

**WAS LESEN SIE DA ÜBERHAUPT? AUTOMATISIERTER VERGLEICH VON  
VERLAGSVARIANTEN IN EINEM GEGENWARTSSPRACHLICHEN ROMANTEXT**

**WHAT YOU EVER READ? AUTOMATED COMPARISON OF PUBLISHING  
VARIATIONS IN A CONTEMPORARY LANGUAGE ROMAN TEXT**

This article describes a project comparing 11 editions of Thomas Mann's novel "*The Magic Mountain*" (1924). The aim hereby is to develop computer-based tools that enable automated text-processing with only a small amount of additional manual work. After scanning a few chapters of the novel the pictures are converted into text-files. The texts are dehyphenated, OCR-errors are corrected automatically. By way of comparison all differences between them are revealed, i.e. punctuation, spelling and lexical modifications. The text of the first edition from 1924 and the variants found in all other editions are combined to a new critical edition.

**Key words:** Thomas Mann, "Magic Mountain", text-processing, comparing texts, critical edition, errors

Wie der Zufall doch manchmal so spielt. Als ich vor Jahren nochmals Thomas Manns *Zauberberg* las, bin ich zu einem Auslandsaufenthalt aufgebrochen. Mein Plan war es, ihn in der Bibliothek am Zielort auszuleihen und im dortigen Exemplar weiterzulesen. Ich hatte den Romanbeginn (Ausgabe 2010) im Kopf:

Die Geschichte Hans Castorps, die wir erzählen wollen, – nicht um seiner willen (denn der Leser wird einen einfachen, wenn auch ansprechenden jungen Mann in ihm kennenlernen), sondern um der Geschichte willen, die uns in hohem Grade erzählenswert scheint (wobei zu Hans Castorps Gunsten denn doch erinnert werden sollte, daß es seine Geschichte ist, und daß nicht jedem jede Geschichte passiert): diese Geschichte ist sehr lange her, sie ist sozusagen schon ganz mit historischem Edelrost überzogen und unbedingt in der Zeitform der tiefsten Vergangenheit vorzutragen.

**Ausdruck 1. [Mann 2010: 9]**

In dem entliehenen Roman-Exemplar (Ausgabe 1987) aber konnte ich lesen:

Die Geschichte Hans Castorps, die wir erzählen wollen – nicht um seiner willen (denn der Leser wird einen einfachen, wenn auch ansprechenden jungen Menschen in ihm kennenlernen), sondern um der Geschichte willen, die uns in hohem Grade er-

\* © Stahl P., 2016

Peter Stahl Peter (peter.stahl@uni-wuerzburg.de), Department of German Linguistics, Julius-Maximilians-University (JMU) Würzburg in Bavaria, Am Hubland, Philosophiegebäude 4.E.12, 97074, Würzburg, Germany.

zählenswert scheint (wobei zu Hans Castorps Gunsten denn doch erinnert werden sollte, daß es *seine* Geschichte ist, und daß nicht jedem jede Geschichte passiert): diese Geschichte ist sehr lange her, sie ist sozusagen schon ganz mit historischem Edelrost überzogen und unbedingt in der Zeitform der tiefsten Vergangenheit vorzutragen.

### Ausdruck 2. [Mann 1987: 5]

Mehrere Dinge fallen auf: Erstens, unterschiedliche Wörter sind getrennt, was am Layout der Editionen liegt. Solche Differenzen sind zu erwarten und deuten nicht auf inhaltliche Unterschiede hin. Zweitens, in der Ausgabe von 2010 steht in der ersten Zeile ein Komma hinter dem Modalverb *wollen*, das in der Ausgabe von 1987 fehlt. Drittens, in Zeile 6 steht das Possessivpronomen *seine* in der Grundschrift, in der anderen Edition in kursiver Schrift; und viertens ist in Zeile 3 in der Ausgabe von 2010 von einem *jungen Mann* die Rede, in der von 1987 von einem *jungen Menschen*. Dieser Befund, der aus dem Vergleich des ersten Absatzes des Romans hervorgeht, hat mich neugierig gemacht. Denn wenn bereits in diesem kleinen Textausschnitt mehrere Unterschiede zwischen zwei Ausgaben festzustellen sind, wie viele mögen es dann noch auf den restlichen eintausend Seiten sein? Und lassen sich noch mehr Unterschiede finden, wenn man zu anderen Ausgaben greift? Vielleicht war es auch naiv zu glauben, dass alle Ausgaben des Romans unverändert dem Text der Urfassung (Ausgabe von 1924) folgen. Immerhin wird sie doch als die „einzige Textvorlage, die als wirklich autorisiert betrachtet werden kann“ [3, S. 47] bezeichnet. Aber „Lektoren griffen mit wechselndem Geschmack in Orthographie und Zeichensetzung ein [...]. Da wurden von Thomas Mann verkürzt wiedergegebene Zitate ergänzt und tatsächliche oder auch vermeintliche, sprachliche wie sachliche Fehler, verbessert“ [3, S. 48]. Ein Vergleich verschiedener Ausgaben erschien lohnenswert.

Hätte ich nicht zufällig die *Zauberberg*-Ausgaben gelesen, in denen *der junge Mensch* zum *jungen Mann* vertauscht worden war, hätte ich das Projekt, von dem ich hier berichte, nicht durchgeführt. Unter diesen Umständen wollte ich herausfinden, inwieweit sich mehrere Ausgaben dieses Romans voneinander unterscheiden. Die Abweichungen sollten in einem möglichst automatisierten Verfahren ermittelt und in einem für kritische Editionen üblichen Variantenapparat dargestellt werden. „Für die Edition eines Textes bieten sich der Theorie nach immer mehrere Modelle an; welches von diesen tatsächlich verwirklicht werden kann, hängt allein von der Quantität und Qualität der [...] Textzeugen ab“ [5, S. 71\*].

Es geht im Projekt nicht um eine inhaltliche Interpretation von Unterschieden in der Textgestalt; es geht nicht darum, ob ein Verlag den Text eines zeitgenössischen Romans gegenüber der Erstausgabe verändern darf oder soll. Ziel des Unternehmens ist es, zunächst eine Bestandsaufnahme vorzunehmen und die Varianten verschiedener Ausgaben zu belegen und damit auf der Textbasis der Erstausgabe 1924 die Vorstufe einer kritischen Edition zu erstellen.

### 1. Textgrundlage

Es galt daher zunächst, möglichst unterschiedliche Ausgaben – hinsichtlich Zeileneinteilung und Seitenlänge – zu beschaffen; unveränderte Nachdrucke sollten also in die Untersuchung nicht mit einfließen. Nach der Akquirierung lagen mehrere Veröffentlichungen von Thomas Manns *Zauberberg* vor, wobei die Zusammenstellung weder den Anspruch erhebt, vollständig zu sein, noch repräsentativ, sondern allein dem exemplarischen Vorgehen dient.

Erscheinungsjahr	Verlag, Auflage	Ausschnitt
1. 1924	S. Fischer Verlag Berlin, 1.-10. Auflage	S. 9–27
2. 1926	S. Fischer Verlag Berlin, 1.-21. Auflage	S. 9–23
3. 1956	G. B. Fischer Verlag Berlin, 2. Auflage	S. 3–14

4. 1968	Aufbau-Verlag Berlin und Weimar	S. 5–21
5. 1970	Fischer Bucherei, Frankfurt am Main und Hamburg, 48.-57. Tausend	S. 5–17
6. 1974	S. Fischer Verlag, Frankfurt am Main, Zweite, durchgesehene Auflage (Gesammelte Werke in 13 Bänden, Band III)	S. 9–24
7. 1981 [?]	Lizenzausgabe mit Genehmigung des S. Fischer Verlages Frankfurt/ M. für [...] die Bertelsmann Club-GmbH, Gutersloh [o. J.]	S. 19–31
8. 1987	Aufbau-Verlag Berlin und Weimar, 6. Aufl.	S. 5–20
9. 1999	<a href="http://www.flc.kyushu-u.ac.jp/~hgmc/">http://www.flc.kyushu-u.ac.jp/ ~hgmc/</a>	Z. #0300907 – #0302426
10. 2002	S. Fischer Verlag Frankfurt a. M., Große kommentierte Frankfurter Ausgabe, Bd. 5.1	S. 9–25
11. 2010	Fischer Taschenbuch Verlag, 3. Auflage	S. 9–24

Sieben Ausgaben sind im Fischer-Verlag erschienen, zwei im DDR-Aufbau-Verlag (Nr. 4, 8), eine im Bertelsmann-Club (Nr. 7) und eine im Internet (Nr. 9), die von der Fakultät für Sprachen und Kulturen der Kyushu Universität (Japan) angeboten wird. „Vom Verlag geduldet, hat sich diese Seite zur zentralen Anlaufstelle für alle Zitatensammler und Forscher entwickelt“ [2]. Mehr zu dieser problematischen Textausgabe unten. In der rechten Spalte wird angezeigt, welche Seiten aus jeder Ausgabe in die Untersuchung eingeflossen sind.

Es wird gezeigt, wie aus den Einzelausgaben Schritt für Schritt eine textkritische Gesamtedition entstehen kann, die alle Varianten aufweist. Das Verfahren wurde in Seminaren mit Studenten der Germanistik und der Digital Humanities an der Uni Würzburg entwickelt und erprobt. Dabei kamen unterschiedliche EDV-Programme zum Einsatz, weswegen es nicht wichtig erscheint, einzelne Programmteile genau zu dokumentieren. Es kommt darauf an, die Vorgehensweise als Strategie aufzuzeigen, so dass auch größere Textabschnitte eines beliebigen Romans mit geringem Aufwand verarbeitet werden können.

Aus allen Exemplaren der *Zauberberg*-Ausgaben wurden *Vorsatz* sowie die beiden ersten Unterkapitel (*Ankunft* und *Nr. 34*) des ersten Kapitels gescannt und mit Hilfe von OCR-Programmen in Textdateien überführt. Etliche derartige kommerzielle (u.a. Omnipage und FineReader<sup>1</sup>) oder nicht-kommerzielle (u.a. Free-OCR, CognitiveForms, FreeOCR<sup>2</sup>) Programme stehen dafür zu Verfügung. Da die Zuverlässigkeit der Aufbereitung in eine Textdatei maßgeblich von der Qualität der Bildvorlage abhängt (Auflösung, Farbe, Druckbild) lassen sich keine allgemeinen Aussagen darüber treffen, wie viele Ausgabefehler erzeugt werden. Es wird berichtet [1, S. 153], dass die Texterkennungsrate bei bis zu 99,8% aller Zeichen liege. Auf Internetseiten wird ausgeführt, „dass 99,9% richtige Erkennung eine Fehlerquote von 1 Promille (1 Fehler bei 1.000 Zeichen) bedeutet. Ganz praktisch heißt das, dass Sie bei deutschsprachigen Texten in rund jedem 150. Wort einen Fehler haben“ [4]. Selbst wenn eine noch höhere Erkennungsrate erzielt wird, muss immer damit gerechnet werden, dass aufgrund der OCR-Erkennung Fehler in den Text hineingelangen, die sich nicht in der gedruckten Vorlage befinden.

Betrachtet man die Datei der Ausgabe von 2010, die nach der OCR-Aufbereitung vorliegt, lassen sich eine Reihe wichtiger Beobachtungen machen, die der Ausschnitt der Zeilen 151 bis 213 zeigt (s. Ausdruck 3):

- Silbentrennungen sind nicht aufgelöst;
- Leerzeilen (Z. 153 und Z. 211), die das OCR-Programm eingesetzt hat, markieren den Beginn eines neuen Abschnitts, der im Buch am Einzug von linken Rand her ersichtlich ist;
- alleinstehende Nummern (Z.164 und Z.200) stellen die Paginierungen dar, die sich jeweils am Seitenende befinden;

- zu Beginn der Zeilen 185 und 192 stehen offensichtlich Auflösungsfehler, die das OCR-Programm verursacht hat.

151 Fernluft ist so ein Trank, und sollte sie weniger gründlich wir-  
152 ken, so tut sie es dafür desto rascher.

153

154 Dergleichen erfuhr auch Hans Castorp. Er hatte nicht beab-  
155 sichtigt, diese Reise sonderlich wichtig zu nehmen, sich innerlich [...]

162 jüngst Zurückliegenden, seinem Examen, und dem unmittelbar

163 Bevorstehenden, seinem Eintritt in die Praxis bei Tunder &

164 12

165 Wilms (Schiffswerft, Maschinenfabrik und Kesselschmiede), be-

166 schäftigt und über die nächsten drei Wochen mit soviel Unge-

167 duld hinweggeblickt, als seine Gemütsart nur immer zuließ. [...]

184 überall und nicht so wie jetzt im Klimmen daran erinnert sein, in

185 v/elchen unangemessenen Sphären man sich befand. Er sah hin-

186 aus: der Zug wand sich gebogen auf schmalem Paß; man sah die [...]

190 Fichten zwischen Felsblöcken gegen einen steingrauen Himmel

191 empor. Stockfinstere Tunnel kamen, und wenn es wieder Tag

192 7/urde, taten weitläufige Abgründe mit Ortschaften in der Tiefe

193 sich auf. Sie schlossen sich, neue Engpässe folgten, mit Schnee- [...]

197 man nicht mehr wußte, wie man fuhr, und sich der Himmelsge-

198 genden nicht länger entsann. Großartige Fernblicke in die heilig-

199 phantasmagorisch sich türmende Gipfelwelt des Hochgebirges,

200 13

201 in das man hinan- und hineinstrebte, eröffneten sich und gingen

202 dem ehrfürchtigen Auge durch Pfadbiegungen wieder verloren. [...]

208 deckte. Das ging vorüber. Er sah, daß der Aufstieg ein Ende ge-

209 nommen hatte, die Paßhöhe überwunden war. Auf ebener Tal-

210 sohle rollte der Zug nun bequemer dahin.

211

212 Es war gegen acht Uhr, noch hielt sich der Tag. Ein See er-

213 schien in landschaftlicher Ferne, seine Flut war grau, und

### Ausdruck 3. OCR-Ergebnis

#### 2. Textaufbereitung

Für unser Projekt ist diese Textdatei so weiterzuverarbeiten, dass alle Abschnittsanfänge markiert, Leerzeilen entfernt sowie Silbentrennungen gekennzeichnet und aufgelöst werden. Die Seitenzahlen sollen verwendet werden, um eine Nummerierung aller Zeilen zu gewinnen. Die erste Zeile hinter der Seitenzahl 13 (Z. 200), die mit *in das man* (Z. 201) beginnt, soll die Seiten-Zeilen-Nummer 14.1, die nächste, die mit *dem ehrfürchtigen* (Z. 202) beginnt, soll die Seiten-Zeilen-Nummer 14.2 u.s.w. erhalten. Für das Aufbereitungsverfahren ist es unerheblich, welche Programme dafür zum Einsatz kommen, können solche Aufgaben doch beispielsweise in PHP, Perl oder Tustep<sup>1</sup>, was hier zum Einsatz kommt, programmiert werden. Entscheidend ist, dass das vorgegebene Ziel einer Vergleichbarkeit mehrerer Fassungen eines Texts am Ende erreicht wird. Sind die Dateisätze zudem wie hier mit Seiten-Zeilen-Nummern versehen, die direkt auf die entsprechenden Stellen in der gedruckten Version verweisen, wird der Zugang zu allen Stellen im Text entscheidend erleichtert. Ausdruck 4 zeigt denselben Ausschnitt wie oben in der neuen Form:

12.24|Fernluft ist so ein Trank, und sollte sie weniger gründlich wir|ken,

12.25|so tut sie es dafür desto rascher.

12.26|<p>Dergleichen erfuhr auch Hans Castorp. Er hatte nicht beab|sichtigt,

12.27|diese Reise sonderlich wichtig zu nehmen, sich innerlich [...]

12.33|Gedankenkreise befangen gewesen, hatte sich mit dem

12.34|jüngst Zurückliegenden, seinem Examen, und dem unmittelbar

12.35|Bevorstehenden, seinem Eintritt in die Praxis bei Tunder &

- 13.1 |Wilms (Schiffswerft, Maschinenfabrik und Kesselschmiede), be|schäftigt  
 13.2 |und über die nächsten drei Wochen mit soviel Unge|duld  
 13.3 |hinweggeblickt, als seine Gemütsart nur immer zuließ. [...]  
 13.20|überall und nicht so wie jetzt im Klimmen daran erinnert sein, in  
 13.21|v/elchen unangemessenen Sphären man sich befand. Er sah hin|aus:  
 13.22|der Zug wand sich gebogen auf schmalem Paß; man sah die [...]  
 13.26|Fichten zwischen Felsblöcken gegen einen steingrauen Himmel  
 13.27|empor. Stockfinstere Tunnel kamen, und wenn es wieder Tag  
 13.28|7/urde, taten weitläufige Abgründe mit Ortschaften in der Tiefe  
 13.29|sich auf. Sie schlossen sich, neue Engpässe folgten, mit Schnee|resten [...]  
 13.33|man nicht mehr wußte, wie man fuhr, und sich der Himmelsge|genden  
 13.34|nicht länger entsann. Großartige Fernblicke in die heilig|phantasmagorisch  
 13.35|sich türmende Gipfelwelt des Hochgebirges,  
 14.1 |in das man hinan- und hineinstrebte, eröffneten sich und gingen  
 14.2 |dem ehrfürchtigen Auge durch Pfadbiegungen wieder verloren. [...]  
 14.8 |Das ging vorüber. Er sah, daß der Aufstieg ein Ende ge|nommen  
 14.9 |hatte, die Paßhöhe überwunden war. Auf ebener Tal|sohle  
 14.10|rollte der Zug nun bequemer dahin.  
 14.11|<p>Es war gegen acht Uhr, noch hielt sich der Tag. Ein See er|schien  
 14.12|in landschaftlicher Ferne, seine Flut war grau, und

#### Ausdruck 4. Markierte Abschnitte sowie Trennstellen

Absätze beginnen mit einem so genannten *P-Tag* (<p>), die ehemaligen Trennstellen sind mit einem Längsstrich markiert. Die Wörter, die einen solchen enthalten, werden im anschließenden Arbeitsschritt aus dem Text herausgezogen und in einer alphabetisch sortierten Liste zusammengefasst. Daraus einige Beispiele in einer verkürzten Wiedergabe:

- beab|sichtigt (1\*) 12.26  
 be|schäftigt (1\*) 13.1  
 er|schien (1\*) 14.11  
 ge|nommen (1\*) 14.8  
 heilig|phantasmagorisch (1\*) 13.34  
 hin|aus (1\*) 13.21  
 Rük|ken (1\*) 25.14  
 Schnee|resten (1\*) 13.29  
 wir|ken (1\*) 12.24

#### Ausdruck 5. Liste der ehemals getrennter Wörter

Die meisten Trennungen werden erwartungsgemäß richtig aufgelöst. Einige jedoch sind problematisch. Bei *heilig|phantasmagorisch* beispielsweise, das nur einmal im Text vorkommt, und zwar auf Seite 13 am Ende der Zeile 34, liegt kein Silbentrennstrich vor, sondern ein Bindestrich, den man erhalten muss. Das wird der Vergleich mit den anderen Ausgaben des Romans noch belegen. Es muss in der Zielform dieser Längsstrich zu einem Bindestrich korrigiert werden. Die Trennstelle *k|k* in *Rük|ken* muss zu *ck* gebessert werden, in *Sak|ko* (was in dieser Auflage zufälligerweise nicht getrennt ist) hingegen nicht. Nach der Überprüfung aller Längsstriche – und gegebenenfalls der Korrektur – können diese Markierungen entfernt werden. Auf diese Weise wird sichergestellt, dass das OCR-Programm bei der Aufhebung der Trennungen keine Fehler erzeugt.

### 3. Textvergleich und Variantenapparat

Vergleicht man nun die Texte zweier Ausgaben miteinander, ergibt sich ein Vergleichsprotokoll, das allein die Unterschiede anzeigt. Wir sehen die Abweichungen der Ausgaben von 2010 (oben) und 1987 (unten) voneinander.

- 9.2 <p>Die Geschichte Hans Castorps, die wir erzählen wollen, - nicht  
 -->

- 5.2 <p>Die Geschichte Hans Castorps, die wir erzählen wollen - nicht  
-----  
9.4 auch ansprechenden jungen Mann in ihm kennenlernen), sondern  
--> \*\*++++  
5.4 auch ansprechenden jungen Menschen in ihm kennenlernen), sondern  
-----  
9.22 der Zeit, - eine Aussage, womit auf die Fragwürdigkeit und  
--> -  
5.22 der Zeit - eine Aussage, womit auf die Fragwürdigkeit und  
-----  
11.29 auch diesen Namen hier gleich zu nennen -, seinem Wintermantel,  
--> --  
7.30 auch diesen Namen hier gleich zu nennen seinem Wintermantel,  
-----  
12.1 einem kleinen grau gepolsterten Abteil; er saß bei niedergelassenem  
--> -  
8.2 einem kleinen graugepolsterten Abteil; er saß bei niedergelassenem  
-----  
12.6 Buch namens >Ocean steamships<, worin er zu Anfang der Reise  
--> \* \*  
8.7 Buch namens "Ocean steamships", worin er zu Anfang der Reise  
-----  
12.19 Gleich ihr erzeugt er Vergessen; er tut es aber, indem er  
--> \*  
8.21 Gleich ihr erzeugt er Vergessen, er tut es aber, indem er  
-----  
12.32 hatte liegen lassen müssen. Noch gestern war er völlig in dem gewohnten  
--> -  
8.35 hatte liegenlassen müssen. Noch gestern war er völlig in dem gewohnten  
-----  
13.21 v/elchen unangemessenen Sphären man sich befand. Er sah hinaus:  
--> \*-  
9.22 welchen unangemessenen Sphären man sich befand. Er sah hinaus:  
-----  
13.28 7/urde, taten weitläufige Abgründe mit Ortschaften in der Tiefe  
--> \*-  
9.29 wurde, taten weitläufige Abgründe mit Ortschaften in der Tiefe  
-----[...]  
-----  
16.21 ein halbes Jahr wird es wohl ziemlich sicher noch dauern."  
--> \*\_  
12.19 ein halbes Jahr wird es wohl ziemlkh sicher noch dauern."

#### Ausdruck 6. Vergleichsprotokoll der Ausgaben 2010 und 1987

Mit einem Pfeil gekennzeichnet befindet sich zwischen beiden Fassungen eine Interpretation des Unterschieds: Zeichen für Zeichen signalisiert ein Minuszeichen, dass in der unteren Zeile etwas fehlt, ein Pluszeichen, dass etwas hinzugefügt ist, und ein Asterisk (\*) steht für einen Unterschied. Als erste Belege für eine Differenz, von denen eingangs die Rede war, findet sich das gesetzte bzw. fehlende Komma hinter *wollen* (9.2; 5.2) sowie die Variante *Mann / Menschen* (9.4; 5.4). Außer unterschiedlicher Zeichensetzung bemerken wir *graugepolsterten* (12.1) gegenüber *grau gepolsterten* (8.2) und *liegen lassen* (12.32) gegenüber *liegenlassen* (8.35). Hierbei handelt es sich um tatsächliche Varianten im Text. Die Zeichenfolge *v/elchen* (13.21) stellt einen OCR-Auflösungsfehler dar; an dieser Stelle (Seite 13, Zeile 1, 1. Wort) enthält die Ausgabe von 1987 das korrekte Relativpronomen

welchen. Anstelle von *7/wurde* steht in der korrespondierenden Position (13.28, 1. Wort) das Hilfsverb *wurde*. Es gibt aber auch den entgegengesetzten Fall: in der Datei des Texts von 2010 steht in 16.21 als 7. Wort der Intensivierer *ziemlich*; *ziemlkh* ist ein OCR-Fehler in TMZB1987.

Werden die Unterschiede zwischen diesen beiden Texten nicht als Vergleichsprotokoll (s. Ausdruck 6), sondern als Wortliste ausgegeben, erhalten wir eine Basis für eine automatische Korrektur der Fehlerstellen des 2010-Texts. Die Liste zeigt – wortbezogen – die Abweichungen:

9.2,8=wollen  
 9.4,4=Menschen  
 9.22,2=Zeit  
 11.29,8-  
 12.1,3-4=graugepolsterten  
 12.6,3="Ocean  
 12.6,4=steamships",  
 12.19,5=Vergessen,  
 12.32,2-3=liegenlassen  
 13.21,1=welchen  
 13.28,1=wurde, [ ]...  
 16.21,7=ziemlkh

#### Ausdruck 7. Liste der Textunterschiede zwischen den Ausgaben 2010-1987

Wenn es nun darum geht, die OCR-Fehler zu beheben, kann die Liste auf zwei Zeilen, nämlich auf 13.21,1=welchen und 13.28,1=wurde, reduziert werden. Mit Hilfe des Tustep-Programm-Bausteins zur automatisierten Korrektur von Dateien werden an den angegebenen Seiten-Zeilen-Positionen die Zeichenfolgen eingetragen, die sich hinter dem Gleichheitszeichen befinden. Diese Vorgehensweise hat den entscheidenden Vorteil, dass keine Handarbeit erfolgt und gesichert die richtige Textstelle geändert wird. Eine erneute Fehlerquelle wird also vermieden. Auf diese Weise werden alle Texte untereinander verglichen und korrigiert, um ORC-Fehler zu tilgen. Der Internet-Text aus dem Jahr 1999 hingegen enthält keine OCR-bedingten Abweichungen; vielmehr sind hier die Probleme ganz anders gelagert. Ausdruck 8 veranschaulicht es eindrücklich:

10.15 <p>Wir werden sie ausführlich erzählen, genau und gründlich, -  
 --> \*           \*+       \*+ +       \*+ +

1.9 <p>wir werden sie ausfuehrlich erzaehlen, genau und gruendlich, -

10.16 denn wann wäre je die Kurz-       oder Langweiligkeit einer Geschichte  
 -->       \*+           \*+++++++

1.9 denn wann waere je die Kurzweilichkeit oder Langweiligkeit einer Geschichte

10.17 abhängig gewesen von dem Raum und der Zeit, die  
 --> \*+                   +

1.9 abhaengig gewesen von dem Raum und der Zeit, die

10.18 sie in Anspruch nahm? Ohne Furcht vor dem Odium der  
 -->                   - + \*

1.9 sie in Anspruch nahm! ohne Furcht vor dem Odium der

10.19 Peinlichkeit, neigen wir vielmehr der Ansicht zu, daß nur das  
 -->       +                   + + +\*

1.10 Peinlichkeit , neigen wir vielmehr der Ansicht +zu , dass nur das

10.20 Gründliche wahrhaft unterhaltend sei.  
 --> \*+                   +

1.10 Gruendliche wahrhaft unterhaltend sei .

#### Ausdruck 8. Vergleichsprotokoll der Ausgaben 1924 und 1999

Zum Zweck der Vergleichbarkeit ist die Edition von 1999 nur um die Zeilennummerierung bereinigt und mit <p>-Tags versehen. U.a. lässt sich an Unterschieden feststellen:

- ein Spatium vor und nach Satzzeichen,
- Großschreibung nur bei Substantiven,
- Umlaute sind aufgelöst zu *ae, oe, ue* und *ß* zu *ss*,
- Auslassungsstriche sind durch Vollformen ersetzt (*Kurzweilichkeit* in 1.9), hier sogar fehlerhaft.

Angesichts dieser weitreichenden Eingriffe darf die Frage gestellt werden, ob ernsthaften Thomas-Mann-Forschern ein solcher Text aus dem Internet wirklich von Nutzen ist. Zumindest müssen für unser Projekt die systematischen Änderungen beim weiteren Vergleich aller Ausgaben unterdrückt werden, so dass nur die tatsächlichen Text-Varianten übrig bleiben. Ein Vergleich aller verwendeten Ausgaben ergibt diesen Befund:

9.1 Vorsatz

[GKFA] VORSATZ  
[1926] VORSATZ  
[1956] VORSATZ  
[1968] VORSATZ  
[1970] =====  
[1974] VORSATZ  
[OJ ] VORSATZ  
[1987] =====  
[WWW ]  
[2010] =====

9.2 Die Geschichte Hans Castorps, die wir erzählen wollen, -

[GKFA] == ===== =  
[1926] == ===== =  
[1956] == ===== wollen =  
[1968] == ===== wollen =  
[1970] == ===== =  
[1974] == ===== =  
[OJ ] == Geschichte, == ===== =  
[1987] == ===== wollen =  
[WWW ] == ===== =  
[2010] == ===== =

9.4 wenn auch ansprechenden jungen Menschen in ihm kennenlernen),

[GKFA] =====  
[1926] =====  
[1956] =====  
[1968] =====  
[1970] =====  
[1974] =====  
[OJ ] =====  
[1987] =====  
[WWW ] =====  
[2010] ===== Mann =====

**Ausdruck 9. Vergleichsprotokoll aller Ausgaben**

Dieses Protokoll zeigt übersichtlich die Verhältnisse in jedem Text an. Die nummerierten Zeilen entstammen der Erstausgabe von 1924, gefolgt vom Text der Großen kommentierten Frankfurter Ausgabe (*GKFA*). *OJ* repräsentiert die Bertelsmann-Ausgabe aus den 1980er Jahren (Erscheinungsjahr fehlt); *WWW* weist auf den Internet-Text hin. Gleichheitszeichen bedeuten identischen Inhalt; die Überschrift *Vorsatz* fehlt bei *WWW*, der Name *Hans Castorps* in *OJ*; die anderen Unterschiede sind schon mehrfach angesprochen worden. Trägt man die Varianten wieder wortpositionsgenau in eine Liste ein, ergibt sich für den Romananfang folgende Aufstellung:

9.1[Vorsatz ]- (WWW )  
 9.1[Vorsatz ]=VORSATZ (GKFA)  
 9.1[Vorsatz ]=VORSATZ (1926)  
 9.1[Vorsatz ]=VORSATZ (1956)  
 9.1[Vorsatz ]=VORSATZ (1968)  
 9.1[Vorsatz ]=VORSATZ (1974)  
 9.1[Vorsatz ]=VORSATZ (OJ )  
 9.2,2[Geschichte ]=Geschichte, (OJ )  
 9.2,3-4[Hans Castorps, ]- (OJ )  
 9.2,8[wollen, ]=wollen (1956)  
 9.2,8[wollen, ]=wollen (1968)  
 9.2,8[wollen, ]=wollen (1987)  
 9.4,5[Menschen ]=Mann (2010)

#### Ausdruck 10. Wortgenaue Unterschiedsliste

Aufgrund dieser Liste kann der Basistext von 1924 so aufbereitet werden, dass zu jedem Wort, zu dem es eine oder mehrere Varianten gibt, ein Eintrag erzeugt wird. In den Basistext werden die Unterschiede mit Kennungen hineinmontiert, so dass damit schließlich das Setzen einer Edition mit Variantenapparat erfolgen kann. Werfen wir einen Blick in diese Textstufe:

```
<h1>VORSATZ</h1>
<a1>Vorsatz <i>fehlt</i><s>WWW</s>; VORSATZ <s>GKFA, 1926, 1956, 1968, 1974,
OJ</s></a1>
<p>Die Geschichte <a1><i>nach</i> Geschichte <i>Komma</i><s>OJ</s></a1>
Hans <a2>Hans Castorps, <i>fehlt</i><s>OJ</s></a2> Castorps, </ee> die wir
erzählen wollen, <a1><i>nach</i> wollen <i>Komma fehlt</i><s>1956, 1968,
1987</s></a1> - nicht um seinetwillen (denn der Leser wird einen einfachen,
wenn auch ansprechenden jungen Menschen <a1>Menschen / Mann s>2010</s></a1>
in ihm kennenlernen), sondern um der Geschichte willen, die uns in hohem
```

#### Ausdruck 11. Basistext mit Varianten

In Anlehnung an HTML-/XML-Auszeichnungskonventionen markieren *Tags* Inhaltliches:

```
<h1>...</h1>  Überschrift
<a1>...</a1>  Apparat, der sich auf ein einziges Wort bezieht
<i>...</i>    kursive Schrift
<s>...</s>   Sigle
<p>          Beginn eines Abschnitts
<a2>...</a2> Apparat, der sich auf eine Wortgruppe bezieht
</ee>       Ende der Wortgruppe, zu der es einen Apparat gibt
```

#### 4. Textedition

Der mit solchen *Tags* versehene Text wird im letzten Verarbeitungsschritt mit dem Tustep-Satzprogramm gesetzt. Hierbei gilt es, Layout-Festlegungen zu treffen u.a. hinsichtlich der Seitengröße, der Grundschriftart und -größe, der Überschriftengröße, des Zeilenzählers. Das Satzergebnis sieht so aus:

#### Vorsatz

Die Geschichte Hans Castorps, die wir erzählen wollen, – nicht um seinetwillen (denn der Leser wird einen einfachen, wenn auch ansprechenden jungen Menschen in ihm kennenlernen), sondern um der Geschichte willen, die uns in hohem Grade erzählenswert scheint (wobei zu Hans Castorps Gunsten denn doch erinnert werden sollte, daß es seine Geschichte ist, und daß nicht jedem jede Geschichte passiert): diese Geschichte ist sehr lange her, sie ist sozusagen schon ganz mit historischem Edelrost überzogen und unbedingt in der Zeitform der tiefsten Vergangenheit vorzutragen.

Das wäre kein Nachteil für eine Geschichte, sondern eher ein Vorteil; denn 10  
Geschichten müssen vergangen sein, und je vergangener, könnte man sagen,  
desto besser für sie in ihrer Eigenschaft als Geschichten und für den Erzähler,  
den raunenden Beschwörer des Imperfekts. Es steht jedoch so mit ihr, wie es  
heute auch mit den Menschen und unter diesen nicht zum wenigsten mit den  
Geschichtenerzählern steht: sie ist viel älter als ihre Jahre, ihre Betagtheit ist 15  
nicht nach Tagen, das Alter, das auf ihr liegt, nicht nach Sonnenumläufen zu  
berechnen; mit einem Worte: sie verdankt den Grad ihres Vergangenseins  
nicht eigentlich der Zeit, – eine Aussage, womit auf die Fragwürdigkeit und  
eigentümliche Zwiennatur dieses geheimnisvollen Elementes im Vorbeigehen  
angespielt und hingewiesen sei. 20

Um aber einen klaren Sachverhalt nicht künstlich zu verdunkeln: die hoch-  
gradige Verflorenheit unserer Geschichte rührt daher, daß sie vor einer ge-  
wissen, Leben und Bewußtsein tief zerklüftenden Wende und Grenze spielt...  
Sie spielt, oder, um jedes Präsens geflissentlich zu vermeiden, sie spielte und  
hat gespielt vormals, ehemals, in den alten Tagen, der Welt vor dem großen 25  
Kriege, mit dessen Beginn so vieles begann, was zu beginnen wohl kaum  
schon aufgehört hat. Vorher also spielt sie, wenn auch nicht lange vorher.  
Aber ist der Vergangenheitscharakter einer Geschichte nicht desto tiefer, voll-  
kommener und märchenhafter, je dichter »vorher« sie spielt? Zudem könnte  
es sein, daß die unsrige mit dem Märchen auch sonst, ihrer inneren Natur 30  
nach, das eine und andre zu schaffen hat.

Wir werden sie ausführlich erzählen, genau und gründlich, – denn wann  
wäre je die Kurz- oder Langweiligkeit einer Geschichte abhängig gewesen  
von dem Raum und der Zeit, die sie in Anspruch nahm? Ohne Furcht vor  
dem Odium der Peinlichkeit, neigen wir vielmehr der Ansicht zu, daß nur das 35  
Gründliche wahrhaft unterhaltend sei.

1 Vorsatz fehlt WWW; VORSATZ GKFA, 1926, 1956, 1968, 1974, OJ 2 nach Geschichte  
Komma OJ Hans Castorps, fehlt OJ nach wollen Komma fehlt 1956, 1968, 1987 4 Men-  
schen / Mann 2010 6 nach ist Komma fehlt 1926, 1956 17 Worte: / Wort: 1956 18 nach Zeit  
Komma fehlt 1956, 1968, 1987 29 spielt? / spielt! WWW 30 inneren fehlt 1956 31 andre /  
andere 1926 32 nach gründlich Komma fehlt 1956 33 Kurz- / Kurzweiligkeit WWW  
34 nahm? / nahm! WWW

## Ausdruck 12. Kritische Edition

Abgedruckt ist hier die erste Seite (von 12) der neuen Edition, die den Text der Erstausgabe von 1924 von Anfang an bis zum Ende des zweiten Unterkapitels bietet, und zwar mit allen Varianten, die sich durch einen Vergleich mit zehn anderen Ausgaben des Romans *Zauberberg* ergeben. Es sei darauf hingewiesen, dass die Darstellung der Varianten vollautomatisch erzeugt wurde und sich nun eine philologische Überprüfung anschließen müsste, ob die jeweilige Form und Schreibweise eines Apparateintrags der Sache nach angemessen ist. Ohne eine solche Nachbereitung könnte eine fundierte, textkritische Edition nicht entstehen.

Auch wenn an dieser Stelle kein vollständiger Abdruck aller 12 Seiten der neuen Edition erfolgen kann, sollen dennoch alle ermittelten Varianten dargeboten werden (s. Ausdruck 13). Um dies zu ermöglichen, wird der Text von 1924 in derselben Seiten- und Zeileneinteilung gesetzt, die der *GKFA*-Ausgabe von 2002 zugrunde liegt und die in Bibliotheken und in privatem Besitz zugänglich ist. Von der Edition mit Variantenapparat, die sich daraus ergibt, werden hier allein die Apparateinträge, ohne Romantext also, auf den nächsten beiden Seiten gezeigt; es wird verzeichnet, auf welche *GKFA*-Seite und -Zeile sich die Apparate beziehen. Die Sigle *GKFA* ist zur besseren Übersicht grau unterlegt, denn an diesen Stellen zeigt sich, worin sich die Erstausgabe und die kommentierte Edition von 2002 unterscheiden, obwohl dort doch betont wird, dass deren Texteinrichtung buchstaben- und zeichengenau der Erstausgabefolge [3, S. 53].

Wie wir sehen, ist dem nicht so. Es werden in der Ausgabe von 2002, die von M. Neumann „textkritisch durchgesehen“ wurde, zurecht fehlerhaft gesetzte Satzzeichen des Erstdrucks gebessert; dennoch finden sich 2 auffällige Abweichungen: Mag man die Versalien der Überschriften (9.1, 10.1-2) noch als typographische Eigenheit ansehen, ist die Schreibung von *Scaletta-Gletscher* (19.9) statt

*Skaletta-Gletscher* überraschend, obwohl beschrieben wird, „Fremdwörter, Eigen- und Ortsnamen folgen der Schreibung“ des Erstdrucks [3, S. 53].

Auf eine weitere Besonderheit sei noch hingewiesen: In den DDR-Ausgaben, deren Textrevision auf E. Neumann zurückgeht (1968 und 1987), werden die Syntagmen *liegen lassen* (S.12, 33), *gut geschnittenen* (S. 16, 10) und *weiß lackierten* (S.22, 15) teilweise auch im Einklang mit anderen Editionen jeweils zu einem einzigen Wort zusammengefasst. Hier darf die Strategie vermutet werden, eine Zusammenschreibung der Getrennschreibung vorzuziehen – eine Strategie, die sich vielleicht durch den gesamten Romantext hindurch zieht. An die Überprüfung der Varianten könnten sich nun sprachwissenschaftliche oder literaturwissenschaftliche Analysen anschließen.

Aber auch ohne solche Untersuchungen muss sich der Leser immer vergegenwärtigen, dass eine beliebige Ausgabe eines zeitgenössischen Romans nicht die Urform, die der Autor verfasst hat, wiedergibt, sondern dass mit Abweichungen zu rechnen ist. Und dass die Frage, „Was lese ich hier überhaupt?“, auch im editorischen Sinn durchaus ihre Berechtigung hat.

Seite: 9

1 Vorsatz *fehlt* WWW; VORSATZ **GKFA**, 1926, 1956, 1968, 1974, OJ 3 nach Geschichte *Komma* OJ Hans Castorps, *fehlt* OJ nach wollen *Komma* *fehlt* 1956, 1968, 1987 5 Menschen / Mann 2010 8 nach ist *Komma* *fehlt* 1926, 1956 22 Worte: / Wort: 1956 23 nach Zeit *Komma* *fehlt* 1956, 1968, 1987

Seite: 10

4 spielt? / spielt! WWW 6 inneren *fehlt* 1956 andre / andere 1926 7 nach gründlich *Komma* *fehlt* 1956 8 Kurz- / Kurzweilichkeit WWW 10 nahm? / nahm! WWW

Seite: 11

1 Erstes / ERSTES **GKFA**, 1926, 1956, 1968, 1974, OJ Kapitel / KAPITEL **GKFA**, 1926, 1956, 1968, 1974, OJ 1 f. Erstes Kapitel Ankunft *fehlt* WWW 2 Ankunft / ANKUNFT 1926 6 dort hinauf, / dorthinauf, 1956 13 solange / so lange 1926, 1956 30 nach nennen *Komma* *fehlt* 1956, 1968, 1970, 1974, 1987, WWW, 2010 nach – *Komma* 1956, 1968, 1970, 1974, 1987, WWW, 2010

Seite: 12

1 nach kleinen *Komma* 1956 grau gepolsterten / graugepolsterten 1926, 1968, 1987 6 »Ocean / 'Ocean WWW; »Ocean 1970, 1974, 2010 steamships«, / steamships', WWW; steamships<, 1970, 1974, 2010 13 Sorgen, / Sorgen,0 WWW 20 Vergessen; / Vergessen, 1968, 1970, 1974, 1987, WWW 22 nach versetzt *Komma* *fehlt* 1956 27 erfuhr / frufuhr WWW 32 wieder aufzunehmen, / wiederaufzunehmen, 1956 33 liegen lassen / liegenlassen 1926, 1956, 1968, 1970, 1974, 1987, WWW

Seite: 13

5 nach Kesselschmiede) *Komma* *fehlt* OJ 6 soviel / so viel 1926 12 nach herrschten *Komma* *fehlt* 1956 16 nach Unbekanntem *Komma* 1956 22 haben? / haben! WWW 33 schlössen / schlossen WWW

Seite: 14

1 Schründen / Schruenden WWW 4 nach fuhr *Komma* **GKFA**, 1926, 1968, 1970, 1974, 1987, WWW, 2010 7 hinan- / hinanstrebte WWW 20 stiegen / steigen WWW 24 hörte, / hörte; 1926

Seite: 15

4 nach verwirrt *Komma* *fehlt* 1956 5 nach Wiedersehens *Komma* 1956 Handtasche / Handschuhe OJ 7 »Ocean / 'Ocean WWW; »Ocean 1970, 1974, 2010 steamships« / steamships' WWW; steamships< 1970, 1974, 2010 10 die / der 1970 zwischen: Zusatz den OJ 22 »Berghof« / 'Berghof' WWW; »Berghof< 1970, 1974, 2010 23 »Platz« / 'Platz' WWW; »Platz< 1970, 1974, 2010 27 Kriegsveteran? / Kriegsveteran! WWW 31 nach sich *Komma* 1968, 1970, 1974, 1987, WWW, 2010 33 mir: Zusatz aber 1956, 1968, OJ, 1987 aber *fehlt* 1956, 1968, OJ, 1987

Seite: 16

10 gut geschnittenen / gutgeschnittenen 1926, 1956, 1968, 1970, 1974, OJ, 1987, WWW 23 Gedanken«, / Gedanken?» 1956

Seite: 17

7 bereit stand, / bereitstand, 1926 9 du / Du 1926, 1956, 1968, OJ, 1987, 2010 10 so viel / soviel 1956 15 schon / noch 1970, 2010 23 besser«, / besser,« 1926 er; / er: WWW 24 klingt / kling 1970 27 Geräusche. / Geräusche.« WWW; Geräusche.« **GKFA**, 1926, 1956, 1968, 1970, 1974, OJ, 1987, 2010 29 »Ja, / Ja, 2010

Seite: 18

4 Es / »Es **GKFA**, 1926, 1956, 1968, 1970, 1974, OJ, 1987, WWW, 2010 12 trotteten: Zusatz sanft 1956, OJ sanft *fehlt* 1956, OJ 25 nach Lehnen *Komma* *fehlt* 1956, 1968, 1987 nach auslud *Komma* *fehlt* 1926, 1956 33 Firmen / Firne 1926

Seite: 19

1 Bergesriesen? / Bergriesen? 1956, OJ 2 scheint.« / scheint. WWW 9 Skaletta-Gletscher. / Scaletta-Gletscher.  
GKFA, 1968, 1970, 1974, 1987, WWW, 2010 25 »Ausgezeichnet!« / »ausgezeichnet!? WWW

Seite: 20

6 dritten- / drittenmal WWW irgendeine / irgend eine OJ 13 nach sind.« Punkt .. OJ

Seite: 21

5 nach Schritt Komma fehlt 1970 6 nach Anfahrt Komma 1926, 1956 Berghof / »Berghof« 1968, 1987;  
'Berghof' WWW; »Berghof: 1970, 1974, 2010 8 Nr. / NR. 1926 Nr. 34 fehlt WWW 11 Telephon / Telefon 1956  
24 Stimme... / Stimme. 1956 26 wasch' / wasch 1926, 1956, 1968, 1970, 1974, OJ, 1987, WWW, 2010 28 Im /  
in WWW 31 soviel / so viel 1956

Seite: 22

6 Kokosläufer / Kokoslaeufener WWW 10 nach irgendwo Komma fehlt 1956 15 weiß lackierten / weißlackierten  
1968, 1970, 1974, 1987; weisslackierten WWW 18 Vierunddreißig. / vierunddreißig. 1956, 1968, 1970, 1974,  
1987, 2010 19 nach Ehepaar Komma fehlt 1956, 1968, 1987 25 nach friedlich Komma fehlt 1956 31 die / der  
1970, 1974, WWW nach gestellt Komma fehlt 1956, 1968, 1987

Seite: 23

4 es fehlt 1970, 2010 8 könntest. / konntest. WWW 12 Cold-cream / Cold Cream 1926; Coldcream 1956, 1968,  
1970, 1974, OJ, 1987, WWW 26 nach aus Komma fehlt 1956 27 Selbstverständlich / Selbstvers-tändlich 1926  
29 »Selbstvers-tändlich« / »selbstvers-tändlich« 1956

Seite: 24

1 macht / machte WWW nach Dingern Komma fehlt GKFA, 1956, 1968, 1970, 1974, 1987, WWW, 2010 5 da /  
das 1956 Cold-cream / Cold Cream 1926; Coldcream 1956, 1968, 1970, 1974, OJ, 1987, WWW 8 nach habe  
Komma fehlt 1926, 1956, 1968, 1970, 1974, 1987, WWW, 2010 9 nach – Komma 1926, 1956, 1968, 1970, 1974,  
1987, WWW, 2010 12 daran zu legen... / daranzulegen... 1956 18 nach echauffiert Komma fehlt 1956 da, / ja,  
WWW 27 nach kurzsichtig Komma 1956

Seite: 25

1 nach Husten Komma fehlt 1956 nach offenbar Komma fehlt 1956 3 gehört / geh=rt WWW 5 nach war  
Komma fehlt 1956 9 »Ja.« / »Ja«, 1956, 1968, 1970, 1974, OJ, 1987, WWW, 2010 »da / da WWW 14 Husten /  
Fusten WWW 15 bedenken.« / bedenken«, 1956, 1968, 1970, 1974, OJ, 1987, WWW, 2010 19 Jugend / Jugend«  
GKFA, 1968, 1970, 1974, 1987, WWW, 2010 20 nach Jugend« Komma GKFA, 1968, 1987 Bräune / »Braeune  
WWW; »Bräune GKFA, 1968, 1970, 1974, 1987, 2010 21 es / er 1926 23 nach nicht Komma fehlt 1956  
26 nach aussieht Komma fehlt 1956 – fehlt 2010 28 »Na.« / »Na«, 1956, 1968, 1970, 1974, OJ, 1987, WWW,  
2010

### Ausdruck 13. Nur Variantenapparat

#### Notes

<sup>1</sup> Webseiten: [omnipage.de.softonic.com](http://omnipage.de.softonic.com) a [www.abbyy.com/de-de/finereader](http://www.abbyy.com/de-de/finereader) [besucht am 01.06.2016].

<sup>2</sup> Webseiten: [www.free-ocr.com](http://www.free-ocr.com), [cognitiveforms.com](http://cognitiveforms.com), [www.paperfile.net](http://www.paperfile.net) [besucht am 01.06.2016].

<sup>3</sup> Webseiten: [php.net](http://php.net), [www.perl.org](http://www.perl.org), [www.tustep.de](http://www.tustep.de) [besucht am 01.06.2016].

#### References

1. Gulbins J. et al. *Elektronische Archivierungssysteme – Image-Management-Systeme, Dokument-Management-Systeme*. Berlin-Heidelberg: Springer, 1993 [in German].
2. Justus M. *Thomas Mann im Internet: Eine kommentierte Linksammlung*. Retrieved from: [http://www.thomasmann.de/thomasmann/im\\_internet/](http://www.thomasmann.de/thomasmann/im_internet/) (accessed 22.05.2016) [in German].
3. Neumann M. *Thomas Mann – Der Zauberberg – Roman. Kommentar*. Frankfurt am Main: S. Fischer, 2002 [in German].
4. Ring R. *OCR-Tricks*. Retrieved from: <http://www.experto.de/computer/ocr-tricks-beste-ergebnisse-bei-der-optischen-zeichenerkennung-am-pc.html> (accessed: 15.05.2016) [in German].
5. Steer G. et. Al. *Die Rechtssumme Bruder Betholds. Eine deutsche abecedarische Bearbeitung der Summa Confessorum des Johannes von Freiburg. Synoptische Edition der Fassungen B, A und C*. Tübingen: Niemeyer, 1987, Bd. 1 [in German].

### ЧТО ВЫ ЧИТАЕТЕ? ИЗДАТЕЛЬСКИЕ ВАРИАНТЫ ТЕКСТОВ СОВРЕМЕННЫХ РОМАНОВ

В данной статье описывается проект по сравнению 11 различных изданий известного романа Томаса Манна «Волшебная гора» (1924 г.). Целью является разработка компьютеризированного инструментария, который дает возможность автоматизированной обработки текстов только с небольшой долей их дополнительной ручной обработки. После сканирования секвенций из нескольких глав романа изображения преобразуются в текстовые файлы. В текстах исправлены дефисные написания, а т. н. OCR-ошибки исправляются автоматически. Для сравнения текстов все различия между ними выявляются: различия пунктуации, орфографии и лексические модификации. Текст первого издания романа Т. Манна 1924 года и варианты, найденные во всех других переизданиях, объединены в новое критическое издание текста.

**Ключевые слова:** Томас Манн, роман «Волшебная гора», машинная обработка текстов, сравнение текстов, критическое издание, ошибки.

Статья поступила в редакцию 06/VI/2016.  
The article received 06/VI/2016.

---

\* Шталь Петер (peter.stahl@uni-wuerzburg.de), кафедра немецкого языкознания, Баварский университет им. Юлиуса-Максимилиана (Вюрцбург, ФРГ), Am Hubland, Philosophiegebäude 4.E.12, 97074, Würzburg, Germany.