



Bibliothek des Trinity College, Dublin.

# Was kommt ans Licht, wenn Texte und Bilder digital analysiert werden?

## »Digital Humanities« – die empirische Wende in den Geisteswissenschaften

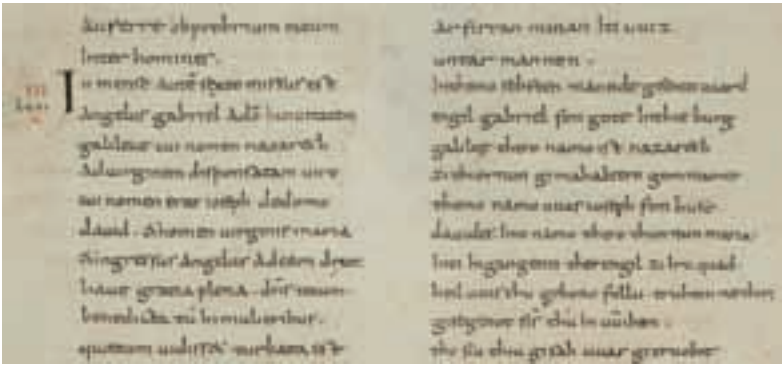
Dass Ludwig Börne seine langjährige vertraute Freundin Jeanette Wohl in einem seiner Briefe scherzhaft mit »Moppel« ansprach, brachte erst eine Spezialaufnahme ans Licht, denn die Stelle war eigentlich für die Nachwelt geschwärzt geworden. Nur ein kleines Beispiel, was die digitale Bearbeitung literarischer Originaltexte zutage fördern kann. Immer mehr Dokumente stehen auch in mehrschichtigen Aufnahmen für die wissenschaftliche Analyse zur Verfügung. Bald werden die bekanntesten schriftlichen Quellen vom Altertum bis zur Gegenwart nachhaltig erfasst und gespeichert und somit digital und online aufrufbar sein. Wie nutzen die Frankfurter Geisteswissenschaftler diese neuen fast grenzenlosen Chancen? Im LOEWE-Schwerpunkt »Digital Humanities« geht es darum, Textcorpora mit digitalen Methoden auszuwerten, zu vergleichen und mit Bilddaten zu verknüpfen.

von Jost Gippert

Als ich vor 15 Jahren an gleicher Stelle<sup>11/</sup> versucht habe, die Möglichkeiten zu umreißen, die die Digitalisierung textueller Quellen im Hinblick auf linguistische Untersuchungen eröffnen kann, war noch nicht abzusehen, wie schnell sich die elektronischen Medien in den Geisteswissenschaften durchsetzen und wie sehr sie unsere alltägliche Arbeit eines Tages dominieren würden. Während es damals noch recht mühsam war, per E-Mail mit Kollegen im Ausland zu korrespondieren, und man im Internet noch kaum irgendwelche brauchbaren Informationen finden konnte, ist die weltweite Kommunikation via Datennetz heute für uns alltäglich geworden, und Suchmaschinen wie Google erschließen uns selbst die obskursten In-

halte in Millisekunden. Neben den elektronisch aufbereiteten Versionen der großen Enzyklopädien steht uns die täglich weiter wachsende und ständig aktualisierte Wikipedia zur Verfügung, und immer mehr gedruckte Bücher und Aufsätze, die man früher in Bibliotheken einsehen oder per Fernleihe bestellen musste, sind nun online abrufbar. Tatsächlich ist es heute keine Utopie mehr, anzunehmen, dass die Überlieferungsmasse der uns bekannten schriftlichen Quellen vom Altertum bis zur Gegenwart in baldiger Zukunft sämtlich digital aufbereitet vorliegen und für Wissenschaftler wie Laien benutzbar sein werden.<sup>12/</sup>

Die Verfügbarkeit großer Datenmengen in elektronischer Form hat aber nicht nur unsere tägliche Ar-



<i>aufferre obprobrium meum</i>	<i>afirran minan itiuiz</i>
<i>inter homines.</i>	<i>untar mannon.</i>
<i>In mense autem sexto missus est</i>	<i>In themo sehsten manode gisentit uuard</i>
<i>angelus Gabriel a deo in civitatem</i>	<i>engil gabriel fon gote In thie burg</i>
<i>Galilee, cui nomen Nazareth,</i>	<i>galilee, thero namo ist nazareth,</i>
<i>ad virginem disponsatam viro,</i>	<i>zi thiornun gimahaltero gommanne,</i>
<i>cui nomen erat Ioseph, de domo</i>	<i>themo namo uuas Ioseph, fon huse</i>
<i>David, et nomen virginis Maria.</i>	<i>dauides. Inti namo thero thiornun maria.</i>
<i>Et ingressus angelus ad eam dixit:</i>	<i>Inti Ingangenti ther engil zi Iru quad.</i>
<i>habe gratia plena! dominus tecum,</i>	<i>heil uuis thu gebono follu. trutin mit thir,</i>
<i>benedicta tu in mulieribus.</i>	<i>gisegenot sis thû in uufban.</i>
<i>Quae cum vidisset, turbata est ...</i>	<i>Tho siu thiu gisah, uuas gituobit...</i>

schen archivierten Materialien wie Handschriften und Abbildungen und ihrer elektronischen Aufbereitung zum Zwecke wissenschaftlicher Ausgaben betreffen, wobei gezielt eruiert werden soll, wie und inwieweit man im Rahmen einer virtuellen (»online«-)Edition über die Möglichkeiten gedruckter Ausgaben hinauskommen kann. Letztlich soll auch die alltägliche Sprache, wie sie sich in den Datenstrukturen des Internets niederschlägt, im Rahmen des LOEWE-Schwerpunkts thematisiert werden, und zwar sowohl hinsichtlich der Analyse bestehender Corpora des Deutschen und Englischen als auch in Bezug auf sich stetig wandelnde, kollaborativ erstellte Textprodukte wie die Wikipedia (»Web 2.0«). Zielsetzungen und Lösungsansätze des LOEWE-Schwerpunkts »Digital Humanities« seien im Folgenden an zwei Beispielen illustriert.

**Lateinischer Einfluss auf das Althochdeutsche – Neue Einblicke durch digitalen Vergleich**

Es ist wohlbekannt, dass das älteste verfügbare Sprachmaterial des Deutschen, das in den althochdeutschen Texten vorliegt, fast ausschließlich aus Übersetzungen besteht, deren Vorlage lateinisch war. In manchen Fällen erkennt man dies unmittelbar an der Struktur der Handschriften, die den betreffenden Text

█ Lukas 1,23–26 in der lateinisch-althochdeutschen Evangelienharmonie Tatians (Cod. Sang. 56, S. 28). Die Textausgabe ist in der TITUS-Datenbank unter <http://titus.uni-frankfurt.de/texte/etcs/germ/ahd/tatianx/tatia.htm> online verfügbar; sie soll im LOEWE-Schwerpunkt im Hinblick auf die Abhängigkeit des althochdeutschen vom lateinischen Text ausgewertet werden.

beitsweise beeinflusst, sondern auch inhaltlich zu einem erheblichen Wandel in den geistes- und kulturwissenschaftlichen Fächern geführt, den man als eine »empirische Wende« bezeichnen kann. Während es früher für Linguisten, Literaturwissenschaftler oder Historiker kaum denkbar war, Hypothesen und Theorien am gesamten einschlägigen Material zu ver- oder falsifizieren, ist dies durch den Einsatz elektronischer Verfahren, das heißt durch die Schaffung von Datenpools und hierauf bezogenen digitalen Auswertungsmethoden, nunmehr möglich geworden. Gleichzeitig hat sich dadurch geradezu die methodologische Forderung ergeben, derartige Verfahren stetig weiterzuentwickeln und fachübergreifend anzuwenden, da sie zu einer wesentlichen Verbesserung der Forschungsergebnisse im Hinblick auf ihre Verlässlichkeit und ihre Tragfähigkeit führen.

**Im Visier von LOEWE: Corpusanalyse von der Antike bis zum Web 2.0**

Diesem Ziel hat sich jetzt ein Forschungsverbund beschrieben, der gemeinschaftlich von insgesamt zwölf Wissenschaftlern der Goethe-Universität, der Technischen Universität Darmstadt und des Freien Deutschen Hochstifts (Frankfurter Goethe-Museum) getragen und ab 2011 durch das Hessische Ministerium für Wissenschaft und Kunst als LOEWE-Schwerpunkt unter dem Titel »Digital Humanities – Integrierte Aufbereitung und Auswertung textbasierter Corpora« mit insgesamt 3,8 Millionen Euro gefördert wird. Gegenstände des Schwerpunkts sind unter anderem Corpora älterer Sprachen, die im Hinblick auf die wechselseitige Beeinflussung von Texten (als »Parallelcorpora«, etwa im Falle von Vorlage und Übersetzung), auf die Beziehungen zwischen den in ihnen repräsentierten Sprachen und Varietäten sowie auf den Wandel charakteristischer Begriffe und ihrer Bedeutung in der Geschichte (»Historische Semantik«) untersucht werden sollen. Des Weiteren wird sich der LOEWE-Schwerpunkt mit Fragen auseinandersetzen, die das Verhältnis zwi-



Palimpsest-Handschriften aus dem Katharinenkloster auf dem Sinai konnten Prof. Dr. Gippert und sein Team (hier Prof. Dr. Manana Tandaschwili) mit der Spezialkamera entschlüsseln. Eine ähnliche Kamera wird nun auch im Rahmen des LOEWE-Schwerpunkts »Digital Humanities« für vergleichbare Forschungsarbeiten eingesetzt.

enthalten. Dies gilt zum Beispiel für die sogenannte Evangelienharmonie des Syrer Tatian, deren althochdeutsche Fassung in einer heute in Sankt Gallen in der Schweiz aufbewahrten Handschrift des neunten Jahrhunderts dem entsprechenden lateinischen Text parallel gegenübergestellt ist.<sup>131</sup> █

Der althochdeutsche Übersetzer hat sich in diesem Fall sichtlich bemüht, den lateinischen Text so exakt wie möglich wiederzugeben. Dennoch sind, bedingt durch die unterschiedlichen Strukturen der beiden Sprachen, signifikante Unterschiede in der Textgestaltung erkennbar, die es zu untersuchen gilt, wenn man den Einfluss des Lateinischen auf das Althochdeutsche analysieren will. Das betrifft beispielsweise die Tatsache, dass das Lateinische über keinen Artikel verfügte, das Althochdeutsche aber durchaus (im obigen Beispiel *themo = dem, thie = die, thero = der, ther = der* █). Einen ähnlichen Fall stellt der Gebrauch der sogenannten Kopula dar, also des Verbs »sein« und seiner Formen.

Im Lateinischen kann dieses Verb häufig fehlen, insbesondere dann, wenn Aussagen über die Gegenwart gemacht werden (Präsens); im Althochdeutschen sind die entsprechenden Formen jedoch zu erwarten. ❸

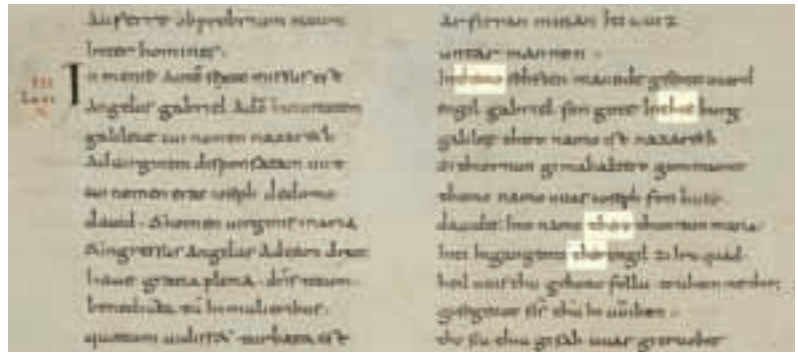
Ein weiterer charakteristischer Unterschied zwischen dem lateinischen und dem deutschen Text besteht darin, dass das Deutsche seit althochdeutscher Zeit zwei verschiedene Bildungen eines Passivs aufzuweisen hat, denen im Lateinischen nur eine Bildung entgegensteht. So finden wir im vorliegenden Beispiel *gisentit uuard* = »wurde gesendet« (mit Auxiliär »werden«) an der Stelle von lateinisch *missus est*, aber *uuas gitruobit* = »war betrübt« (eigentlich »getrübt«, mit Auxiliär »sein«) als Entsprechung des lateinischen *turbata est*; hinzu kommt die Verwendung unterschiedlicher Zeitformen (*uuard* = »wurde« und *uuas* = »war« gegenüber dem lateinischen *est* = »ist«). ❹

Noch weiter können die beiden Textversionen auseinandergehen, wenn man die Anordnung der einzelnen Wortformen in den Sätzen beziehungsweise Zeilen vergleicht. So zeigt das Lateinische zum Beispiel oft die Nachstellung von Attributen, die aus Adjektiven, Possessivpronomen oder Ordinalzahlen bestehen, wie in *obprobrium meum*, wörtlich »Schmach meine«, oder *in mense sexto*, wörtlich »in Monat sechstem«, während das Althochdeutsche die Voranstellung bevorzugt (*minan itiuuiz* beziehungsweise *in themo sehsten manode*). Auch innerhalb von analytischen Passivformen kann die Reihenfolge zwischen dem Partizip und dem Auxiliär divergieren (*turbata est* gegenüber *uuas gitruobit*). Letztlich gibt es bisweilen Elemente, die völlig unübersetzt geblieben sind, wie im gegebenen Beispiel die Konjunktion *autem*, die neuhochdeutsch »aber« entsprechen würde, in der Tatianübersetzung jedoch kein Gegenstück findet. ❺

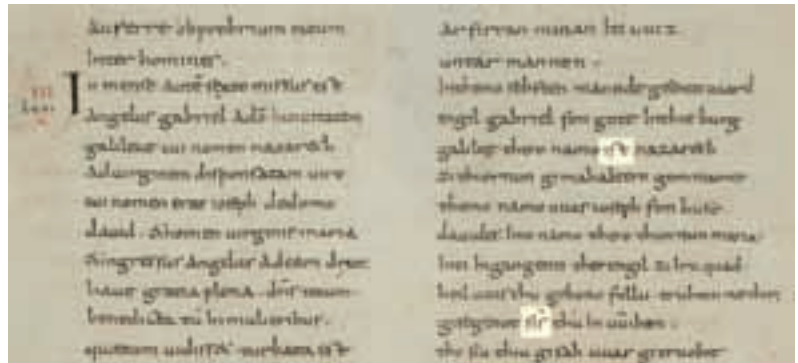
**»Parallelcorpora«: Grammatische Phänomene in wechselseitiger Abhängigkeit**

Alle derartigen Divergenzen sind nun von entscheidender Aussagekraft für das Sprachsystem: Gerade wenn ein Übersetzer versucht hat, eine Vorlage möglichst wortgetreu wiederzugeben, deuten die bestehenden Unterschiede an, welche Grenzen ihm dabei durch die grammatischen Regeln der Zielsprache gesetzt waren, und können so dazu dienen, diese Regeln – auch in einer nur spärlich überlieferten Sprache wie dem Althochdeutschen – klarer zu erfassen. Will man derartige Abweichungen aber systematisch untersuchen, so setzt dies weit mehr voraus als eine bloße digitale Erfassung der beiden Versionen.<sup>141</sup> Natürlich kann man einen elektronisch vorliegenden Text leicht etwa auf sämtliche Belege von lateinisch *autem* hin durchsuchen und sich notieren, was der althochdeutsche Text an der entsprechenden Stelle bietet. Dazu bedarf es streng genommen nicht einmal einer Digitalisierung, denn man kann dies auch an einer gedruckten Fassung des Textes vornehmen, und so sind entsprechende Untersuchungen teilweise auch schon vor hundert Jahren unternommen worden. Wenn es jedoch darum geht, die wechselseitige Abhängigkeit verschiedener lexikalischer und grammatischer Phänomene systematisch zu eruieren, stößt man mit derartigen Verfahren bald an seine Grenzen.

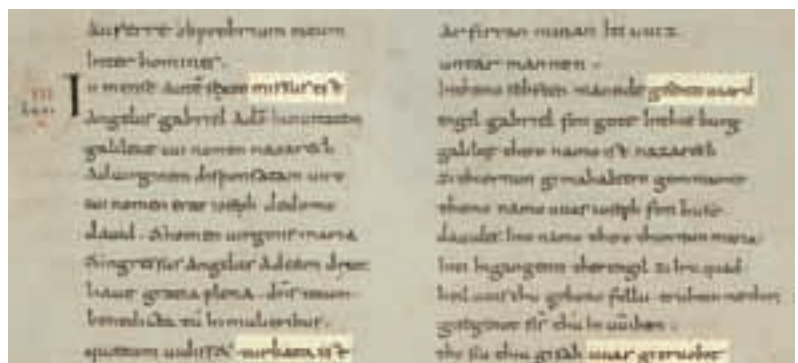
Was stattdessen erforderlich ist, ist zunächst eine durchgehende sogenannte Annotation der digitalen



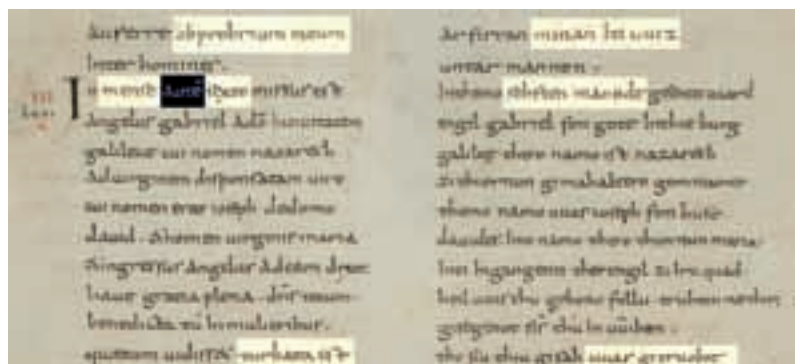
❷ In diesem Text sind Formen des bestimmten Artikels im Althochdeutschen (*themo*=dem, *thie*=die, *thero*=der, *ther*=der) hervorgehoben.



❸ Hervorgehoben werden hier die Formen des Verbs »sein« im althochdeutschen Text, die keine Entsprechung in der lateinischen Vorlage haben (*ist*=ist, *sifis*=seiest).



❹ Im Text markiert sind hier die Passivformen unterschiedlicher Bildung (althochdeutsch *gisentit uuard* – lateinisch *missus est* »wurde gesendet«, *uuas gitruobit* ~ lateinisch *turbata est* »wurde betrübt«).



❺ Sonstige Differenzen zwischen der lateinischen und der althochdeutschen Textgestalt (Wortstellung, Gebrauch von Konjunktionen) sind in dieser Variante herausgestellt.

Textfassungen, bei der jeder einzelnen Wortform die dazugehörigen grammatischen (morphologischen und syntaktischen) Angaben beigefügt werden, so dass zum Beispiel eine pauschale Auswertung der Passivformen möglich wird; dieser Aufgabe widmet sich derzeit das Verbundprojekt »Referenzkorpus Altdeutsch«,

Die Textstelle aus Lukas 1,23–26 in der lateinisch-althochdeutschen Evangelienharmonie Tatians ist hier mit interlinearen Annotationen versehen: Jeder Wortform werden die dazugehörigen grammatischen Informationen in tabellarischer Form zugeordnet, um komplexe Abfragen zu ermöglichen.



an dem neben der Goethe-Universität die Humboldt-Universität Berlin und die Schiller-Universität Jena beteiligt sind.<sup>15/</sup> In einem weiteren Schritt, der jetzt im LOEWE-Schwerpunkt angegangen werden soll, geht es dann darum, Verfahren zu entwickeln, um die gram-

matischen Annotationen des Lateinischen und des Althochdeutschen aufeinander beziehbar zu machen und damit eine gezielte Auswertung von Unterschieden und Gemeinsamkeiten zu ermöglichen, etwa im Sinne der Fragestellung: Besteht ein signifikanter Zusammenhang zwischen der vom Lateinischen abweichenden Stellung des Auxiliars im althochdeutschen Passiv und der Struktur des Satzes (Haupt- oder Nebensatz), in dem es steht?

Es dürfte sich von selbst verstehen, dass eine derartige Aufbereitung auch dort zu neuen – und verifizierbaren – Erkenntnissen führen kann, wo die Vorlage eines Übersetzungstextes nicht klar ist. Im LOEWE-Schwerpunkt betrifft dies neben vielen Bibelübersetzungen aus dem Orient, wo etwa griechische oder syrische Vorlagen in Betracht kommen, auch zahlreiche Übersetzungstexte anderer, nicht christlicher Traditionen, wie diejenigen des Buddhismus im Alttürkischen, die auf Vorlagen in anderen an der Seidenstraße gesprochenen Sprachen (wie zum Beispiel Chinesisch, Sogdisch oder Tocharisch) zurückgehen können. Eine systematische Auswertung im Hinblick auf Gemeinsamkeiten und Divergenzen verspricht hier weit mehr Klarheit zu schaffen als die – notgedrungen meist eklektizistisch vorgenommenen – Einzelbeobachtungen früherer Forschergenerationen.

**Handschrift und Inhalt – Bild und Text**

Das elektronische Zeitalter eröffnet den Geisteswissenschaften nicht nur in Form von Textcorpora neue Perspektiven, sondern auch durch die Möglichkeit einer Verknüpfung textueller Informationen mit digitalen Bildern. Welchen immensen Fortschritt es für alle, die sich mit handschriftlich überliefertem Textmaterial beschäftigen, bedeutet, dass täglich immer mehr Bibliotheken ihre Manuskriptbestände online verfügbar machen, lässt sich leicht an einem weiteren Beispiel aus

**Anmerkungen**

<sup>11/</sup> TITUS – Von der Keilschrifttafel zur Textdatenbank. Forschung Frankfurt 4/1995, Seite 46–56.

<sup>12/</sup> Die an der Goethe-Universität beheimatete TITUS-Datenbank umfasst zum Beispiel schon jetzt vollständige Corpora verschiedener alt- und mitteliranischer Sprachen, des Tocharischen und anderer Überlieferungen (siehe <http://titus.uni-frankfurt.de/texte/texte2.htm>).

<sup>13/</sup> Die Handschriften der Klosterbibliothek Sankt Gallen sowie anderer schweizerischer Bibliotheken werden derzeit durch

das Projekt »e-codices« digitalisiert und in mustergültiger Form online zur Verfügung gestellt (siehe <http://www.e-codices.unifr.ch/>). Die hier (ausschnittsweise) abgebildete Seite, die das Ende des zweiten und den Beginn des dritten Kapitels des Lukas-Evangeliums umfasst, findet sich unter <http://www.e-codices.unifr.ch/de/csg/0056/28/medium>. Die Textausgabe Tatians ist in der TITUS-Datenbank unter <http://titus.uni-frankfurt.de/texte/etcs/germ/ahd/tatianx/tatia.htm> verfügbar.

<sup>14/</sup> Die in der TITUS-Datenbank vorliegende digi-

tale Fassung geht bereits auf die frühen 1990er Jahre zurück; sie wurde in Kooperation verschiedener Sprachwissenschaftler in Spanien und Deutschland erarbeitet.

<sup>15/</sup> Siehe <http://www.e-codices.unifr.ch/en/list/one/csg/0911>.

<sup>16/</sup> Zuerst bei Heinrich Hattemer, Denkmahle des Mittelalters, I. Bd., St. Gallen 1844, S. 324. – Frühere Erwähnungen des Textes nennen die Überschrift nicht, ebenso wenig wie der Handschriftenkatalog von Gustav Scherrer (Verzeichniss der Handschriften der Stiftsbibliothek

von St. Gallen, Halle 1875, S. 330).

<sup>17/</sup> Siehe <http://www.e-codices.unifr.ch/en/csg/0911/321/medium>.

<sup>18/</sup> Zum multispektralen Verfahren siehe: Jost Gippert, The Application of Multispectral Imaging in the Study of Caucasian Palimpsests, Bulletin of the Georgian National Academy of Sciences 175/1 (2007), S. 168–179 (online: <http://www.science.org/ge/2007-vol1/168-179.pdf>).

<sup>19/</sup> Brief Nr. 156; siehe <http://edocs.ub.uni-frankfurt.de/volltexte/2007/999999/pdf/0156.pdf>.

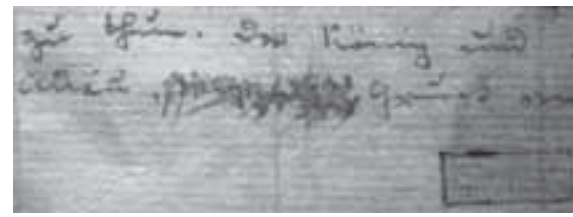
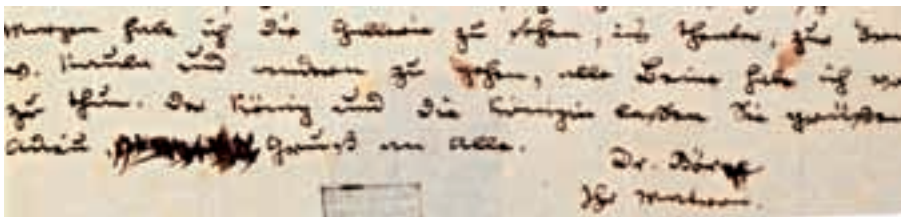
der Sankt Galler Klosterbibliothek illustrieren. Auf einer seiner letzten Seiten enthält der aus dem achten Jahrhundert stammende Codex Nr. 911 dieser Bibliothek<sup>61</sup> eine althochdeutsche Fassung des Glaubensbekenntnisses unter einem lateinischen Titel, der seit dem 19. Jahrhundert durchweg in der grammatisch absonderlichen Form *Credo in deo* (mit Ablativ statt Akkusativ nach *in*) zitiert wird.<sup>171</sup> Das vom e-codices-Projekt auf meinen Vorschlag hin seit März 2010 online zur Verfügung gestellte Digitalisat der Handschrift<sup>181</sup> **7** beweist nun, dass es sich um eine einfache Fehllesung handelt: Wie die Vergrößerung **8** zeigt, ist stattdessen *credo in dñm* zu lesen, wobei letzteres Wort die abgekürzte Variante des regulären Akkusativs *deum* repräsentiert.

Die Bereitstellung hochauflösender digitaler Aufnahmen ist nun aber nur ein erster Schritt im Hinblick auf eine wirkliche Integration von Text- und Bilddaten, wie sie im Rahmen der »Digital Humanities« angestrebt werden. Tatsächlich zielen nur die wenigsten



**7** Das althochdeutsche Glaubensbekenntnis aus dem Cod. Sang. 911, S. 321 (Ausschnitt).

**8** Überschrift des althochdeutschen Glaubensbekenntnisses, nicht *credo in deo* sondern *credo in dñm* (mit *dñm* als Abkürzung für den Akkusativ *deum*) zu lesen.



der zahlreichen Digitalisierungsprojekte, die zurzeit ablaufen, darauf ab, die elektronischen Bilder von Handschriften unmittelbar mit editorischen Bearbeitungen ihrer textualen Inhalte zu verknüpfen, obwohl gerade in einer solchen Verknüpfung der entscheidende Mehrwert im Sinne der Verifizierbarkeit gegenüber herkömmlichen gedruckten Ausgaben, die sich meist auf den reinen Text beschränken, liegen würde. Hier will der LOEWE-Schwerpunkt ansetzen, indem er an ausgewählten Beispielen die Visualisierung handschriftlicher Überlieferungszusammenhänge im Rahmen elektronischer Editionen zu erproben und text-technologische Erschließungsinstrumentarien für die teilautomatische Verarbeitung und Verlinkung von Archivmedien zu entwickeln anstrebt, die über digitalisierte Textcorpora hinaus Manuskript- und Bildsammlungen umfassen.

### Durchblicke mit der Multispektralanalyse – Getilgte Stellen im Fokus

Je nach der Art der zugrunde liegenden Bildmaterialien wird es dabei erforderlich sein, spezielle Digitalisierungsverfahren anzuwenden. Dies betrifft vor allem solche Handschriften, bei denen der ursprüngliche Text aufgrund von Defekten oder durch menschliche Eingriffe wie Tilgungen, Streichungen oder Überschreibungen unkenntlich geworden ist. Tatsächlich bieten bestimmte Verfahren der digitalen Fotografie die Möglichkeit, Textteile, die mit dem bloßen Auge nicht zu erkennen sind, zu erschließen. Im LOEWE-Schwerpunkt soll insbesondere das Verfahren der Multispektralanalyse angewendet werden, bei dem mehrere Aufnahmen desselben Objekts in unterschiedlichen Spektralbereichen elektronisch miteinander verglichen werden, um beispielsweise die untere Schicht von Palimpsest-Handschriften oder getilgte Stellen in Autografen wieder lesbar zu machen.<sup>191</sup> So wird es auch da-

rum gehen, die zahlreichen Schwärzungen, die sich in den in der Frankfurter Universitätsbibliothek Johann Christian Senckenberg gesammelten Briefen Ludwig Börnes finden, herauszufiltern, um den darunterliegenden Text zu eruieren. Die Abbildung **9** zeigt eine derartige Schwärzung im Brief des Journalisten an Jeanette Wohl vom 11.11.1820<sup>191</sup> die Spektralaufnahme im Infrarot-Bereich (980 nm) lässt deutlich erkennen, dass das überschriebene Wort die freundschaftliche Anrede *Moppel* **10** war, die offenbar als nicht für die Nachwelt geeignet erachtet wurde. Ob die Schwärzung vom Autor selbst oder von der Adressatin des Briefs stammt, bleibt noch zu ermitteln. ♦

**9** Börne-Brief Nr. 156, Ausschnitt mit geschwärtzter Anrede. **10** Die Anrede »Moppel« in Börnes Brief wird in der Spektralaufnahme bei 980 nm (Infrarotbereich) lesbar.

## Der Autor



**Prof. Dr. Jost Gippert**, 54, studierte Vergleichende Sprachwissenschaft, Indologie, Japanologie und Sinologie an der Universität Marburg und an der Freien Universität Berlin. Nach der Habilitation an der Universität Bamberg (1991) übernahm er 1994 die Professur für Vergleichende Sprachwissenschaft an der Goethe-Universität. Zu seinen Arbeitsschwerpunkten gehören neben den indogermanischen Sprachen und ihrer Geschichte sowie der allgemeinen Sprachtypologie auch die Sprachen des Kaukasusraums, denen in der jüngeren Zeit verschiedene internationale Kooperationsprojekte unter seiner Federführung gewidmet sind. Als Gründer und Leiter des in Frankfurt beheimateten TITUS-Projekts (Thesaurus Indogermanischer Text- und Sprachmaterialien) engagiert er sich seit 1987 für den Aufbau von Textcorpora älterer Sprachen. In dem 2011 anlaufenden LOEWE-Schwerpunkt »Digital Humanities«, den Gippert gemeinsam mit Partnern der Goethe-Universität, der Technischen Universität Darmstadt und des Freien Deutschen Hochstifts initiiert hat, sollen derartige Corpora in den Fokus der Forschung gerückt werden. [siehe auch Interview mit Prof. Gippert und Prof. Jussen auf Seite 70]

[gippert@em.uni-frankfurt.de](mailto:gippert@em.uni-frankfurt.de)