

Henri Béhar

Le golem et moi

Au plus loin que je puisse remonter dans ma mémoire, c'est l'histoire du Golem que je retrouve. Pas banal, me direz-vous, et vous aurez raison. Pourquoi n'ai-je pas été impressionné par la légende de saint Nicolas sauvant les petits enfants du saloir ? Pourquoi, au fin fond de mes souvenirs, ne trouvai-je point, comme tout le monde, la cruauté de la Barbe bleue ? Si mes frères et sœurs, plus âgés que moi, se plaisaient à m'effrayer avec les contes de *Ma Mère l'Oye*, je les prenais pour tels : des histoires qu'on raconte à la veillée, pour passer le temps. Mais le Golem, c'était une tout autre affaire !

Paris libéré, nous, les enfants juifs du XI^e arrondissement, nous pouvions enfin traîner partout, sans risque de nous faire arrêter. Les cours d'histoire sainte reprenaient au Talmud Torah, le jeudi évidemment. L'un des enseignants, que ma mémoire persiste à identifier à Élie Wiesel au rebours de toute vraisemblance, s'échappait de la Bible et de la chronologie attendue pour s'assurer l'attention de son auditoire par des histoires à sa façon.

Dans la tradition de la kabbale, l'homme est en mesure de créer, à partir du limon, un être à sa ressemblance, qui a toute l'apparence et les propriétés du vivant. C'est le *golem*, destiné à le servir. Pour l'animer, on lui imprime sur le front, en hébreu, le mot *émet* (vérité). Il ne faut pas craindre la supériorité du mécanique sur l'homme car, dès que le golem tend à lui échapper, son créateur peut le stopper net en supprimant la première lettre du mot gravé sur son front, ce qui donne *met*, la mort !

Les visiteurs de la « vieille nouvelle » synagogue de Prague (bel oxymore) entendent réciter l'histoire de Rabbi Loew, qui sauva ses concitoyens des persécutions habituelles en acceptant, à la demande de l'Empereur, de détruire sa créature. Le Golem étant très grand, bien plus grand que le rabbin, celui-ci lui demanda de lui lacer les chaussures, et il profita de ce que l'automate s'était penché vers lui pour supprimer cette lettre de vie. Il me semble qu'aujourd'hui tout le monde connaît la légende mise en forme par certains auteurs fantastiques, et surtout par les concepteurs de jeux numériques. Mais ce qui m'a toujours frappé, c'est le pouvoir d'une lettre, interrupteur infailible, dont un informaticien s'est servi pour supprimer ou restaurer un fichier numérique. D'où le fait que j'ai nommé golem les différents micro-ordinateurs à ma personne attachés, et mes livres concernant l'usage du numérique dans les lettres : *La Littérature et son Golem* (t. I, Champion, 1996 ; t. II, Garnier, 2010). Après publication, on m'a dit que le premier ordinateur géant conçu en Israël en 1965 fut baptisé *Golem* par le grand kabbaliste Gershom Scholem. Vous comprenez pourquoi ! Une chose est certaine à mes yeux : c'est bien cette légende qui m'a toujours incliné vers l'usage de la mécanique dans le discours littéraire.

Au cours de cette anamnèse, je voudrais dire ici comment j'ai conçu un attachement particulier pour les micro-ordinateurs en qualité d'outils d'analyse littéraire, et pourquoi je ne peux plus lire un texte sans en nourrir mon Golem.

Le dinosaure parle... Aujourd'hui, ceux qui téléchargent un roman sur une tablette, ou bien un recueil de poésie capturée sur *Gallica*, ne peuvent imaginer les étapes par où il a fallu passer pour arriver à ce raffinement de civilisation. Je suis de l'époque de la carte perforée : une opératrice tapait le texte, mot par mot, sur des cartes perforées, lesquelles, après divers traitements, devaient nous en donner une version appauvrie (ni capitales, ni accents, aucun des enrichissements formant le texte imprimé). On doit en trouver trace dans la version numérique du *Surréalisme au Service de la Révolution*, revue saisie de cette manière par des bénévoles, et révisée de manière automatique dans les années 2000.

À l'origine, ceux qui pensaient que l'ordinateur (un gros Gamma 60 de Bull-France) pouvait nous aider à avancer dans l'étude des textes littéraires, se divisaient en deux catégories. Les uns pensaient qu'un texte numérisé devait s'accompagner de caractérisants grammaticaux pour chaque terme ; autrement dit, il devait être préalablement encodé par l'opérateur. Les autres, dont je faisais partie, estimaient qu'il fallait saisir le texte brut, en laissant le soin à chaque utilisateur de spécifier la nature du vocabulaire, selon ses besoins. L'historien des idées, par exemple, a rarement l'occasion de s'intéresser à l'emploi du subjonctif ou à la graphie de *néuphar* dans tel document historique. Cherchant à caractériser le vocabulaire des surréalistes, je m'étais, dès 1970, rapproché du laboratoire de Saint-Cloud, parce qu'il privilégiait le texte brut.

En même temps, il m'apparut que la plus grosse banque de textes, à l'époque, *Frantext*, vaste projet gaullien constitué à Nancy pour servir à l'élaboration du *Trésor de la Langue Française* (TLF), dictionnaire destiné à prolonger le Littré au XX^e siècle avec des exemples pris dans la littérature contemporaine, ne pouvait se limiter à cela. À mes yeux, les littéraires devaient pouvoir tirer parti pour leurs propres études de tous ces textes soigneusement numérisés. De son côté, Etienne Brunet produisait ses analyses des grands ensembles, le vocabulaire de Giraudoux, Proust, Zola, etc.

Malheureusement, je ne pouvais me livrer à de telles analyses avec le matériel de *Frantext*, le recteur Imbs, fondateur de ce laboratoire du CNRS, ayant privilégié ce qu'il considérait comme les grands textes de la littérature moderne et contemporaine, en excluant les surréalistes, par définition. Je m'employai donc à faire saisir, outre le *Surréalisme au Service de la Révolution* déjà nommé, *La Révolution surréaliste*, les Poésies de Benjamin Péret, les Romans de Crevel, les *Œuvres complètes* de Tristan Tzara (que par ailleurs j'étais de façon traditionnelle chez Flammarion), que l'on trouve désormais sur le site de Mélusine, par exemple : http://melusine-surrealisme.fr/site/Peret/Peret_Poesies-completes.htm.

Je me reproche souvent d'avoir entraîné mes étudiants, comme autrefois Étiemble avec son *français*, sur des chemins par trop mécaniques, au gré de l'avancement des techniques. Ils n'y perdirent pas leur temps, cependant, puisqu'ils apprirent à maîtriser la transcription numérique, et surtout se confrontèrent aux œuvres en un « tête à texte » que bien peu de leurs condisciples pratiquaient.

Que chacun, où qu'il se trouve, puisse lire tel poème de Saint-Amand dont il se souvient vaguement, qu'il « relise » la confrontation, entre deux gendarmes, de Jean Valjean et de Mgr Myriel pour en repérer, par-delà le pathos, les traces de l'idéologie hugolienne, c'est déjà magnifique. De même qu'autrefois le lecteur du dictionnaire passait son dimanche en allant d'un mot à l'autre, sans but préconçu, et ainsi se constituait une

culture, en quelque sorte, le lecteur actuel peut s'engager dans des navigations infinies, non seulement de la vedette du dictionnaire à l'autre, mais aussi de tout mot du texte à tout autre (ce que l'on nomme le plain-texte), puis de rechercher ceux qui ont utilisé tel terme unique (*hapax*), sans autre limite que sa patience et sa curiosité.

Quelles que soient les critiques que l'on adresse à Google (que j'ai toujours pris pour un paronyme de Golem), il faut reconnaître que son programme de numérisation en langue française a dépassé tous les espoirs, toutes les capacités de nos systèmes institutionnels ou privés.

Ainsi, ayant découvert, par hasard et un peu par méthode, l'un des plagiats de Lautréamont, j'écrivais en 1990, qu'il serait possible, à l'avenir, de confronter l'œuvre d'Isidore Ducasse à la presse de son temps pour épuiser l'investigation et retrouver tous les textes contemporains qu'il avait manipulés pour les mettre dans ses *Chants de Maldoror* : « Pourtant, il n'est pas interdit de rêver au jour où la future Bibliothèque de France, ayant numérisé un grand nombre de textes contemporains, il sera possible de procéder, automatiquement, aux recoupements souhaités, avec le secours d'un programme approprié¹. » Cette utopie me fut aussitôt reprochée, pour son caractère non-scientifique. Aujourd'hui, c'est devenu banalité (mais je dois reconnaître, en le regrettant, que personne n'ait pris mes propos au pied de la lettre et procédé à la confrontation suggérée).

Les Chants de Maldoror ont été numérisés par mes soins. Cette nouvelle édition se bornerait-elle à fournir un index total de l'œuvre que ce serait déjà un progrès considérable. En outre, elle donne tout le texte, ce qui a conduit certains étudiants à fournir un dictionnaire (en ligne) du vocabulaire ducassien, renvoyant automatiquement aux diverses occurrences du mot-vedette et à ses contextes. Se demandant si les calculs de fréquence allégués par Bachelard dans son livre sur Lautréamont étaient exacts, à une époque où les comptages ne pouvaient se faire qu'à la main, telle autre thésarde a procédé à une vérification exhaustive, prouvant sans conteste que le philosophe Bachelard savait compter avant de proposer une interprétation. Plus près de nous, la numérisation des œuvres de Claude Simon étant interdite par son éditeur, deux thésards ont jugé utile d'en fournir un index général, ce que la loi autorise et la pratique exige !



Les œuvres complètes d'Isidore Ducasse et diverses modalités d'interrogation se trouvent désormais sur le site d'Hubert de Phalèse, qui bénéficie de deux adresses (www.cavi.univ-paris3.fr/phales ; et <http://www.phalese.fr/spip/>). Il convient de révéler ici (encore qu'il n'y eut jamais de secret à ce sujet) qui est Hubert de Phalèse. C'est un nom collectif (à l'image de N. Bourbaki pour les mathématiciens), adopté en 1989 par un petit groupe de chercheurs décidés à produire des travaux, notamment des livres (papier) démontrant l'intérêt d'employer les outils informatiques dans l'analyse des œuvres littéraires. Visant au plus haut, ils commencèrent par un ouvrage pour les candidats à l'agrégation, traitant d'*À rebours* de Huysmans, publié chez le plus conservateur des éditeurs, lequel s'emballa pour la nouveauté. *Comptes à rebours* (Nizet, 1990) fournit le patron d'une série d'ouvrages, de manuels, devrais-je dire, exploitant toutes les ressources que procurait alors l'informatique personnelle (PC). Il en découla une sorte de manifeste, un discours de la méthode Hubert de Phalèse².

La préparation au concours d'agrégation exclut l'invention et le génie. Mais elle n'interdit pas un travail méthodique, l'élaboration de fiches thématiques qui, combinées entre elles, couvraient à peu près tous les champs des dissertations ou des leçons possibles. Instrument redoutable et redouté par les jurys ! Il m'est arrivé, par la suite, de rencontrer de jeunes professeurs, très fiers de me révéler qu'ils avaient préparé le concours, Phalèse en main, si l'on peut dire.

De fait, le nouveau Phalèse prit naissance un matin de juillet 1989, alors que les ordinateurs, situés au 5^e étage du centre Censier n'avaient pu supporter la chaleur de la nuit. Face au désastre (la perte de toutes nos données), je réagis en emportant un ordinateur valide et deux portables. Nous étions trois, qui mîmes les outils au frais dans une petite maison bretonne. Après les avoir calmés, rassurés, nous pûmes faire tourner les machines et produire nos premières sorties : index du vocabulaire, des noms propres, concordances lemmatisées, calculs de fréquence, etc.

On a vu que la première tâche de ces littéraires dévoyés était la constitution de ces index, indispensables au lecteur sérieux, dont, hélas, les meilleures éditions sont toujours dépourvues. Nous ajoutâmes l'explication des mots difficiles, à l'aide des définitions disponibles dans les dictionnaires numérisés (il y avait, à l'époque, floraison de CD-Rom chez les éditeurs spécialisés) et surtout en prélevant dans la base de données *Frantext* les attestations complémentaires. Suivaient des études de lexicométrie, partant de la statistique du vocabulaire global, par chapitres, etc., et ce que nous appelâmes des fiches thématiques proposées à partir de ce qui ressortait mathématiquement du texte. La critique littéraire n'ayant pas de nez, j'eus, par la suite, l'occasion de proposer un protocole pour l'étude des odeurs dans les textes, à partir des obsessions du héros. Ce fut l'article « Des essences pour des Esseintes », que je présentai aux USA devant un public fort étonné, tant il lui paraissait impossible d'allier technologie et critique littéraire. Ce protocole peut s'adapter à d'autres champs lexicaux, les couleurs notamment.

C'est ainsi que j'ai constitué un relevé personnel du vocabulaire relatif à la nourriture dans l'œuvre d'Albert Cohen, à partir duquel j'ai pu écrire mon essai de littérature par l'estomac, *À table avec Albert Cohen* (éd. Non Lieu, 2015), illustré de 40 recettes traditionnelles. Curieux comme les anorexiques sont les premiers à parler de ce qu'ils ont détesté à un moment de leur existence !

Ici, je veux souligner un point en particulier : que les mots soient présentés dans l'ordre alphabétique ou dans l'ordre des fréquences, la lecture tabulaire, outre qu'elle se caractérise par une notable économie de moyens qui aurait dû intéresser les écolos de tous poils depuis longtemps (en moyenne 1/8^e du texte), permet un repérage immédiat de certains éléments, y compris ce que nous avons nommé les *nullax*, ces mots absents d'une isotopie donnée.

À côté de la critique impressionniste, qui a de beaux jours devant elle, j'ai toujours rêvé d'une critique littéraire irréfutable, à l'image d'un théorème mathématique ou d'une loi physique. Ainsi les termes beaucoup, peu, presque, me sont insupportables. Je veux savoir combien exactement, et même plus, la quantité relative et la quantité absolue. Avec un simple traitement de texte, on peut savoir en deux clics si tel mot fait partie du vocabulaire de l'auteur. Ainsi, commentant le *Voyage au bout de la nuit*, un critique émérite (ils le sont tous !) nous informe que Céline s'évertue à ne pas employer le mot

« victoire » (ce qui, dans ce cas, formerait un beau *nullax*). Bien, mais je ne suis pas obligé de le croire sans vérifier. Or le texte présente 5 occurrences de ce terme, dont la *Place des Victoires*. Ce n'est pas la victoire, dira mon contradicteur. Exact, mais c'est encore mieux ! Car Céline avait tout loisir de faire passer son personnage à quelques mètres près, Place de la Banque, par exemple, ou bien place des Petits-Pères, chère au *Canard enchaîné*.

Piètre victoire, me direz-vous, si vous n'avez trouvé que cela ! Soit, mais ce n'est qu'un exemple, le point de départ d'une critique refusant toute approximation. Lorsque chaque commentateur, chaque lecteur, aura pris le réflexe de vérifier ce qu'il écrit par un retour systématique au texte, à la phrase, au mot dans son contexte, alors nous pourrons parler de critique littéraire.

Grâce à la quantité d'informations procurées par le réseau, dans tous les domaines, je peux désormais reconstituer les niveaux de culture propres à chaque auteur, à chaque texte. De sorte que je ne me permets plus de dire que telle œuvre, tel texte est « difficile ». Je crois en avoir fait la démonstration avec l'œuvre de Claude Simon, dont j'ai pu montrer qu'elle contenait peu de mots rares ou difficiles, et que la culture de notre prix Nobel de littérature ne dépassait pas le niveau de connaissance autrefois exigé d'un élève du secondaire (« La culture de l'auteur et la compétence du lecteur dans *La Route des Flandres* », *La Littérature et son golem*, t. II).

On sait la fortune que l'informatique a connue dans le domaine bibliographique, à telle enseigne que les bibliographies imprimées chaque année ont totalement disparu de l'édition française. Bien plus, chacun peut constituer sa propre base de données, tant pour les livres lus, par exemple, que pour ce qui relèverait de l'histoire des œuvres, des lectures, des auteurs, des genres, des thèmes, etc. C'est ce que j'ai cherché à faire, dès 1985, sur une vaste échelle, pour ce qui constitue le domaine des lettres. C'est la *Banque de données d'histoire littéraire* (BDHL, <http://www.phalèse.fr/bdhl/bdhl.php>) qui s'abstient de tout jugement de valeur, et dont les principes et les choix méthodologiques ont été repris par Wikipedia, entre autres. Mais c'est là une aventure, en partie collective, dont la relation dépasserait largement les cadres de la carte blanche qui m'a été si généreusement accordée, de sorte que je peux ranger mon golem, lui ôter son initiale jusqu'à la prochaine *Secousse*.

¹ Article paru dans les *Cahiers Lauréamont* (n°15-16, 2^e sem. 1990, p. 51-55) ; partiellement repris dans l'édition des *Chants de Maldoror* procurée par Jean-Pierre Goldenstein (*Presses Pocket*, 1992, coll. Lire et voir les classiques, p. 400-405) ; enfin paru intégralement dans Henri Béhar, *La Littérature et son golem* (Champion, 1996, p. 139-142).

² Henri Béhar, « La méthode Hubert de Phalèse » in *La littérature et son golem* (Champion, 1990).

Henri Béhar est né en 1940 à Paris. Professeur, historien de la littérature française, spécialiste de l'avant-garde, il a écrit des ouvrages de référence sur Alfred Jarry, Dada et le Surréalisme. Il dirige la revue *Mélusine* (Cahiers du Centre de recherche sur le surréalisme). Fondateur de l'équipe de recherche Hubert de Phalèse, il est l'un des pionniers de l'utilisation de l'informatique dans les études littéraires. Il est depuis 2003 professeur émérite de l'Université Paris III, qu'il a dirigée de 1982 à 1986. Parmi ses nombreux ouvrages : *La Littérature et son golem* (t. I, Champion, 1996 ; t. II, Garnier, 2010).