moderne. Cependant, face à l'immensité du dépouillement des sources, un premier projet avait été abandonné peu après 1900, tandis que le centre de recherche créé en 1944 à l'université de Hambourg, sous ce même nom de *Thesaurus Linguae Graecae* et avec la mission d'élaborer un dictionnaire exhaustif, avait restreint ses ambitions[19]. Le second de ces projets était dirigé par Bruno Snell. C'est l'une de ses élèves, Marianne McDonald, qui suggéra plus tard à ses enseignants de l'université de Californie l'entreprise savante qui nous intéresse, en 1971. Cette étudiante avait pour père l'un des fondateurs de la société Zenith, spécialisée dans les technologies de la radio et de la télévision. Un million de dollars accompagnait sa proposition. Rédigeant à cette époque une thèse sur le lexique du bonheur dans les tragédies d'Euripide, elle imaginait en connaissance de cause l'utilité d'une vaste banque de données *(« databank») :* il fallait pouvoir étudier un réseau de termes et de thèmes à travers leurs occurrences et non pas seulement consulter une synthèse lexicographique.

24 La conférence internationale qui marqua l'établissement du *TLG* en octobre 1972 confirma que le but du projet n'était pas de constituer un thésaurus au sens traditionnel[20], mais un corpus de textes, ce qui élargissait le spectre des usages possibles. Devait y figurer la totalité des textes grecs conservés depuis l'épopée homérique jusqu'à l'an 200 de notre ère ; à partir de 1988, ce programme fut porté à vingt-deux siècles de littérature en incluant le monde byzantin, avec pour borne conventionnelle la date de 1453 ; l'extension au grec moderne est maintenant l'horizon déclaré du projet. Corrélativement, les textes transmis par les inscriptions et par les papyrus étaient exclus. Enfin, on prit le parti de ne pas reproduire les annotations des éditions savantes. Nous reviendrons sur ce fait lourd de conséquences, mais compréhensible en l'état des techniques.

25 Devenu une marque déposée aux États-Unis, le nom du *TLG* désigne donc à la fois une institution savante, rattachée au département d'Études anciennes de l'université de Californie à Irvine, et le corpus qu'elle a pour fonction de constituer et de diffuser[21]. Quatre versions successives furent commercialisées sous la forme de disques : les CD-Rom A (1985), C (1988), D (1992) et E (2000). Depuis 2001, le *TLG* – *« a digital library of Greek literature »* – est un site en ligne[22]. Ces états du corpus décrivent déjà une progression quantitative du simple au triple. On trouvait 27 millions de mots dans la version A et plus du double dans la version E, encore largement utilisée, qui proposait 6 625 œuvres appartenant à 1 823 auteurs ou corpus distincts. Régulièrement augmenté, le site contenait en 2009 plus de 100 millions de mots, 10 000 œuvres et 4 000 auteurs ou corpus.

26 Cette histoire est inséparable d'expérimentations et de choix antérieurs à la standardisation et à la généralisation des technologies

informatiques. Dès 1976, un nombre restreint de chercheurs put utiliser un corpus embryonnaire inscrit sur des bandes magnétiques. Cependant, ce corpus fut conçu avant que les ordinateurs ne soient véritablement « personnels » et le « *TLG* A » est réputé avoir été le premier disque compact à contenir autre chose que de la musique. Grâce au concours de David Packard, on adapta au traitement du grec ancien un ordinateur mis au point quelques années auparavant dans la Silicon Valley par la firme Hewlett-Packard[23]. Ce système, dénommé Ibycus, était nécessaire à la consultation des textes jusqu'à la fin des années 1980 et il continua à assurer l'élaboration du *TLG* jusqu'en 1999. Au tournant du siècle, l'ensemble du *TLG* migra vers un environnement Unix, conforme aux standards récents. Par la même occasion, le disque fut abandonné au profit d'une publication sur la toile (voir fig. 1) et le site intégra un moteur de recherche qui lui était propre.

Figure 1. Fig. 1. Le frontispice du *TLG* en ligne (juin 2010).

27 Résumons cette évolution. Dans sa phase initiale, la conception et l'exploitation du *TLG* étaient liées à ce que l'on peut nommer un prototype ; matériel et logiciel demeuraient indissociables. Profitant de la diffusion des machines et d'un support de mémoire externe à la fois portatif et de grande capacité, le *TLG* devint à la fin des années 1980 un auxiliaire indispensable pour bien des spécialistes et l'un des plus importants projets informatiques hors des sciences dites exactes. L'institution ne fournissait que les disques contenant la base de données dont elle était propriétaire ; divers chercheurs ou centres développèrent des outils de consultation, tels les logiciels Diogenes, Lexis, Musaios, Pandora et SNS Greek & Latin. La troisième phase est celle que nous connaissons encore : seuls l'abonnement et la consultation en ligne font bénéficier du dernier état du corpus. Ainsi, l'équipe actuelle du *TLG* a fait le choix d'une forte indépendance. Or, la dialectique de l'autonomie et de la collaboration est capitale lorsqu'un projet participe au renouvellement des instruments de la recherche avec une telle influence. On le comprendra mieux une fois notre aperçu complété.

28 L'exigence fondamentale du *computer*, ou « calculateur », est que les informations utilisées soient explicites et systématiques. Un texte numérique consiste en une suite de nombres binaires ; la base de toute opération est le maniement des chaînes de caractères que forment ces nombres. Comme tout objet numérique, le texte résulte donc d'un codage, d'une représentation informatique que la machine traduit

visuellement pour l'utilisateur humain. Celui-ci modifie ces données en leur appliquant, en fonction des résultats qu'il souhaite obtenir, certaines des règles de transformation prévues par les programmes, c'est-à-dire, plus concrètement, par les systèmes et les logiciels qu'il emploie. Grâce à ces notions très générales, nous pouvons examiner les difficultés rencontrées par les concepteurs du *TLG* et les solutions qu'ils ont inventées.

29 Tout d'abord, l'usage de l'alphabet grec et des signes diacritiques anciens – accents et esprits, mais aussi ponctuation et autres signes savants – était encore un problème. Par ailleurs, une bibliothèque numérique se devait de conserver la structure usuelle des textes, dont nous avons rappelé la genèse et l'importance scientifique. Pour répondre à ces objectifs, un jeu de conventions d'encodage spécifique fut mis au point : les fichiers de texte du *TLG* respectent ce *Beta Code*, qui n'est pas fait pour une lecture suivie, bien qu'il reste compréhensible et puisse servir à formuler des requêtes (fig. 2). Ce système a été utilisé par de nombreux projets pour traiter le grec, surtout avant l'élaboration de l'Unicode (fig. 3), dont la vocation universelle repose sur l'identification des caractères par des codes alphanumériques complexes. La constitution du corpus demande la saisie des textes suivant le Code bêta, qui est toujours effectuée en Asie par des opérateurs ignorant les alphabets grec et latin. Selon un paradoxe bien connu, cette procédure est non seulement la plus économique, mais aussi la plus efficace. Des outils de correction automatique accroissent ensuite la fiabilité des résultats.

Figure 2. Fig. 2. L'incipit de l' *Odyssée* affiché en Code bêta sur le site du *TLG*.

30 À cette structure invisible s'est adjoint un index intégral du corpus, qui est l'un des « sous-systèmes » facilitant sa consultation. Il met instantanément à la disposition de l'usager, en fonction de ses besoins, les listes d'occurrences dont la compilation et la publication nécessitaient naguère des années. En outre, l'organisation du *TLG* présuppose ce que nous avons nommé plus haut l'inventaire de l'héritage grec. Pour classer chaque document selon les catégories en vigueur et rendre viable son traitement automatisé, il fallait décider de son identification, de son attribution, de sa datation, etc., même en l'absence de consensus des spécialistes. C'est le rôle du *Canon* diffusé à partir de 1977 [24]. Son caractère exhaustif et sa commodité ont tendu à lui donner la valeur d'une norme.

Figure 3. Fig. 3. L'incipit de l' *Odyssée* affiché dans une police unicode sur le site du *TLG*.

31 L'autorité d'une telle base de données devrait naturellement être évaluée à chacune de ses utilisations. En proposant une édition numérique de travaux antérieurs, le *TLG* réalise la conversion du savoir philologique accumulé au fil des siècles. De ce point de vue, il peut passer pour la mise en œuvre la plus aboutie du projet né avec la bibliothèque d'Alexandrie. Pourtant, les principes mêmes qui l'ont rendu viable, et durablement utile, limitent sa pertinence. En effet, en ne retenant qu'une seule édition de chaque œuvre, avec d'éventuelles substitutions d'une version à l'autre, et en la dépouillant de l'apparat critique qui rendait compte de sa transmission, cette archive de la littérature grecque renvoie le chercheur à un état pré-alexandrin de la science des textes : elle passe sous silence les variantes et la distinction entre leçons manuscrites et conjectures, c'est-à-dire l'épaisseur de la tradition et, partant, les bases de toute discussion. L'image qui convient est celle d'une bibliothèque où figurerait de chaque ouvrage un exemplaire sans notes ni présentation. Ce défaut de dimension critique est d'autant plus grave que les écrits antiques forment, comme on le sait, un corpus historiquement préconstruit – par les intérêts des groupes et des individus qui en ont été les relais ou par le jeu du hasard – dont les textes ont généralement connu plusieurs copies à des époques différentes. En pratique, le *TLG* ne remplace donc jamais les livres sur lesquels il se fonde[25]. Somme d'érudition destinée à favoriser les études, il offre des textes qui doivent être constamment consultés ou contrôlés dans les éditions imprimées – tout en permettant des recherches à peine imaginables avant lui.

32 L'avènement de la toile et l'habitude des catalogues électroniques ont sans doute fait de la requête lexicale une compétence mieux partagée que jamais. On comprendra intuitivement ce que signifie le couplage d'une base de données et d'un logiciel d'interrogation : un corpus comme le *TLG* n'offre pas seulement le résultat figé d'un vaste effort de documentation, il s'adapte aux requêtes formulées. Le principe de son usage est par conséquent de repérer un ou des « lieux » au moyen de coordonnées explicites et de les faire apparaître sous la forme traditionnelle d'un texte statique, éventuellement par l'intermédiaire d'une liste de passages plus ou moins étendus répondant aux critères de recherche. Le « lecteur » applique ainsi des filtres d'une complexité variable afin de vérifier une référence ou prendre connaissance d'un texte, rarement d'une façon continue, en faisant jouer au corpus le rôle d'une bibliothèque portative, par souci de rapidité ou pour avoir

accès à un texte rare. Il peut extraire un passage, voire un texte, pour en disposer sous une forme électronique en vue d'un travail en cours ; c'est alors l'homogénéité du milieu numérique et l'accès à un texte fiable qui importent. Enfin et surtout, d'une manière plus élaborée, il peut constituer un corpus en sélectionnant des éléments dans des listes closes (celles que reprend le *Canon*, répertoire d'autorité du *TLG*) : nom de l'auteur, genre littéraire pratiqué par l'auteur, sexe de l'auteur, genre littéraire du texte, lieu associé au texte, intervalle chronologique incluant ou non les textes composés à des époques diverses ou de date incertaine. Il lui est ensuite possible de consulter l'index de ce corpus ou d'y effectuer des requêtes lexicales, qui ont trait soit à l'occurrence d'une chaîne de caractères, soit à la co-occurrence de plusieurs chaînes ; la présentation des résultats comporte le nombre total des réponses. Ces opérations élémentaires de la linguistique de corpus, accessibles à chacun et reproductibles, sont une révolution pour la philologie, dans la mesure où elles sont susceptibles de modifier notre perception des textes et les questions que nous nous posons à leur sujet.

33 Les limites du *TLG*, signalées au fil des années[26], tiennent notamment à son mode d'élaboration, à son fonctionnement et à la politique de sa diffusion. Sans entrer dans le détail, soulignons que les critiques sont aussi des hommages, car le projet et les données du *TLG* ont été le ferment de travaux innombrables. Mais ce corpus n'est ni immobile ni isolé.

34 Ses évolutions ont levé certains obstacles majeurs. Nous en citerons trois, introduites dans la dernière décennie. En ce qui concerne les polices de caractères grecs, les responsables du *TLG* ont fait mieux qu'adopter la norme Unicode, puisqu'ils ont soumis des propositions au consortium mondial qui l'administre. Le problème de la lemmatisation est particulièrement aigu pour une langue flexionnelle : des procédures automatiques autorisent désormais à appeler ensemble les formes d'un même lexème ; la logique de la recherche d'une chaîne de caractères se combine avec une analyse préexistante des structures de la langue. Enfin, l'affichage du texte grec, qui est la vocation du *TLG*, s'accompagne parfois d'un lien vers une traduction anglaise disponible sur un autre site.

35 De telles innovations sont précieuses, mais elles reflètent également les tendances générales d'un champ de recherche. Les avancées d'autres projets, qui ne portent pas nécessairement sur l'Antiquité, mettent en évidence des lacunes et font apercevoir de nouvelles possibilités. Mentionnons quelques perspectives proches ou lointaines, mais nullement utopiques – encore que l'utopie ait ici droit de cité. Au-delà du progrès des travaux lexicologiques, l'étude de la langue et des œuvres profitera de l'exploration du trésor grec sous ses aspects syntaxiques et textuels, étant entendu que ces paliers d'analyse sont en réalité interdépendants. Plus largement, la

quantification des faits linguistiques ou littéraires paraît être une dimension indispensable de l'interprétation : les hellénistes commencent seulement à disposer d'instruments statistiques, dont l'usage fait l'objet d'une réflexion depuis plus d'un demi-siècle dans certaines disciplines. L'apparat critique des éditions savantes imprimées est le socle de la méthode philologique. Il est certes impensable de ne pas l'intégrer dans notre environnement de travail numérique, mais il est probable que cette transposition implique une réinvention : forme concise de commentaire, l'apparat traditionnel est un dispositif hypertextuel, dont la fonction évolue quand les livres forment un réseau et communiquent entre eux. La mise en relation débute avec la comparaison de textes parallèles, ou avec l'affichage d'un dictionnaire en regard du texte lorsque l'on clique sur un mot (fig. 4). « Hypertexte » et « réseau » sont les mots que nous employons pour désigner l'horizon de cette évolution, qu'il appartient aux praticiens clairvoyants, ou visionnaires, de préciser.

Figure 4. Fig. 4. Le passage d'Athénée cité plus haut et l'entrée *thêsauros* du Liddell-Scott-Jones [LSJ] dans le logiciel Diogenes (juin 2010).

36 On peut alors se demander si le *TLG*, dans son état actuel, n'est pas à la fois une réussite exemplaire et un modèle dépassé. Plusieurs des fonctions citées proviennent des travaux du projet *Perseus*, qui nous fournit un contrepoint particulièrement pertinent, car sa politique scientifique est tout à fait différente. Fondé en 1987, il a pris le parti de rassembler et, autant que possible, de mettre en correspondance sources, traductions, commentaires, grammaires et dictionnaires. En attendant la réinvention de l'apparat critique, c'est-à-dire des instruments d'une interprétation informée, c'est un pas pour s'éloigner du trésor sans mémoire[27]. En outre, cette équipe a pour objectif de produire des textes et des outils plus élaborés, qui sont ensuite librement mis à la disposition de la communauté des chercheurs et de toute personne intéressée[28]. Nous vivons l'époque des « incunables numériques », ne serait-ce que parce que l'analogie avec la page imprimée est encore prégnante, et le *TLG* en est un symbole[29]. Du point de vue des services rendus aux spécialistes et à la société dans son ensemble, le *TLG* et *Perseus* illustrent deux modèles intellectuels et économiques opposés : le premier repose sur un accès restreint aux fruits du travail universitaire et de l'expertise technique, tandis que le second favorise un libre accès aux données.

37 Le *TLG* est cependant encore un corpus irremplaçable par sa taille.

Insistons, pour finir, sur cette question de l'échelle, qui peut déterminer un autre maniement des textes. Dans la conservation et l'exploitation d'un héritage culturel, diverses stratégies et diverses priorités sont envisageables. Elles expriment une tension, qui dépasse de loin le cas des études anciennes, entre une approche intensive et une approche extensive. Bien que la philologie vive de précision et que l'ordinateur ait des capacités de mémoire et de calcul incomparables à celles de l'homme, il serait évidemment déraisonnable de les situer à des pôles opposés. Disons simplement ici que les technologies informatiques sont un nouvel arsenal pour l'étude qualitative des textes[30] et que le traitement des données quantitatives met au jour des faits qui échappent à toute autre appréhension[31]. Notre présentation du *TLG* – et de ce qu'il annonce – a peut-être pour premier effet de réactiver le vertige borgésien des parcours infinis dans la bibliothèque. Personne ne lira jamais la totalité des textes grecs de ce corpus. Cependant, de nouveaux protocoles de recherche enrichiront ceux qui existent déjà. Telle est la réponse inattendue à l'une des questions de notre époque : que faire avec un million de livres[32] ?

Vers une philologie numérique

38 Au-delà d'une nouvelle lexicographie et de l'organisation d'une philologie dite computationnelle[33], nous voyons l'émergence de ce que l'on peut nommer la « philologie numérique[34] ». Le passage au médium numérique – après le rouleau de papyrus et le codex, après l'écriture et l'imprimerie – suscitera d'autres types de lecture[35] et d'autres modes de recherche. Dans la coalescence des pratiques et des technologies, nous vivons sans doute un interrègne. Les textes antiques paraissent être un bon terrain d'expérimentation. Alors même que la philologie lutte pour sa survie[36], c'est une dimension nouvelle de sa pratique et de son épistémologie qui s'ouvre à elle. Aussi de nombreux spécialistes de l'Antiquité prennent-ils une part active à la construction des humanités numériques[37].

39 En 1989, le directeur scientifique des *Lieux de savoir* achevait ainsi sa présentation du *TLG :* « Contrairement à une idée répandue par ses détracteurs, l'informatique invite à penser et à réfléchir. [...] Il est sans doute trop tôt pour faire le bilan des découvertes rendues possibles par le *TLG* [38]. » Vingt ans plus tard, on ne peut que confirmer et étendre cette conclusion. Dans bien des domaines, il est devenu impossible ou intellectuellement ruineux d'ignorer l'informatique. Cependant, à l'heure où les technologies de l'information s'imposent dans la vie privée et dans la vie scientifique, la plus grave menace réside sans doute dans une perspective purement instrumentale. Si la

bibliothèque numérique ne servait qu'à reconduire les automatismes du travail avec les textes imprimés, le *TLG*, ses avatars futurs et ses équivalents dans d'autres champs de savoir seraient comparables aux esclaves savants qu'un riche Romain utilisait pour produire telle ou telle citation grecque au moment opportun[39]. Comme toute base de données qu'une communauté exploite, ou plutôt comme tout corpus qu'elle se donne pour résultat et pour matériau de ses recherches[40], le « thésaurus » grec est une création continuée. Affaire de maniement et d'interprétation, son usage doit stimuler l'esprit d'invention.

40 Dans les sciences de la culture, le développement d'une compétence numérique partagée passera sans doute par un certain degré d'hybridation des spécialités et par l'instauration d'un meilleur dialogue entre les disciplines. La distinction du technique et de l'intellectuel paraît ici moins pertinente que jamais. Trop bien assimilé, devenu invisible, un outil nouveau se réduit à ses propriétés les plus banales et perd sa puissance heuristique. On comprend mieux le paradoxe qui fait déclarer à un *computing humanist*, dans une leçon inaugurale : « Est-ce que je sais de quoi je parle ? Non[41]. » Car personne ne sait encore ce qu'est une édition numérique. Qui peut imaginer ce que les pratiques de lecture et de recherche à venir nous feront percevoir, analyser et comprendre ? Histoire à suivre.

Notes

1. *Scholies à l'Iliade*, IX, 443 (T).
2. Willard McCarty, auteur de la première grande réflexion sur ce qu'il nomme *humanities computing* (McCarty, 2005 ; voir aussi McCarty, 2003), a mis en chantier une étude des relations compliquées qu'ont entretenues l'informatique et la littérature. En ce qui concerne les *Classics*, on peut se référer à un article de Theodore Brunner, qui fut le directeur du *TLG* pendant un quart de siècle : Brunner, 1993.
3. L'orthographe et la présentation originales des sources sont respectées. Les lecteurs qui ne sont pas familiers avec le latin et le grec n'en percevront que plus rapidement les structures, qui sont l'objet de nos remarques.
4. Dendien, Pierrel et Quemada, 2004 *(TLFi)*, version informatisée de Imbs et Quemada, 1971-1994 *(TLF)*.
5. Ce thésaurus conçu comme une hiérarchie de notions est à rapprocher de l'usage actuel du mot « ontologie » en informatique. Dans les deux cas, c'est la langue anglaise qui introduit des innovations sémantiques reflétant les transformations des technologies de la connaissance.
6. Estienne, 1531.
7. Estienne, 1572.

8. Voir Pfeiffer, 1976, et Cazes, 2003.
9. Estienne *et al*, 1831-1865, est le résultat de cette ultime amélioration.
10. Ces définitions sont celles de Chantraine, 1999.
11. Sur l'histoire de la « mise en texte » et de la « capitulation », voir par exemple Martin, 2003, p. 41, et Martin, 1995. Les presses de la famille Estienne ont favorisé ce mouvement.
12. Le mot « trésor » n'a évidemment pas de sens technique ici, mais il nous a fourni un critère pour sélectionner un passage au moyen du *TLG.*
13. Le texte grec est cité d'après Kaibel, 1887-1890, qui reprend la numérotation des pages et leur subdivision en sections (de *a* à *f*) introduites en 1598 dans l'édition publiée par Isaac Casaubon, gendre d'Henri Estienne. La traduction est la nôtre.
14. *« A book is a machine to think with » :* les acteurs des *digital humanities* n'ont pas manqué de souligner la signification nouvelle de la formule d'Ivor Richards (*Principles of Literary Criticism* , 1924), citée par exemple dans Burrows, 2004.
15. « *Referring* », « faire référence », est l'un des gestes fondamentaux de l'activité savante selon John Unsworth, dont la liste souvent citée des « *scholarly primitives* » est essentiellement ordonnée à une réflexion sur les nouveaux outils de la recherche (Unsworth, 2000).
16. L'adresse du *TLG Online* contient le nom latin de l'auteur du *Thesaurus graecae linguae* ou *TGL :* Stephanus.
17. L' *Index thomisticus* de Roberto Busa a été compilé à partir de 1949, sur cartes perforées, en collaboration avec l'International Business Machines Corporation (IBM). Ce travail précurseur n'a cependant porté ses fruits qu'avec la publication des 56 volumes de sa version imprimée entre 1974 et 1980 ; un CD-Rom a été diffusé en 1994 et l' *Index* est en ligne depuis 2005.
18. Depuis l'aperçu de Jacob, 1989, p. 131, qui bénéficiait d'informations recueillies à titre personnel, le *TLG* a commencé à publier sa propre histoire : nous renvoyons à Pantelia, 2000, et aux pages « Project History » et « Timeline » du site, qui sont les meilleures sources d'information à ce jour.
19. Il a produit un index hippocratique et un dictionnaire de l'épopée grecque archaïque, récemment achevé.
20. La relative rapidité de l'élaboration du *TLG* provient notamment de l'abandon de la perspective lexicographique. Citons pour comparaison le cas du *Thesaurus Linguae Latinae*, dictionnaire fondé sur dix millions de fiches : commencée en 1900, sa publication a atteint la lettre P, soit les deux tiers de l'ouvrage. Quant au dictionnaire historique du sanskrit en préparation à Pune, il n'a pas encore exploité toutes les fiches correspondant à la lettre A.
21. *TLG*, 1972-.
22. *TLG Online*, 2001-.
23. Fondé en 1987, le Packard Humanities Institute, ou PHI, a diffusé d'autres corpus électroniques, dont certains sont complémentaires du *TLG :* une collection de textes littéraires latins et un ensemble de

papyrus et d'inscriptions grecs.

24. Sa troisième et dernière publication sous forme de livre est Berkowitz, Squitier et Johnson, 1990. Le *Canon* électronique se développe maintenant au gré des mises à jour du *TLG* en ligne.

25. Pour savoir quand se fier au *TLG*, il faut avoir l'expérience de chaque tradition textuelle. Un guide comme Dickey, 2007, reconnaît avec pragmatisme la valeur de ce corpus et se donne notamment pour fonction d'éviter le recours naïf à des éditions dépassées ou problématiques.

26. Voir par exemple Heslin, 2001.

27. Selon la formule radicale de Jean Bollack, en matière d'interprétation, « c'est l'inventaire ou rien » (Bollack, 1997, p. 27).

28. Sont notamment issus de cette politique la version électronique du LSJ, des outils de lemmatisation et de nombreux fichiers encodés dans le métalangage XML et distribués sous un contrat Creative Commons. Voir *Perseus*, 1987-.

29. Sur cette expression et sur son application au *TLG*, voir Crane, 2004, et Crane, Bamman et Babeu, 2007.

30. Il suffit de penser aux techniques fines d'encodage des textes, qui constituent une pratique intermédiaire entre l'écriture et le commentaire, avec pour finalité une description formelle et sémantique que la machine sache exploiter : voir les recommandations de la Text Encoding Initiative fondée en 1987 (Burnard et Bauman, 2007), ainsi que Burnard, O'Brien O'Keeffe et Unsworth, 2006.

31. Par exemple grâce à la statistique lexicale, à l'analyse des réseaux ou à la représentation visuelle de phénomènes de distribution.

32. *« What do you do with a million books ? »* demande-t-on souvent à la suite du directeur du projet *Perseus*, Gregory Crane (voir Crane, 2006), que ce soit en référence à l'Universal Library et à son projet initial « Million Book Digital Library » ou aux entreprises de numérisation mondiales (voir les analyses de Dacos et Mounier, 2010).

33. Quemada, 2004, et Bozzi, 2004.

34. Voir les propositions de François Rastier dans Rastier, 2001. L'auteur rend justement hommage à Roberto Busa, qui a consacré sa vie de savant à une « philologie électronique » (voir Busa, 1998, article au titre savoureux).

35. Chartier, 1995.

36. Sur cette crise, voir les réflexions d'un indianiste dans Pollock, 2009.

37. Pour un panorama concernant l'étude de l'Antiquité dans les pays anglophones, voir Crane et Terras, 2009, ou Terras et Crane, 2010. Les *digital humanities* ont déjà leurs *companion books*, qui donnent une idée de la diversité et de la vivacité du domaine : Schreibman, Siemens et Unsworth, 2004, et Siemens et Schreibman, 2007. Voir également les références collectées par l'auteur de ces lignes dans Berra, 2010-.

38. Jacob, 1989, p. 135.

39. Sénèque, *Lettres à Lucilius*, III, 27. Cette satire d'un parvenu inculte prêterait à une réflexion sur le rôle de la mémoire auxiliaire dans la

pratique des philosophes et des philologues, jadis et aujourd'hui.

40. Avec justesse et humour, Bruno Latour propose de rebaptiser « obtenues » les « données » scientifiques, qui n'ont en effet rien de naturel (Latour, 2007). La « numérisation généralisée » met particulièrement en relief la dimension matérielle de la vie des savoirs, constate-t-il, ainsi que la convergence entre les sciences de l'interprétation et les sciences de la nature, qui toutes travaillent avec des corpus et, *mutatis mutandis*, avec des textes.

41. McCarty, 2010. Pour cet observateur participant des humanités numériques, l'inconnu vaut promesse : « *I smell food on the wind*. »

Bibliographie

- Berkowitz, Squitier et Johnson, 1990 : Luci Berkowitz, Karl A. Squitier et William A. Johnson, *Thesaurus Linguae Graecae. Canon of Greek Authors and Works* [1977], 3e éd., New York.
- Berra, 2010- : Aurélien Berra, *Philologie à venir*, carnet de recherche en ligne [http://philologia.hypotheses.org].
- Bollack, 1997 : Jean Bollack, *La Grèce de personne. Les mots sous le mythe*, Paris.
- Bozzi, 2004 : Andrea Bozzi, « Verso una filologia computazionale : la prima Euroconferenza della European Science Foundation », *Euphrosyne*, 32, p. 127-138.
- Brunner, 1993 : Theodore F. Brunner, « Classics and the Computer : The History », in Jon Solomon (éd.), *Accessing Antiquity. The Computerization of Classical Databases*, Tucson, p. 10-33.
- Burnard et Bauman, 2007- : Lou Burnard et Syd Bauman (éd.), *TEI P5. Guidelines for Electronic Text Encoding and Interchange*, TEI Consortium [http://www.tei-c.org].
- Burnard, O'Brien O'Keeffe et Unsworth, 2006 : Lou Burnard, Katherine O'Brien O'Keeffe et John M. Unsworth (éd.), *Electronic Textual Editing*, New York [http://www.tei-c.org/About/Archive_new/ETE].
- Burrows, 2004 : John Burrows, « Textual Analysis », in Susan Schreibman, Raymond G. Siemens et John M. Unsworth (éd.), *A Companion to Digital Humanities*, Oxford.
- Busa, 1998 : Roberto Busa, « Concluding a Life's Safari from Punched Cards to the World Wide Web », in Lou Burnard, Marilyn Deegan et Harold Short (éd.), *The Digital Demotic. Selected Papers from DRH86, Digital Resources for the Humanities Conference, St. Anne's College, Oxford, September 1997*, Londres, p. 3-12.
- Cazes, 2003 : Hélène Cazes, « Étude introductive. La passion et les épreuves d'Henri II Estienne, imprimeur humaniste », in Jean Céard (éd.), *La France des humanistes. Henri Estienne, éditeur et écrivain* , Turnhout, p. XI-XLVIII.
- Chantraine, 1999 : Pierre Chantraine, *Dictionnaire étymologique de la langue grecque. Histoire des mots. Avec un Supplément sous la direction*

d'Alain Blanc, Charles de Lamberterie et Jean-Louis Perpillou [1968], Paris.

- Chartier, 1995 : Roger Chartier, « Lecteurs dans la longue durée : du *codex* à l'écran », in Roger Chartier (éd.), *Histoires de la lecture. Un bilan des recherches*, Paris, p. 271-283.
- Crane, 2004 : Gregory Crane, « Classics and the Computer : An End of the History », in Susan Schreibman, Raymond G. Siemens et John M. Unsworth (éd.), *A Companion to Digital Humanities*, Oxford, p. 46-55.
- Crane, 2006 : Gr. Crane, « What Do You Do with a Million Books ? », *D-Lib Magazine. The Magazine of Digital Library Research* , 12, 3 [http://www.dlib.org/dlib/march06/crane/03crane.html].
- Crane, Bamman et Babeu, 2007 : Gr. Crane, David Bamman et Alison Babeu, « ePhilology : When the Books Talk to their Readers », in Raymond G. Siemens et Susan Schreibman (éd.), *A Companion to Digital Literary Studies*, Oxford, p. 29-34.
- Crane et Terras, 2009 : Gr. Crane et Melissa Terras (éd.), « Changing the Center of Gravity : Transforming Classical Studies Through Cyberinfrastructure », *Digital Humanities Quarterly*, 3, 1 [http://www.digitalhumanities.org/dhq/vol/3/1].
- Dacos et Mounier, 2010 : Marin Dacos et Pierre Mounier, *L'Édition électronique*, Paris.
- Dendien, Pierrel et Quemada, 2004 : Jacques Dendien, Jean-Marie Pierrel et Bernard Quemada, *Trésor de la langue française informatisé*, Paris [http://atilf.atilf.fr et http://www. cnrtl.fr/lexicographie].
- Dickey, 2007 : Eleanor Dickey, *Ancient Greek Scholarship. A Guide to Finding, Reading, and Understanding Scholia, Commentaries, Lexica, and Grammatical Treatises, from their Beginnings to the Byzantine Period*, Oxford.
- Estienne, 1572 : Henri Estienne, Θησαυρὸς τῆς ἑλληνικῆς γλώσσης. *Thesaurus graecae linguae*, Genève.
- Estienne, 1531 : Robert Estienne, *Dictionarium, seu latinae linguae thesaurus*, Paris.
- Estienne *et al.*, 1831-1865 : Henri Estienne, Charles Benoît Hase, Wilhelm Dindorf et Ludwig August Dindorf, Θησαυρὸς τῆς ἑλληνικῆς γλώσσης. *Thesaurus graecae linguae*, Paris.
- Heslin, 2001 : Peter J. Heslin, compte rendu du *Thesaurus Linguae Graecae* (disque E), *Bryn Mawr Classical Review*.
- Imbs et Quemada, 1971-1994 : Paul Imbs et Bernard Quemada (éd.), *Trésor de la langue française. Dictionnaire de la langue du xixe et du xxe siècle (1789-1960)*, Paris.
- Jacob, 1989 : Christian Jacob, « Le *Thesaurus Linguae Graecae* sur CD-Rom », *Préfaces. Les idées et les sciences dans la bibliographie de la France* , 14, p. 131-135.
- Kaibel, 1887-1890 : Georg Kaibel, *Athenaei Naucratitae Dipnosophistarum libri XV*, Leipzig.
- Latour, 2007 : Bruno Latour, « Pensée retenue, pensée distribuée », in Christian Jacob (éd.), *Lieux de savoir. I. Espaces et communautés* , Paris, p. 605-615.

- Martin, 1995 : Henri-Jean Martin, « Lectures et mises en texte », in Roger Chartier (éd.), *Histoires de la lecture. Un bilan des recherches* , Paris, p. 249-259.
- Martin, 2003 : H.-J. Martin, « Du livre à la lecture », in Christian Jacob (éd.), *Des Alexandries II. Les métamorphoses du lecteur*, Paris, p. 35-45.
- McCarty, 2003 : Willard McCarty, « Humanities Computing », in Miriam Drake (éd.), *Encyclopedia of Library and Information Science* , 2e éd., New York, p. 1224-1235 [http://staff.cch.kcl.ac.uk/ ~ wmccarty/cv.html].
- McCarty, 2005 : W. McCarty, *Humanities Computing*, Basingstoke.
- McCarty, 2010 : W. McCarty, « Attending From and To the Machine », leçon inaugurale, Center for Computing in the Humanities, King's College, Londres [http://staff.cch.kcl.ac.uk/ ~ wmccarty].
- Pantelia, 2000 : Maria C. Pantelia, « " *Noûs*, into Chaos". The Creation of the Thesaurus of the Greek Language », *International Journal of Lexicography*, 13, 1, p. 1-11.
- *Perseus*, 1987- : *Perseus*, Tufts University [http://www.perseus.tufts.edu/hopper].
- Pfeiffer, 1976 : Rudolf Pfeiffer, *History of Classical Scholarship. From 1300 to 1850*, Oxford.
- Pollock, 2009 : Sheldon Pollock, « Future Philology ? The Fate of a Soft Science in a Hard World », *Critical Inquiry*, 35, 4, p. 931-961 [http://www.columbia.edu/cu/mealac/faculty/pollock].
- Quemada, 2004 : Bernard Quemada, « Nouvelle lexicographie ? », *Euphrosyne*, 32, p. 45-54.
- Rastier, 2001 : François Rastier, « Philologie numérique », in *Arts et Sciences du texte*, Paris, p. 73-97.
- Schreibman, Siemens et Unsworth, 2004 : Susan Schreibman, Raymond, George Siemens et John M Unsworth (éd.), *A Companion to Digital Humanities*, Oxford [http://www.digitalhumanities.org/companionDLS].
- Siemens et Schreibman, 2007 : Raymond George Siemens et Susan Schreibman (éd.), *A Companion to Digital Literary Studies* , Oxford.
- Terras et Crane, 2010 : Melissa Terras et Gregory Crane (éd.), *Changing the Center of Gravity. Transforming Classical Studies Through Cyberinfrastructure*, Piscataway.
- *TLG*, 1972- : *Thesaurus Linguae Graecae* , université de Californie à Irvine.
- *TLG Online*, 2001- : *Thesaurus Linguae Graecae. A Digital Library of Greek Literature*, université de Californie à Irvine [http://stephanus.tlg.uci.edu].
- Unsworth, 2000 : John M. Unsworth, « Scholarly Primitives : What Methods do Humanities Researchers Have in Common, and How Might Our Tools Reflect This ? », colloque *Humanities Computing : Formal Methods, Experimental Practice* , King's College, Londres [http://www3.isrl.illinois.edu/ ~ unsworth/Kings.5-

00/primitives.html].

Nos partenaires

Le projet *Savoirs* est soutenu par plusieurs institutions qui lui apportent des financements, des expertises techniques et des compétences professionnelles dans les domaines de l'édition, du développement informatique, de la bibliothéconomie et des sciences de la documentation. Ces partenaires contribuent à la réflexion stratégique sur l'évolution du projet et à sa construction. Merci à eux !

CONCEPTION : ÉQUIPE *SAVOIRS*, PÔLE NUMÉRIQUE RECHERCHE ET PLATEFORME GÉOMATIQUE (EHESS). DÉVELOPPEMENT : DAMIEN RISTERUCCI. DESIGN : WAHID MENDIL.