

Notes et remarques sur l'HTML du premier tome de correspondance de Lamennais

Par Lucie Slavik, stagiaire 19 mai-29 juillet 2020, deux jours par semaine en télétravail.

Mardi 16 juin 2020, J9

REMARQUES DU CAHIER DES CHARGES

MARQUEURS DE LA LETTRE

Début des lettres : Chiffre arabe suivi d'un point et d'un tiret + adresse au destinataire en petites capitales sur le mode « A.... » (sur la même ligne)

+ lieu et date en minuscule au dessous

Fin de la lettre : les lettres sont rarement signées.

- Correspondance active **PAS QUE !**
- Très longue préface dans le volume 1 intitulée « Notes et souvenirs » à enlever (jusqu'à la page 140) **C'EST FAIT**
- Distinction entre les lettres grâce à des chiffres arabes
- Alinéas marquent des paragraphes
- Le destinataire est indiqué en majuscule. Lettres à la suite adressées aux mêmes destinataires s'ouvrent sur « AU MÊME » ou « À LA MÊME » (voir p. 147)
- En principe date et lieux à droite sur le modèle (« Lieu, jj mois AAAA. »). Dans certains cas Lamennais mentionne également l'heure (p. 147)
- Structure de la lettre : les formules de politesse de début et de fin sont en grande majorité intégrées au corps de la lettre.

MES REMARQUES

1) Remarques d'ordre général sur la correspondance de Lamennais, tome 1 et ses marqueurs, en vue de l'extraction

a) Remarques sur l'HTML, distinction des lettres

-l'HTML comporte 12 204 lignes au départ, 8680 après nettoyage, soit le double de l'HTML de Lamartine après nettoyage.

- **Le premier tome regroupe trois volumes, ce qui explique son importance.**

Le premier volume regroupe la correspondance de 1818 à 1823.

Le deuxième couvre les années 1826-1827.

Le troisième est entièrement consacré à l'année 1828.

-il y a 197 lettres, soit 100 de plus que Lamartine.

- Les lettres sont précédées d'un chiffre arabe (de 1 à 197) puis suivies du destinataire avec son nom, et si plusieurs lettres s'adressent au même destinataire, on retrouve ces expressions : « A la même », « Au même », « Le même au même », ce qui arrive assez fréquemment.

Les chiffres arabes sont suivis d'un point et d'un tiret :

Exemple, l. 7682 « 169. — A MADAME LA COMTESSE DE SENFFT. »

- Jamais il n'est écrit cher untel : pas de <salute/> dans l'<opener/>

- On trouve en revanche le lieu puis la date (JJ Mois AAAA) suivis directement du début de la lettre.

-Parfois l'indication du lieu est absente, comme à la ligne 7920 <p>179. - A MADAME LA COMTESSE DE SENFFT. </p><p>Le 27 août 1 828. </p>

- Parfois, l'heure est ajoutée, après la date comme à la ligne 4912 : <p>« A la Chenaie, mardi 31, 3 heures de l'après-midi. </p>

+ ligne 153 + 4919.

Manifestement, cela arrive maximum 3 ou 4 fois.

Comment le signaler dans XML ? Avec la date ?

- **Les balises ** indiquent les doutes sur l'OCR. On en trouve 510 ouvrantes et fermantes (donc 255). Ce sont pour les erreurs quasiment avérées, mais beaucoup d'autres erreurs se sont glissées dans le texte, sans être signalées par une balise. Une relecture sera donc nécessaire.

Exemple de span l. 7806 : La°viê de ce pays-ci...

Exemple de fautes non signalées l. 7825 : Leur conscience leur a répondu qu'il valait mieux obéir à Dieu qu'(.ux hommes... : le a est remplacé par une parenthèse ouvrante et un point.

Parfois, les balises sont là alors qu'aucune faute n'est à signaler, comme par exemple à la ligne 1263 où le guillemet est juste alors que signalé comme douteux : « Regardez bien, vous autres, afin que je ne perde pas mon temps, et que vous puissiez légalement déposer du fait! >>

- **Les balises <hr>** mentionnent un changement de page. Elles sont en général précédées et suivies d'une balise fermante puis ouvrante <p/>

Exemple : changement de page 344-345 dans le pdf, ligne 5063 dans le HTML.

voire </p><hr /><p>si douce amitié;

→ La phrase est divisée en paragraphes distincts.

Dans l'idéal, il faudrait les supprimer pour qu'il n'y ait pas de paragraphes inutiles, qui divisent une phrase et rompent l'unité.

- Fort heureusement, contrairement à la correspondance de Lamartine, les en-têtes à chaque page (« CORRESPONDANCE » page de gauche, « DE LAMENNAIS » page de droite) ont été automatiquement supprimés par l'OCR et ils ne polluent donc pas l'HTML.

Seules les divisions par livre sont encore visibles.

b) Les destinataires

- les destinataires sont au nombre de 7. On y trouve :

Mlle Cornulier de Lucinière

Mlle de Tremereuc

M. l'abbé Carron

M. Berryer

M. le marquis de Coriolis d'Espinouse

Mme la comtesse de Senfft-Pilsach /Mme la comtesse Louise de Senfft (est-ce la même?)/ la comtesse Louise... (lettres 124, 167)

M. le comte de Senfft-Pilsach

- les destinataires sont écrits en majuscules, ce qui peut être un moyen d'extraire les lettres avec des regex.

PAS TOUJOURS CEPENDANT.

- Certaines lettres ne sont pas de Lamennais ni destinées à lui

→ Les lettres 77, 80, de M. l'abbé Guerbet à M. de Senfft

→ La lettre 173, de l'abbé Jean de Lamennais (frère de Félicité de Lamennais) à M. de Senfft

Doit-on les garder ? On verra je pense en fonction de la cohérence interne (si ce sont des lettres jointes à celles à Lamennais, comme par exemple la lettre 160 : si elles ont été publiées, c'est que leur « insertion [...] paraît utile »¹...).

- D'autres sont insérées dans les notes de bas de page, comme une lettre de M. de Vitrolles à Lamennais, p. 423. Que faire dans ce cas-là ?

- D'autres sont des lettres écrites à Lamennais = CORRESPONDANCE PASSIVE

Comme les lettres 120, 134, 140, 156, 182, 183, 190, 195 de M. de Coriolis à Lamennais

c) Notes

- on y trouve de fréquentes notes. Elles sont précédées d'un chiffre arabe, ce qui est contrariant puisque ce sont également les chiffres arabes qui délimitent les lettres.

d) Formules en fin de lettre ou <salute/>

On trouve des formules de fin de lettre, comme

Mille tendresses (lettre 151)

Mille et mille tendresses (lettre 169)

Mille et mille tendres hommages... (lettre 175)

Mille et mille hommages (lettre 179)

J'aurais encore à vous dire une infinité de choses [...] Adieu... (lettre 154)

Adieu... (lettres 158, 173, 194)

Tout à vous bien tendrement (lettre 163)

Ne soyez point surprise si je ne vous écris pas d'ici à quelque temps, car je vais être excédé d'embarras de toute espèce, et de fatigue. A bientôt. (lettre 168)

J'embrasse... (lettre 180)

e) Signature

- la plupart du temps, les lettres ne sont pas signées.

- Sinon on y trouve, en guise de signature :

- F. M. (lettres 3, 4, 18)

- F. (lettres 15)

- F. L. (lettres 46)

En bref, sur 197 lettres, 5 sont signées de façon sûre, par de simples initiales.

2) Nettoyage de l'HTML

-Je ferme les balises <meta/>,
 et <hr/> qui empêchent une bonne indentation et un html bien formé.

- le pdf comporte 518 pages, dont 151 sont des notes et souvenirs, et non de la correspondance (même si d'abondants passages de correspondance y sont cités). Ces passages sont séparés par des chiffres romains. Par E. D. FORGUES.

Je les supprime de l'HTML (lignes 56 à 3416).

- MKME remplacé par MÊME : 4 occurrences.

- Suppression de la table des matières, de la ligne 8681 à 8840.

Remarques pour le code python qui permet d'extraire les lettres de l'HTML et de les mettre dans des fichiers xml

Mercredi 17 juin 2020, J10

sur le fichier nommé `extraction-elicom_lamennais.py`

1) Choix des REGEX

a) Pour matcher les débuts de lettres : chiffres arabes (de 1 à 3) suivis d'un point, d'un espace et d'un tiret

RE_CHIF_AR = re.compile(r">[0-9]{1,3}\.") #pour matcher les débuts de lettres. 153 matches au lieu de 197

Pourquoi des matches en moins ?

→ Lettre 14 signalée dans l'OCR l. 618 par un H <p>H. - A MADEMOISELLE DE TREMEIiEIH". </p>

→ Lettre 15, même problème l. 648 <p>l.l. — A L \ MÊME.</p>

→ Lettre 19 l. 751 <p>I!>. - A MADEMOISELLE DE THEMEREUC. </p>

→ Lettre 20 l. 773 <p>-20. - A L \ MÈMK. </p> Il y a un tiret après la balise donc le chiffre arabe n'est pas matché.

→ Lettre 21 l.793 <p>-2). - A MADEMOISELLE DE LUCINIÈRE. </p>

→ Lettre 22 l.824 <p>2*2. \ LA MK.M I. </p>

→ La lettre 23 est signalée comme étant la lettre 25. Il y a donc deux lettres 25 dans l'OCR.

→ Lettre 26 l.885 <p>:!i. - A LA)1 È 31 E. </p>

→

Mardi 23 juin 2020

Suite aux conseils d'Arthur, je modifie manuellement les chiffres arabes mal matchés, même si ce n'est pas l'idéal, car sinon cela va être l'horreur pour l'extraction des lettres et on perdra encore plus de temps.

Pour les regex pour extraire les lettres, éviter de toucher à l'HTML.

Si ce sont des corrections qui peuvent être facilement être répercutées sur les autres tomes de la correspondance, pas de problème, c'est quelque chose qu'on peut facilement réutiliser.

Si c'est que pour cette correspondance, c'est dommage.

Si ça ne prend pas trop de temps, le faire.

Si ça facilite l'extraction, c'est mieux.

Liste des fautes les plus courantes qui font que les chiffres arabes ne sont pas reconnus :

lettre

5 ci

6 (i

14 H

15 l.'l

19 Il>

-20 J'enlève le tiret

21 -2)

22 2*2

23 prise pour 25

24 :li

29 -

31 vu comme 51 Je laisse parce que ce qui compte, ce sont les chiffres arabes et non leur ordre. De toute façon, il seront extraits dans l'ordre normalement (j'espère!)

La plupart des chiffres de la trentaine sont reconnus comme des chiffres de la cinquantaine, mais ce qui compte c'est qu'ils soient reconnus comme chiffres arabes.

38 7-8 J'enlève juste le tiret

39 -

40 411.

41 a.

44 t4

45 .

47 i7

50 bU,

51 M.

53 53,

59 f.VJ

60 CO

63 ftj

64 tH

65 C5

66 titi

71 j'l,

73 loi

74 "4

75 7a

84 S4

85 Sa

86 8(5

89 S9

104 6104

125 12,

156 15G

164 16i

166 KiC

171 17 1

175 17a

178 17S

190 19»

Le 5 est souvent reconnu comme un a

Le 5 et le 3 sont souvent inversés

Le 8 est pris pour un S

=> **Voir si avec ces repérages on ne pourrait pas faire par la suite des expressions régulières qui marchent pour les lettres suivantes (le volume suivant).** Après, au pire, il n'y a que deux volumes de Lamennais.

Attention aux dates qui ne sont pas toujours bien reconnues dans l'HTML.

Par exemple, ligne 2791, pour la lettre 73, la date est reconnue comme 21 juillet alors que c'est le 27 juillet.

=> Maintenant, 197 matches.

Il faut compter près d'une heure pour ce repérage

b) Pour matcher les destinataires

RE_DEST = re.compile(r"(A [A-Z][A-Z.]+)|^(?:A[A-Z.]+)") # Pour matcher les destinataires. 215 matches non complets

Il serait bon d'affiner, ce résultat est loin d'être idéal.

c) Pour matcher les notes de bas de page

#RE_NOTES = re.compile(r">1 .+ -|>1 .+—") #pour matcher les notes. Pour l'instant 20 matches dont un non pertinent

RE_NOTES = re.compile(r"1 .+ -|>1 .+—|1 [A-Z].+") #pour matcher les notes. Pour l'instant 191 matches dont plusieurs non pertinents

#2 .+ -|>1 .+—|1 [A-Z].+ ICI 190

Mercredi 1^{er} juillet 2020

Remarques sur le code python

=> Comme pour Lamartine, toutes les lettres moins une sont extraites, je ne comprends pas pourquoi.

=> Problème sent et received mal identifiés (surtout received : il y a la date à la place)

=> date when mal matché : il y a du texte à la place

→ à la ligne 52, j'enlève date : cela évite les problèmes de dates.

=> Les notes parsèment le texte.

Mardi 7 juillet 2020, J15

- Je rencontre au début un problème d'environnement que j'ai oublié de mettre:/

- Je demande à Arthur le script de Pasteur pour m'aider.

- Noms de lieux desquels Lamennais écrit ses lettres (quelques dizaines de lettres sans nom de lieu) :

- Saint-Brieuc, A la Chenaie, Paris, Genève, Rome, Dinan, Rennes, Lyon, Versailles, Saint-Sauveur, Château de Fleury, Château de Verneuil, A Vitrolles

Pour matcher Dinan hors du texte, écrire :

>Dinan, 1 match
>Saint-Brieuc 3 matches
>Paris 44
>Genève 1
>Rome
>Lyon 1
>La Chenaie, 12
A la Chenaie 48
>Versailles 5
>Saint-Sauveur, 1
Château de Fleury, 1
Château de Verneuil, 1
A Vitrolles, 2
>Rennes 1

A Vitrolles,|>Rennes|Château de Verneuil,|Château de Fleury,|>Saint-Sauveur,|>Versailles|A la Chenaie|>La Chenaie,|>Lyon|>Rome|>Genève|>Paris|>Saint-Brieuc|>Dinan,

- Des vers sont écrits dans certaines lettres, tout au moins la lettre 40. Il faudra être attentive à cela au moment de la correction des fichiers XML.

-J'enlève les balises <hr /> qui me paraissent être inutiles.

Normalisation des destinataires :

les destinataires sont au nombre de 7. On y trouve :

Mlle Cornulier de Lucinière

Je pensais Hélène

1820-1891

[https://gw.geneanet.org/bernardimhaus?](https://gw.geneanet.org/bernardimhaus?lang=en&pz=bernard+didier+hugues&nz=imhaus&ocz=0&p=helene+anne+marie&n=de+cornulier+luciniere)

[lang=en&pz=bernard+didier+hugues&nz=imhaus&ocz=0&p=helene+anne+marie&n=de+cornulier+luciniere](https://gw.geneanet.org/bernardimhaus?lang=en&pz=bernard+didier+hugues&nz=imhaus&ocz=0&p=helene+anne+marie&n=de+cornulier+luciniere)

<https://www.decarne.com/gencar/pag632.html#1>

Mais cela n'y paraît pas.

Je ne vois pas...

Mlle de Tremereuc

Aucune info précise trouvée

M. l'abbé Carron Guy-Toussaint-Julien Carron (1760-1821)

https://data.bnf.fr/fr/10691491/guy-toussaint-julien_carron/ Probablement lui. Ref BNF

<https://bibliotheques-specialisees.paris.fr/ark:/73873/FRCGMSUP-751045102-NF01/BHPNF010374>

M. Berryer

Pierre Antoine Berryer (1790-1868)

https://data.bnf.fr/fr/12343838/pierre_antoine_berryer/

M. le marquis de Coriolis d'Espinouse

https://data.bnf.fr/fr/10736321/charles_louis_alexandre_coriolis_d_espinouse/

Charles Louis Alexandre Coriolis d'Espinouse (1770?-1841)

Mme la comtesse de Senfft-Pilsach /

Femme du comte

Mme la comtesse Louise de Senfft /a comtesse Louise... (lettres 124, 167)

Fille de la comtesse [https://books.google.fr/books?](https://books.google.fr/books?id=tarJIWg8irwC&pg=PA153&dq=comtesse+de+Senfft-Pilsach&hl=fr&sa=X&ved=2ahUKEwiz6e7Vmr3qAhUvxYUKHeQQAzUQ6AEwCHoECAAQAg#v=onepage&q=comtesse%20de%20Senfft-Pilsach&f=false)

[id=tarJIWg8irwC&pg=PA153&dq=comtesse+de+Senfft-](https://books.google.fr/books?id=tarJIWg8irwC&pg=PA153&dq=comtesse+de+Senfft-Pilsach&hl=fr&sa=X&ved=2ahUKEwiz6e7Vmr3qAhUvxYUKHeQQAzUQ6AEwCHoECAAQAg#v=onepage&q=comtesse%20de%20Senfft-Pilsach&f=false)

[Pilsach&hl=fr&sa=X&ved=2ahUKEwiz6e7Vmr3qAhUvxYUKHeQQAzUQ6AEwCHoECAAQ](https://books.google.fr/books?id=tarJIWg8irwC&pg=PA153&dq=comtesse+de+Senfft-Pilsach&hl=fr&sa=X&ved=2ahUKEwiz6e7Vmr3qAhUvxYUKHeQQAzUQ6AEwCHoECAAQAg#v=onepage&q=comtesse%20de%20Senfft-Pilsach&f=false)

[Ag#v=onepage&q=comtesse%20de%20Senfft-Pilsach&f=false](https://books.google.fr/books?id=tarJIWg8irwC&pg=PA153&dq=comtesse+de+Senfft-Pilsach&hl=fr&sa=X&ved=2ahUKEwiz6e7Vmr3qAhUvxYUKHeQQAzUQ6AEwCHoECAAQAg#v=onepage&q=comtesse%20de%20Senfft-Pilsach&f=false)

M. le comte de Senfft-Pilsach

Le comte de Senfft-Pilsach (1777-1853)[https://fr.wikisource.org/wiki/Une_Correspondance_in_%C3%A9dite_de_Lamennais - Lettres %C3%A0 Vuarin/01#cite_note-7](https://fr.wikisource.org/wiki/Une_Correspondance_in_%C3%A9dite_de_Lamennais_-_Lettres_%C3%A0_Vuarin/01#cite_note-7)

Frédéric Christian Louis de Sefft-Pilsach [https://books.google.fr/books?](https://books.google.fr/books?id=TwwaAgAAQBAJ&pg=PA105&lpg=PA105&dq=comtesse+de+Senfft-Pilsach+lamennais&source=bl&ots=ZjxD9Lt7zC&sig=ACfU3U32aHjIrNLZSS5dNDsofsZ6hvk5ww&hl=fr&sa=X&ved=2ahUKEwignPngmb3qAhVJzRoKHTBqA0EQ6AEwAnoECAoQAQ#v=onepage&q=comtesse%20de%20Senfft-Pilsach%20lamennais&f=false)

[id=TwwaAgAAQBAJ&pg=PA105&lpg=PA105&dq=comtesse+de+Senfft-](https://books.google.fr/books?id=TwwaAgAAQBAJ&pg=PA105&lpg=PA105&dq=comtesse+de+Senfft-Pilsach+lamennais&source=bl&ots=ZjxD9Lt7zC&sig=ACfU3U32aHjIrNLZSS5dNDsofsZ6hvk5ww&hl=fr&sa=X&ved=2ahUKEwignPngmb3qAhVJzRoKHTBqA0EQ6AEwAnoECAoQAQ#v=onepage&q=comtesse%20de%20Senfft-Pilsach%20lamennais&f=false)

[Pilsach+lamennais&source=bl&ots=ZjxD9Lt7zC&sig=ACfU3U32aHjIrNLZSS5dNDsofsZ6hvk5](https://books.google.fr/books?id=TwwaAgAAQBAJ&pg=PA105&lpg=PA105&dq=comtesse+de+Senfft-Pilsach+lamennais&source=bl&ots=ZjxD9Lt7zC&sig=ACfU3U32aHjIrNLZSS5dNDsofsZ6hvk5ww&hl=fr&sa=X&ved=2ahUKEwignPngmb3qAhVJzRoKHTBqA0EQ6AEwAnoECAoQAQ#v=onepage&q=comtesse%20de%20Senfft-Pilsach%20lamennais&f=false)

[ww&hl=fr&sa=X&ved=2ahUKEwignPngmb3qAhVJzRoKHTBqA0EQ6AEwAnoECAoQAQ#v=](https://books.google.fr/books?id=TwwaAgAAQBAJ&pg=PA105&lpg=PA105&dq=comtesse+de+Senfft-Pilsach+lamennais&source=bl&ots=ZjxD9Lt7zC&sig=ACfU3U32aHjIrNLZSS5dNDsofsZ6hvk5ww&hl=fr&sa=X&ved=2ahUKEwignPngmb3qAhVJzRoKHTBqA0EQ6AEwAnoECAoQAQ#v=onepage&q=comtesse%20de%20Senfft-Pilsach%20lamennais&f=false)

[onepage&q=comtesse%20de%20Senfft-Pilsach%20lamennais&f=false](https://books.google.fr/books?id=TwwaAgAAQBAJ&pg=PA105&lpg=PA105&dq=comtesse+de+Senfft-Pilsach+lamennais&source=bl&ots=ZjxD9Lt7zC&sig=ACfU3U32aHjIrNLZSS5dNDsofsZ6hvk5ww&hl=fr&sa=X&ved=2ahUKEwignPngmb3qAhVJzRoKHTBqA0EQ6AEwAnoECAoQAQ#v=onepage&q=comtesse%20de%20Senfft-Pilsach%20lamennais&f=false)

→ Je matche les noms de lieux pour rien.

→ Je finis plus ou moins le script python

→ Je lis la doc sur Elicom

→ Je commence à rechercher des infos pour normaliser les destinataires... Pas évident.

Continuer demain et corriger les XML

20 juillet 2020

Correction des XML

→ Nécessité de se reporter au PDF, notamment pour la suppression des notes, et la vérification de l'identité du destinataire (à la même : qui est-ce?)

- fichier 104 : Il manque tout le dernier paragraphe que je dois copier manuellement.

<p>J'attends mon frère le samedi de la semaine de la Pentecôte. Il ne sera que peu de jours ici. De Trémigon, point de nouvelles ; mais je pense qu'on s'y porte bien. J'embrasse la bonne de Villiers, votre excellent ami, ma petit Hélène, enfin tout le monde, et vous même aussi, from all my heart. </p>

Question : comment rendre l'anglais ? La dernière phrase est sensée être en italiques.

Comme pour Lamartine, je fais le choix de séparer le paragraphe et de mettre dans le salut la formule de salutations :

<p>J'attends mon frère le samedi de la semaine de la Pentecôte. Il ne sera que peu de jours ici. De Trémigon, point de nouvelles ; mais je pense qu'on s'y porte bien. </p>

<closer>

<salute>J'embrasse la bonne de Villiers, votre excellent ami, ma petit Hélène, enfin tout le monde, et vous même aussi, from all my heart. </salute>

*103

Idem, le dernier paragraphe manque !

Je dois le copier.

J'ajoute ce qui est mentionné entre parenthèses dans l'édition et en italique. Voir comment il faut le rendre pour notre édition numérique.

Seulement, il s'est perdu par entêtement et par présomption,
n'ayant jamais voulu écouter personne, même le chef de la compagnie à laquelle il appartenait... (La fin de la lettre manque.)</p>

*102

Il faut supprimer le renvoi vers les notes dans le corps du texte
retrait1.</p>

*101

Toujours le même problème : le dernier paragraphe a disparu...

Je dois le copier manuellement.

Pour la phrase de salute, c'est du latin. Voir comment rendre le latin et les italiques.

<p>Vous voyez, mon cher comte, j'abuse de votre amitié ; mais vous êtes si bon, que vous me pardonnerez. Il s'agit d'ailleurs d'avancer mes travaux que je crois importants, et de soutenir un peu ma santé qu'altère une solitude trop absolue.</p>

<closer>

<salute>Tuissimus in Christo.</salute>

<signed/>

*100

Latin à rendre...

Je copie le dernier paragraphe.

Suppression de notes qui étaient dans le corps du texte parce que c'était les suites des notes de la page précédente.

CAS INTÉRESSANT à mentionner pour le mémoire (impossibilité de tout matcher!).

Et importance d'avoir le PDF sous les yeux...

*10

Ici il y avait le dernier paragraphe.

Ce n'est donc pas systématique. C'est à n'y rien comprendre...

*1 Je dois copier le dernier paragraphe...

Comment rendre les abréviations ?

Mlles de Tremereuc et de Villiers, Mme de Cougnac, Mlle d'Ozonville, MM. de Rumedon, Murinais.

Mlles → M^{lle} dans le texte

*0 Là encore je dois copier le dernier paragraphe

<salute>On m'interrompt. Mille choses bien tendres à tous nos amis. Embrassez, s'il vous plaît, pour moi, nos chères petites filles. Tout à vous pour la vie.</salute>

Mardi 21 juillet 2020

*108

Voir ce que signifient les points dans l'édition papier en début de lettre.
Je les retranscris ainsi. Voir s'il faut les supprimer.

<p>.....</p>

*105

Vers italiens, comment le signaler en TEI ?

<lg>

<l>Ecco Dite...</l>

<l>Lo'imperador del doloroso regno...</l>

<l>S'ei t'u si bel, com' egli è ora brutto,</l>

<l>E contra'l suo Fattore alzo le ciglia,</l>

<l>Ben dee da lui procedere ogni lutto</l>

</lg>

Attention aux fichiers XML 116 et 117 !

Il y a plusieurs lettres de dedans et les destinataires sont vagues.

Il faudra être vigilant.

Je n'ai pas fait le fichier 11.

!!!!!!!!!!!!!! Suite aux remarques d'Arthur, j'ai modifié le script dans un dossier lamennais_python (en minuscules)

!!!!!!!!!!!!!! Attention à ne pas faire de confusions entre le dump de LAMENNAIS_PYTHON (avec les premières corrections pour les trois premières colonnes) et le dump de lamennais_python qu'il faudra que je corrige.

> repenser le nommage des fichiers

Vendredi 24 juillet 2020

*116

Je mets une autre div pour mettre une note de bas de page, mais c'est une solution provisoire.

<signed/>

</closer>

<postscript>

<p/>

</postscript>

</div>

<div>

<p>La correspondance de Lamennais subit ici une lacune qu'explique la grave maladie à laquelle il faillit succomber, en juillet-août 1827. Nous pensons qu'on ne lira pas sans intérêt les lettres de l'abbé Gerbet qui en rendent compte à la famille de Senfft</p>

</div>

</body>

Mercredi 29 juillet 2020

*123

Je me permets de mettre une note pour la traduction des vers italiens glissés dans la lettre de Lamennais. Voir si on garde cette balise pour l'ODD

<lg>

<l>Poi che voi ed io più volte abbiam provato</l>

<l>Come '1 nostro sperar torne fallace ;</l>

<l>Dentr' a quel sommo ben che mai non spiace</l>

<l> Levate '1 core a più felice stato.</l>

<note resp="editor">Puisque vous et moi bien des fois avons éprouvé - combien notre espoir se montre décevant, - en attendant ce bien suprême qui jamais n'est déplaisant, - élevez votre coeur vers une condition plus heureuse.</note>

</lg>

Je mets des tirets comme pour l'édition papier. Mais pour l'édition numérique, il serait peut-être plus judicieux de mettre des <lb/>

C'est une note de l'éditeur papier, j'ajoute un attribut resp pour le mentionner.

*125 Citation de lettre dans la lettre.

Comment la rendre ?

Je pense mettre <floatingText> mais cela rend l'XML invalide.

Je mets alors <quote>

Voir si on garde ce balisage.

<quote>" Votre lettre et les journaux m'ont fait grand plaisir, -en me rassurant sur la santé de votre illustre ami, cet homme que je regarde comme étant incontestablement le premier écrivain de l'époque; ils ne m'ont pas profondément affligé en me confirmant le décès de Georges Canning. J'étais tout consolé par anticipation, et les Débats n'ont pas pu parvenir à me faire partager leur douleur, pas plus que M" Dupin à m'empêcher de hausser les épaules de sa pitoyable proposition de médaille. Si la grande ombre du chevalier d'Harmensen revenait, ne pensez-vous pas qu'elle réclamerait une potence? "</quote>

*126 Pas de traduction disponible

<l>Ahi, null'altro che pianto al mondo dura !</l>

*127 Exemple de note mal matchée : j au lieu de 1 : donc non effacé par regex python

<p>j Il est évident que cette dénomination singulière s'appliquait à Mgr Lambruschini, le nonce du Pape, et que Lamennais parle ici du prélat par lequel Mgr Lambruschini avait été remplacé.</p>

*128 Problème : la note explique les étoiles

Je crois voir le **, le *** et le ****,

Je choisis donc de laisser une note pour la compréhension. Sinon cela risque d'être compliqué pour le lecteur.

Je crois voir le **, le *** et le ****,<note resp="editor">

>Les noms supprimés sont ceux des personnes engagées dans les malheureuses affaires où Lamennais perdait le plus clair de sa fortune. Il était inutile, pour le moins, de laisser sur ces noms le fâcheux reflet d'un ressentiment passager, puisque, d'ailleurs, il ne s'agit ici que d'intérêts purement privés.</note> avec des intérêts divers,

C'est une note de l'éditeur papier, j'ajoute un attribut resp pour le mentionner.

* 14

Je me rends compte (un peu tard!) que quand date et lieux sont divisés et mis au sein de deux balises différentes, la virgule qui lui le lieu et précède la date est supprimée.

<placeName>A la Chenaie</placeName>

<date>le 10 mai 1822.</date>

Désormais je l'ajoute manuellement, mais cela n'avait pas été fait pour les précédentes lettres. Première lettre signée par un F. !

*153 Importance des italiques : comment les rendre ?

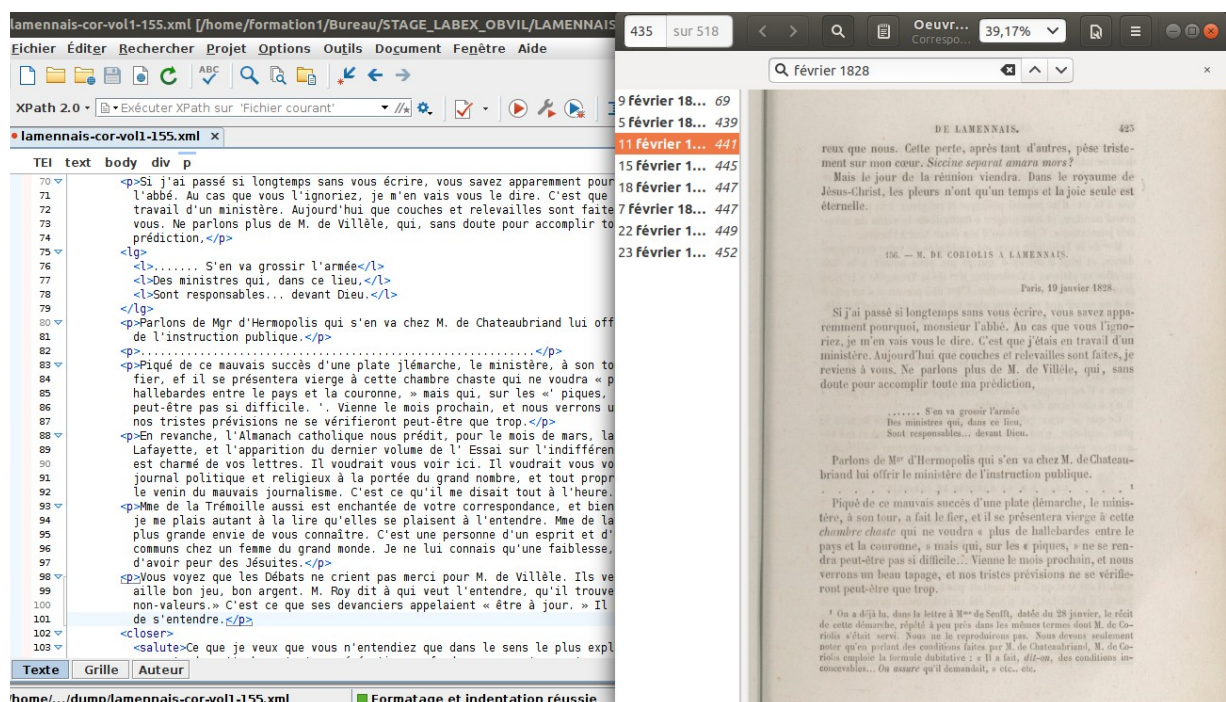
*156 Les 3 sont souvent pris pour des 5

<date when="le 51 janvier 1828."/>

Normalisation : <date when="1828-01-31"/>

*156b Je nomme ainsi un fichier XML qui est une réponse de Coriolis à Lamennais en note. Pour l'instant, j'ai traité les lettres publiées dans les notes comme des lettres à part entière, au même titre que celles de Lamennais. Voir si à long terme on les supprime ou si elle nous intéressent.

*155 Comment traiter et interpréter la suite (ligne) de points faite par l'éditeur ? Est-ce pour montrer une coupure ou est-ce ainsi chez Lamennais.



*161 Guillemets. Italiques : comment rendre la citation ?

Je mets en note la traduction.

On s'effraye de la moindre action, on tremble de

remuer, on sacrifie l'avenir à la peur qu'on a de la plus légère résistance immédiate. « E

natura comune degli uornini, dit Guichardin, temere prima i pericoli piit vicini, e stirnare piit che non conviene le cose pre senti, e tenere minor conto che non si debbe delle future, è lontane, perche a quelle si possano sperate molli remedj, dagli accidenti e dal tempo. » <note resp="editor">"Il est ordinaire aux hommes de craindre en première ligne les périls les plus proches, d'attacher aux choses présentes plus d'importance qu'elles n'en méritent, moins, en revanche, aux chances de l'avenir, surtout éloigné, parce qu'à celles-ci on peut espérer bien des remèdes inattendus, fruits du hasard et du temps."</note>Un des

Il faut encore corriger les fichiers :

164 à 195

2 à 9

17 à 99

J'ai fait 10 colonnes de 8 fichiers donc 80 fichiers

Il reste 15 colonnes de 8 fichiers et 1 col. de 4 donc 124 fichiers.