

1. Vorbemerkung

Die Textauszeichnung wird auf der Grundlage der TEI-Richtlinien beschrieben. Die Eingabe von Elementen und Attributen wird soweit möglich und sinnvoll durch eine entsprechende Konfiguration des Editors unterstützt. Dies gilt insbesondere für die Kennzeichnung des für das jeweilige Textelement Verantwortlichen (Schüler oder Lehrer); dabei wird das Attribut `@hand` verwendet, das die Werte `#student` oder `#teacher1` (rot hinterlegt)/`#teacher2` (blau hinterlegt)/`#teacher3` haben kann.

Anmerkung zu den Schreibweisen in diesem Dokument: TEI-Elemente stehen in Schreibmaschinenschrift in spitzen Klammern, Attribute in Schreibmaschinenschrift mit einem `@` am Anfang, das nicht Bestandteil des Attributs ist.

2. Transkription

2.1. Grundsätze der Textwiedergabe

Die Textwiedergabe erfolgt grundsätzlich unter Beibehalt der originalen Orthographie und Zeichensetzung, ohne Normalisierungen oder Anpassungen an heutige (normierte) Schreibweisen (diplomatische Transkription).

2.2. Transkription auf Zeichenebene

Teilweise normalisiert wird die Wiedergabe von *Leerzeichen*: Worttrennungen und Zusammenschreibungen werden übernommen, auch wenn sie der heutigen Orthographie widersprechen, aber Leerzeichen bei Interpunktion folgen dem aktuellen Usus, und größere Abstände werden normalisiert.

Groß- und Kleinschreibung wird getreu der Vorlage wiedergegeben, auch bei Verwendung von Großbuchstaben innerhalb zusammengeschiedener Wörter, nur in Zweifelsfällen gemäß heutiger Rechtschreibung normalisiert. Gleiches gilt für *Getrennt- und Zusammenschreibung* von Wörtern.

Spezielle Graphien der Kurrentschrift werden auf die im Antiqua-Satz üblichen Zeichen zurückgeführt. Dies gilt insbesondere für das lange S (#), für den Doppelstrich (=) als Bindestrich und für den Geminationsstrich zur Verdoppelung der Konsonanten m und n. Distinktionszeichen, etwa zur Kennzeichnung von u oder y (ÿ), werden fortgelassen.

Zahlzeichen werden vorlagegetreu wiedergegeben, es sei denn, dass gewichtige Gründe für eine Auflösung sprechen. Bei Ordnungszahlen wird die konsequente Hinzufügung eines Punktes empfohlen. Punkte nach Grundzahlen werden nicht wiedergegeben.

3. Textauszeichnung

3.1. Grundstruktur des Textes

Für die Grundstruktur eines Klassensatzes von Abituraufsätzen wird eine Vorlage bereitgestellt. Das oberste Element dieser Vorlage ist `<teiCorpus>`, gefolgt von `<teiHeader>`. In diesem `<teiHeader>` werden alle Metadaten für den gesamten Klassensatz definiert, wie z.B. Namen der Lehrer, Datum der Prüfung, usw. Jeder Aufsatz wird umfasst von einem Element `<tei>`, wo wieder ein `<teiHeader>` definiert wird. Hier werden die Metadaten für einen Aufsatz beschrieben. Wichtig ist, dass möglichst viele Metadaten eingetragen werden. Bei der URN gibt es eine für die gesamte Akte, die muss zu Beginn stehen, und es gibt weitere für jede einzelne Seite, die auch bei jeder einzelnen Seite (auch bei der ersten) genutzt werden müssen.

Die Textstrukturierung erfolgt durch `<div>`-Elemente innerhalb von `<body>`, die in Absätze `<p>` unterteilt sind. Jeder Aufsatz in einem Klassensatz steht in einer `<div type="student-essay">` (ohne `@hand`-Attribut) und in der es eine oder mehrere `<div type="teacher-assessment">` mit `@hand` gibt. Eine Gliederung steht in einer `<div type="outline">`.

Überschriften werden übernommen und mit dem `<head>`-Element ausgezeichnet; es werden keine Überschriften ergänzt.

Zeilenumbrüche der Vorlage werden nicht wiedergegeben, *Worttrennungen* am Zeilenende ohne Kennzeichnung zur Normalform zusammengezogen. Eine Ausnahme kann gelten, wenn ein Lehrerkommentar sich auf die (falsche) Trennung bezieht.

Einrückungen der ersten Zeile eines Absatzes bleiben unberücksichtigt.

Der *Seitenwechsel* wird durch `<pb>` markiert. Ein Link auf die URN der Seite wird mit `@facs` erfasst. Dies ermöglicht einen direkten Link von Forschungsumgebung bzw. TEI-Datei in den Goobi-Viewer, um das Original aufzurufen. Beispiel: `<pb facs="urn:nbn:de:0111-bbf-spo-16940200"/>`

3.2. Schrift- und Sprachwechsel

Die Grundschrift eines Aufsatzes (Kurrent oder lateinisch) wird in den Metadaten auf Aufsatzebene vermerkt. Einzelne signifikante Passagen in der jeweils anderen Schriftart werden mit `<hi>` und dem Attribut `@latin/@kurrent` gekennzeichnet.

Der Wechsel der *Sprache* wird nicht gekennzeichnet; fremdsprachige Passagen werden im Original wiedergegeben.

3.3. Hervorhebungen im Text

Unterstreichungen werden mit `<hi rend="underline">` (bzw. `"doubleunderline"`) und dem `@hand`-Attribut ausgezeichnet. Dazu steht der Button "U" zur Verfügung.

3.4. Textänderungen durch den Schüler

Textlöschungen durch den Schüler werden mit `<del hand="#student">` erfasst. Das Attribut `@rend` kann dabei folgende Werte annehmen:

- `striketrough` (durchgestrichen)
- `overwritten` (überschrieben)
- `erasure` (gelöscht)
- `none` (wenn die Löschung implizit erfolgt und durch Vorhandensein der verbesserten Version zu erschließen ist).
- Der Button "delete" erlaubt die Auswahl von Verantwortlichem und Art der Löschung.

Innerhalb von `` können auch `<unclear>`, `<gap>` und `<supplied>` verwendet werden, wenn der gestrichene Teil schlecht oder gar nicht zu lesen ist (siehe unten).

Textergänzungen durch den Schüler stehen in `<add hand="#student">`, dessen Attribut `@place` folgende Werte annehmen kann

- `inspace` (überschrieben)
- `above` (über der Zeile)
- `below` (unter der Zeile)
- `margin` (für alle Ränder)
- `inline` (direkt nach der Löschung)
- `dots` (Aufhebung einer Löschung durch Unterpunktung)

Der Button "add" erlaubt die Auswahl von Verantwortlichem und Art der Ergänzung.

Eindeutig zusammengehörige Paare von `` und `<add>` werden mit dem umgebenden Element `<subst>` vereinigt. Beispiel: `<subst><del rend="overwritten">wollte<add place="inline">musste</add></subst>`.

3.5. Korrekturen und Anmerkungen des Lehrers

Randbemerkungen des Lehrers werden mit `<add hand="#teacher">` ausgezeichnet. Sie stehen stets unmittelbar nach dem Wort oder der Passage, auf die sie sich beziehen. **BEZUGSTELLE KENNZEICHNEN** Das `@place`-Attribut wird verwendet wie bei `<add>`.

Senkrechte Striche, mit denen ein Lehrer die Erstreckung der Passage kennzeichnet, auf die sich die Randbemerkung bezieht, werden mit `<hi rend="vertical">` gekennzeichnet (Button).

KATEGORIEN: FORMAL UND TEXTLICH, KURZ CODIEREN, NICHT ALS SEMANTISCHER TEXT

3.6. Bewertung am Ende des Aufsatzes

Der Bewertungsblock am Ende eines Aufsatzes wird als eigene `<div hand="teacher1">` erfasst. Der Name des Lehrers am Ende wird als `<signed>` ausgezeichnet.

3.7. Textkorrekturen bei der Transkription

Rechtschreib- und Grammatikfehler werden nicht korrigiert. Schreibversehen (wie z.B. versehentliche Wortwiederholungen) und ungewöhnliche Schreibweisen werden vorlagegetreu übernommen. Bei Bedarf können sie mit `<sic>` markiert werden. Offenkundig fehlende Wörter, die zum Textverständnis erforderlich wären, werden nicht ergänzt.

3.8. Schwierig zu lesende Passagen

Wörter oder Passagen, deren Lesung unsicher ist, werden mit `<unclear>` markiert (in der Anzeige grün hinterlegt). Mit dem Attribut `@reason` kann der Grund für die Unsicherheit spezifiziert werden (z. B.: `illegible` = unlesbar; `stain` = Tintenfleck; `damage` = mechanische Beschädigung des Papiers).

Endsilbenverschleifungen, bei denen nicht eindeutig zu erkennen ist, ob es sich um ein lediglich undeutlich ausgeschriebenes Wort (etwa bei den Endsilben: „-nen“, „-nem“, „-ung“) oder eine „Kürzelschleife“ handelt, die aus einem dem deutschen Buchstaben „h“ oder „l“ ähnlichen Gebilde besteht, können stillschweigend aufgelöst werden.

Kann eine Stelle gar nicht gelesen werden, ist das leere Element `<gap>` statt `<unclear>` zu verwenden (mit zusätzlichem Attribut `@extent` für die Angabe des ungefähren Umfangs der Lücke).

3.9. Abkürzungen

Abkürzungen werden beibehalten und nicht aufgehört.

PRÜFEN: DELEATUR-ZEICHEN, ANDERE SONDERZEICHEN

3.10. Sonstiges

Listen und Tabellen werden in schematisierter Form transkribiert, ohne zu versuchen, die genaue typographische bzw. handschriftliche Gestaltung nachzubilden.

Der Umgang mit Abbildungen, die zum Text gehören, ist im Einzelfall abzusprechen.

4. Struktur des Dokuments

Grobstruktur des Dokuments:

Wir nutzen TEI-Corpus als gliederndes Element. Die einzelnen Texte werden unter `tei` getaggt. Somit enthält ein Dokument „TEI-Corpus“ mehrere Unterdokumente. Die einzelnen Absätze werden mit `p`-Tags ausgezeichnet

Wichtig: unser Dokument muss immer mit TEI kompatibel sein. Prüft dazu das Dokument in entsprechenden Abständen

5. Metadaten

Unter dem Label TEI-Corpus werden alle Metadaten erfasst, die für den ganzen Klassensatz gelten. Dies sind z.B. Datum, Personen, usw. # eine Übersicht könnt ihr einfach kopieren

Datum bitte als YYYY-MM-DD angeben: `<date when="1950-12-30">` angeben. Zu Beginn einer neuen Seite wird ein Pagebreak angegeben: Wichtig ist dabei auch, dass ihr den Link mit URN und DNB-Identifizier angebt. Nur die sind eindeutig und ermöglichen die Rückverlinkung in der Forschungsumgebung. `<pb facs="urn:nbn:de:0111-bbf-spo-18205088" n="12r"/>`

`idno` gibt die ID der Person an. Hier ist kein weiterer Text nötig.

`State` gibt das Bundesland an, während `"Country"` den Staat angibt. Beim Ort auch gern den Stadtteil dazuschreiben

Die Metadaten für die Aufsätze sind entweder in Goobi schon erfasst oder müssen neu erfasst werden.

Die Vornote ist auch die Klassenleistung. Die Klausurnote ist nur die Note. "Noch Befriedigend" wird zu "Befriedigend".

Es werden verschiedene Metadaten im Block "allgemeine Metadaten" eingegeben, wobei das Attribut `"ana"` anzeigt, um welches Metadatum es sich handelt

Die Themen der Aufsätze werden im Block `<list type="Aufsatzthemen">` angegeben. Die Anzahl der noch kompletten gewählten Aufsätze wird im Attribut `"n"` erfasst, die reale Anzahl wird berechnet

6. Anmerkungen von Transkriptoren und Bearbeitern

Anmerkungen bei der Transkription werden in ein Kommentarfeld vor Textbeginn (TEI-Header) geschrieben.

neu: margin wohin schreiben? muss innerhalb eines `p`-Tags geschrieben werden!!

7. Titel bei Personen

Diese werden mit `<roleName type="honorific">Dr.</roleName>` dargestellt, vgl. <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ND.html>

8. Aufsatzthemen

Es sollen alle Aufsatzthemen erfasst werden, auch wenn diese nicht genommen wurden