

ANNOTATIONSGUIDELINES ZUR NORMORTHOGRAPHISCHEN KORREKTUR

A. Grundsätzliches:

Die Normalisierung findet mithilfe des Tools CAB (Cascaded Analysis Broker), das vom DTA (Deutsches Text Archiv) entwickelt wurde, statt. Diese automatische Normalisierung wird manuell kontrolliert und korrigiert. Das CAB erkennt von der heutigen Schreibweise abweichende Token und korrigiert diese. Token, die nach aktuellen Regeln korrekt geschrieben wurden, müssen in diesem Layer nicht annotiert werden.

Die Korrektur richtet sich nach der aktuellen Rechtschreibung und nachfolgenden weiteren Kriterien: Einheitlichkeit, Verständlichkeit, automatische Verarbeitbarkeit und Durchsuchbarkeit. Einheitlichkeit bezieht sich nicht nur auf das einzelne Drama, sondern auf das gesamte Korpus. Daher ist es wichtig, dass sich alle an die nachfolgenden Regeln und Vorgehensweisen halten. Verständlichkeit bezieht sich jedoch nicht auf eine komplette syntaktische Einheit, sondern einzelne Wörter oder Token.

Die Kriterien automatische Verarbeitbarkeit und Durchsuchbarkeit führen in einigen Fällen dazu, dass das CAB Fehler produziert, die sorgfältig korrigiert werden müssen (FABRIC. – Abkürzung für die Figur Fabricius wurde zu Fabrik; Schaw-Platz zu Schaf-Platz → dies muss selbstverständlich korrigiert werden).

Die Eintragungen im Layer Norm sollen überprüft und ggf. korrigiert und ergänzt werden, so dass manchmal auch eine neue Annotation angelegt werden muss. Dies erfolgt als freie Annotation, d.h. es gibt kein festgelegtes Tagset. Da sich die Lemmatisierung auf die Orthographie bezieht und auf diese wiederum das PoS, sollte die Korrektur so nah wie nur irgend möglich an der Wortart des Tokens bleiben (übermassen → übermäßig; und nicht zu über die Maßen).

B. Regeln:

1.	Grundsätzlich wird „Wort-für-Wort“ korrigiert.
2.	In Einzelfällen wird der Kontext zur Entscheidung herangezogen. Insbesondere bei: <ul style="list-style-type: none"> • Fragen der Zusammen- und Getrennschreibung, • Fragen der Flexionsendung bei attributiven Adjektiven • Fragen der Erkennung von Wortarten.
3.	Ziel ist nicht der wohlgeformte, neuhochdeutsche Satz.
4.	Im Bereich <u>Graphematik</u> korrigieren wir hin zu einer orthographischen Anpassung nach Duden (<i>Káyser</i> → <i>Kaiser</i>). Hierbei gilt: Orientierung ausschließlich am

	<p>Duden, auch wenn bspw. für die Gryphius-Editionen von Mannack andere Schreibweisen vorgeschlagen werden.</p> <p>Wenn laut Duden zwei orthographische Realisierungsformen zulässig sind (<i>zu Grunde</i> / <i>zugrunde</i>, <i>mit Hilfe</i> / <i>mithilfe</i>), wird sich an der von Duden <u>empfohlenen</u> Schreibweise orientiert.</p>
5.	<p>Im Bereich <u>Phonologie</u> sind die frühneuhochdeutsche Lautwandelprozesse zu beachten – also Diphthongierung, Monophthongierung, Synkope, Apokope, Nukleussenkung, etc. (<i>duppel</i> → <i>doppel</i>, <i>hellsche</i> → <i>höllische</i>, <i>suchet</i> → <i>sucht</i>).</p>
6.	<p><u>Lexikologie</u>: ausgestorbenes lexikalisches Material wird: wenn es z. B. im Duden oder DWB nachgewiesen werden kann, nach aktuellen Orthographie-Regeln angeglichen (<i>Hofgewend</i> → <i>Hofgewende</i>); wenn es nirgendwo anders nachgewiesen werden kann, übertragen (<i>schlappermentsch</i> → <i>Sakrament</i>– Erläuterung: „Verwünschendes Beiwort mit verhüllender Entstellung von „Sakrament““).</p>
7.	<p><u>Wortbildung</u>: ausgestorbene Wortbildungsmorpheme werden nach Möglichkeit durch neuhochdeutsche Bildungen ersetzt (<i>selbte</i> → <i>selbe</i> oder <i>Zuthuung</i> → <i>Zutun</i>). Häufig wird das Präfix „zu-“ bei Wörtern verwendet, die im Neuhochdeutschen mit „zer-“ beginnen (<i>zustückt</i> → <i>zerstückt</i>).</p>
8.	<p><u>Morphologie</u>: wird dann normalisiert, wenn keine Flexion vorhanden ist oder wenn Flexion vorhanden ist, die heute so nicht mehr existiert. In diesem Fall verändern wir nicht den Kasus selbst, sondern nur die Form. Dies ist besonders zu beachten bei:</p> <ul style="list-style-type: none"> • attributiven Adjektiven (<i>groß Register</i> → <i>großes Register</i>); • Flexionsendungen von Nomen (<i>Vatern Aug‘</i> → <i>Vaters Auge</i>); • häufiger Ausfall von Präfixen bei Partizipien (<i>funden</i> → <i>gefunden</i>).
9.	<p>Ansonsten wird die Grammatik nicht korrigiert (also in Bezug auf Tempus etc.).</p>
10.	<p>Es werden auch keine Übersetzungen gemacht.</p>
11.	<p>Ausnahme zu den Punkten 9. und 10.: Im Scherzspiel <i>Die gelibte Dornrose</i> (Gryphius) wird anders als in den anderen Dramen die Mundart und damit das nicht mehr verständliche Wortmaterial ins Neuhochdeutsche übertragen.</p>
13.	<p>Abkürzungen werden ausgeschrieben → Es sei denn, es handelt sich um Sprecherangaben.</p>
14.	<p>Komposita, die auseinandergeschrieben wurden (ob mit oder ohne Bindestrich) werden durch eine Spannenannotation verbunden. (<i>zweiffels ohn</i> → <i>zweifelsohne</i>, <i>Schaw-Platz</i> → <i>Schauplatz</i>).</p>

15.	Bei Komposita, die durch einen Trennstrich abgekürzt werden, werden die ersten beiden Token zusammengefasst und in der Trennstrichschreibweise normalisiert. (<i>Vor – und Hinter – Hof</i> → „Vor-“ „und“ „Hinterhof“).
16.	Wörter, die nach heutiger Orthographie nicht zusammengeschrieben werden, werden auch in der Annotation getrennt geschrieben. Dies gilt auch für Klitisierungen, bei denen ein oder mehrere Buchstaben wegfallen. (<i>lebenlang</i> → <i>Leben lang</i> , <i>bistu</i> → <i>bist du</i>).
17.	Klitisierungen mit „es“ werden mit Apostroph realisiert (<i>umbs</i> → <i>um 's</i>).
18.	Apostrophe werden aufgelöst (<i>Har'</i> → <i>Haare</i> , <i>tritt'</i> → <i>trete</i>).
19.	Die Schreibweise der Namen wird unter Rückgriff auf die jeweilige Realisation korrigiert. Es sei denn, es handelt sich um Sprecherangaben – Hier wird nur das Umlautzeichen aufgelöst. (<i>Bullabutáin</i> → <i>Bullabutaein</i> ; <i>Bulla Butáin</i> → <i>Bulla Butaein</i> ; <i>Bullabután</i> → <i>Bullabutaen</i> ; ABER: „LIBE.“ → Keine Korrektur).
20.	Sollte ein lateinischer Name (z.B.: <i>Pyramum</i> , <i>Caroli</i> etc.) <u>nicht</u> im Druck in Antiqua gesetzt sein, so wird er nach deutscher Regelung flektiert (<i>Pyramum</i> → <i>Pyramus</i> , <i>Caroli</i> → <i>Carolus</i>).
21.	Ortsnamen sind keine Eigennamen und werden daher nach aktueller Schreibweise korrigiert.
22.	Lateinische Zahlen werden zu arabischen vereinheitlicht.

C. Technisches Vorgehen:

1.	Einloggen in INCEpTION mit Benutzerkennung.
2.	Gryphius- bzw. Dramen-Projekt auswählen und auf <i>Annotation</i> klicken.
3.	Das Dokument öffnen, für dessen Korrektur/Annotation man eingeteilt wurde (Dokument anklicken).
4.	Gegebenenfalls unter <i>Settings</i> (Zahnrad-Symbol) für den Layer <i>Orthography Correction</i> ein Häkchen setzen. In der Regel werden alle Layer automatisch ausgewählt.
5.	Auf der rechten Seite den Layer <i>Orthography Correction</i> auswählen.

6.	Das zu annotierende Token durch Markierung auswählen und eine neue Annotation hinzufügen bzw. in die schon vorhandene Annotation klicken, um diese zu korrigieren.
6a.	Neue Annotation: Das / Die zu annotierende/n Token wird durch Markieren ausgewählt. Danach erscheint dann über dem Token, die Annotationsebene <i>Orthography Correction</i> und ist (dunkel-)orange umrahmt. Auf der rechten Seite kann dann die Korrektur im Feld <i>Correction</i> durch freie Annotation erfolgen, d.h. das Wort in der aktuellen Schreibweise eingegeben werden. Durch Enter wird diese dann eingefügt und erscheint über dem / den jeweiligen Token.
6b.	Korrektur einer schon vorhandenen Annotation: Die schon vorhandene Annotation wird ausgewählt, indem der Eintrag über dem zu korrigierenden Token auf dem Layer <i>Orthography Correction</i> per Doppelklick angeklickt wird. Der Eintrag erhält einen orangefarbenen Rahmen, sodass die Auswahl ersichtlich wird. Auf der rechten Seite erscheint dann das Annotationsfeld <i>Correction</i> . Nachdem der schon vorhandene Eintrag korrigiert und durch Enter bestätigt wurde, erscheint dieser über dem / den jeweiligen Token.
6c.	Löschen falscher / unnötiger Korrekturen: Die schon vorhandene Annotation wird ausgewählt, indem der Eintrag über dem zu korrigierenden Token auf dem Layer <i>Orthography Correction</i> per Doppelklick angeklickt wird. Der Eintrag erhält einen (dunkel-)orangefarbenen Rahmen, sodass die Auswahl ersichtlich wird. Nun wird die Annotation über den <i>Delete</i> -Button auf der rechten Seite gelöscht.

Die vorliegende Version ist eine Anpassung der Grundlagenguidelines aus: Eggert, Lisa; Müller, Melissa (2021): „NORM – Normorthographische Korrektur“. In: „Guidelines. Interaktionale Sprache bei Gryphius“ URL: <https://gryphiusprojekt.wordpress.com/guidelines/orthography-correction/>