

Protokoll zur konstituierenden Sitzung der TEI-Arbeitsgruppe „Briefe und Tagebücher“ am 14. September 2007 in der Akademie der Wissenschaften in Mainz

Teilnehmer:

Anwesend:

Dr. Gabriele Buschmeier (Union der Akademien der Wissenschaften, Mainz)
Dr. Thomas Burch (Kompetenzzentrum für elektron. Erschließung in den Geisteswiss., Trier)
Dr. Andreas Kuczera (Akademie der Wissenschaften und der Literatur, Mainz)
Dipl. Wirt. Inf. Daniel Röwenstrunk (Edirom-Projekt, Detmold/Paderborn)
Peter Stadler M. A. (Edirom-Projekt, Detmold/Paderborn)
Prof. Dr. Joachim Veit (Weber-Gesamtausgabe, Detmold/Paderborn)
Prof. Dr. Werner Wegstein (Kompetenzzentrum EDV-Philologie, Würzburg)

Entschuldigt:

Dr. Roland Kamzelak (Deutsches Literaturarchiv Marbach)
Dr. Andrea Rapp (Kompetenzzentrum für elektron. Erschließung in den Geisteswiss., Trier)

Gabriele Buschmeier wies in ihren einleitenden Bemerkungen auf die Bedeutung der 1997 veröffentlichten *Richtlinien-Empfehlungen zur Edition von Musiker-Briefen* für neuere Editionsunternehmen hin. Vergleichbare, aber sinnvollerweise die eigenen Fachgrenzen überschreitende Empfehlungen sollten nun auch im Bereich digitaler Briefeditionen angestrebt werden, um neu beginnenden Editionen eine Orientierung zu bieten. Werner Wegstein hob hervor, daß die gegenwärtigen Konstellationen die einmalige Chance eröffnen, fächerübergreifend einen breiten Konsens für die Auszeichnung dieser Textkorpora zu finden.

Bereits in Angriff genommene oder geplante Briefeditionen im Bereich der Literaturwissenschaft sind häufig im weitesten Sinne als Retro-Digitalisierung angelegt. Thomas Burch führte am Beispiel der Heine-Brief-Edition aus, daß hier als Vorgabe für die Textpräsentation die Übereinstimmung von gedruckter und digitaler Ausgabe galt, was dazu führte, daß viele äußere Details der früheren Edition mit codiert werden mußten (etwa Seitenumbrüche, Zeilenfälle oder leere Zwischenräume, so daß Textstellen in beiden Medien ohne Probleme auffindbar und referenzierbar bleiben). Im Sinne effektiver Arbeitsteilung wurden dabei Brieftexte, Kommentare und Meta-Informationen in getrennten Dateien abgelegt, so daß von unterschiedlichen Gruppen gleichzeitig am Material gearbeitet werden konnte. Dort, wo bereits etablierte Ausgaben vorliegen (z. B. bei Mozart), wird man vermutlich ähnlich verfahren wollen.

Dagegen sind Musikerbrief-Editionen – wie z. B. die geplanten Ausgaben der Briefe von Brahms, Mendelssohn, Reger oder Weber – in der Regel eher Erst-Editionen (zumindest wenn Maßstäbe der neueren Editionsphilologie angelegt werden), so daß hier genuine Brief-Auszeichnungs-Systeme ohne Rücksicht auf sonstige äußere Vorgaben verwendet werden können. Solche Editionen sollten daher Ausgangspunkt der Überlegungen sein, wobei in den Diskussionen deutlich wurde, daß sich ggf. ein mehrstufiges Konzept empfiehlt, das von einer Grundausszeichnung ausgeht und eine sukzessive Vertiefung der Erfassung der Inhalte (u. U. auch von unterschiedlichen Erkenntnisinteressen ausgehend) erlaubt.

Joachim Veit stellte die Probleme heraus, die sich bei den Versuchen einer zunächst auf XML basierenden und schließlich an TEI angelehnten Auszeichnung der inhaltlich vielfach verflochtenen Brief- und Tagebuchtexte der Weber-Gesamtausgabe (WeGA) ergaben:

- Es wurden (und werden nach wie vor) zu wenige ausführliche Beispiele für die Codierung von Briefen im Netz angeboten; oft waren (sind) dabei Struktur und Inhalt der Texte nur sehr oberflächlich erfaßt.
- TEI erlaubt in vielen Fällen mehrere Wege, die zu ähnlichen Zielen führen, was zu Unsicherheiten bei weniger erfahrenen Benutzern führt.
- Es gibt keine zentrale Informationsstelle und kaum Erfahrungsaustausch zwischen Projekten.
- Es fehlen (von XML-Editoren abgesehen, die in der Regel für Außenstehende eher abschreckend wirken) vernünftige Eingabe-Oberflächen, die das Auszeichnen in einer der gewohnten Arbeitsumgebung entsprechenden Form erlauben.

Unter Berücksichtigung zugänglicher Beispiele und basierend auf den eigenen Erfahrungen beim Auszeichnen von Dokumenten wurde daher von der WeGA und den Mitarbeitern des Edirom-Projekts auf der Grundlage von TEI P5 ein Muster für die Auszeichnung von Briefen entwickelt, das im Gegensatz zu den verdienstvollen, auf P4 beruhenden Modellen des holländischen DALF-Projekts (Digital Archive of Letters in Flanders, <http://www.kantl.be/ctb/project/dalf/>) von Edward Vanhoutte zunächst mit den vorhandenen Elementen und Attributen von P5 auszukommen sucht – was nur möglich ist, weil einige der Anregungen des DALF-Projekts bereits in modifizierter Form in P5 eingegangen sind.

Ein aus den *Richtlinien-Empfehlungen* abgeleitetes Papier mit der Zusammenstellung beschreibender Kategorien für Briefe und dieser *Entwurf zur Auszeichnung von Briefen in der Carl-Maria-von-Weber-Gesamtausgabe* (vgl. <http://www.adwmainz.de/index.php?id=492>) bildeten die Grund-

lage der Mainzer Diskussionen. Die Ergebnisse sollen jeweils in den angegebenen *Entwurf* eingearbeitet und an Ort und Stelle kommentiert werden. Im folgenden sind lediglich die grundsätzlicheren Erwägungen dieser Diskussionen festgehalten.

Gegenstand waren zunächst die **BRIEFTEXTE** selbst. *Adresse* und *Absender* (auf dem Briefbogen oder auf einem separaten Umschlag) sind Bestandteile des Kommunikationsmittels Brief und sollten daher als zum Text gehörig betrachtet werden, d. h. aus dem <front>-Bereich in den eigentlichen <body>-Bereich verlegt werden. Sie sind dort in einem eigenen <div> mit entsprechender Attributierung (type="address" oder type="address envelope") abzulegen.

Als mögliche strukturelle Einheiten wurden in der Diskussion vor allem die „Schreibeinheiten“ (Tage, Vormittag/Nachmittag o. ä.) bezeichnet, die mit entsprechenden Attributen versehen werden können (z. B. <div type="writingDate">). Da überlappende <div>-Hierarchien nicht möglich sind, müßten in einer solchen Einheit z. B. vorkommende Wechsel der Quellengrundlage (etwa von Autograph zu Drucküberlieferung) in anderer Form gekennzeichnet werden (z. B. als <milestones>).

Entsprechend sollten auch *Nachschriften* (PS) nicht in den <back>-Bereich ausgelagert werden, sondern als eigene <div>-Abschnitte Teil des Textes bleiben (z. B. als <div type="ps">. Damit lassen sich die in DALF in eigene Elemente gefaßten Teile innerhalb des Elementvorrats von P5 abbilden.

(Denkbar wäre, Brief-*Beilagen* im <back>-Bereich zu erfassen; diese Frage blieb jedoch noch ungeklammert.)

W. Wegstein schlug vor, für die grundlegenden Auszeichnungen immer möglichst schlichte Formen vorzugeben, so daß z. B. in einer *Adresse* <address> nicht jeder Zeilenumbruch als <addrLine> bezeichnet werde, sondern die Adresse als Einheit aufgefaßt werden könnte. Wenn jedoch der Zeilenumbruch verzeichnet werden soll, empfiehlt sich statt des <lb/>-Elements eine Wiederholung der <addrLine> (oder in anderen Abschnitten auch des <salute> oder ähnlicher Elemente). Das <address>-Element muß innerhalb des <div>-Elements leider in <p>...</p> eingeschlossen werden – eine logische Begründung für diesen Zwang fehlt.

Einigkeit herrschte auch darüber, daß eventuell für die Anzeige/Wiedergabe der Texte verwendete Informationen zur räumlichen Anordnung als Merkmale des Textes und nicht im Hinblick auf deren spätere Darstellung codiert sein sollten. So könnten z. B. besondere Anordnungen der Zeilen innerhalb einer *Grußfloskel* als „<salute rend="left | center | right"> vermerkt werden; sollen erste Absätze des jeweiligen <div> gekennzeichnet werden, liegt die Verwendung von „n=1“ (statt des *rend*-Attributs) näher.

Eine Markierung der **Zeilenumbrüche** im Original (durch <l>...</l>) wurde empfohlen, wenn geplant ist, Übertragungen und Faksimile in der Edition zu kombinieren. [Alternativ <lb/>, was auch Auszeichnungen (kursiv etc.) über die Zeilengrenze hinweg erlaubt.] Außerdem erleichtert diese Auszeichnung – auch wenn sie aufwendig ist – die Korrekturvorgänge.

Trennungen bei Zeilenwechsel sollten zumindest dann gekennzeichnet werden, wenn sich daraus Interpretationsprobleme ergeben. Dabei wird der Trennungsstrich als <c type="hyphenation">-</c> markiert; ggf. kann die ungetrennte Wortform für eine Suche zusätzlich angegeben werden (mit <reg>?).

Übereinstimmung herrscht auch hinsichtlich der Verwendung der Elemente <emph> und <hi>. Für die Erfassung von Textauszeichnungen wie kursiv, hochgestellt, unterstrichen, gesperrt, Frakturschrift usw. wird einheitlich <hi/> verwendet, da die Unterscheidung von graphischem Befund und inhaltlicher Bedeutung oft schwierig erscheint und eine doppelte Auszeichnung nur in sehr speziellen Fällen sinnvoll ist.

Personennamen sollten nur dann als <persName> ausgezeichnet werden, wenn tatsächlich Namen genannt sind (ggf. auch nur Vornamen); bei allen indirekten Nennungen („die Tante“; „er“) sollte das <rs>-Element Verwendung finden.

Lediglich angesprochen wurde das Problem der Auszeichnung von **Namen** (Rollen, Werke usw.). Hier empfiehlt sich für eine weitergehende semantische Erschließung von Texten die Einführung von Typologien – dieses Thema sollte einer der Gegenstände der nächsten Sitzung sein.

Bei der Diskussion der **METADATEN** erwies sich zwar die vorgegebene Gesamtstruktur als unproblematisch, unklar blieb aber in vielen Fällen, in welchem der Container die Informationen zum Brief jeweils am günstigsten unterzubringen wären. Insbesondere zahlreiche der für einen Briefkatalog üblichen Informationen (Adressat, Schreibort, Empfangsort, Provenienzen, Erstdruck, sonstige Drucküberlieferung, Faksimiles, Abschriften; Kurzinhaltsangabe usw.) haben innerhalb des <teiHeader> keinen eindeutigen Ort. Ob einige dieser Informationen als <note> innerhalb des <notesStmnt> oder besser innerhalb der <profileDesc> unterzubringen wären, blieb zunächst offen. Ebenso zu klären wäre, ob die jetzige Verwendung des <textClass>-Elements innerhalb der <profileDesc> den Intentionen von TEI entspricht.

Als weitgehend unproblematisch konnte die Frage der **Quellenangaben** bei Brieftexten, die sich aus mehreren jeweils fragmentarisch überlieferten Zeugen zusammensetzen, gelten. Sie werden zu Beginn der Quellenangaben in einer <witList> zusammengefaßt und dann jeweils einzeln beschrieben.

Ohne Eingriffe übernommen können auch die Angaben zur <encodingDesc> und <revisionDesc>, dagegen sollen Details der <fileDesc> und der <profileDesc> noch eingehender erörtert werden.

Die <title>-*Angabe* innerhalb des Headers müßte sich deutlicher auf die vorliegende Veröffentlichung beziehen, also z. B. lauten: „Weber-Gesamtausgabe, Digitale Edition“, gefolgt von Details zum einzelnen Brief. Vermutlich wären hier auch die bisher noch im <head> des <front>-Bereichs hinterlegten Informationen besser unterzubringen. Um eindeutige Identifikationsnummern für die Briefe festlegen zu können, könnte eine ISBN-ID für die digitale Briefausgabe beantragt werden.

In den Bereichen <physDesc> und <history> sollten wahlweise schlicht ausformulierte Formen (innerhalb von <p>-tags) oder formalisierte Varianten angeboten werden; in ähnlicher Weise offen können auch bibliographische Angaben erfolgen.

Die in der ersten Arbeitssitzung besprochenen Änderungen sind in die Neufassung des Papiers eingearbeitet worden und stehen ebenfalls auf der Homepage der Mainzer Akademie der Wissenschaften (<http://www.adwmainz.de/index.php?id=492>) zur Verfügung.

Eine Fortsetzung der Diskussionen ist während der Paderborner Tagung vom 6.-8. Dezember 2007 geplant. Zur Zeit geplante neue Musikerbriefausgaben sollen angesprochen werden, um ihre Wünsche und Bedürfnisse in die geplanten Briefmuster mit einbringen zu können.

Detmold, 1. Oktober 2007