



**HAL**  
open science

# La modélisation des données de l'édition numérique “ Lettres et textes. Le Berlin intellectuel des années 1800 ” et son approche de la variance dans les lettres

Anne Baillot

## ► To cite this version:

Anne Baillot. La modélisation des données de l'édition numérique “ Lettres et textes. Le Berlin intellectuel des années 1800 ” et son approche de la variance dans les lettres. La variance en philologie et dans la critique génétique, Nov 2013, Paris, France. halshs-01132693

**HAL Id: halshs-01132693**

**<https://shs.hal.science/halshs-01132693>**

Submitted on 17 Mar 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



## La modélisation des données de l'édition numérique "Lettres et textes. Le Berlin intellectuel des années 1800" et son approche de la variance dans les lettres

Anne Baillot, novembre 2013

La question que j'aimerais aborder dans le cadre de cette conférence est celle du statut de la variance dans les textes de nature épistolaire. Pourquoi s'intéresser aux variantes dans les lettres ? Comment juger de la pertinence et de la signification de ces variantes pour l'histoire littéraire et comment en rendre compte en tant qu'éditeur ? Avant d'en venir à ces questions en elles-mêmes, je commencerai par présenter dans les grandes lignes les questions de recherche et la structure de l'édition numérique à partir de laquelle j'argumenterai mon propos : *Lettres et textes. Le Berlin intellectuel autour de 1800*<sup>1</sup>.

### I. Présentation de l'édition

#### 1) Multiplicité des genres et des auteurs

L'édition numérique *Lettres et textes. Le Berlin intellectuel autour de 1800* n'est construite ni autour d'un auteur singulier, ni autour d'un type de source unique. C'est une première originalité, qui s'explique par la question de recherche qui a présidé à son élaboration et à laquelle elle doit fournir des éléments de réponse. Cette édition fait partie d'un projet plus large dans lequel il s'agit de reconstruire les réseaux intellectuels gravitant autour du Berlin intellectuel sur une période qui s'étend des années 1760, donc du milieu du règne de Frédéric II, aux prémices de la révolution de 1848, en passant par les guerres napoléoniennes, la fondation de l'Université de Berlin et les réformes de l'État prussien autour de 1810, la restauration, et les soulèvements de 1830.

Sont ainsi édités différents types de textes qui permettent de reconstruire autant que possible les facettes significatives des processus de genèse et de réception liés à certaines œuvres ou à des moments intellectuels marquants. Nous pouvons ainsi retracer la réception de Rousseau ou de ses œuvres, qui sont diffusées, lues, reprises, critiquées, traduites, évoquées sous différentes formes à la période qui nous intéresse. Mais c'est également une caractéristique de ce projet de recherche que de ne pas se limiter à un corpus d'œuvres et d'auteurs *majeurs*.

Son principe est de faire défiler l'ensemble des étapes de constitution du texte et de sa réception. Cela commence par exemple à l'évocation, dans une lettre, du travail à un brouillon, passera par le manuscrit de l'ouvrage, et ira jusqu'au compte-rendu critique qui en sera fait. Cette reconstruction passe aussi parfois par des documents encore plus secs comme les catalogues de bibliothèque personnelle. Bien évidemment, il n'est pas possible de séparer strictement les lignes de vie des différentes œuvres - ce qui est genèse d'un ouvrage est aussi souvent réception d'un autre. La modélisation des données joue donc dans cette édition un rôle essentiel dans la mesure où c'est elle qui permet de faire le lien entre ces différents « nœuds » propres aux réseaux qui sont au cœur du propos.

<sup>1</sup> <http://tei.ibi.hu-berlin.de/berliner-intellektuelle/> (17.03.2015)



Deux motifs principaux président au choix de mêler les corpus d'auteurs différents. Le premier est pragmatique : la conservation des documents au cours de deux derniers siècles fait que, sur la période et dans l'espace linguistique qui nous intéressent, il n'y a que pour un Goethe qu'il soit envisageable de recourir à la quasi-totalité des documents de genèse et de réception. La deuxième raison est méthodologique : le cas, singulier s'il en est, de Goethe, ne nous en apprendrait pas tant que cela sur la situation de la masse des acteurs du monde littéraire et savant. Bien que chaque auteur obéisse à des impératifs (conscients et inconscients) singuliers, le fait de croiser les corpus d'auteurs différents permet de mettre en évidence des stratégies de publication propres à un ensemble d'auteurs. Il est ainsi possible de dédramatiser et désingulariser un certain nombre d'habitus d'écriture que l'on est, sinon, aisément tenté de tenir pour spécifiques d'un auteur précis.

De plus, ce n'est pas Weimar qui nous intéresse ici, mais Berlin. Dans le contexte du Berlin romantique qui fait le cœur de notre corpus (autour de 1810), la pratique de l'écriture fonctionne en partie comme une activité collective. Un relatif flou entoure encore la notion d'auteur. La reconstitution de réseaux soit s'agençant au gré de querelles scientifiques, littéraires, voire même politiques, soit s'organisant autour de sociétés littéraires ou savantes singulières, permet dans ce contexte de retracer des lignes de force à travers les écrits d'auteurs différents. Cette manière de procéder permet aussi en partie de reprendre à nouveaux frais les grands moments de l'histoire littéraire des deux derniers siècles et de faire se croiser "à plat" Minores et Majores, reprenant et interrogeant ainsi les hiérarchies établies.

Les lettres jouent un rôle crucial d'une part dans la mesure où elles permettent de reconstituer des points de jonction entre les corpus littéraires et d'autre part dans la mesure où elles sont elles-mêmes, en ce début du XIX<sup>ème</sup> siècle, le lieu de stratégies d'écriture conçues en lien étroit avec des stratégies de publication et d'auto-promotion. J'y reviens après avoir présenté une deuxième singularité de notre édition.

2

## 2) Multiplicité des modalités de visualisation

Les éditions numériques de lettres portant sur la période "autour de 1800" sont en plein essor en Allemagne, portées par des institutions et des intentions différentes. Si toutes ces éditions s'affranchissent, dans leur composante numérique, de l'exigence historico-critique, cela ne permet pas encore de faire émerger un standard commun ou unique. Au niveau de l'encodage, il y a bien une unité technique (XML/TEI-P5 ; indexage en référence aux fichiers d'autorité ; travail sur la base de licences CC-BY simple ou -SA), qui permettra de mettre en place des passerelles entre ces éditions. Les financements de ces éditions sont plus ou moins conséquents, mais ils sont tous limités, souvent même très limités, dans le temps, contrairement à ce qui a longtemps été la pratique des grandes éditions. Même une fois limité dans ses options par les contraintes matérielles, il reste encore une marge de manœuvre considérable dans la conception d'une édition numérique - et cela vaut également pour les correspondances. Je résume les grandes tendances particulièrement fortes en Allemagne : S'agit-il de présenter une transcription diplomatique ou une version de lecture? Éditer, est-ce présenter le résultat du travail sur le manuscrit, ou présenter aussi le manuscrit? Faut-il s'appliquer à rendre signe par signe, ou lettre par lettre, ce qui figure sur le manuscrit? Les



limitations proviennent en bonne partie, dans ce domaine, de ce qui est matériellement ou techniquement réalisable avec les browsers actuels.

Les choix que nous avons faits pour notre part sont les suivants. Tout d'abord, nous avons renoncé à représenter, dans la transcription, tous les éléments à l'endroit exact où ils se trouvent dans le manuscrit. Nous offrons un scan de très bonne qualité en regard de la transcription, qui a notamment pour but de permettre au lecteur de s'orienter dans la transcription. Fournir un scan signifie aussi que nos éventuelles erreurs de transcription sont visibles et donc corrigibles (merci de nous envoyer un e-mail si vous repérez des erreurs, que nous nous ferons un plaisir de corriger si nous trouvons votre proposition de correction plausible). Quelqu'un qui ne s'intéresse qu'au résultat (et donc à la version de lecture) a la possibilité de réduire l'écran à une seule colonne ou d'afficher une version pdf qui reprend pour l'essentiel la version de lecture, mais fait figurer en bas de page les éventuelles corrections visibles dans la transcription diplomatique.

Les caractéristiques principales de notre interface sont les suivantes. Nous ne reproduisons pas signe par signe, mais caractère par caractère. Nous ne normalisons pas l'orthographe, même dans la version de lecture. Nous y ajoutons en revanche des hypothèses de lecture entre crochets droits dans le cas par exemple où le manuscrit est déchiré mais que la logique syntaxique et sémantique de la phrase permet de reconstruire avec une certitude assez élevée la partie manquante ; le degré de certitude est précisé. Nous présentons 5 versions affichables sur une ou deux colonnes (facsimilé, transcription diplomatique, version de lecture, métadonnées, entités, XML), toutes générées à partir du même fichier XML/TEI-P5. La navigation permet d'évoluer facilement au sein d'un corpus, les corpus étant composables selon 5 critères (auteurs, genres, thèmes, provenance, dates). Les requêtes, soit plein texte soit partant des catégories encodées (entités nommées : personnes, œuvres, groupes, lieux, ainsi que dates), portent sur l'ensemble du corpus ainsi que sur la base de données des manuscrits d'August Boeckh qui est adossée à l'édition numérique. Enfin, les différentes mains intervenant sur le manuscrit sont signalisées. La main de l'auteur est en noir, toutes les autres mains sont en rouge, le orange étant réservé aux éléments cliquables. Lorsqu'on déplace le curseur sur les caractères qui sont en rouge, une fenêtre apparaît qui explicite l'identité de la main en question (destinataires, éditeur, archiviste, etc.).

Il me semble que l'une des choses les plus importantes, mais aussi les plus difficiles, pour toute édition, est de dire d'elle-même où elle se situe, de se définir d'un point de vue épistémologique. Cela suppose tout d'abord une formulation claire de la question de recherche qui la sous-tend. L'édition numérique *Lettres et textes. Le Berlin intellectuel autour de 1800* s'articule autour de deux axes :

- d'une part, il s'agit de mettre en relation des personnes et des œuvres pour éclairer à la fois en finesse et en largeur les réseaux intellectuels à l'œuvre dans la capitale prussienne à un moment clef de son histoire. Un travail de visualisation de ces données vient d'être engagé avec un spécialiste ;
- d'autre part, elle s'attache aux stratégies de publication développées par les auteurs en valorisant la présence de corrections, ainsi que de formes de constitution d'autorité textuelle par la coprésence de plusieurs mains sur un même manuscrit destiné à publication. Cet aspect devra être facilité par l'implémentation d'un outil de collation courant 2014.



Situer son édition suppose aussi d'être capable de se reconnaître (ou non) d'une tradition scientifique. Sur ce point, il reste encore beaucoup d'interrogations. Les textes que nous éditons sont certes pour l'essentiel bien allemands, mais pour une partie d'entre eux, étroitement liés à la France. Certains corpus sont partiellement ou totalement en français quoique rédigés dans un contexte prussien.

Dans ses prémisses aussi, cette édition est à cheval, à la fois héritière d'une certaine aspiration de l'édition génétique française et chargée de contextualisations historiques et littéraires très allemandes, que l'on trouve dans des commentaires situés en bas de page (dont la mise en valeur pose problème dans un contexte scientifique où tout ce qui n'est pas historico-critique est toujours considéré par de nombreux chercheurs, et davantage encore par de nombreux éditeurs, comme insatisfaisant). Mais elle réinterroge aussi les questions de constitution du texte dans une perspective textologique proche des travaux anglo-saxons.

Cette édition ne réinvente certainement rien qui n'ait déjà été pensé, mais sa singularité est de le faire en se nourrissant d'inspirations diverses qu'il faut savoir faire répondre à une question tout en tenant dans un budget et un temps de projet limités. C'est notamment pour ces raisons que nous avons renoncé à utiliser le module "apparat critique" de la TEI, très complexe, et préféré développer un schéma génétique allégé propre.

## II. La variance dans les textes épistolaires

4

Les textes épistolaires sont difficiles à classer du fait de l'ambiguïté de leur statut documentaire. S'agit-il de textes privés ou bien sont-ils destinés à être rendus publics, et de ce fait à considérer comme des ouvrages de facture littéraire ? Livrent-ils des informations considérables dans leur objectivité ou ne sont-ils que mise en scène d'une personnalité ou d'un dialogue entre deux personnes (ou plus) ? S'agit-il, en somme, d'une fiction, et si oui, laquelle ? Le statut à accorder à la variance dans les textes de facture épistolaire dépend de la réponse à ces questions. Or celle-ci est loin d'être universelle ; elle varie selon les contextes historiques, personnels, ou de publication.

Qu'en est-il en Allemagne autour de 1800 ? Il s'y trouvera sans doute un certain nombre de contre-exemples, mais on peut dire que, de manière assez générale, il était d'usage d'éditer la correspondance de toute personnalité un tant soit peu intellectuellement éminente. Les trois cas de figure éditoriaux les plus fréquents sont les suivants :

- soit la personne en question édite sa propre correspondance (souvent sur le tard ou avec un point fort thématique du type "Goethe et moi") ;
- soit elle en charge quelqu'un qui s'attelle à la tâche après sa mort ;
- soit encore quelqu'un, le plus souvent un proche ou un ami, s'en charge spontanément, le plus souvent, là aussi, à titre posthume.

Très tôt dans la vie d'un journaliste, d'un bibliothécaire, d'un professeur d'université, d'un écrivain, d'un éditeur, émergeait l'idée que cette carrière en train de se constituer allait être, dans un avenir plus ou moins proche, documentée dans ce qu'elle avait de significatif par le moyen d'une édition de



correspondance à caractère fortement biographique, basée plus ou moins fidèlement sur le texte original des lettres échangées. On peut donc admettre comme prémisse de départ que ce groupe de personnes rédige sa correspondance en sachant qu'elles contribuent à écrire une histoire (*a personal narrative*) destinée à publication.

La principale restriction à cette assertion concerne les femmes. Même pour celles d'entre elles qui faisaient une carrière littéraire (souvent dans l'ombre d'un homme ou sous couvert d'un pseudonyme), la conscience de se construire comme objet littéraire semi-public par l'intermédiaire de la correspondance est structurellement beaucoup moins marquée que chez leurs pendants masculins.

Les conséquences de cette conscience d'écrire sa propre histoire pour la postérité par l'intermédiaire de ses lettres sont multiples. Elles sont d'abord cruciales au regard des stratégies de conservation adoptées par les personnes elles-mêmes et par leurs proches. La seule manière de s'assurer qu'une information ne passe pas à la postérité, c'est de la brûler. Dans les cas extrêmes, la censure s'étend jusqu'à l'autocensure, dans une mesure qu'il est impossible de reconstruire dans toute son ampleur : pour être vraiment sûr qu'une information ne passe pas à la postérité, il ne faut tout simplement pas la formuler, ou du moins pas la formuler de manière transparente.

Par ailleurs cependant, on ne part pas, à l'époque, du principe que c'est l'intégralité du texte de correspondance qui sera édité. Il va de soi qu'un travail d'élagage et de réécriture sera effectué. C'est précisément pour s'assurer que ce travail appelé "rédaction" (*redigieren*) sera bien fait dans le sens qu'ils souhaitent que certains hommes de lettres désignent explicitement de leur vivant un chargé de publication. Chamisso écrit à Hitzig : « Garde mes lettres comme je garde les tiennes », signalant bien la mise en scène de la relation épistolaire comme moment de conservation réciproque. Parfois sont laissées des consignes explicites sur la manière dont il s'agira d'éditer ces lettres.

Enfin, cette situation éditoriale génère chez les éditeurs de ces correspondances un sentiment de co-autorité mis en avant dans la publication elle-même. Les couvertures des ouvrages comportent ainsi un agencement, voire une ribambelle de noms dont la hiérarchie est variablement transparente, mais, le plus souvent du moins, assez parlante sur les conditions de genèse et la reconnaissance de l'autorité dans le volume concerné.

Tous ces phénomènes sont lisibles à même les manuscrits. Pour réaliser leurs copies pour impression, les éditeurs n'avaient aucun scrupule à modifier les originaux, rayer des passages, les réécrire, souvent à l'encre rouge ou au crayon sepia. La liste des types de modifications auxquelles ont recours les éditeurs est longue : transformation des signes de ponctuation (il arrive à Hitzig de changer un point d'interrogation en point d'exclamation), ajout ou suppression de signes de ponctuation (Tieck les pratique dans ses propres lettres et en ajoute dans la version imprimée des lettres des autres qu'il édite), remplacement du nom d'une personne par un tiret ou par la réduction à une seule lettre l'anonymisant plus ou moins ("S" pour Schleiermacher, "F" pour Fichte dans le contexte de l'Université de Berlin en 1812 restent pour l'essentiel transparents), correction de noms ou de dates (soit parce que l'éditeur croit y voir une erreur, soit parce qu'il fait lui-même preuve d'inattention), suppression de parties de textes (souvent pour des raisons morales), réécriture de demi-phrases ou de phrases entières pour des raisons esthétiques, grammaticales ou morales, ajout



d'explications de contexte, suppression d'en-tête ("Mon très cher ami"). Par ailleurs, des lettres entières sont délibérément laissées de côté, tandis que pour certaines, on ne sait pas si l'éditeur avait ou non le manuscrit sous les yeux (sans compter tous les cas de figure où il avait un manuscrit que nous n'avons plus).

Cette censure pour la publication, qui nous apparaît aujourd'hui comme une sorte de violation d'une entité quasi sacrée, était alors pratique courante. Dans le cas de lettres, comme les éditeurs étaient eux-mêmes le plus souvent correspondants de la personne qu'ils éditaient, on peut en un sens voir dans ces corrections portées à même l'original une sorte de prolongation du dialogue épistolaire interrompu par la mort, et non une dégradation/dévaluation de l'archive papier originale. Une fois faites les modifications destinées à l'édition, le manuscrit original était en effet pieusement conservé, souvent par la personne même qui avait préparé l'édition. Sur la base du manuscrit ainsi annoté était réalisée une copie, faite cette fois par une petite main, et remise ensuite à la maison d'édition. Ces copies n'échappaient quant à elles que bien rarement à la destruction systématique qui prévalait chez les éditeurs, encombrés de papiers. Il n'est donc pas possible de restituer précisément, uniquement de spéculer, sur les modifications à l'impression imputables aux erreurs de copie de ces petites mains. Il ne fait pas de doute qu'un certain nombre de modifications supplémentaires s'ajoutaient à cette étape du processus, en particulier en ce qui concerne l'ordre dans lequel les feuillets avaient à être publiés.

Joue également un rôle dans la constitution du texte le rôle joué par la factualité historique. Certains des aspects dont traite une lettre renseignent de manière documentaire, recoupable avec d'autres sources historiques (Tieck écrivant "Je viens d'arriver à Karlsbad pour ma cure d'été et ai croisé Goethe"). D'autres en revanche ne renseignent qu'en finesse sur les relations avec l'une ou l'autre personne. Un même événement peut être rapporté par le même scripteur de manière différente (voire antithétique), selon la personne à laquelle il s'adresse, parfois même le même jour puisqu'il n'était pas rare que l'on réserve une journée ou une demi-journée à sa correspondance en souffrance. Est-ce à dire qu'une information est vraie et l'autre fausse ? Pas nécessairement. Mais il est certain qu'elles ne disent pas la même chose.

Il est ainsi fréquemment difficile de définir la frontière entre le factuel et le fictionnel, et les documents susceptibles de lever bon nombre de ces incertitudes sont rarement suffisamment abondants pour pouvoir faire des recoupements sur une quantité représentative de données. On ne peut se passer d'extrapoler et d'interpréter.

Or, qu'avons-nous à notre disposition qui puisse nous aider dans ce travail d'interprétation, qui puisse nous permettre de dégager des aspérités du texte au moment de sa constitution ? Il arrive que d'autres ego-documents éclairent les lettres (correspondances d'auteurs connexes, journal de bord des partenaires du dialogue épistolaire) : ce serait la première source possible. On peut aussi se référer, souvent avec grand profit, aux paratextes des éditions de correspondance, qui énoncent de manière plus ou moins transparente les principes de sélection suivis au cours de la constitution du texte à publier : deuxième source. Mais parfois (souvent), il n'y a rien de plus que la matérialité du texte sur laquelle s'appuyer, les ratures, les corrections, les commentaires. Il faut bien sûr être prudent et ne pas céder à la tentation de la sur-détermination de spécificités en réalité locales ou pas forcément significatives. Cela dit, les différents moments dans la constitution du texte correspondent au développement de variantes associées à des publics différents. Dans sa fonction



primaire de communication, la lettre doit générer une connivence avec un ou plusieurs destinataires. Dans sa fonction d'archivage biographique, elle doit perpétuer une image de la personne dont il est question à destination d'un public qui n'est plus intime. Faut-il considérer qu'il s'agit du même texte ou d'un texte différent, entre celui de la lettre originale et celui de la lettre publiée ? Si l'on considère que les éditeurs ont en charge une partie de la fonction auctoriale, il ne s'agit en effet que de variantes d'un même texte. Preuve en est que, partant d'un manuscrit original O, un éditeur d'une édition  $N=O+1$  pourra proposer un texte qui ne sera pas repris par l'éditeur d'une édition  $N=O+2$ , qui pour sa part préférera revenir à O.

Cela soulève un certain nombre de questions concernant la conception d'une édition numérique de ces corpus, aujourd'hui. Accepter que le texte épistolaire se construit autour d'un appareil de variantes développé à n mains suppose de l'éditer également comme tel. La question de la relation effort/effet se pose alors immédiatement. Même si, en théorie, l'édition numérique permet mieux que toute autre forme d'édition de mettre en relation des moments de constitution du texte et en regard les unes des autres les variantes et la façon dont elles se rapportent les unes aux autres, cela reste un défi technique d'en proposer une édition clairement lisible.

L'exemple des mains le montre. Qu'on distingue sur un manuscrit la main de l'auteur, celle de l'éditeur, et celle de l'archiviste (le plus souvent un numéro dans un coin de page), c'est certes spéculation au sens où rien n'est officiellement attesté, mais c'est dans la logique de la pratique connue de l'archivage et de l'édition. On pourrait alors aisément attribuer une couleur à chacune des trois mains.

Mais comment faire si c'est l'auteur lui-même qui reprend ses lettres plusieurs années plus tard ? Et si l'on marque la différence dans ce cas, que faire des lettres écrites en plusieurs fois ? Elles ne le sont qu'à quelques jours d'intervalle et l'on peut sans doute arguer qu'en termes de fonctions, la position de l'auteur ne change pas. Cela dit, lorsqu'il reprend sa lettre pour la poursuivre, il la relit et la corrige. Comment pouvons-nous être sûrs que cette correction n'est pas faite avec l'idée d'une publication ultérieure en tête, comme le suggèrent les nombreuses améliorations des périodes rhétoriques repérables ? Et comment pouvons-nous, du même coup, être sûrs qu'un éditeur a procédé en une seule fois à la rédaction du texte, qu'il ne s'y est pas repris à des moments différents dans le temps, changeant là aussi éventuellement de rôle, reprenant les lettres pour faire ressortir davantage un aspect ou davantage un autre ? Enfin, comment montrer tout ceci en ligne ? La multiplication des couleurs pour désigner les différentes temporalités et les différentes fonctions des différentes mains contribuant à constituer les différentes variantes du texte est une mauvaise solution. Cela devient vite trop coloré et illisible. De plus, certaines couleurs se voient mieux que d'autres, ou sont plus ou moins bien rendues selon les browsers. Enfin, il ne faut pas négliger le fait que les couleurs elles-mêmes sont connotées.

Avec cette question simple : de quelle couleur notons-nous les variantes, on ne peut plus se satisfaire de reléguer la question auctoriale derrière une simple équation auteur+éditeur. Il faut déployer l'ensemble des fonctions et des strates temporelles présentes sur le manuscrit et en filigrane du manuscrit. Lire et éditer les lettres comme des textes s'articulant autour d'une série de variantes est une belle idée – mais pas si simple à mettre en pratique.