

Editionsmodell

Teil II Textkonstitution

Inhalt

1. Editionsprinzipien
2. Darstellung philologischer Befunde durch typographische Auszeichnungen in integraler Wiedergabe
3. Darstellungskonventionen
 - 3.1 Materialität der Schriftträger
 - 3.2 Positionierungs- und Grössenangaben
 - 3.3 Mehrschriftigkeit und Schriftschnitte
 - 3.4 Einfügungen und Ersetzungen, Umstellungen, Tilgungen und Textverluste
 - 3.5 Nichtedierte Zeichen und Bspurungen der Schriftträger
4. Transkriptions- und Transliterationsregeln
 - 4.1 Autorspezifische Orthographie und Satz-/Druckfehler
 - 4.2 Gross-/Kleinschreibung und Buchstabenverwechslungen
 - 4.3 Endsilbenverschleifungen und Suspensionen
 - 4.4 Abkürzungen, Kontraktionen und <ℓ>/<ℓ̃>-Suspensionen
 - 4.5 Geminations- und Nasalstriche
 - 4.6 Akzente, Diakritika, Superscript-Schreibungen und Apostrophierungen
 - 4.7 Ligaturen
 - 4.8 Sonderzeichen, Währungsangaben und Masseinheiten
 - 4.9 Wiedergabe von <ſ>/<f>/<ß>
 - 4.10 Interpunktion und Zeichenabstände
 - 4.11 Viertelgeviert-, Halbgeviert- und Geviertstriche

1. Editionsprinzipien

Zeichen und Zeichenstand werden vorlagengetreu ediert, es gibt bei der Wiedergabe der Orthographie, Interpunktion und zeitgenössischer Schreibungen von Eigen- oder Ortsnamen keine Korrekturen. Die Dokumente werden dabei so einfach und intuitiv wie möglich wiedergegeben, Durchstreichungen (auch mehrmalige) durch durchstrichene Grapheme, Überschreibungen und spätere (besonders interlineare) Einfügungen durch hochgestellte Grapheme angezeigt. Die Edition folgt damit dem Prinzip, die edierten Quellen nicht mit einem komplizierten Zeichen-Apparat zu ummanteln und zuzustellen. Sie ist diplomatisch im Anspruch, alle bedeutungstragenden Graphemvorkommnisse exakt zu transkribieren, nicht aber hyperdiplomatisch darin, deren Positionierungen auf den Schriftträgern und deren Grössenverhältnisse zueinander mimetisch abbilden zu wollen. Die Textanordnung wird linearisiert und XML-konform arrangiert, was validere Suchergebnisse und bessere Datenbankperformances verspricht, aber auch einfachere Nachnutzung der Daten. Hieraus resultiert, dass manche Elemente der zu edierenden Dokumente nicht wiedergegeben werden (etwa Graphempositionen auf der Seite), andere in den XML-Dateien kodiert, in der Transkriptionsansicht nicht sichtbar werden (etwa Zeilenumbrüche), die wesentlichen aber XML-kodiert sowie im Frontend mit den Faksimiles synoptisch dargestellt werden sollen.

2. Darstellung philologischer Befunde durch typographische Auszeichnungen in integraler Wiedergabe

Die integrale Wiedergabe der edierten Dokumente und Materialien mit einem Minimum an Sonderzeichen und zuvor zu definierenden (und entsprechend zu berücksichtigenden) speziellen editorischen Zeichen und -verwendungen sieht die folgenden typographischen Auszeichnungen und Darstellungskonventionen vor. Dies ermöglicht für die handschriftlichen sowie für die gedruckten Quellen (Briefe und Rezensionen) ein einheitliches Editionsmodell:

Grundschrift recte Schriftwechsel	Autorschrift/Grundschrift des gedruckten Dokuments Schriftwechsel ohne weitere Spezifizierung: <hi rendition="#Schriftwechsel">anderssschriftige(s) Zeichen</hi>
ABC	Graphem(e) in Versalien (ohne spezielle Auszeichnung, Darstellung in Grossbuchstaben)
ABC	Graphem(e) in Kapitälchen: <hi rendition="#sc">Kapitälchen</hi>
ABCabc	fett gesetzte(s) Graphem(e): <hi rendition="#b">fette(s) Zeichen </hi>
<i>ABCabc</i>	kursiv gesetzte(s) Graphem(e): <hi rendition="#i">kursive(s) Zeichen</hi>
A B C a b c	gesperrt gesetzte(s) Graphem(e): <hi rendition="#sp">gesperrte/s Zeichen</hi>
ABCabc	höherpositionierte(s) Graphem(e) aus geringerem Schriftgrad als Grundschrift: <hi rendition="#sup">hochgestellte(s) Zeichen</hi>
ABCabc	tieferpositionierte(s) Graphem(e) aus geringerem Schriftgrad als Grundschrift: <hi rendition="#sub">tiefgestellte(s) Zeichen</hi>
<u>ABCabc</u>	ein-, zwei- oder mehrfach unterstrichene(s) Graphem(e): <hi rendition="#u">unterstrichene(s) Zeichen</hi>
ABCabc	ein-, zwei- oder mehrfach durchstrichene(s) leserliche(s) Graphem(e): <hi rendition="#lt">durchstrichene(s) Zeichen</hi>
XXXxxx	getilgte(s) unleserliche(s) Graphem(e); zusätzlich kann textkonstitutionelle Anmerkung erfolgen (<note type="annotText" n="1*"><p>getilgte(s) unleserliche(s) Zeichen zwischen ABCabc und ABCabc </p></note>)
XXXxxx	unleserliche(s) Graphem(e); zusätzlich kann textkonstitutionelle Anmerkung erfolgen (<note type="annotText" n="1*"><p>unleserliche(s) Zeichen zwischen ABCabc und ABCabc</p></note>)
ABCabc	Einfügung/Addition: Einfügung zwischen Graphem(e) oder Zeilen, am Rand oder auf der Rückseite durch höherpositionierte(s) Graphem(e) aus geringerem Schriftgrad als Grundschrift; zusätzlich kann textkonstitutionelle Anmerkung erfolgen (<note type="annotText" n="1*"><p>eingefügte(s) Zeichen ABCabc zwischen ABCabc und ABCabc</p></note>)
ABCabc ABCabc	Ersetzung/Substitution in der Zeile: durchstrichene(s) Graphem(e) durch Graphem(e) aus Grundschrift; zusätzlich kann textkonstitutionelle Anmerkung erfolgen (<note type="annotText" n="1*"><p>ersetzte(s) Zeichen ABCabc durch ABCabc zwischen ABCabc und ABCabc</p></note>)
ABCabc ABCabc	Ersetzung/Substitution: durchstrichene(s) Graphem(e) durch höherpositionierte(s) Graphem(e) aus geringerem Schriftgrad als Grundschrift; zusätzlich kann textkonstitutionelle Anmerkung erfolgen (<note type="annotText" n="1*"><p>ersetzte(s) Zeichen ABCabc durch ABCabc zwischen ABCabc und ABCabc</p></note>)
ABCabc ^{ABCabc} ABCabc	Umstellung/Permutation: tieferpositionierte(s) Graphem(e) durch höherpositionierte(s) Graphem(e) aus geringerem Schriftgrad als Grundschrift; zusätzlich kann textkonstitutionelle Anmerkung erfolgen (<note type="annotText" n="1*"><p>umgestellte(s) Zeichen ABCabc durch ABCabc zwischen ABCabc und ABCabc</p></note>)
[~]	Herausgeberzusatz, der mit UNICODE nicht wiedergebbare/s Sonderzeichen oder speziellen Textbefund anzeigt; zusätzlich kann textkonstitutionelle Anmerkung erfolgen (<note type="annotText" n="1*"><p>Anmerkung zur Textkonstitution</p></note>)

3. Darstellungskonventionen

3.1 Materialität der Schriftträger

Die materialen Besonderheiten der verwendeten Schriftträger (Hader- oder Schreibpapier, Wasserzeichen etc.) oder auch Poststempel und andere spätere Bespurungen werden in aller Regel nicht vermerkt, ebensowenig Faltungen und Gebrauchsspuren oder auch Schreibgeräte und Tinten. Die Briefe werden komplett so faksimiliert, wie sie sich aktuell (nachdem Haller selbst sie einbinden liess und sie später wieder ausgelöst und neu gruppiert wurden) präsentieren. Wie sich der heutige Bestand zeigt, ist daher abhängig von bereits durch Haller (und hierauf) vorgenommenen Veränderungen. Nicht angegeben wird das Format, da ohnehin kaum ein Brief ohne Beschnitt überliefert ist. Von Farbänderungen abgesehen gibt es kaum Varianz bei Tinten zu verzeichnen, die Farbdigitalisate werden ohne Mass- und Farbeinheiten wiedergegeben, um bei einfachem und unkompliziertem

Zugang keine weiteren Schwellen zu errichten und auch zu ihrer Wahrnehmung als lesbare Dokumente und Medien der Kommunikation einzuladen. Sie werden daher vollständig mit Vorder- und Rückseiten abgebildet, sogar für den Fall, dass diese unbeschrieben geblieben sind. Auch gedruckte Quellen werden ohne weitere typographische, druckanalytische und quantifizierende Zusätze und ohne Bemassungen und Farbangaben faksimiliert.

3.2 Positionierungs- und Grössenangaben

Auf Positionierungen von Graphemen relational zur Seite (links/zentriert/rechts sowie oben/mittig/unten) wird ebenso verzichtet wie auf die Dokumentation der Abstände zwischen Zeilen und Zeichengruppierungen in der Zeile; interlineare Abstände sind daher immer einzeilig, als Standardwortabstand und Leerzeichen findet ausschliesslich Spatium/ Leerschlag Verwendung (U+0020). Auch Einzüge und Einrückungen werden nicht nachgebildet. Damit sollen nicht zuletzt White Space-Probleme bei einer Nachnutzung der Daten vermieden werden. Zeilenanfänge nach Punkten und Absätzen (auch bei fehlender Punktierung) werden nicht ergänzt. Nicht markiert werden Grössenunterscheidungen von Schrift (dabei auch Initialschreibungen oder Schreibungen einzelner Buchstaben). Entsprechend werden auch bei den gedruckten Texten die teilweise erheblichen Schriftgradunterschiede, die funktionale Differenzierungen markieren, in der Transkription nicht nachgeahmt. Als wichtigeres Ziel erschien es, eine Textausgabe zu erhalten, die Zeichen und Zeichenstand der Vorlage gut lesbar und komfortabel in allen Ausgabeeinstellungen auf unterschiedlichen elektronischen Geräten und Displays wiederzugeben erlaubt.

Bei Wiedergabe gedruckter Materialien werden Seitenzahlen, Kolumnentitel, Kustoden, Druckersignaturen und dergleichen nicht mitediert; bei den Briefen hingegen alles, was zum Dokument zählte, als es vom Schreibenden aus der Hand gegeben bzw. aufgegeben und speditiert wurde. Die Edition verzichtet auf Herausgebereingriffe so weit als möglich. Entsprechend bleiben auch Eigenheiten bestimmter Briefe und ihrer Verfasser (etwa handschriftliche Kustoden oder Dittographien) erhalten. Die Markierung von Zeilen- und Seitenumbrüchen sowie von Absätzen durch entsprechende Diakritika in der Transkription wird durch die synoptische Darstellung redundant; sie werden auf XML-Ebene aber codiert (<lb/>, <pb/> und <p/>) und bleiben bei Exporten erhalten (siehe Editionsmodell I). Wo klärende philologische Zusätze nötig sind, erscheinen diese in eckigen Klammern und werden in der XML-Notierung durch <note type="annotText"> vermerkt (dort auch mit Auflösungen von Abkürzungen und anderen Erläuterungen, wo dies angeraten sein sollte).

3.3 Mehrschriftigkeit und Schriftschnitte

Schriftwechsel und typographische Abweichungen von der Standardschrift werden in der XML-Datei vermerkt und im Frontend sichtbar gemacht. Wo im Unterschied zur sonst vorherrschend verwendeten Schrift Schriftwechsel anzutreffen ist, etwa im Umfeld deutscher Kurrent lateinische Buchstaben (auch bei Zitaten) vorkommen, werden die entsprechenden Grapheme markiert (<hi rendition="#Schriftwechsel">), es erfolgt eine Darstellung in leicht unterscheidbarer Schrift. Wo sich fremde Schreiberhände eindeutig erkennen lassen, werden solche Dritten zuweisbaren Grapheme durch den Hinweis «Fremde Hand» (<handNote>) gekennzeichnet (es gibt eine Anmerkung zur Textkonstitution <note type="annotText">); regulär werden spätere Hinzufügungen Dritter aber nicht mittranskribiert.

Alle kursiv ausgezeichneten Grapheme (<hi rendition="#i">) in Drucktexten werden als kursiviert gekennzeichnet, Fettmarkierung (<hi rendition="#b">) markiert Schwabacher im Umfeld deutscher Schrift bzw. besonders auffälligen Schriftwechsel in Antiqua im Drucktext. Zusätzlich können so aber auch stark markierte spätere Zusätze, auffällig geschriebene Abweichungen von der Grundschrift durch betont breiten Federzug oder mehrfache Überschreibung etc. auch in handschriftlichen Dokumenten ausgewiesen werden. Da die GGA-Rezensionen allesamt aus Frakturen gesetzt wurden, indiziert ein Schrifttypenwechsel dort die Verwendung einer Antiquaschrift; Sperrungen werden gesperrt wiedergegeben (<hi rendition="#sp">), Vorkommnisse aus weiteren Schriften (etwa in der *Bibliothèque raisonnée*) ebenfalls durch Schriftwechsel. Griechische und hebräische Grapheme werden

griechisch und hebräisch wiedergegeben und nicht lateinisch transliteriert (Vorkommnisse aus anderen Schriftsystemen ebenfalls). Allerdings werden besondere handschriftliche Idiosynkrasien einzelner Schreibender wie Schriftmischungen (etwa lateinisches «e» im Umfeld deutscher Kurrent) nicht immer akkurat nachgebildet und gekennzeichnet.

3.4 Einfügungen und Ersetzungen, Umstellungen, Tilgungen und Textverluste

Einfügungen (Addition/Interpolation) und Ersetzungen (Substitution) werden durch Höherpositionierung und Durchstreichung besonders hervorgehoben; dies kann im Schreibprozess ein einzelnes sofort überschriebenes Graphem sein (etwa bei Sofortrevision in der Zeile: Durchstreichung und Grundschrift (<note type="annotText" n="1*"><p>ersetzte(s) Zeichen ABCabc durch ABCabc zwischen ABCabc und ABCabc</p></note>)), aber auch eine spätere Einfügung (meist Spätrevision zwischen bestehenden Zeilen, am Rand oder auf der Rückseite bzw. einer anderen Seite: Durchstreichung und Hochstellung (<note type="annotText" n="1*"><p>ersetzte(s) Zeichen ABCabc durch ABCabc zwischen ABCabc und ABCabc</p></note>)). Ebenso werden Hallers handschriftliche meist als Marginalien realisierte Annotationen in seinem Handexemplar der *GGA* (dort auch explizite Korrekturen im Text) transkribiert und als handschriftliche Ergänzungen ausgewiesen. Umstellungen (Permutation) werden durch eine Kombination aus tieferpositionierten und höherpositionierten Graphemen aus geringerem Schriftgrad als Grundschrift angezeigt; zusätzlich kann auch hier eine textkonstitutionelle Anmerkung erfolgen (<note type="annotText" n="1*"><p>umgestellte(s) Zeichen ABCabc durch ABCabc zwischen ABCabc und ABCabc</p></note>)).

Nachgeahmt werden einfache (wie mehrfache) Durchstreichungen durch einfache Durchstreichung, zumindest soweit noch leserlich ist, was durchstrichen wurde (<hi rendition="#it">). Tilgungsvarianten, die zu Unleserlichkeit führen wie Rasuren oder Schwärzungen, werden durch durchstrichene Gross-X und Klein-x (⌘ und ⌘) graphemweise dargestellt und können in den Annotationen (<note type="annotText">) explizit vermerkt werden, ebenso Textverluste infolge beschädigter Zeichenträger. Dabei wird bei Tilgungen (Deletion) das folgende Vokabular verwendet: «Ausschneidung», «Bleichung», «Expungierung», «Schwärzung», «Streichung», «Radierung», «Rasur» und «Überklebung»; bei vollständigen Textverlusten und Textverderbnissen «Papierausbruch», «Siegelauriss», «Tintenfrass», «Verbleichung» und «Wasserschaden». Unsichere Lesungen werden durch den Hinweis «Unsichere Lesung» entsprechend annotiert ((<note type="annotText">), im Frontend aber nicht zusätzlich markiert; dies gilt auch für alternative Lesungen (<choice>/«Alternativlesung»). Unleserliche Elemente, bei denen weder Tilgung noch Textverlust vorliegen, werden als Gross-X und Klein-x (⌘ und ⌘) graphemweise transkribiert und als «unleserlicher Text» annotiert ((<note type="annotText">)).

3.5 Nichtedierte Zeichen und Bespurungen der Schriftträger

Die vorliegende Edition ist eine genuin digitale (*born digital*), die von den diesbezüglichen Auszeichnungsmöglichkeiten und editionsphilologischen Möglichkeiten Gebrauch macht und entsprechende Recherchen und Suchmöglichkeiten, aber auch weitere Verwendungen eröffnet. Die Transkriptionen geben die Dokumente so wieder, wie sie heute vorzufinden sind. Dennoch werden Nutzungsspuren und spätere Zusätze wie Ordnungsziffern und archivalische Signaturen, die mitunter korrigiert und überschrieben wurden, aber auch zeitgenössische Zusätze, Siglen oder andere Vermerke regulär nicht mittranskribiert; sie sind im Fall der Korrespondenzen Teil des heutigen Dokuments, nicht aber des ursprünglichen Briefs, wie er vom Schreibenden aufgegeben wurde und werden (wenn überhaupt) lediglich in den Metadaten bzw. in der XML-Codierung im Feld <address> mit dem generischen Vermerk <PostallInfo> versehen. Vermerke von fremder Hand werden regulär nicht mittranskribiert. Eine Ausnahme bilden Hallers eigenhändige Marginalien und Korrekturen seines Handexemplars der *GGA*. Exakte und vollständige Beschreibungen der materialen Eigenschaften der Schriftträger (Wasserzeichen und Papierprovenienz, Faltungen und Faltpuren, Siegelreste und Poststempel) müssen späteren Ergänzungen und Forschungen vorbehalten bleiben.

4. Transkriptions- und Transliterationsregeln

Aus Darstellungsgründen orientieren sich die Transkriptionen am gängigen UNICODE-Schrifteninventar und verwenden – abgesehen von der eckigen Klammer (⟦/U+005B und ⟧/U+005D) und der Tilde (◌~/U+223C) als Platzhalter für mittels UNICODE nichtdarstellbare Zeichen – bewusst keine editionsphilologischen Sonderzeichen. Bei der Wiedergabe von Graphemen gelten die folgenden Vereinheitlichungen und Normierungen bezüglich Graphie, Orthographie und Interpunktion:

4.1 Autorspezifische Orthographie und Setz-/Druckfehler

Orthographische bzw. graphematische Besonderheiten bestimmter Korrespondierender beim Schreiben (etwa Weglassung von ◌i>-Punkten oder von ◌t>-Querstrichen, von Umlaut-Anzeichen, fehlende ◌u>-Striche oder -Bögen) werden grundsätzlich nicht nachgeahmt. Orthographiefehler hingegen bleiben beibehalten, ebenso Verschreibungen, fehlerhafte Schreibungen und andere Fehler; getreu den vorliegenden Dokumenten («vorlagengetreu»), nicht getreu einer historischen Person oder deren konstruierten Absichten und Intentionen. Daher werden auch intendierte Abkürzungen im transkribierten Text nicht aufgelöst, können im Kommentar aber, wo notwendig, erläutert werden (<note type="annotText">). Anders als beim Briefschreiben sind Drucktexte als arbeitsteilige Produkte anfälliger für Fehler in Satz und Druck. Solche nichtintendierten eigentlichen Fehler im typographischen Produktionsprozess sind nicht erhaltenswert und werden (ausgewiesen und dokumentiert) emendiert («Emendation»). Dazu zählen vertauschte oder falsch positionierte Lettern (etwa ◌u>/◌n> oder ◌d>/◌b> oder ◌p>/◌q>-Vertauschungen) ebenso wie Zwiebelische, falsche Abstände, Fehler bei Getrennt- und Gross-/Kleinschreibung. Kommentiert wird nur bei sinnentstellenden Fehlern.

4.2 Gross-/Kleinschreibung und Buchstabenverwechslungen

Gross- und kleingeschriebene Buchstaben ergeben sich speziell in deutschen Kurrentschriften häufig aus der Positionierung und sind mitunter kaum voneinander zu unterscheiden (etwa ◌D>/◌d>). Wo eindeutig lesbar, wird nach der Vorlage transkribiert. In Zweifelsfällen wird auf Briefebene vereinheitlicht, sonst nach den Massstäben des Autors, wo eine Tendenz zu erkennen ist; wo keine eindeutige Lesart möglich ist und unentschieden gewechselt wird, werden Eigennamen, Substantive und Werktitel gross geschrieben, ebenso bei neuen Sätzen und Absätzen gross eingesetzt. Ebenso wird bei ◌I>/◌J> und ◌i>/◌j>-Unentschiedenheiten vereinheitlicht, da auch ◌I> und ◌J> in deutscher Kurrent häufig nicht differenziert werden können; in aller Regel wird entsprechend des Lautwerts vor Vokalen ◌J> transkribiert, vor Konsonanten ◌I>. Allerdings werden eindeutige Vorkommnisse niemals redigiert, etwa ◌J> nicht zu ◌I> emendiert, wenn es eindeutig vorzufinden ist. Ebenfalls sind lateinische ◌u>/◌v>- und ◌U>/◌V>-Schreibungen oft schwer zu unterscheiden und werden, wo das nicht möglich ist, stillschweigend nach den Massstäben des Autors vereinheitlicht, sobald eine Tendenz zu erkennen ist; wo nicht, wird etwa ◌vna> nach modernen Usancen zu ◌una>, ◌reuulsio> zu ◌revulsio> vereinheitlicht. «Unklare Getrennt- und Zusammenschreibung» wird graphisch nicht angezeigt, kann aber gegebenenfalls so annotiert werden.

4.3 Endsilbenverschleifungen und Suspensionen

Endsilbenverschleifungen und Suspensionen werden in aller Regel stillschweigend aufgelöst und ausgeschrieben. Sie ergeben sich aus dem Briefschreiben als dynamischer, schneller und mitunter wenig akkurat ausgeübter Tätigkeit oft von selbst, etwa im Deutschen bei Wortendungen wie ◌-e>, ◌-em>, ◌-en>, ◌-er>, ◌-ung> usw. Sie werden daher als von den Schreibenden intendierte Ausschreibungen im edierten Text ohne Kennzeichnung realisiert wie eine unpräzise Schrift. Dazu zählen auch einer erkennbaren Systematik folgende regelmäßige Suspensionen bei Wortendungen (wie im Lateinischen bei ◌-que> oder ◌-tur>), sobald die fehlenden Wortelemente graphematisch deutlich realisiert wurden und sichtbar ist, dass an dieser Stelle etwas zu ergänzen ist (im Lateinischen etwa ◌at> zu ◌atque> oder ◌metiunt> zu ◌metiuntur>). Sie werden bei Bedarf in XML mit <abbr type="suspension"> kodiert und in diesem Fall im Frontend als kursivierte(s) Zeichen dargestellt.

4.4. Abkürzungen, Kontraktionen und <ℓ>/<ℓ̃>-Suspensionen

Abkürzungen bleiben allermeist bestehen, besonders durch Punktierung oder andere Zeichen wie Semikolon oder Doppelpunkt als intentionale und intendierte ausgewiesene. Auch bei Anreden und Grussformen werden sie in aller Regel so wiedergegeben, wie sie vorfindlich sind und ohne Auflösung im Rahmen der mittels UNICODE gegebenen Möglichkeiten zeichengetreu transkribiert. Abkürzungen stellen meist keine besonderen Anforderungen an das Leseverständnis, können aber mit <abbr> versehen in einer textkonstitutionellen Anmerkung aufgelöst werden, wenn dies für nötig angesehen wird. Bestimmte aus Bequemlichkeit heraus verwendete Kurzverschriftungen (etwa für <das> oder <den> ein symbolisches Kürzel für <d>/<dies> bzw. <d.> in Datumsangaben) werden im edierten Text nicht aufgelöst, können aber gleichfalls mit <abbr> versehen werden.

Bei Kontraktionen erfolgt keine mimetische Nachahmung von Schleifen, Haken und anderen Diakritika, wie sie etwa bei <eè> die Auflösung <esse> oder bei <poè> die Auflösung <posse> anzeigen. Diese Grapheme werden ergänzt, mit <abbr type="contraction"> kodiert und erscheinen im Frontend kursiviert. Schliesslich werden verlängerte Minuskel-ℓ- oder -ℓ̃-Vorkommnisse und Suspensionsschleifen, die nicht nur für einzelne Buchstaben, sondern auch für komplexere Abkürzungen stehen können (normierter Suspensionsbuchstabe etwa bei <Hℓ>, <hℓ>, oder <Wolgeℓ̃>) mit <abbr type="suspension"> kodiert, im Frontend als kursivierter Normalbuchstabe (<ℓ> oder <ℓ̃>) ohne tiefere Unterlänge dargestellt, sobald es sich eindeutig nicht um ein reguläres <ℓ>, <ℓ̃> oder einen anderen Buchstaben handelt.

4.5 Geminations- und Nasalstriche

Geminationen, die durch Geminationsstriche (Verdopplungsstriche/Reduplikationsstriche) meist über <ṁ> und <ṛ> angezeigt sind, werden ausgeschrieben und nicht nachgeahmt, weil sie über UNICODE nicht leicht zu realisieren wären. Fehlen sie hingegen, wird auch nicht verdoppelt, d.h. etwa <komt> nicht zu <kommt> erweitert, wenn es kein Anzeichen graphisch realisierter Gemination gibt. Dazu zählen auch vergessene Doppelpunkte oder u-Bögen (<¨>). Längen über Vokalen (<ˉ>), die als Nasalstriche <m> oder <n> oder mit Nasal endende Silben anzeigen, werden entsprechend aufgelöst (etwa <fratrū> zu <fratrum>, <taṁ> zu <tamen> etc.).

4.6 Akzente, Diakritika, Superscript-Schreibungen und Apostrophierungen

Akzente und Diakritika werden so wiedergegeben wie in der Vorlage, auch wenn sie (nach damaligen wie heutigen Usancen) als falsch gesetzt angesehen werden müssen. Fehlen Akzente oder Diakritika, werden diese nicht ergänzt, bei nicht eindeutigen Positionen (etwa Unentscheidbarkeit zwischen <é> und <è>) aber orthographisch korrekt gesetzt. Dazu zählen auch Doppelpunkt bzw. Trema über dem <ï> und <ı> (<ï̈>/U+00EF und <ı̈>/U+00CF) oder <e> und <E> (<ë>/U+00EB und <Ë>/U+00CB) sowie andere sprachenspezifische Akzentuierungen. Diese werden wiedergegeben, wenn sie vorhanden sind, aber nicht nachgetragen (im Französischen auch <c> und <s> neben <ç>/U+00E7). Eine Ausnahme bilden handschriftliche <ÿ>/<ÿ̃>-Vorkommnisse, die durchgehend als <y> und <Y> transkribiert werden. Sie können aber (auch in gedruckten Texten) als Alternativverschriftungen von <ij> bzw. <ii> vorkommen und werden dann vorlagengetreu exakt so wiedergegeben (<ÿ>/U+00FF).

Bei Umlauten werden Doppelpunkte durchgehend gesetzt, also auch bei Vokalen mit Superskript-e (<ᵉ>/U+0364) in gedruckten Texten; eine Wiedergabe von Vokalen mit Superscript-e wäre mit UNICODE problematisch. Apostrophierungen werden – wo nicht eindeutig entscheidbar – auf Briefebene vereinheitlicht. Wo dies nicht möglich ist, erfolgt eine Vereinheitlichung nach der Praxis des Schreibenden. Lässt sich auch eine solche nicht ermitteln, wird nach damaligen Usancen normalisiert. Es werden aber niemals Apostrophe nachgetragen, wo damalige wie heutige Lesegewohnheiten deren Verwendung erforderlich machen. **[Problem des Suchens/Sortierens]**

4.7 Ligaturen

Bei den handschriftlichen Dokumenten werden konsonantische Ligaturen (wie <ch>, <ck>, <ct>, <ff>, <fi>, <th>, <sch>, <sz>, <st>, <tz>, <ts>) aufgelöst, sie sind mit UNICODE nicht bequem wiederzugeben, hingegen die vokalischen Ligaturen für <ae> und <oe> (<Æ>/U+00C6 und <æ>/U+00E6, <Œ>/U+0152 und <œ>/U+0153) beibehalten, sobald eine Quelle sie enthält (d.h. sie werden aber nicht nachgetragen, wo sie fehlen). Bei den Drucktexten werden auch konsonantische Ligaturen im XML dokumentiert. **[Problem des Suchens/Sortierens]**

4.8 Sonderzeichen, Währungsangaben und Masseinheiten

Währungsangaben und andere in den Dokumenten vorfindliche Sonderzeichen, auch solche für mathematische oder astronomische Notationen und Bemassungen, werden mit dem entsprechenden UNICODE-Zeichen wiedergegeben, wenn es existiert. Ist dies nicht der Fall und es gibt kein UNICODE-Zeichen, wird stellvertretend die Tilde (<~>/U+223C) als Platzhalter für nichtdarstellbare Zeichen gesetzt und auch im Frontend angezeigt, im XML erfolgt die Codierung durch <g> **[Problem des Suchens/Sortierens]**.

4.9 Wiedergabe von <s>/<ſ>/<ß>

Die unterschiedlichen <s>-Schreibweisen als Rund-s (<s>/U+0073), Lang-s (<ſ>/U+017F), oder Scharf-s (<ß>/U+00DF), auch in Verbindungen wie <ss>, <st>, <ft>, <ßt>, <sf> oder <fs>, werden beibehalten. Das gilt auch für Fälle, wo sie regelwidrig gesetzt wurden, etwa bei zusammengesetzten Wörtern, aber auch bei <daſ>/<deſ> oder <desfen>/<dasſ> (statt <das>/<des> oder <dessen>/<deſſen>/<deßen> bzw. <dass>/<dafs>/<daß>). Wo sie nicht zu unterscheiden sind, werden sie auf Briefebene vereinheitlicht. Sollte dies nicht möglich sein, erfolgt eine Vereinheitlichung nach der grundsätzlichen Praxis des Schreibenden. Lässt sich auch eine solche nicht ermitteln, wird nach damaligen Usancen normalisiert. **[Problem des Suchens/Sortierens]**

4.10 Interpunktion und Zeichenabstände

Interpunktion wird grundsätzlich beibehalten und vorlagengetreu transkribiert, d.h. (wenn fehlend) nicht ergänzt und auch nicht emendiert. Dazu gehört auch ein punctus elevatus (<Mittelpunkt>), wenn eindeutig als solcher zu lesen (<·>/U+00B7). Die Abstände zwischen den Zeichen und dem Wort, an das sie anschliessen, werden nicht nachgebildet; wenn ein Spatium zwischen Doppelpunkt oder Semikolon und angeschlossenem Zeichenvorkommnis sichtbar ist, wird die Zeichensetzung nach den sprachenspezifischen Regeln angeschlossen. Entsprechend werden etwa zweiteilige Interpunktionszeichen (<:>, <:>, <?>, <!>) in allen französischsprachigen Texten spationiert und mit geschützten Leerzeichen (<·>/U+00A0) versehen wiedergegeben, was bei den deutsch- oder lateinischsprachigen nicht geschieht, egal wie die Schreibenden dies jeweils gehandhabt hatten. Große und übergroße Abstände in der Zeile zwischen Zeichen und Textpartien, besonders nach Punkten und vollständigen Sätzen, bleiben unberücksichtigt oder werden als Absätze interpretiert, wenn es in der Vorlage keine anderen Absatzkennzeichnungen gibt.

4.11 Viertelgeviert-, Halbgeviert- und Geviertstriche

Wort-Trennzeichen vor Zeilenumbrüchen werden mit <pc type="hyphenation"/> codiert. Der nachfolgende Zeilenumbruch wird als <inWord> ausgezeichnet. Andere Trennungs- und Bindestriche werden nach Möglichkeit so wiedergegeben, wie sie anzutreffen sind, d.h. als Divis wird <÷>/U+002D verwendet, neben Halbgeviertstrichen (<↯>/U+2013) werden sonst noch Geviertstriche (<—>/U+2014) gesetzt. Weitere Trenn- und Kopplungsstriche sind über diese drei Strichvarianten (<÷>, <↯> und <—>) hinaus nicht vorgesehen. Ebenso werden als Divis verwendete <⇒>-Grapheme (oder Doppelpunkte und andere Zeichen) nicht als solche wiedergegeben, auch wenn sie in manchen gebrochenen Druckschriften mitunter Standard als Wort-Trennzeichen sind (<⇒> kann allerdings als mathematischer Operator vorkommen).