



**HAL**  
open science

## Les archives, de l’analogique au numérique : le projet “ Renan Source ”

Domenico Paone

► **To cite this version:**

Domenico Paone. Les archives, de l’analogique au numérique : le projet “ Renan Source ”. LE COURS DE LINGUISTIQUE GÉNÉRALE, 1916-2016. L’ÉMERGENCE, LE DEVENIR, CERCLE FERDINAND DE SAUSSURE, 2017, Genève, Suisse. hal-03908097

**HAL Id: hal-03908097**

**<https://hal.science/hal-03908097>**

Submitted on 23 Jan 2023

**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.

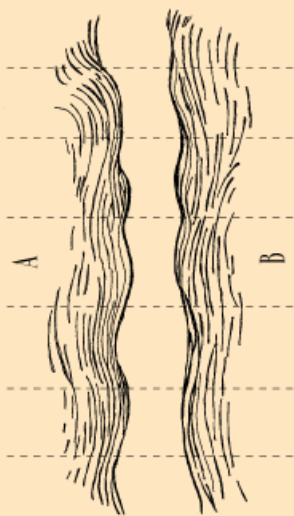
L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License

# Le Cours de Linguistique Générale 1916-2016

Genève - Paris • 2016 - 2017



## TRAVAUX DES COLLOQUES LE COURS DE LINGUISTIQUE GÉNÉRALE, 1916-2016. L'ÉMERGENCE, LE DEVENIR

Éditeurs scientifiques : Daniele  
GAMBARARA, Fabienne REBOUL.

### **Domenico PAONE, Les archives, de l'analogique au numérique : le projet « Renan Source »**

Communication donnée dans l'atelier de Giuseppe COSENZA, Giuseppe D'OTTAVI et Luka NERIMA, Les manuscrits de Saussure parmi d'autres. Problèmes, solutions, stratégies d'édition pour les archives numériques, au colloque **Le Cours de Linguistique Générale, 1916-2016. L'émergence**, Genève, 9-13 janvier 2017.

CERCLE FERDINAND DE SAUSSURE

N° D'ISBN : 978-2-8399-2282-1

Pour consulter le programme complet de l'atelier de Giuseppe Cosenza, Giuseppe D'OTTAVI et Luka NERIMA,

***Les manuscrits de Saussure parmi d'autres. Problèmes, solutions, stratégies d'édition pour les archives numériques :***

<https://www.clg2016.org/geneve/programme/ateliers-libres/les-manuscrits-de-saussure-parmi-dautres-problemes-solutions-et-strategies-dedition-pour-les-archives-numeriques-saussures-manuscripts-among-others-problems-strategies-and-publishing-solutions-for-digital-archives/index.html>



**CERCLE  
FERDINAND  
DE SAUSSURE**

## Les archives, de l'analogique au numérique : le projet « Renan Source »

Domenico Paone  
ITEM (CNRS / ENS), Paris  
domenico.paone@ens.fr

Le projet « Renan Source » se propose de créer la première édition numérique du corpus manuscrit d'Ernest Renan. Le projet, que je dirige au sein de l'ITEM, l'Institut des Textes et Manuscrits Modernes à Paris, a démarré en 2015 et se terminera en 2019. Nous sommes donc presque à mi-parcours et c'est cette expérience de mi-chemin, inévitablement partielle et inachevée, que je partagerai aujourd'hui avec vous. Je ne m'attarderai pas sur l'aspect strictement génétique de l'édition, pourtant au cœur de notre travail, mais je privilégierai une perspective plus générale, susceptible d'intéresser le public le plus vaste possible.

### 1. *Les manuscrits*

La presque totalité des manuscrits d'Ernest Renan est conservée dans deux fonds :

1) Les *Papiers Renan* de la Bibliothèque nationale de France, fonds créée en 1894 grâce au legs de ses fils. C'est l'un des fonds les plus importants de l'âge d'or des manuscrits d'écrivain en France, qui s'ouvre par la célèbre donation de Victor Hugo vers la fin du XIX<sup>e</sup> siècle, geste qui introduit « une institutionnalisation de la valeur culturelle du manuscrit littéraire »<sup>1</sup>. Renan sera le troisième, après Hugo en 1885 et Quinet en 1890, à léguer tous ses manuscrits à la BnF.

Il se compose de 105 volumes de manuscrits, carnets et notes autographes, 14 volumes d'épreuves corrigées, 52 volumes d'œuvres de Renan annotées par l'auteur et 62 volumes faisant partie de sa bibliothèque personnelle, glosés et annotés.

2) La Collection *Scheffer-Renan* du Musée de la Vie Romantique. L'Hôtel de la Rue Chaptal à Paris fut la résidence de la famille Renan jusqu'en 1983, quand il devient d'abord « Musée Scheffer-Renan », puis, en 1988, l'actuel « Musée de la Vie Romantique ». En 1991, le musée entre en possession de la *Collection Scheffer-Renan*, léguée à la Ville de Paris par Corrie Psichari-Siohan, petite-fille d'Ernest Renan. On y conserve environ 60 cartons de lettres, documents personnels, notes et manuscrits.

Le fonds de la BnF possède un catalogue, accessible en ligne depuis peu, dont la granularité, c'est-à-dire le degré de détail, n'est pas assez fine. Cela pose des problèmes, puisque tous les papiers ont été reliés en volume et que ces volumes renferment souvent des manuscrits et notes concernant plusieurs œuvres et parfois plusieurs périodes de la vie de Renan. Autrement dit, on ne connaît pas toujours exactement le contenu de chaque volume. Il existe un inventaire beaucoup plus détaillé d'une partie seulement du fonds, mais le catalogue est à son tour manuscrit et, depuis quelque temps, il n'est consultable que sous forme de microfiche. Par ailleurs, certains volumes n'étaient pas référencés à cause d'une faute de catalogage : un de ces volumes, que nous avons découvert récemment, fera justement l'objet de notre édition numérique.

La Collection *Scheffer-Renan* ne possède pas encore de catalogue, mais il existe un inventaire consultable sur place.

---

<sup>1</sup>Almuth GRÉSILLON, *Éléments de critique génétique. Lire les manuscrits modernes*, Paris, PUF, 1994, p. 85 (rééd. CNRS éditions, Collection : Textes et manuscrits, Paris, 2016).

## 2. *Quoi éditer ?*

Le projet d'éditer les manuscrits de Renan est né après la soutenance de ma thèse de doctorat, consacrée à plusieurs volumes de ses notes, et il s'est concrétisé à la suite de la découverte des volumes de manuscrits partiellement inédits et non classés<sup>2</sup>. Au début, la tentation d'échafauder une grande édition en vue de tout publier était forte, d'autant plus qu'avec le numérique on a l'impression que les éditions ne coûtent pas grand-chose ; en réalité, on apprend vite qu'une édition numérique, bien qu'ayant des avantages par rapport à une édition papier, n'est pas sans frais, tant s'en faut. Réaliser une édition « intégrale », « complète » et « définitive » – et j'emploie ici une série d'adjectifs qui s'appliquent en réalité assez mal à l'environnement numérique – pose toute une série de problèmes, dont l'un des premiers est celui des « grandes attentes » du numérique. Depuis quelques années, le « Web 2.0 » nous a habitués à accéder à une masse énorme de contenus dans un temps toujours plus réduit, si bien que nous avons la sensation, en naviguant sur internet, de pouvoir accéder à tout, tout de suite. Évidemment, cela n'est qu'une impression, mais presque tout le monde partage la même croyance que tout ce qui est numérique, finalement, est facile d'accès, rapide et gratuit.

Lancer une édition numérique –en ligne, de surcroît – du corpus manuscrit renanien nous aurait exposé fatalement aux « grandes attentes » des utilisateurs du web 2.0. Pour restreindre le champ trop vaste d'une édition complète des manuscrits de Renan, irréalisable dans un bref délai, nous avons donc décidé de fixer des objectifs intermédiaires, bien définis et reconnaissables en tant que tels. En gestion de projets informatiques, on appelle cela des « livrables », c'est-à-dire des produits qui suivent un calendrier bien précis et dont les résultats sont vérifiables à chaque étape. Notre « livrable » de lancement du projet « Renan Source » sera donc l'édition de deux romans de jeunesse inachevés, *Ernest et Béatrix* et *Patrice*, avec leur dossier génétique corrélé, et des *Notes de voyage en Italie*, qui sont aussi un des avant-textes des deux romans.

## 3. *Comment éditer ?*

Plusieurs types d'éditions numériques de manuscrits sont possibles ; « Renan source » en présentera trois : une édition en fac-similé, une édition critique, une édition génétique. Un cadre unique, un seul lieu virtuel hébergera ces trois éditions : le logiciel d'édition [Sources](#), un environnement numérique reposant sur une plateforme de publication qui utilise plusieurs logiciels open source (comme par exemple EVT - *Edition Visualization Technology*). L'histoire de cet environnement remonte au projet européen *Discovery Project. Digital semantic corpora for virtual research in philosophy* (2006-2009), qui a donné naissance à la bibliothèque « Philosource », une fédération de bibliothèques numériques dont les deux projets de plus grande longévité sont « [Wittgenstein Source](#) » et « [Nietzsche Source](#) ». Le projet « Renan source » a vocation à s'insérer dans cette lignée d'éditions numériques, en prolongeant une tradition et en profitant aussi de l'expérience passée des autres projets.

## 4. *Edition fac-similé*

La première fonction d'une édition en fac-similé est de permettre un accès direct aux manuscrits, sources primaires de l'étude approfondie d'un auteur. Phase préliminaire à tous travaux, l'acquisition des images numériques présente deux cas de figure – avec leurs propres avantages et inconvénients –, selon que l'archive est déjà numérisée ou non. Le *fonds Renan* de la BnF rentre dans le premier cas.

---

<sup>2</sup> Domenico PAONE, Francesco PETRUZZELLI, « [Deux volumes de manuscrits d'Ernest Renan retrouvés](#) », Item [En ligne], Mis en ligne le: 30 octobre 2012. Disponible sur: <http://www.item.ens.fr/articles-en-ligne/deux-volumes-de-manuscrits-dernest-renan-retrouves/>

Depuis quelques années, la BnF a lancé un grand chantier de numérisation de ses manuscrits, diffusés via le site *Gallica*, la bibliothèque numérique qui compte aujourd'hui presque 5 millions de documents disponibles en ligne. Tous les 233 volumes qui composent le fonds ne sont pas encore numérisés, loin de là, mais nous avons obtenu que priorité soit donnée aux manuscrits qui font l'objet de notre premier « livrable ». Si la BnF a fait le travail de numérisation (avantage évident qui nous a fait gagner du temps et de l'argent), elle a numérisé ces manuscrits selon ses propres spécifications techniques (format et résolutions des fichiers, modalité des prises de vue, fonds, etc.), qui n'étaient pas forcément compatibles avec celles de notre édition ; un travail supplémentaire d'adaptation en aval a donc été nécessaire. La version numérique de ces manuscrits étant disponible en même temps sur Gallica, nous nous sommes trouvés face à un choix : embarquer les images en local ou bien se contenter de faire un lien vers Gallica ? Des considérations d'ordre théorique et pratique, sur lesquels je reviendrai tout à l'heure, nous ont fait opter finalement pour la première solution.

Le fonds du Musée de la Vie Romantique n'étant pas numérisé, nous avons lancé une campagne de numérisation des manuscrits qui nous intéressaient. Nous avons pu ainsi intervenir méthodiquement sur le processus de fabrication des copies numériques, afin d'obtenir des fichiers adaptés à notre édition. Mais suivre de très près chaque phase de la numérisation a demandé beaucoup de temps. Il y a eu des réunions avec les responsables du Musée et de l'établissement qui le gère (Paris Musées) , on a dû passer par un marché public, mettre en place un cahier de charge pour la numérisation, accompagné d'un tableau de plus de 1500 lignes décrivant les manuscrits page par page (il s'agit de manuscrits assez complexes, présentant parfois de bandelettes de papier collées qui cachent un premier état de rédaction) et préconisant, pour chaque page, les prises de vue adaptées. Nous avons aussi organisé des réunions avec le prestataire auquel l'établissement public a confié la numérisation, avant, pendant et à la fin du travail. Heureusement, la présence dans notre équipe de l'archiviste du Musée de la Vie Romantique, Marie-Claude Sabouret, a facilité les choses.

Le choix de garder une copie des manuscrits sur notre site repose tout d'abord sur une question théorique : la différence essentielle entre la version d'un document présentée dans une *bibliothèque numérique* et l'*édition numérique* de ce même document. Le but principal d'une *bibliothèque numérique* est de mettre *une copie la plus fidèle possible* d'un document – présenté généralement avec très peu de détails – à disposition d'un potentiel lecteur.

Gallica

Rechercher...

dans tout Gallica

Accueil > Consultation

Œuvres d'Ernest Renan (1823-1892). Manuscripts autographes et éditions annotées.

Informations détaillées

NOTICE

Titre : Œuvres d'Ernest Renan (1823-1892).  
Manuscripts autographes et éditions annotées.  
Manuscripts autographes. XIII Vie de Jésus.  
Date d'édition : 1801-1900  
Type : manuscrit  
Langue : Français  
Format : 113 f  
Droits : domaine public  
Identifiant : ark:/12148/btv1b53003147d  
Source : Bibliothèque nationale de France,  
Département des manuscrits, NAF 11448  
Provenance : Bibliothèque nationale de France  
Date de mise en ligne : 09/01/2012

Aide

Source gallica.bnf.fr / Bibliothèque nationale de France

Première page du manuscrit de la Vie de Jésus de Renan, présenté dans Gallica. On note que les informations détaillées sur le document sont très sommaires : l'unité pour la datation est le siècle.

Une *édition numérique*, au contraire, tout comme son ancêtre analogique, vise à placer le document dans un contexte (historique, littéraire, biographique) suivant une démarche philologique. Le document se trouve ainsi disposé sur une ligne diachronique qui en fait émerger les liens avec la production de l'auteur – les œuvres imprimées ou bien d'autres documents. Des détails codicologiques peuvent enrichir sa description et nous aider souvent à retracer son histoire. Une édition numérique peut aussi parvenir à rétablir l'ordre de certains manuscrits classés erronément (ce qui a été le cas pour plusieurs volumes du *fonds Renan* de la BnF, créés par les conservateurs de la BnF à la fin du XIX<sup>e</sup> siècle, en montant sur onglets des feuilles disparates et en reliant le tout).

Ces deux façons de présenter un même document (et le fait que ce soit *exactement* le même document est la seule différence avec l'ère analogique : car auparavant, on avait un *original* conservé dans les bibliothèques et une *copie* de cet original publiée dans les éditions) représentent deux rôles et deux fonctions qu'il faut maintenir séparés : *la bibliothèque* ou *l'archive numérique* – et *l'édition savante numérique*. Bien évidemment, tout comme à l'âge analogique, ces deux rôles sont complémentaires : la coopération et la collaboration des archivistes et des éditeurs seules peuvent nous restituer une édition scientifiquement rigoureuse.

D'un côté plus pratique, les images embarquées permettent de travailler sur des portions de page, pour isoler une note ou un fragment spécifique. Cela est très important dans le cas d'une édition génétique, dont les unités fondamentales ne correspondent souvent pas aux feuillets.

### 5. *Édition génétique*

Je ne m'attarderai pas ici sur les détails de l'édition génétique ; le logiciel en question fera bientôt ses débuts dans l'édition génétique de deux œuvres de Nietzsche, *Le voyageur et son ombre* et *Aurore*, qui

paraîtront dans « Nietzsche Source ». Cette édition permettra de visualiser et suivre les parcours génétiques entre les notes, les brouillons, les plans de travail et les manuscrits<sup>3</sup>.

#### 6. *Édition critique*

Les deux romans de jeunesse *Ernest et Béatrix* et *Patrice*, aussi bien que les *Notes d'Italie*, ont déjà été publiés, mais dans des éditions très incomplètes et scientifiquement faibles (et qui sont par ailleurs indisponibles depuis des années). Une édition critique s'est donc imposée comme une évidence. L'équipe a déjà transcrit plus de la moitié des manuscrits. L'édition permettra de visualiser les différentes couches de réécriture (biffures, ajouts marginaux, surcharges, substitutions, etc.) et ses fichiers seront compatibles XML-TEI.

#### 7. *Le droit d'auteur sur les éditions numériques*

Notre édition se développe dans un cadre éminemment public : les porteurs du projet sont des laboratoires de recherche financés par l'État (CNRS-ENS), en partenariat avec d'autres institutions publiques (BnF, Ville de Paris). Le choix de l'*Open Access*, de l'archive ouverte en libre accès (choix qui est culturel, politique et aussi, dans un certain sens, moral), nous a donc paru naturel. Tout le contenu de « Renan source » sera distribué sous la licence *Creative Commons Attribution – Pas d'Utilisation Commerciale – Pas de Modification 4.0 International* (CC BY-NC-ND 4.0).

Le choix de l'interdiction des modifications a donné lieu à de longues discussions. Plusieurs acteurs dans le monde de l'*Open Source* – et notamment la [Definition of Free Cultural Works](#) utilisée par la Wikimedia Foundation – soutiennent qu'un véritable « libre accès » doit prévoir la possibilité de modifier l'objet mis à disposition du public (et aussi la possibilité d'en faire une utilisation commerciale). Finalement, a prévalu la considération que non seulement toute édition comporte un choix, mais elle coïncide même avec ce choix. Dans un contexte savant, il est important que la responsabilité éditoriale soit bien claire et définie. Une diffusion avec attribution de paternité et interdiction de modification, nous a donc paru le meilleur choix pour garantir cette responsabilité éditoriale. Rien n'empêche, évidemment, d'autres chercheurs de faire d'autres éditions fondées sur d'autres choix, ce qui est la base d'une saine confrontation critique.

#### 8. *Droit d'auteur et droit des archives : les conventions*

L'œuvre d'Ernest Renan est entrée depuis longtemps dans le domaine public. En France, toutefois, les droits de reproduction et de diffusion des manuscrits restent toujours l'apanage des établissements qui les conservent, même si ce droit est de plus en plus débattu et est considéré par certains comme une utilisation abusive, ce qu'on appelle en anglais une *copy fraud*<sup>4</sup>. Avant toute édition de manuscrits, il faut donc obtenir les droits à la publication. L'ITEM a signé à cet effet une convention avec Paris-Musées, qui concerne une partie des manuscrits de la *Collection Scheffer-Renan* du Musée de la Vie Romantique et il est en train de signer une convention avec la BnF (une démarche un peu plus complexe, vu la taille de l'établissement).

---

<sup>3</sup> Cf. Paolo D'TORIO, « L'edizione genetica dell'opera di Nietzsche », *Autografo*, 57 (2017), p. 131-146 ; id, « Die Schreib- und Gedankengänge des Wanderers. Eine digitale genetische Nietzsche-Edition », *Editio. Internationales Jahrbuch für Editions-wissenschaft*, 31 (2017), p. 191-204.

<sup>4</sup> Alain GUERREAU, [L'Avenir de la philologie. Textes anciens et domaine public](#), 2015 <halshs-01112213> [en ligne : <https://halshs.archives-ouvertes.fr/halshs-01112213>].



## 9. Hébergement des données

L'hébergement des données est une phase considérée, souvent à tort, secondaire ou trop technique. Savoir où nous entreposerons nos données est, au contraire, une question de première importance. En effet, les serveurs hébergeront non seulement les sources primaires et les données de notre recherche (images numériques, transcriptions, documents), mais aussi les logiciels qui forment l'infrastructure numérique de l'édition (bases de données, logiciels de visualisation, moteurs de recherche, etc.). Selon la volumétrie des données, l'espace nécessaire— qui doit être sécurisé et permettre un accès permanent — peut se révéler important. Mais le stockage ne suffit pas : il faut aussi considérer un archivage à long terme et la sauvegarde périodique sur plusieurs serveurs : personne n'aimerait voir des années de travail perdus à cause d'une panne d'électricité ou d'un autre accident !

Il y a plusieurs choix possibles pour l'hébergement d'une édition numérique : serveurs locaux ou distants du laboratoire, serveurs privés extérieurs, serveurs collectifs. Notre projet a pu bénéficier des services de la TGIR [Huma-Num](#), par le biais de l'ITEM. Les services de cette Très Grande Infrastructure de Recherche du CNRS sont fiables et surtout gratuits, quitte à respecter des critères qui sont :

- avoir des besoins spécifiques sur le traitement des données (enrichissements et annotations, calculs, signalements, etc.) ;
- s'engager sur l'interopérabilité des données de la recherche (protocole OAI-PMH - *Open Archives Initiative Protocol for Metadata Harvesting*) ;
- avoir une démarche d'archivage à long terme des données.

## 10. Conclusions

La rapidité de mon exposé n'a sans doute pas restitué la complexité des procédés nécessaires à la mise en place de l'édition numérique d'une archive manuscrite. Dans notre expérience, la phase la plus chronophage a été le début. C'est le moment où tout est conçu, sur le papier, dans les moindres détails, du budget à la division du travail dans l'équipe ; mais la réalité a souvent tendance à ne pas rentrer dans le cadre qu'on lui a si soigneusement préparé. Les maillons de la chaîne éditoriale étant liés les uns les autres, il se peut que l'attente d'une réponse, d'une réunion d'un conseil d'administration, mette en stand-by le reste du processus. Et tout cela sans compter avec les imprévus qui vous obligent à adapter le projet aux nouvelles données et parfois à le repenser de fond en comble. Un des aspects les plus difficiles à gérer est le rôle d'« officier de liaison » entre plusieurs acteurs, qui vont de votre collègue de l'équipe aux techniciens informatiques, de l'association savante à la très grande institution publique. Ces acteurs n'ont pas du tout les mêmes temps de réaction, et votre tâche principale consiste, le plus souvent, en un effort pour accorder tous les rouages de la machine et les faire fonctionner ensemble.

Tout compte fait, malgré les nombreuses difficultés, je trouve que la réalisation d'une édition numérique reste une entreprise très passionnante et qui peut ouvrir de belles perspectives à la recherche. Je ne peux donc que vous inviter de tout cœur à vous engager dans ce parcours.

## Bibliographie

Paolo D'ORIO, *HyperNietzsche. Modèle d'un hypertexte savant sur Internet pour la recherche en sciences humaines. Questions philosophiques, problèmes juridiques, outils informatiques* / sous la direction de Paolo D'Iorio. Paris, PUF, 2000.

Paolo D'IORIO, « [Qu'est-ce qu'une édition génétique numérique ?](#) », in *Genesis. Manuscrits, recherche, invention*, n° 30, 2010, p. 49-53.

Paolo D'IORIO, « L'edizione genetica dell'opera di Nietzsche », *Autografo*, 57 (2017), p. 131-146.

Paolo D'IORIO, « Die Schreib- und Gedankengänge des Wanderers. Eine digitale genetische Nietzsche-Edition », *Editio. Internationales Jahrbuch für Editions-wissenschaft*, 31 (2017), p. 191-204.

Almuth GRÉSILLON, *Éléments de critique génétique. Lire les manuscrits modernes*, Paris, PUF, 1994 (rééd. CNRS éditions, Collection : Textes et manuscrits, Paris, 2016).

Élisabeth GRESLOU et al. : [L'édition numérique de corpus d'auteurs – aspects juridiques](#) [en ligne : <https://cahier.hypotheses.org/corpus-auteurs-aspects-juridiques>]

Alain GUERREAU, [L'Avenir de la philologie. Textes anciens et domaine public](#), 2015 <halshs-01112213> [en ligne : <https://halshs.archives-ouvertes.fr/halshs-01112213>]

Domenico PAONE, « La bibliothèque personnelle d'Ernest Renan », *Bulletin du bibliophile*, année 2016, n. 1, p. 99-130.

Domenico PAONE, Francesco PETRUZZELLI, « [Deux volumes de manuscrits d'Ernest Renan retrouvés](#) », Item [En ligne], Mis en ligne le: 30 octobre 2012. Disponible sur: <http://www.item.ens.fr/articles-en-ligne/deux-volumes-de-manuscrits-dernest-renan-retrouves/>