

Documentation of Modern Group, TranscriboQuest 2024 - Lyon, France

Team: *Béatrice Chaix Rouchon, Serena Crespi, Stéphanie Dord-Crouslé, Emmanuelle Morlock, David Rabouin, Arilès Remaki, Alyx Taounza-Jeminet*

Team Leader: *Simon Gabay*

The repository the group uploaded on HTR United hosts the HTR ground truth created as part of the TranscriboQuest training in 2024, Lyon, France. The dataset contains information on modern manuscripts in Italian, French, and Latin, dating from the 17th to the 19th century. These manuscripts include working drafts, diaries, notebooks, and preparatory works. The manuscripts in this dataset exhibit a wide variety of handwriting styles, differing significantly in both the type of ink used and the time periods in which they were produced, as well as the individual hands that created them.

The transcriptions were manually completed in two rounds by a group of seven contributors from various disciplinary backgrounds (engineers, PhD students, researchers) and different universities across France.

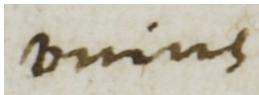
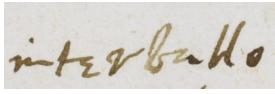
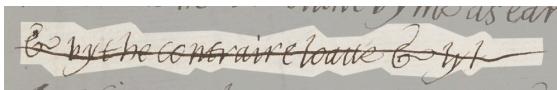
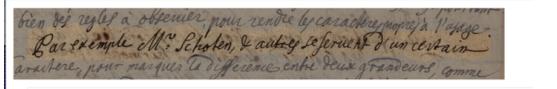
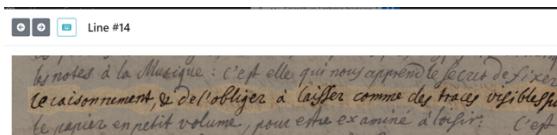
Rules we adopted for the transcription

The working group responsible for the modern manuscript dataset has chosen to detail the decisions made throughout the training and transcription of the various manuscript witnesses. These choices were the result of collaborative discussions aimed at identifying the most effective solutions. Any unresolved challenges, for which no satisfactory solution could be agreed upon, have been carefully documented for further review.

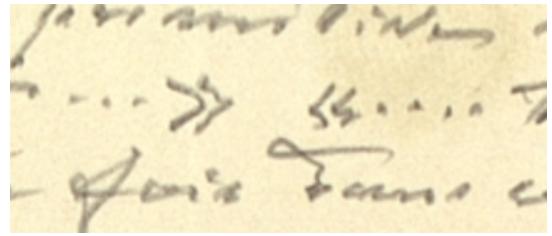
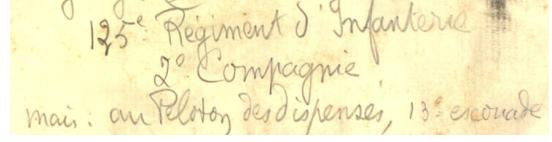
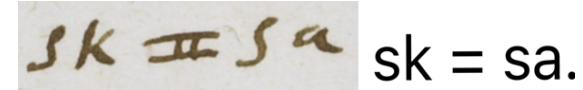
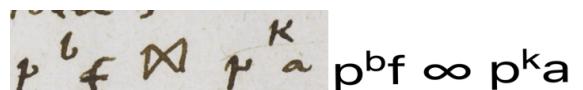
See also the DEVELOPED RULES AND EXAMPLES.

Synthesis of the choices and difficulties

| Case | Rule explanation | Image |
|-------------|---|--|
| V/v and U/u | <p><i>V and v are normalized to U and u, respectively.</i></p> <p><i>U and u are normalized to V and v,</i></p> |  |

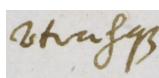
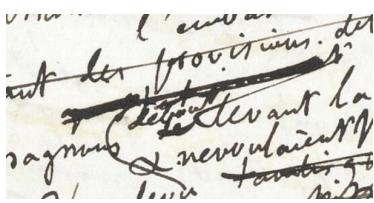
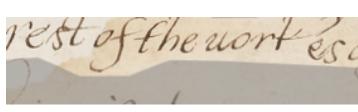
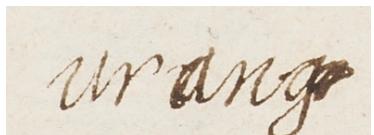
| | | |
|-------------------------------|--|--|
| | <i>respectively.</i> |   |
| Strikethrough | <i>Strikethroughs present in the manuscript are faithfully reproduced within double square brackets [[]].</i> |  [[& by the contraire loave & lyt]]  |
| Illegible word | <i>When a word is illegible, include what can be understood, and place the rest between [[]].</i> |  Una Cotorrice Un[[]] Il Ca'e DoAo Si Desinò In Camp^o Ci Vennen Il Sign^r Prior by crespi_serena (description) on Thu Sep 12 2024 14:58:49 GMT+0100 |
| Symbols & and α | <i>The symbols "&" and "α" are preserved as they appear in the transcription.</i> |  α α Line #20  Par exemple M. Schoten, & autres se servent d'un certain avantage, pour marquer la difference entre deux grandeurs comme by Rabouin (import) on Fri Sep 13 2024 09:30:39 GMT+0100 |
| "long s" and "round s" | <i>No distinction is made between forms, for example, between the long "s" and the round "s".</i> |  le notes à la Musique : c'est elle qui nous apprend le fait de fixer ce raisonnement, & de l'obliger à laisser comme des traces visibles sur le papier en petit volume, pour être examiné à loisir. C'est le raisonnement, & de l'obliger à laisser comme des traces visibles sur by Rabouin (import) on Fri Sep 13 2024 09:30:39 GMT+0100 |
| « i » and « long i » | <i>There is no distinction between « i » and « long i »</i> |  |

| | | |
|---|--|--|
| Capital letters | <i>Capital letters are preserved, even if they do not appear at the beginning of a sentence.</i> | <p>Lun Una Lepre Il Cav'e Ignatio Segue La Pioggia</p> <p>by crespi_teresa (eScriptorium) on Thu Sep 12 2024 13:08:46 GMT+0100</p> |
| Series of points | <i>Addition of as many points in the transcription as are visible in the original document</i> | |
| Change in handwriting | <i>When the handwriting changes in the manuscript, only the main hand is transcribed, and the secondary hand is not taken into account.</i> | |
| (/) | <i>Slashes (/) are preserved as they appear in the transcription.</i> | <p>rest of the uorkes</p> |
| Parenthesis encompassing several lines | <i>Parenthesis encompassing several lines are only integrated to the first line if they are not specifically aligned with the first and last line.</i> | |

| | | |
|------------------------------|--|--|
| Quotation mark | <i>There is no space after an opening quotation mark or after a closing quotation mark.</i> |  |
| Abbreviations | <i>Textual superscripts for textual abbreviations (ordinals, Saint in St, etc.) preceded by a ^ sign</i> |  |
| Symbols | <i>For specific symbols, we decided to transcribe it by the actual equivalent when it was available.</i> |  |
| Mathematical exponent | <i>For mathematical exponent, we used Unicode exponent characters.</i> |  |

Difficulties

| Cas | | |
|-----|-------------------------------------|--|
| | <i>We decided to transcribe the</i> |  |

| | | |
|--|--|--|
| Ligatures | <i>ligature when it was available (first image) but we developed it when it was not (second image)</i> |  utrasque |
| Undecipherable crossed words or sentences (ratures in French) | <i>As the piece of advice was to put blank spaces for each character between double square brackets [] (Ariane Pinche), it's difficult to decide the numbers of blanks space to insert when you can't count them. See recommendation for "Strikethrough"</i> |  |
| u for w (esp. in English) | <i>We kept the u when used as a w (it is too difficult to choose whether w works as a semi-vowel or a semi-consonant)</i> |  rest of the uordes rest of the uorkes  vrang (for wrong) |

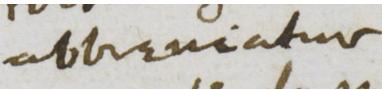
DEVELOPED RULES AND EXAMPLES

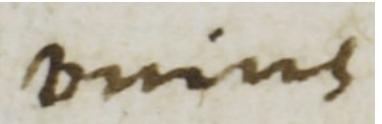
Summary of the highly intellectual discussions held by the team of modernists, who, of course, worked on this project with great passion.

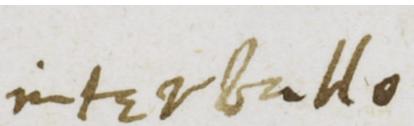
- Flaubert - use of alpha for "et" ("and"). We transcribe by the ampersand "&"
Example : page 10 (folio Nouvelles Acquisitions Françaises) 17599 f°8r = <https://gallica.bnf.fr/ark:/12148/btv1b60000866/f22.item>



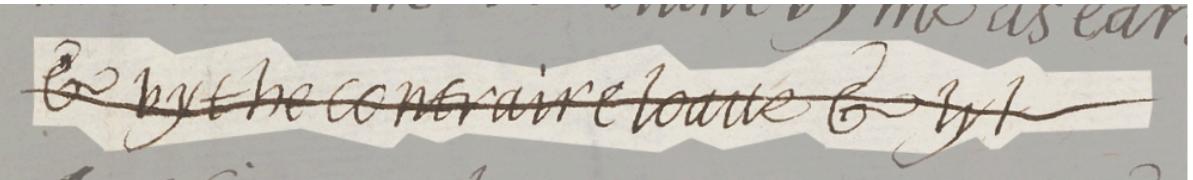
2. Normalization of the V/v and U/u for the XVIIth Century hands

 **abbreviatur**

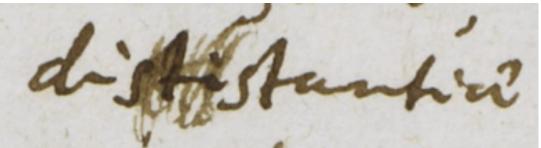
 **unius**

 **intervallo**

3. We transcribe the crossed passages into double brackets :

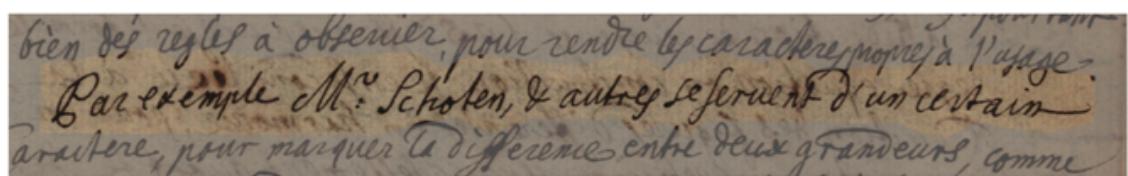


[[& by the contraire loave & lyt]]

 **di [[sti]] stantiae**

4. We keep the ampersand “&”

 Line #20



Par exemple M. Schoten, & autres se servent d'un certain

by Rabouin (import) on Fri Sep 13 2024 09:30:39 GMT+0100

5. No distinction between long and rond “S”

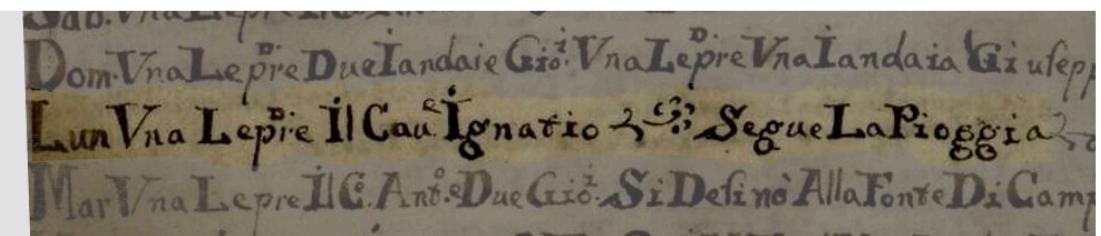
Line #14

les notes à la Musique : c'est elle qui nous apprend le secret de fixer
le raisonnement, & de l'obliger à laisser comme des traces visibles sur
le papier en petit volume, pour être examiné à loisir: C'est

le raisonnement, & de l'obliger à laisser comme des traces visibles sur

by Rabouin (imnort) on Fri Sep 13 2024 09:30:39 GMT+0100

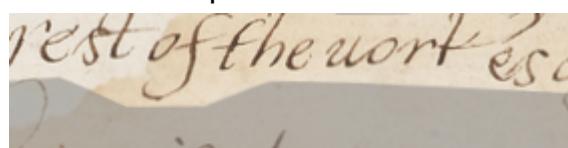
6. We keep capital letters, even if they are not at the beginning of the sentences



Lun Una Lepre Il Cav^e Ignatio Segue La Pioggia

by crespi_serena (eScriptorium) on Thu Sep 12 2024 13:58:46 GMT+0100

7. We keep u for w when it is in the manuscript:



rest of the uorkes

8. We keep the slash inside sentences

ande/vritt yea/sk

9. When the word is not readable, we put what we succeed reading and the remaining part into double square brackets []

Line #30

x

Una Cotornice Un[] II Ca^e Do^o Si Desinó In Camp^o Ci Venner II Sign^r Prior

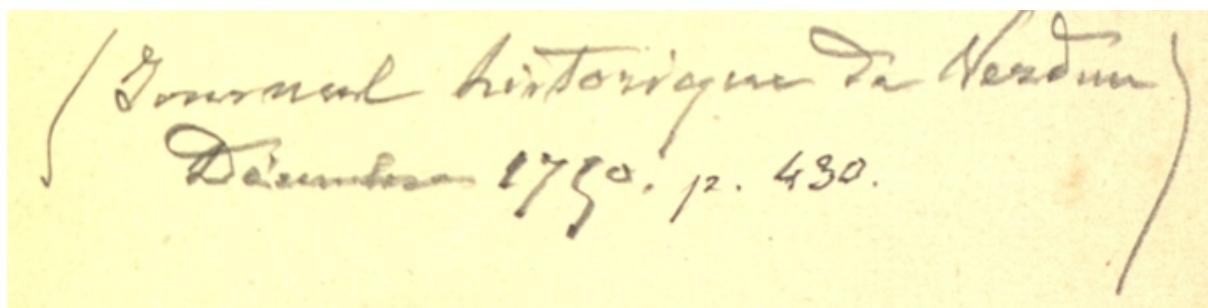
by crespi_serena (eScriptorium) on Thu Sep 12 2024 14:56:49 GMT+0100

10. When there is a changing of hands, we did not transcribe the second hand if it was not related to our corpus:

11. As my dots in the transcription as there are visible dots in the original document

Transcription : pillaient..... (Ancy

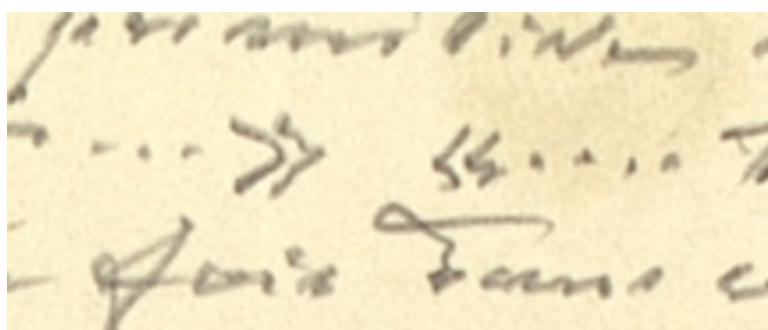
12. Parenthesis encompassing several lines are only integrated to the first line if they are not specifically aligned with the first and last line.



Transcription:

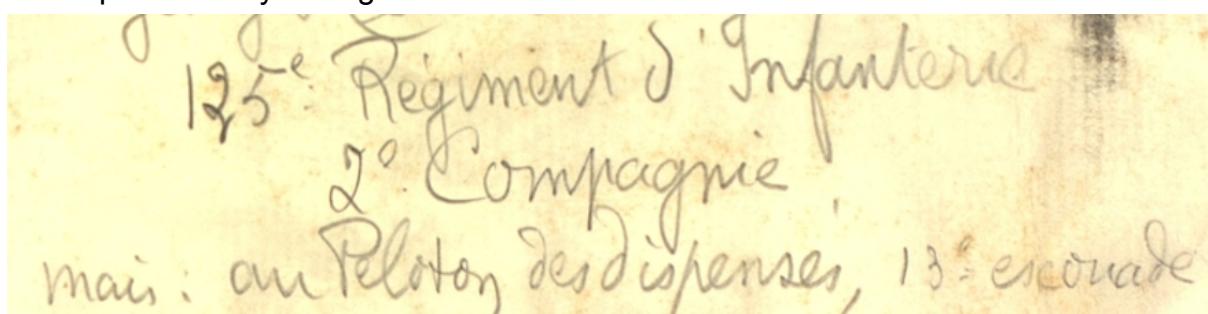
(Journal historique de Verdun)

Décembre 1750. p.430.



13. Pas d'espace après un guillemet ouvrant ni après un guillemet fermant

14. Textual superscripts for textual abbreviations (ordinals, Saint in St, etc.) are preceded by a ^ sign

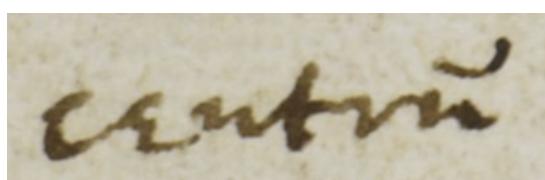


Transcription: 125^e Régiment d'Infanterie

2^e Compagnie

mais : au Peloton des dispensés, 13^e escouade

15. We decided to transcribe the ligature when it was available :

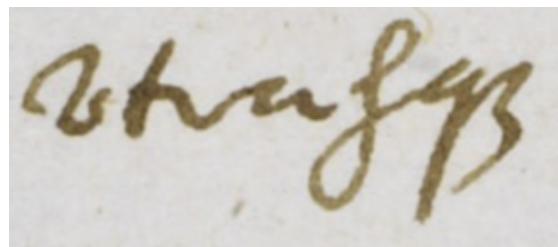


centrū

But we developed it when it was not :

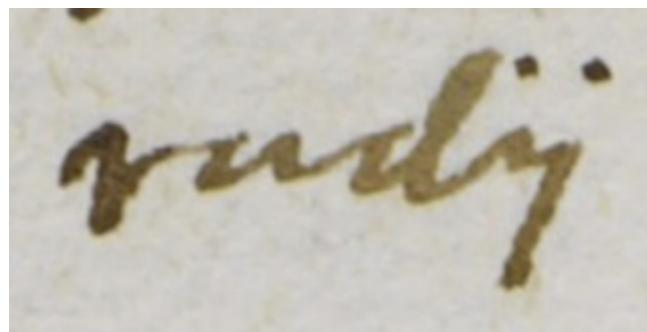
Documentation GroupModern TranscriboQuest 2024 - Lyon, France

S.C. Crespi ; B.C. Rouchon ; E. Morlock ; D. Rabouin ; A. Remaki ; A. Taounza-Jeminet



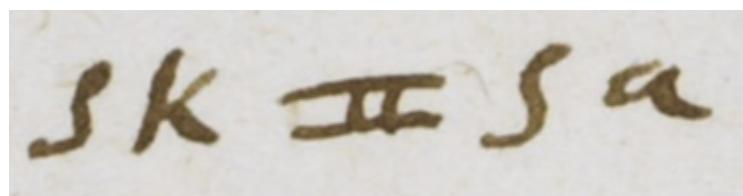
utrasque

16. We did not make distinction between « i » and « long i » :



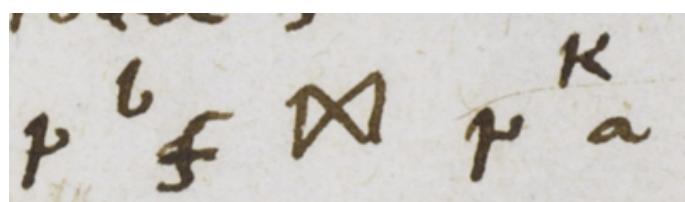
radii

17. For specific symbols, we decided to transcribe it by the actual equivalent when it was available. For example : the equal sign



sk = sa.

18. For mathematical exponent, we used Unicode exponent characters :



p^bf ∞ p^ka

ANNEXE

Here are presented the manuscripts and the corresponding folios that were used for the transcription process.

Flaubert Manuscript :

page 1 : Bibliothèque Historique de la Ville de Paris (BHVP). Rés. Ms 98 f°1r =
<https://gallica.bnf.fr/ark:/12148/btv1b105370062/f17.item>

page 2 : BHVP. Rés. Ms 99 f°1r =
<https://gallica.bnf.fr/ark:/12148/btv1b10537007h/f11.item>
manuscrit de la main du copiste (pas de Flaubert donc !)

page 9 : Nouvelles Acquisitions Françaises (NAF) 17599 f°7r =
<https://gallica.bnf.fr/ark:/12148/btv1b60000866/f20.item>

page 10 : NAF 17599 f°8r = <https://gallica.bnf.fr/ark:/12148/btv1b60000866/f22.item>

Leibniz Manuscript :

page 1 : Leibniz-Handschriften zur Philosophie LH 4, 5, 10, fol. 11r
https://dfg-viewer.de/show?id=9&tx_dlf%5Bid%5D=https%3A%2F%2Fdigitale-sammlungen.gwlb.de%2Fcontent%2F00069349%2F00069349.xml&tx_dlf%5Bpage%5D=21

page 2 : Leibniz-Handschriften zur Philosophie LH 4, 5, 10, fol. 11v
https://dfg-viewer.de/show?tx_dlf%5Bdouble%5D=0&tx_dlf%5Bid%5D=https%3A%2F%2Fdigitale-sammlungen.gwlb.de%2Fcontent%2F00069349%2F00069349.xml&tx_dlf%5Bpage%5D=22&cHash=d54f9a0c7bb757249e2d8bf446629b83

page 3 : Leibniz-Handschriften zur Philosophie LH 4, 5, 10, fol. 15r
https://dfg-viewer.de/show?id=9&tx_dlf%5Bid%5D=https%3A%2F%2Fdigitale-sammlungen.gwlb.de%2Fcontent%2F00069349%2F00069349.xml&tx_dlf%5Bpage%5D=29

page 4 : Leibniz-Handschriften zur Philosophie LH 4, 5, 10, fol. 15v
https://dfg-viewer.de/show?tx_dlf%5Bdouble%5D=0&tx_dlf%5Bid%5D=https%3A%2F%2Fdigitale-sammlungen.gwlb.de%2Fcontent%2F00069349%2F00069349.xml&tx_dlf%5Bpage%5D=30&cHash=d6f1090afb3700cdfd5b0651bcfce20b

Harriot Manuscript :

page 1 : Harriot's paper : Ms. 6784, Fol.200
<https://echo.mpiwg-berlin.mpg.de/ECHOdocuView?tocMode=thumbs&url=%2Fmpiwg%2Fonline%2Fpermanent%2Flibrary%2FXT0KZ8QC%2F&viewMode=image&tocPN=1&searchPN=1&characterNormalization=req&query=&pn=399&queryType=&start=390>

page 2 : Harriot's paper : Ms. 6784, Fol.198

<https://echo.mpiwg-berlin.mpg.de/ECHOdocuView?tocMode=thumbs&start=351&url=/mpiwg/online/permanent/library/XT0KZ8QC/&viewMode=image&tocPN=1&searchPN=1&characterNormalization=reg&query=&pn=395&queryType=>

page 3 : Harriot's paper : Ms. 6784, Fol.187

<https://echo.mpiwg-berlin.mpg.de/ECHOdocuView?tocMode=thumbs&start=351&url=/mpiwg/online/permanent/library/XT0KZ8QC/&viewMode=image&tocPN=1&searchPN=1&characterNormalization=reg&query=&pn=373&queryType=>

English Manuscript

page 1 : Poetical and prose works of king James I of England. MS. Bodl. 165, fol. 16r

<https://digital.bodleian.ox.ac.uk/objects/7f4fc1a8-99dd-4d0d-ac66-00948f98af21/surfaces/15065ba1-c943-41c2-a2ab-78688a26ac99/>

1690

page 2 : MS. Bodl. 165, fol. 16v

<https://digital.bodleian.ox.ac.uk/objects/7f4fc1a8-99dd-4d0d-ac66-00948f98af21/surfaces/7498680b-3e07-4136-b894-d0392061a564/>

page 3 : MS. Bodl. 165, fol. 17r

<https://digital.bodleian.ox.ac.uk/objects/7f4fc1a8-99dd-4d0d-ac66-00948f98af21/surfaces/2532d4f0-5f7b-4867-9d10-b927a463ba4c/>

page 4 : MS. Bodl. 165, fol. 18v

<https://digital.bodleian.ox.ac.uk/objects/7f4fc1a8-99dd-4d0d-ac66-00948f98af21/surfaces/1740657c-b145-44a6-907f-80b99ad2c9e1/>

page 5 : MS. Bodl. 165, fol. 19r

<https://digital.bodleian.ox.ac.uk/objects/7f4fc1a8-99dd-4d0d-ac66-00948f98af21/surfaces/f6a89700-b73d-4dd9-8bf8-ba7c830a2687/>

Italian Manuscript

page 1: Ricordi di villeggiatura di caccia e d'altro dal 1 ottobre 1665 all'11 settembre, Ms Rossi-Cassigoli 380, BNCF, Florence, fol. 51

<https://teca.bncf.firenze.sbn.it/ImageViewer/servlet/ImageViewer?idr=BNCF0003478716#>

page 2: Ricordi di villeggiatura di caccia e d'altro dal 1 ottobre 1665 all'11 settembre, Ms Rossi-Cassigoli 380, BNCF, Florence, fol. 52

page 3: Ricordi di villeggiatura di caccia e d'altro dal 1 ottobre 1665 all'11 settembre, Ms Rossi-Cassigoli 380, BNCF, Florence, fol. 53

page 4: Ricordi di villeggiatura di caccia e d'altro dal 1 ottobre 1665 all'11 settembre, Ms Rossi-Cassigoli 380, BNCF, Florence, fol. 54

page 5: Ricordi di villeggiatura di caccia e d'altro dal 1 ottobre 1665 all'11 settembre, Ms Rossi-Cassigoli 380, BNCF, Florence, fol. 55

Documentation GroupModern TranscriboQuest 2024 - Lyon, France

S.C. Crespi ; B.C. Rouchon ; E. Morlock ; D. Rabouin ; A. Remaki ; A. Taounza-Jeminet

Fonds du Père de la Croix :

<https://archives-deux-sevres-vienne.fr/ark:/28387/vta86206f35fbd43d12>

Copies de documents de 1812 à 1890, documents originaux (1853-1856), fichier de notes historiques de 1807 à 1902, notes historiques et bibliographiques.

Pages 5, 6, 7, 10, 11, 12, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32