

Schriften des Instituts für Dokumentologie und Editorik — Band 2

Kodikologie und Paläographie im digitalen Zeitalter

Codicology and Palaeography in the Digital Age

herausgegeben von | edited by

Malte Rehbein, Patrick Sahle, Torsten Schaßan

unter Mitarbeit von | in collaboration with

Bernhard Assmann, Franz Fischer, Christiane Fritze

2009

BoD, Norderstedt

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de/> abrufbar.

Leicht veränderte Fassung für die digitale Publikation (siehe Vorwort).

Slightly modified version to be published digitally (see preface).

Publication réalisée avec le soutien d'Apices
Association Paléographique Internationale
Culture – Écriture – Société
Dotation J.M.M. Hermans.
<http://www.palaeographia.org/apices/>



© 2009

Herstellung und Verlag: Books on Demand GmbH, Norderstedt
ISBN: 978-3-8370-9842-6
Einbandgestaltung: Katharina Weber
Satz: X_YT_EX und Bernhard Assmann

Forschung am Rande des paläographischen Zweifels: Die EDV-basierte Erfassung individueller Schriftzüge im Projekt *DAmals**

Wernfried Hofmeister, Andrea Hofmeister-Winter,
Georg Thallinger

Zusammenfassung

Das Pilotprojekt *DAmals* (Datenbank zur Authentifizierung mittelalterlicher Schreiberhände) hat es sich zum Ziel gesetzt, neue Kriterien zur Unterscheidung von Schreiberhänden in mittelalterlichen deutschsprachigen Handschriften aufzustellen und zuverlässigere Methoden und Werkzeuge für diese Aufgabe zu entwickeln. *DAmals* beruht auf den drei Säulen einer elementgetreuen Basistransliteration in XML, computerbasierten graphetischen Analysen und einem neuartigen Verfahren der bildorientierten Mustererkennung. Diese Säulen sind in eine Datenbankstruktur integriert, welche sowohl die Archivierung als auch die technisch hochkomplexe Verarbeitung der Bild- und Textdokumente leistet. Auf diese Weise bietet *DAmals* eine Art Brille, durch die paläographische ExpertInnen bei ihrer Schriftbegutachtung unterstützt werden. In einer weiteren Ausbaustufe soll das Projekt *DAmals* in ein neues Projekt namens MOSES (Musterorientiertes System zur Erfassung von Schriftindividualität) eingebettet werden, welches sich auf neuzeitliche und aktuelle handgeschriebene Materialien ausdehnen lässt, um dann z. B. auch für forensische Zwecke hilfreich zu sein.

Abstract

In order to provide objective criteria for distinctions between presumably different writing hands in medieval German vernacular manuscripts the project *DAmals* (Datenbank zur Authentifizierung mittelalterlicher Schreiberhände/Database for the Authentication of Medieval Writing Hands) has developed new methods and tools: three pillars—a palaeographically extremely detailed XML transliteration, manifold graphetical statistics and image-based pattern recognition—have been integrated into an innovative database as complex and highly interrelated techniques for the analysis of handwritten documents. By these means *DAmals* also offers “virtual spectacles”

* *DAmals* steht für ›Datenbank zur Authentifizierung mittelalterlicher Schreiberhände‹. Die nachfolgenden Ausführungen lassen sich wie folgt grob den drei Beiträgern zuordnen: Wernfried Hofmeister: Abschnitt 1, 4.1 u. 5; Andrea Hofmeister-Winter: 2.1 u. 3.1; Georg Thallinger: 2.2, 3.2 u. 4.2.

through which palaeographic experts may look and thereby be supported in their challenging judgements. In a further step *DAmals* is to be incorporated into a new project called MOSES (Musterorientiertes System zur Erfassung von Schriftindividualität/Pattern Orientated System for the Detection of Individuality in Handwriting) and e. g. also be helpful in solving current forensic problems.

1 Vorbemerkungen – *DAmals* im Kontext einer neuen Überlieferungsphilologie

Die Frage nach der Anzahl der Hände, die an einem Überlieferungsträger gearbeitet haben, gehört für die Paläographie seit jeher zu den wichtigsten, dabei zugleich zu den herausforderndsten, und sie hat in jüngster Zeit noch an Brisanz gewonnen: Gestärkt durch die Bestrebungen der ›New Philology‹, bemüht man sich in zahlreichen historischen Textfächern um eine umfassende Würdigung jedes einzelnen Aufzeichnungsprozesses rund um die vielfältigen Aspekte von ›Produktion und Kontext‹.¹ Motiviert wird diese erhöhte Aufmerksamkeit für die individuelle Genese und für die nur scheinbar banale Text-Materialität² aller Schriftdokumente durch ein neues *Werkverständnis*, welches sich nicht zuletzt darin ausdrückt, dass man umfassender als zuvor bereit ist, eine prinzipielle *Werk-→Offenheit* anzuerkennen: Anstatt weiterhin eine imaginäre ›Urfassung‹ in den Mittelpunkt zu stellen, welche es anhand aller präsumtiv minderwertigen Überlieferungsträger zu rekonstruieren gelte, hat man erkannt, dass viele der erhaltenen Überlieferungen als ernst zu nehmende *Varianten* zu sehen sind, oft als ganz eigenwertige *Fassungen*, welche an veränderte Rezeptionsbedingungen angepasst worden waren³ und eine eingehende Neubewertung verdienen. Alle *Werk-Handschriften* sind im Grunde für uns damit von philologischem Rohmaterial zu aussagekräftigen Dokumenten geworden. In besonderem Maße gilt dies dort, wo bei näherer Betrachtung schon in historischer Zeit eine systematische Überlieferungs*strategie* sichtbar wird, mithin ein fast präphilologisch zu nennendes Konzept, welches sich in dem Maße ausdrückt, in welchem von den (ab-)schreibenden Persönlichkeiten unter Bedachtnahme auf diverse

¹ Unter diesem Titel fand 1998 in Den Haag die 7. internationale Tagung der Arbeitsgemeinschaft für germanistische Edition statt: Vgl. *Produktion und Kontext*.

² »Materialität in der Editionswissenschaft« lautete das Generalthema der 12. internationalen Tagung der Arbeitsgemeinschaft für germanistische Edition vom 13.–16. Februar 2008 an der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW) in Berlin. (Näheres findet sich dazu auf der Website der AG.)

³ Besonders schöne Beispiele für die Eigenwertigkeit von Varianten und Fassungen konnten im Rahmen des ›Parzival-Projekts‹ von Michael Stolz erzielt werden. Aktuelles dazu auf der Projekt-Homepage.

Faktoren wie Verfügbarkeit, Vollständigkeit, Verlässlichkeit der Überlieferungsvorlagen resp. Quellen gearbeitet wurde, sodass im Grunde frühe *Editionen* entstanden.⁴

Es ist also diese neue Sicht auf den potenziellen Eigenwert jeder einzelnen Überlieferung, die auch unser Interesse an den dafür verantwortlichen *SchreiberInnen* hat wachsen lassen, weil ja letztlich *sie* es waren, die sich im Zuge einer solchen Textreproduktion mit eingeschrieben haben – mit all ihren (mehr oder minder gewollten) Textabweichungen in Gestalt von Kürzungen, Wortumstellungen, Begriffsauswechslungen etc. Entsprechend bedeutsam ist es für die gesamte Textforschung zu wissen, ob bzw. wo wir es in einem Codex mit ein und derselben Hand (im Sinn von ein und derselben Person) zu tun haben, weil diese ja gemäß mittelalterlicher ›Überlieferungsmoral‹ eine spezifische *inhaltliche* und *sprachlich-formale* Textverantwortung⁵ trägt und diese auch durch ihre Arbeit zum Ausdruck bringt: Neben den schon angedeuteten Strategien im *Inhaltsbereich* (in Bezug auf die rein ›stoffliche‹ Wiedergabe von schriftlichen Vorlagen oder mündlichen Diktaten) sind es die *linguistisch* relevanten Aspekte, welche sich besonders offensichtlich an ein schreiberspezifisches Sprachvermögen rückgebunden zeigen und uns daher nur unter Bedachtnahme auf diesen Zusammenhang als Evidenz für sprachhistorische Erscheinungen dienen können. Oder anders ausgedrückt: Gerade in der mittelalterlichen Zeit, wo das schriftliche Überliefern von Texten bekanntlich eine Fertigkeit darstellte, die in hohem Maße von frei migrierenden und deshalb an verschiedenen Orten arbeitenden SpezialistInnen erledigt wurde, ist es für unsere Erschließung der historischen Laut- und Formenvielfalt von eminenter Bedeutung, die Grenzen zwischen verschiedenen Schreiberpersönlichkeiten verlässlich ausmachen zu können; erst das ermöglicht eine bewusste Berücksichtigung ihrer schriftsprachlichen Idiosynkrasien und erlaubt in weiterer Konsequenz eine fundierte(re) Debatte über das Heranwachsen von regionalen und überregionalen Sprachnormen! – In diesem Sinne spiegelt eine paläographische *ad fontes*-Bewegung, für welche die individuellen Schriftzüge – wie eine ›Grammatik der Schreiberhände‹ (Hofmeister-Winter 2005) – zum wichtigsten Anker für zahlreiche weitere Erkenntnisse geworden sind, den Urgrund einer neuen, viele Fächer verbindenden *Überlieferungs-Philologie*⁶ wider.

⁴ Der editorische Eigenwert mittelalterlicher Überlieferungen wurde jüngst im Rahmen der mediävistischen Fachtagung »Wege zum Text« betont. Siehe dazu den Beitrag des Tagungsleiters, Wernfried Hofmeister, der darauf in seinem »Problemaufriss zum Stellenwert von Editionen und ihrer Verfügbarkeit« Bezug nahm; diese Ausführungen sind auf der Tagungs-Homepage als Podcast nachzuhören; der Sammelband zur Tagung wird voraussichtlich im Sommer 2009 im Druck erscheinen.

⁵ Die Arbeitsteiligkeit zwischen Text-Autor und Überlieferern erscheint zwar im Lichte diverser Sorgen von mittelalterlichen Autoren bezüglich drohender Textentstellungen durch schlechte Textweitergaben nicht spannungsfrei gewesen zu sein, belegt aber indirekt ebenfalls genau diesen – von den SchreiberInnen offensichtlich gerne genutzten – Freiraum für ihre Textaufzeichnung. Daneben kennen wir freilich auch Signale der Zustimmung zu den unvermeidlich ausbleibenden Texteingriffen, müssen uns jedoch in einigen dieser Fälle fragen, ob nicht eine ironische Brechung einer solchen ›Lizenz zum Verändern‹ vorliegt. Dies steht etwa auch für den Autor Hugo von Montfort zu vermuten; vgl. Hofmeister 2004.

⁶ Für ihre Etablierung wird explizit von Wernfried Hofmeister (Hofmeister 2001) geworben. Aufgegriffen

1.1 Globalziele von DAMaIS

Das nun vorzustellende Projekt DAMaIS⁷ versteht sich ganz bewusst als ein Teil dieser eingangs skizzierten Rückbesinnung auf das Elementare unserer Schriftkultur. Es ist angetreten, um dort, wo bislang mittels herkömmlicher paläographischer Befundungen zwischen einzelnen ForscherInnen kein Konsens über die Anzahl der beteiligten Schreiberhände zu gewinnen war, mit einer plurimethodischen Herangehensweise *relevante Indikatoren* für eine Art *Schriftindividualität* zu ermitteln und diese Spezifika hernach auf höchstem technischen Niveau durch objektive schriftanalytische Messwerte zu dokumentieren. Somit will dieses neue Instrumentarium nicht mehr, aber auch nicht weniger sein als eine Art *Brille*, durch die unser paläographisch geschultes Auge *zusätzlich* blicken kann, um ein intersubjektiv teilbares Urteil fällen zu können. Dies ist nicht nur hilfreich in jenen Fällen, wo verschiedene SchriftexpertInnen unterschiedliche Meinungen vertreten, sondern auch dort, wo unsere Augen nach wiederholtem Befunden einer fraglichen Handschrift zu durchaus schwankenden Ergebnissen gelangen.

Noch hat – um im Bild zu bleiben – diese DAMaIS-Brille nicht ihren letzten Schliff erhalten, da einige ergänzende Optimierungen erst zu leisten sind, doch für ihren zentralen Probefall, nämlich den deutschsprachigen cpg 329⁸ (um 1415, aus dem Besitz Hugos von Montfort), hat sie offensichtlich schon genügend Trennschärfe gezeigt, um uns dabei behilflich zu sein, eine langjährige Streitfrage erfolgreich zu lösen; ohne dazu (im Vorgriff auf die nächsten Beitragsabschnitte) die Details vorwegzunehmen, seien zunächst ganz allgemein und überblickshaft das operationale Design dieses neuen Befundungsinstruments kurz vorgestellt sowie einige seiner methodischen Hintergründe.

1.2 Arbeitsprozess-Schema für DAMaIS: Authentifizierungs-Szenario

Als Orientierungshilfe für die folgenden Ausführungen, welche dem zentralen ›Authentifizierungs-Szenario‹ gelten, dient das ›Arbeitsprozess-Schema‹ (Abb. 1)⁹. Dessen 3-säulige Grundstruktur ist erstmals Anfang 2008 einem internationalen Fach-

und durch neue Beobachtungen an der Materialität ergänzt wurde dieser Beitrag jüngst von Martin Schubert (Schubert).

⁷ Aktuelle Informationen über dieses Projekt bietet die [Projekt-Homepage](#). – Dem Begründer von DAMaIS, Wernfried Hofmeister, wurde zur Förderung dieser Pilotstudie von der Forschungsabteilung des Landes Steiermark im Jahr 2006 ein namhafter Förderungsbetrag zuerkannt. Ein Großteil dieser Förderung dient seinem Projektpartner, der Forschungsgesellschaft Joanneum Research (vertreten durch Georg Thallinger) für die Entwicklung der technischen Projekt-Applikationen. In Summe wirken an DAMaIS (freilich nur zeitweise bzw. auf Werkvertragsbasis) rund 10 Personen mit, darunter VertreterInnen des Zentrums für Informationsmodellierung in den Geisteswissenschaften (ZIMig) an der Karl-Franzens-Universität Graz.

⁸ Siehe die Online-Präsentation dieses Codex sowie zu seiner – für DAMaIS grundlegenden – Einbindung in die neue Hugo von Montfort-Ausgabe von Wernfried Hofmeister die [Editions-Homepage](#).

⁹ Alle im Abbildungsteil gezeigten Ausschnitte aus dem Montfort-Codex cpg 329 stammen von der Online-Präsentation cpg 329 der UB Heidelberg.

plenum bekannt gemacht worden (Hofmeister und Hofmeister-Winter) und hat dabei über die Fachgrenzen hinweg Zustimmung erhalten. Für diesen Sammelband erfährt das soeben erwähnte ›Arbeitsprozess-Schema‹ nun außer einer differenzierteren graphischen Aufbereitung vor allem eine detailliertere Beschreibung seiner *technischen* Komponenten.¹⁰

Generell ist anzumerken, dass DAMaLS von einer zweifachen Datenbankstruktur getragen wird, einer äußeren und einer inneren: Nach *außen* hin formen die einzelnen Überlieferungsträger sowie die darin ermittelten Schreiberhände Datensätze, durch deren Verwaltung (u. a.) das spätere Wiedererkennen von bereits erfassten Daten/Personen ermöglicht wird. Aber auch in ihrem *Inneren* stützt sich jede einzelne Dokumentaufnahme und Befundung auf ein (interrelationales) Datenbank-Design, durch das alle Einträge miteinander verknüpft werden können, um ein sukzessives Sammeln und abschließendes Auslesen aller Informationen, Werte und Daten zu erlauben. Nähere Hinweise zur dafür gewählten Spezialdatenbank *imdas pro*, welche sowohl die äußere als auch die innere DAMaLS-Struktur trägt, finden sich weiter unten in Abschnitt 2.2.

Der gesamte Authentifizierungsablauf ruht im Wesentlichen auf den schon erwähnten drei Säulen resp. Untersuchungsansätzen: Sie bestehen aus der ›klassischen‹ Paläographie, der ›graphetischen Statistik‹ und der ›musterorientierten Schriftbild-Erfassung‹. Diese drei Ansätze sind nur teilweise als sukzessiv zu denken, denn manches läuft parallel ab. Den *Beginn* des Arbeitsprozesses repräsentiert jedoch immer (in Säule 1) die kodikologische Beschreibung des Überlieferungsträgers und seiner *paläographischen* Charakteristika. Am Ende dieser Routine steht als Ziel eine maßgeschneiderte, dabei bis in die unscheinbarsten Teile elementgetreue Basistransliteration in XML-Codierung. Damit ist (für Säule 2) die Möglichkeit zur *statistischen* Bestimmung jener Graphvarianten, -teile und -sequenzen gegeben, die es via Datenbank auszuwerten gilt, um allfällige signifikante Schwankungen erkennbar zu machen. Praktisch simultan dazu findet (in Säule 3) eine messtechnische und algorithmische Befundung des *Schriftbildes* statt; für die hier nötige höchste Präzision sorgen die zuvor schon in die Datenbank integrierten *Bild-Digitalisate* aller Schriftstücke, indem deren Bildoberfläche mit allen Graph-Elementen der XML-Transliteration verknüpft wird. – Alle drei Befundungssäulen zusammen ergeben bereits einen ausgeprägten *plurimethodischen* Ansatz. Ergänzt wird dieser interdisziplinäre Zugriff durch die Einbeziehung kognitionspsychologischer Zusammenhänge (rund um die Verarbeitung und den Ausdruck gestalthafter Einheiten durch feinmotorische Impulse bei handschriftlicher Sprachreproduktion), aber auch der quantifizierenden Linguistik sowie allg. der Textmodellierung, für die es neben bzw. in Kooperation mit TEI neue Wege für noch feinere Abbildungen graphetischer ›Elementarteilchen‹ zu finden galt.

¹⁰ Dies erfolgt in den Abschnitten 2.2, 3.2 und 4.2; bezüglich der namentlichen Zuordnung dieser (und aller übrigen) Beitragsabschnitte vgl. Anm. 1.

3-säuliges DArma/S Arbeitsprozess-Schema für die Authentifizierung von Schreiberhänden

durch die schriftvergleichende Ermittlung allfälliger Divergenz-Zonen u. daran knüpfbarer Schreibergerenzen

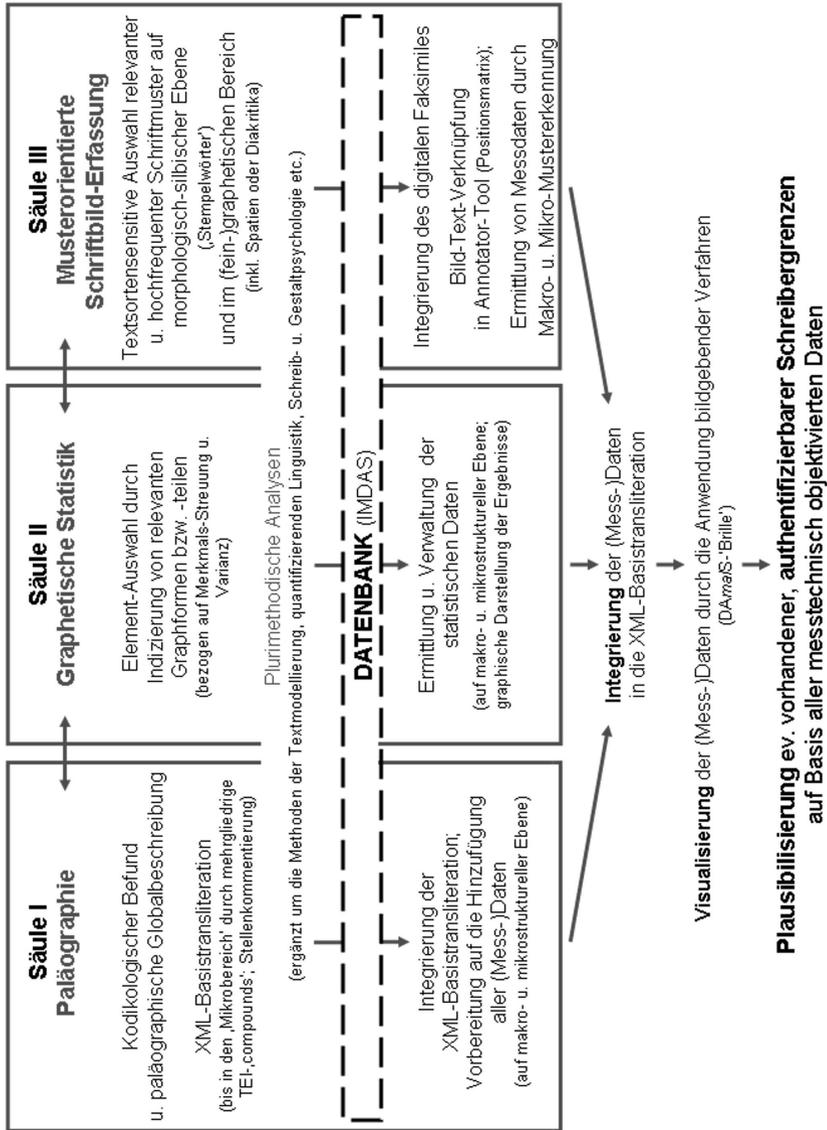


Abbildung 1. Arbeitsprozess-Schema für die Authentifizierung mittelalterlicher Schreiberhände.

Dank dieser konsequenten XML-Codierungsweise können in weiterer Folge alle hilfreichen Informationen (wie paläographisch-textkritische Stellenkommentare, abschnittsbezogene Frequenzwerte oder gestaltvergleichende Konvergenz- bzw. Divergenzparameter zwischen Graphen oder Graphfolgen/Wörtern) direkt im jeweiligen Transliterationsbereich gespeichert resp. dort hinein *integriert* werden. Danach lassen sich unter Einbeziehung der bildgebundenen Positionsdaten sämtliche Werte gleichsam in die Handschrift zurückprojizieren und – gesteuert von allen Vergleichsdaten – punktgenau jene Bildbereiche erkennen, an denen sich auffallende und evtl. nur durch einen Schreiberwechsel erklärbare Veränderungen abzeichnen: Zur Verdeutlichung der abstrakten Vergleichszahlen dienen unterschiedliche *Visualisierungstechniken*; indem sie unser Auge auf die verdächtigen Übergangszonen aufmerksam machen, fungieren sie gleichsam als Brille. – Das letzte Wort hat freilich nie allein die Datenbank, sondern *wir* selbst, denn alles, was DAMaLS bietet, sind objektive Messdaten, die zwar unser Gesamturteil bezüglich der Schreiberauthentizität wesentlich *plausibler* und vor allem intersubjektiv besser argumentierbar machen sollten, aber gemäß ihrer inhärenten ›fuzzy logic‹-Struktur letztlich nur Wahrscheinlichkeitswerte anbieten, welche es zu *interpretieren* gilt.

1.3 Komplementäres ›Eichungs-Szenario‹ in DAMaLS

Die Bewusstheit für diese prinzipielle Interpretationsbedürftigkeit aller vergleichenden Messergebnisse in einem humanwissenschaftlichen Kontext hat schon am Projektbeginn neben das ›Authentifizierungs-Szenario‹, das primär der Ermittlung von allfälligen *Divergenzen* und daran knüpfbaren Schreiberhandgrenzen dient, ein ›Eichungs-Szenario‹ treten lassen. Dieses geht von einer bereits (durch eindeutige Quellenbeweise) identifizierten Schreiberhand aus und bemüht sich an ihrem Beispiel um das Ermitteln jener *Konvergenzen*, durch die – über Jahre und unterschiedliche Textsorten hinweg – eine Hand charakterisiert und wiedererkennbar scheint. Wie dazu in der bereits genannten DAMaLS-Publikation (Hofmeister und Hofmeister-Winter) ausführlich dargestellt wird, haben wir dafür nach der *Bandbreite* des Individuellen im 15. Jahrhundert¹¹ gesucht, denn erst dadurch werden näherungsweise Urteile über die zu erwartenden Merkmalsschwankungen bei noch nicht authentifizierten Händen möglich. Ausgewählt haben wir als ein solches ›Eichmaß‹ die umfangreichen (Ab-)Schriften der Lohnschreiberin Clara Hätzlerin. Noch konnten nicht alle ihre europaweit verstreuten Überlieferungsträger autopsiert und in unsere Datenbank eingespeist werden, aber doch die meisten. Dabei hat eine erste tentative Auswertung ergeben, dass sich anhand signifikanter Buchstaben- und ›Wortbilder‹ quer durch alle Stilisierungsebenen und Gebrauchshände dieser Augsburgsburger Schreiberin in der Tat bereits so etwas wie ei-

¹¹ Und damit in zeitlicher Nähe der zu authentifizierenden Überlieferung der Werke Hugos von Montfort.

ne Schrift->DNA< erkennen lässt, also rekurrente individuelle Schriftzüge. So hoffen wir (nach Abschluss unserer 3-Säulen-Befundung aller erhaltenen Schriftstücke der Clara Hätzlerin) jenen ›Maßstab< zu gewinnen, der – an noch nicht authentifizierte, aber zeitlich benachbarte deutschsprachige Hände angelegt – zumindest erahnen lässt, wo die Grenzen ›normaler< Ähnlichkeits-Streuungen liegen. Die künftige Einbeziehung weiterer identifizierter Schreiberhände, von denen es im Bereich deutschsprachiger Aufzeichnungen freilich speziell in spät- oder gar hochmittelalterlicher Zeit nur wenige gibt, könnte weitere ›Maßstäbe< beisteuern.¹²

2 ›Paläographische Schrifterfassung< (erste Authentifizierungs-Säule)

2.1 Allgemeines zum paläographischen Ansatz

Wie bereits oben erwähnt, wird der Paläographie im Methodenkanon von DAMALS zentrale Bedeutung beigemessen: Ihr kommt der erste prüfende Blick auf ein zu befundenes Schriftstück zu, um anhand formaler Merkmale eine grobe zeitliche Einordnung zu treffen; sie ist es auch, die bei der *Schlussbefundung* quasi das letzte Wort haben muss, wenn es gilt, die mit verschiedenen Methoden gewonnenen Einzelergebnisse in Relation zueinander zu setzen und ihre Plausibilität abzuwägen.

Die wesentliche Kompetenz der Paläographie als Hilfswissenschaft für alle historisch orientierten Disziplinen liegt unbestritten in der Hilfestellung bei der Entzifferung historischer Schriftsysteme, für die sie dank ihrer synchronen und diachronen Sammlungen von Schriftsymbol-Inventaren¹³ sowie durch ›Schlüssel-Werke< aller Art¹⁴ das notwendige Instrumentarium liefert. Erst damit wird es den Textwissenschaften möglich, verschriftlichte Sprache zu decodieren und so vielfältigen wissenschaftlichen Auswertungen zuzuführen.

Um diese Hilfestellung leisten zu können, muss die Paläographie den Blick auf die ›wesentlichen< Merkmale der Schrift einschränken, sie destilliert gleichsam die formalen Grundmuster einer Stilepoche aus einer Vielzahl von einzelnen Schriftzeugnissen durch Abstraktion von individuellen Ausprägungen. Genau auf diese individuellen Ausprägungen, die in mehr oder weniger deutlich erkennbaren *Abweichungen* von zeitlich bzw. regional gebräuchlichen Grundmustern bestehen, gilt es jedoch bei der Unterscheidung von Schreiberhänden zu fokussieren. Auch hier kann die Paläographie unterstützen, indem sie ein – wenn auch nicht allgemein verbindliches, so im Großen

¹² Vgl. die bemühte Dokumentation namentlich bekannter mittelalterlicher SchreiberInnen des europäischen Mittelalters durch Krämer.

¹³ Zur Bestimmung des Alters und der Herkunft einer Handschrift steht eine Fülle von Tafelwerken und Übersichtstabellen zur Verfügung.

¹⁴ Z. B. Cappelli, Grun.

und Ganzen doch brauchbares – *terminologisches Instrumentarium* zur Beschreibung des ›Augenbefundes‹ entwickelt hat.

Das Problem ist allerdings, dass die Beschreibung individueller Schriftzüge mit den Methoden der Paläographie – wie schon die oben erwähnte Aufstellung von ›Normen‹ für bestimmte Epochenabschnitte und Regionen – auf *Verallgemeinerung* des Befundes hinauslaufen muss: Die Suche nach den Charakteristika einer Schreiberhand führt zur Feststellung eines ›Durchschnitts‹ der beschreibbaren Merkmale, was gestützt auf paläographische Methoden zwar auf der Basis von Fachwissen und viel Erfahrung des Begutachters geschieht, sich aber letztlich doch zu einem Gutteil auf den *subjektiven* Eindruck gründet.

Um nun einerseits den paläographischen Befund in *intersubjektiv* nachvollziehbarer Weise festzuhalten, d. h. so *objektiv* wie möglich zu dokumentieren, und auf dieser Basis andererseits exakte Frequenzzählungen u.v.m. vornehmen zu können, ist es unumgänglich, den mit Hilfe der Paläographie decodierten Text zu ›recodieren‹ – idealerweise in Form einer *elementgetreuen Basistransliteration*, wie sie am Grazer Institut für Germanistik im Rahmen von Editionsarbeiten inzwischen zum Standard geworden ist und u. a. für den DAmALS-Beispielfall Hugo von Montfort angewandt wurde.¹⁵ Wie die Bezeichnung ›elementgetreu‹ bereits ahnen lässt, beschränkt sich diese Transliterationsmethode nicht auf die Umsetzung von alphabetischen Schriftsymbolen und allenfalls die Wiedergabe von buchstabenförmigen Superskripten und gängigen Abbiatursymbolen im Stile diplomatischer Abdrucke, sondern sie verfeinert die Dokumentation zumindest so weit, dass sich darin *sämtliche* vertretenen Schriftelemente entsprechend codiert wiederfinden. In Details, die im Verdacht stehen, hinsichtlich der Individualität von Schreiberhänden Relevanz zu besitzen, kann die Transliteration sogar noch tiefer in Richtung einer ›phänomengetreuen Wiedergabe‹ gehen: Hier wird z. B. auch die *Form* von Superskripten speziell berücksichtigt, etwa auffällig variierende Häkchen-Formen. Aber auch den auf Abbildungen oft kaum sichtbaren feinen Haarstrichen, die von SchreiberInnen wohl weniger zur deutlicheren Differenzierung von Buchstabenformen (z. B. bei e, r, t) angebracht werden als aus ästhetischen Gründen, können charakteristisch für Schreiberindividuen sein und werden daher konsequent erfasst – und selbstverständlich später an den originalen Schriftstücken peinlich genau autopsiert. Dass bei den Superskripten konsequenterweise auch das Vorhandensein bzw. das Fehlen des i-Punkts extra verzeichnet wird, versteht sich nach den vorangegangenen Ausführungen fast von selbst.

In Form von *Annotationen* fließt in diese Basistransliteration auch Materielles ein: Mängel des (Be-)Schreibmaterials (Tinte, Papier/Pergament, Abnutzungsspuren etc.)

¹⁵ Entwickelt und erstmals angewandt wurde dieses Verfahren von Andrea Hofmeister-Winter anhand eines umfangreichen Editionsprojektes (Das Brixner Dommesnerbuch). Die ›elementgetreue Basistransliteration‹ bildet in diesem mehrstufig angelegten Editions-konzept die Grundlage für die gesamte weitere editorische Bearbeitung des Textes. Vgl. auch Hofmeister-Winter 2003 sowie ihre [Forschungshomepage](#).

beeinflussen die Brauchbarkeit der Daten, so dass ›physisch beeinträchtigte‹ Belege ggf. aus dem Untersuchungsmaterial auszuschneiden sind. Das Hauptaugenmerk gilt aber dem ›Material‹ der Schriftzüge: Auch hier kann es vorkommen, dass Schriftsymbole durch Korrekturmaßnahmen (Tilgung/Rasur, Nachbesserung/Überschreibung) oder durch Schreiberversehen derart missgestaltet sind, dass sie nicht mehr als intakte Repräsentanten eines bestimmten Graphyps anzusehen sind und daher für bestimmte Vermessungs- und Berechnungsoperationen nicht herangezogen werden können. Diesen Umständen wird durch entsprechende Annotationen Rechnung getragen.

Alle genannten Maßnahmen der objektiven Dokumentation des paläographischen Befundes fallen in den Bereich der ›niederen Textkritik‹ (Schieb), deren Aufgabenbereich nicht nur in der philologischen Editionswissenschaft oft zu wenig Beachtung findet, obwohl hier das Fundament für alle weiteren Analysen bis hin zur Interpretation im Rahmen der ›höheren Textkritik‹ im Sinne Karl Lachmanns und seiner AnhängerInnen gelegt wird (Hofmeister-Winter 2005). War die Herstellung einer deskriptiven Basistransliteration ursprünglich, d. h. im Rahmen einer ›dynamischen Edition‹ (Hofmeister-Winter 2003) in erster Linie als Hilfestellung für den Editor selbst gedacht, als eine Art ›Wahrnehmungsprotokoll‹ zur Schulung des editorischen Auges (Hofmeister 1999 33), erwies sich das solcherart gesicherte Informationsmaterial mittlerweile als vielfältig nutzbar: So konnten etwa die minuziös encodierten i-Punkte in den Schriften Veit Feichters¹⁶ unsere ersten Auswertungen des Hugo von Montfort-Codex bestätigen und stützen (Hofmeister-Winter 2007 108–9).

Zu den Prinzipien dieser ›Mikro-Codierung‹, die im Rahmen von DAMaLS auf XML-Basis erfolgt,¹⁷ gehört es, dass Superskripte getrennt von ihren Basisgraphen codiert werden. Auf diese Weise ist nicht nur die Kombinationsfähigkeit von Superskriptformen mit Basisgraphen besser analysierbar, sondern es können sog. ›verrutschte‹ Superskripte auf einer späteren Editionsstufe mit Hilfe ›höherer Textkritik‹ leichter korrigiert werden; auf der Stufe der Basistransliteration bewahrt diese Maßnahme davor, dass der paläographische Befund vorschnell durch emendierende oder gar konjizierende Interpretation ›verwischt‹ wird.

Welche Schriftsymbole in der Handschrift aufeinander bezogen sind und in welcher schreibräumlichen Relation das geschieht, ist durch die Codierung systematisch dokumentiert und kann daher aus der Transliteration (auch ohne Beiziehung der Handschriftenabbildungen) eindeutig rekonstruiert werden. In der XML-Transformation werden

¹⁶ Außer dem »Dommesnerbuch« (Das Brixner Dommesnerbuch) sind von Veit Feichter auch ein »Urbar« und ein »Inventar« des Brixner Dommesneramtes erhalten; eine Edition auch der letzteren beiden Schriften durch Andrea Hofmeister-Winter ist in Vorbereitung.

¹⁷ Bisher sind unsere Basistransliterationen im vertrauten Programm WinWord (als unformatierte ASCII-Dateien) entstanden und wurden von unseren Projektpartnern vom Zentrum für Informationsmodellierung in den Geisteswissenschaften (s. Anm. 8) für das Projekt DAMaLS mit Hilfe von Transformationsregeln nachträglich in XML überführt; dieses Verfahren wird bis zur ›Marktreife‹ eines anwendungsfreundlichen Eingabetools (das zu den Nebenzielen des Projekts zählt) weiterhin beibehalten werden.

Basisgraph und die dazugehörigen Superskripte als Glyphen (compound graphs) definiert, bestehend aus zwei diskreten Elementen, die in der Ausgabe-Anweisung beliebig (z. B. durch Unicodes) zur Darstellung gebracht werden können.¹⁸ (Siehe Abb. 2: Beispiel für die XML-Codierung zusammengesetzter Zeichen.)

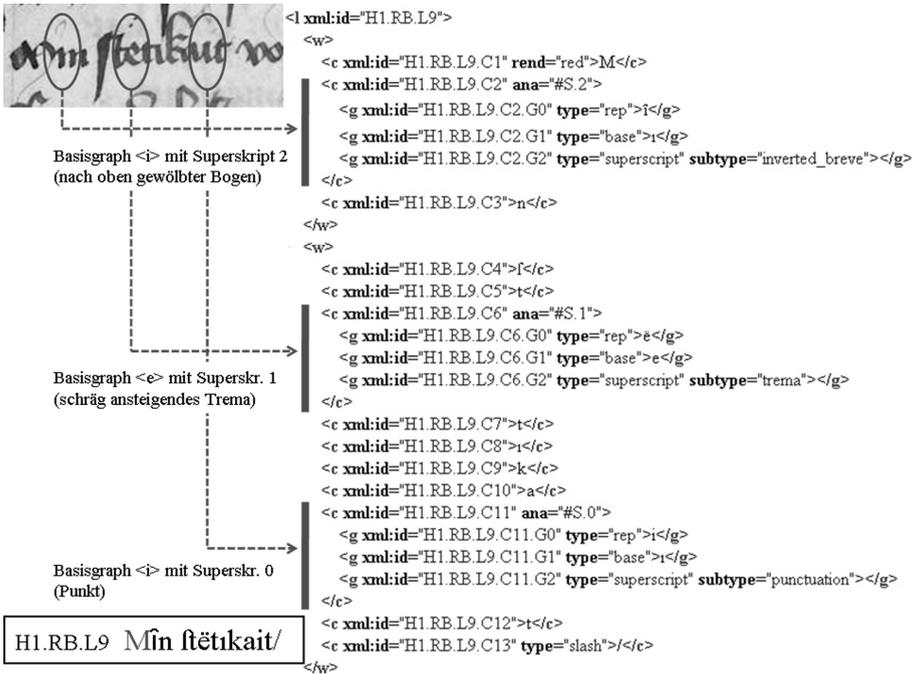


Abbildung 2. Beispiel für die XML-Codierung zusammengesetzter Zeichen und die Darstellung in der ›Augenfassung‹ anhand Heidelberg, UB, cpg 329, fol. 1rb, Zeile 9.

Die elementgetreue Basistransliteration findet ihre obligatorische Ergänzung durch ein vollständiges Graphinventar, das einerseits einen Überblick über den in der Handschrift verwendeten Elementvorrat liefert und andererseits als *Transliterationsschlüssel* fungiert. Selbstredend gilt, dass das Verhältnis zwischen handschriftlichen Phänome-

¹⁸ Zur sog. ›Augenfassung‹, einer leserfreundlichen Online-Synopse des cpg 329 und der elementgetreuen Basistransliteration, vgl. Wernfried Hofmeister: Perspektiven und Auswirkungen des Edierens am Beispiel der neuen Hugo von Montfort-Ausgabe [im Druck].

nen und Transliteration ›umkehrbar eindeutig‹ sein muss, damit der codierte Informationsgehalt 1 : 1 auf die handschriftliche Quelle rückführbar ist.¹⁹

2.2 Technische Umsetzung und Systemintegration der XML-Codierung

Für die Datenhaltung wurde – wie zuvor beschrieben – ein zweistufiger Ansatz gewählt: *imdas pro*, eine Applikation für die Verwaltung von Archivalien und Museumsobjekten, wird für die kodikologische Beschreibung und paläographische Befundung sowie für die Verwaltung der Schriftdokumente verwendet, wobei hier die *Digitalisate* mit eingespeist werden und so bereits in der Übersicht zur schnellen Orientierung zur Verfügung stehen. In der zweiten Stufe wird für jedes Dokument die Basistransliteration in einer TEI-konformen XML-Datei gespeichert. Die mächtigen Codierungsmöglichkeiten von TEI erlauben es, die Informationen zur Basistransliteration vollständig umzusetzen, wobei es im Speziellen aber als innovative Erweiterung dieser Möglichkeiten nötig war, für die Beschreibung von zusammengesetzten Buchstaben (Basisgraph und diakritisches Zeichen) eine Codierung über *compound-Zeichen* durchzuführen.²⁰ Das gestattete es in weiterer Folge, die einzelnen Komponenten getrennt zu markieren und darauf aufbauend die geometrischen Beziehungen zwischen Basisgraph und diakritischem Zeichen zu analysieren. Für jede Seite werden weiters (über ein ergänzendes *Tagging*) Angaben zur Position aller einzelnen Zeichen in einer eigenen Datei abgelegt,²¹ wobei für jedes Zeichen auch die Referenz zum Zeichen in der TEI-Datei gespeichert wird. Die vom TEI-Dokument getrennte Speicherung dieser Informationen sowie die Auftrennung in einzelne Seiten ermöglicht ein gleichzeitiges, verteiltes Arbeiten.

Zur Erfassung dieser detaillierten Informationen wurde ein maßgeschneidertes Annotator-Tool umgesetzt, das die Eingabe der Basistransliteration erlaubt, dessen zentrale Aufgabe jedoch in der Erfassung der Zeichenpositionen liegt. (Das Userinterface des sog. *DAmals-Annotators* ist in Abb. 3 zu sehen; dargestellt ist im oberen Teil ein Ausschnitt aus einer Handschriftenseite, in der ausgewählte Abschnitte annotiert wurden, im unteren Teil die dazugehörige Basistransliteration einschließlich der Zeilenzählung.) Zur Auswertung der erfassten Daten stellt der *DAmals-Annotator* einerseits die Möglichkeit zur Verfügung, gezielt nach Zeichen, Buchstaben oder Buchstabenketten zu suchen und für die Fundstellen automatisch die entsprechenden Regionen aus den

¹⁹ Vgl. die Hugo von Montfort-Editionshomepage (Anm. 9), wo der Transliterationsschlüssel den Basistransliterationen aller Überlieferungsträger unmittelbar beigeschlossen ist.

²⁰ An dieser Stelle sei Hubert Stigler und Petra Steinkellner vom Zentrum für Informationsmodellierung in den Geisteswissenschaften (vgl. Anm. 8) für ihre Entwicklung einer TEI-Codierung (mit speziellen Tags z. B. für Basisgraph-Diakritikum-Verbindungen) herzlichst gedankt.

²¹ Diese Informationen sind in SVG codiert, je Zeichen wird ein Polygon (im einfachsten Fall ein Rechteck) gespeichert.

Digitalisaten auszuschneiden und gemeinsam mit der Positionsinformation als Einzelbilder abzuspeichern, womit – quasi als Abfallprodukt – der Aufbau eines graphischen Zeichen- bzw. Wortinventars möglich ist. Weiters sind die in Abschnitt 4.2 ausgeführten *Bildverarbeitungsmethoden* zur Unterstützung der Suche nach Schreiberhandwechseln – mit entsprechender visueller Darstellung – integriert.

3 ›Graphetische Statistik‹ (zweite Authentifizierungs-Säule)

3.1 Allgemeines zur graphetischen Frequenzanalyse

Statistische Verfahren der Korpusanalyse sind aus der modernen Linguistik nicht mehr wegzudenken: Sie erst führen durch die Möglichkeit der exakten *Quantifizierung* von rekurrenten Phänomenen zu einer Objektivierung des Befundes. Voraussetzung dafür ist die schon im Abschnitt 2.1 beschriebene Sicherung des Datenmaterials durch entsprechende Codierung: Welche Fakten zur Auswertung gelangen können, hängt maßgeblich von der ›Informationstiefe‹ des Datenmaterials ab. Speziell für die Untersuchung von Schreiberhänden ist es wichtig, dass die Transponierung einer Handschrift nicht erst auf der Graphemebene ansetzt, wo ein Großteil der paläographischen Beobachtungen bereits weggefiltert wurde, sondern möglichst nahe an der Handschrift, um den *maximalen Informationsgehalt* zu sichern.²²

Der Mangel an solcherart aufbereitetem Material scheint ein plausibler Grund dafür zu sein, warum sich die statistische Methode mit ihrer exakten *Frequenzzählung* in der Paläographie bis heute nicht durchgesetzt hat; hier werden die ›Haupttendenzen‹ einer Hand, die aufgrund der Häufigkeit des Auftretens bestimmter Merkmale ins Auge stechen, nach wie vor durch relative und damit entsprechend unscharfe Frequenzangaben festgemacht.²³ Hemmen mag den Einsatz statistischer Methoden in der Paläographie auch die Befürchtung, dass z. B. die Berechnung von Durchschnittsmaßen von Buchstabenhöhe und -breite eher zur Nivellierung individueller Eigenarten führen könnte, anstatt diese sichtbar werden zu lassen (Schlögl 264–5).

Diese Gefahr besteht tatsächlich, wenn lediglich Mittelwerte für ganze Handschriften oder großräumige Abschnitte berechnet werden. Um Merkmalwechsel oder Brüche im Verlauf einer Handschrift erkennen und nachweisen zu können, ist es daher notwendig, die Berechnungen seiten- oder spaltenbezogen durchzuführen oder gar noch kleinere Zonen bis hin zu einzelnen Zeilen unter die Lupe zu nehmen. Es liegt auf der Hand, dass das nur mittels solcher Merkmale sinnvoll ist, die entsprechend häufig vorkommen. Daher konzentriert sich *DAMALS* eben nicht in erster Linie auf seltene, auffällige Merkmale, sondern auf Buchstabenformen, -kombinationen bis hin zu Silben und Wörtern,

²² Vgl. Hofmeister-Winter 2005 6 (Graphische Darstellung der Relation von Befund und Deutung).

²³ Vgl. z. B. die jüngste Arbeit von Schneider 2007, die rein auf relativen Frequenzangaben basiert; erste Ansätze von statistischen Methoden in der Paläographie referiert Bromm.

die möglichst zahlreich aufscheinen und einigermaßen homogen über den gesamten Text gestreut sind (vgl. Abschnitt 4.1).

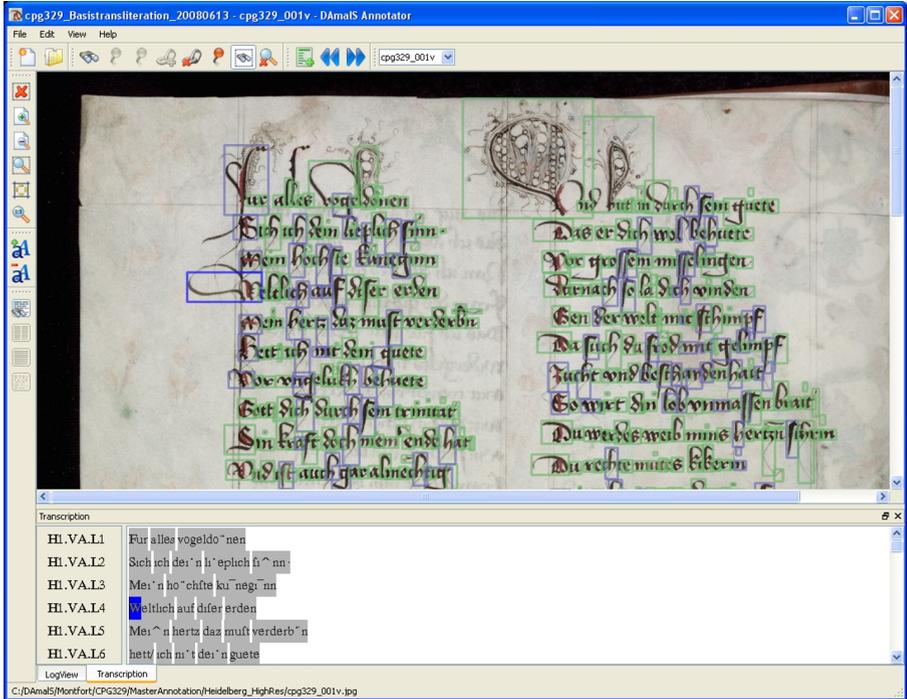


Abbildung 3. DAMaS-Annotator Screenshot.

Welcher Erkenntnisfortschritt auf der Materialbasis einer elementgetreuen Basistransliteration zu erzielen ist, sei im Folgenden anhand der Heidelberger Montfort-Handschrift cpg 329 exemplarisch vorgeführt: Bereits 1881 hatte Josef Wackernell in Zusammenhang mit seiner Edition der Texte des adligen Dichters (Hugo von Montfort 1881) ausführliche Untersuchungen zu den beteiligten Schreiberhänden angestellt, mit denen er sich als ›Kronzeuge‹ des gesamten interdisziplinären Methodenrepertoires seiner Zeit erwies und ein großartiges Lehrstück der mediävistischen Textforschung lieferte. Seine verdienstvollen paläographischen und sprachwissenschaftlichen Befunde bildeten einerseits die Ausgangsbasis, andererseits eine Art ›Reibebaum‹ für die Analysen des Codex im Rahmen des Projekts DAMaS. Wackernells Ergebnisse galt es in einem ersten Schritt auf ihre Stichhaltigkeit zu überprüfen und – da sein Resultat nicht restlos überzeugte – durch weitere Untersuchungen zu ergänzen.

Wackernell ging bei seinen Untersuchungen systematisch ›von außen nach innen‹ vor und bemühte sich sichtlich um Glaubwürdigkeit seiner Behauptungen, indem er (im Rahmen seiner bescheidenen technischen Möglichkeiten) Häufigkeitsangaben machte. Frequenzzählungen gestalteten sich in den Anfängen der Sprachwissenschaft denkbar mühsam, konnten sie doch lediglich auf der Basis händischer Auszählung erfolgen. Absolute Frequenzzählung findet sich daher nur bei selten auftretenden Phänomenen, während ansonsten relative Häufigkeitsangaben bevorzugt werden: ›stets – meist – häufig – selten – nie‹, so lautet in etwa die gängige Skala dieser intuitiven Befundungsmethode. Selbst wenn gelegentlich konkrete Zahlenangaben angeführt werden, sind diese oft nicht nachvollziehbar – nicht so sehr aufgrund von Irrtum (der gerade in diesem Bereich im wahrsten Sinn des Sprichworts *Errare humanum est* verzeihlich wäre), sondern weil nicht immer deutlich genug deklariert ist, nach welchen Kriterien Belege gezählt oder ausgeschieden wurden (Hofmeister-Winter 2007 94–5).

Als Erstes begutachtete Wackernell das Layout des Codex und stellte Diskontinuitäten beim Seitenspiegel resp. bei der Schreibraumeinteilung fest, die sich konkret an der durchschnittlichen Zeilenzahl pro Spalte festmachen ließen und ihre Ursache in einer geringfügig abweichenden Schriftgröße haben. Bei der paläographischen Schriftanalyse konzentrierte sich Wackernell auf Einzelmerkmale wie z. B. die Schaftform von Langs, f und p oder bestimmte Verzierungen an Ober- bzw. Unterlängen der ersten/letzten Zeile einer Seite sowie auf das Auftreten von Graphvarianten wie etwa Ligatur-r. Im Bereich der Graphie-Unterschiede untersuchte Wackernell die Verwendung von Majuskel <R>, die (in Text Nr. 38) nicht nur obligatorisch am Zeilenanfang, sondern auch im Zeileninneren auftritt. Graphieunterschiede boten sich aber auch im Bereich der graphischen Umsetzung bestimmter Phoneme für eine genauere Untersuchung an; so erscheint z. B. die Wiedergabe des frühneuhochdeutschen Diphthongs [ei] in den von Wackernell ausgemachten Zonen teils in den moderneren Diphthong-Graphien <ei, ey, ai, ay>, teils in der konservativen (weil noch nicht diphthongierten) Schreibung mittelhochdeutscher Handschriften als <i> (ggf. mit Superskript).

Aus der Summe dieser und weiterer Beobachtungen zog Wackernell den Schluss, dass an der Niederschrift des cpg 329 insgesamt vier Schreiberhände beteiligt gewesen seien, denen er mangels Identifizierbarkeit die Bezeichnungen A, B, C und D zuwies. Er legte auch die Zonen fest: Demnach habe Hand A fol. 1r–12v geschrieben (diese Grenze ist zugleich eine Lagengrenze, wenn sie auch mitten durch Text Nr. 12 verläuft), Hand B fol. 13r–46v, Hand C fol. 47r–48va und Hand D (als einzige deutlich und zweifelsfrei unterscheidbar von den übrigen angenommenen Händen) fol. 48vb–52va (Hugo von Montfort 1881 CXII–CXX; vgl. Hofmeister-Winter 2007 93, Tab. 1).

Die von Wackernell sorgfältig gesammelten Indizien zeigen leider einen Mangel: Sie sind größtenteils nicht wirklich tragfähig – zu wenig ist über die Entstehungsumstände des Codex, die Qualität (Homogenität/Heterogenität) der Vorlage(n) und folglich über den

Umgang der von Hugo beauftragten Kopisten mit ihren Vorlagen bekannt.²⁴ Daher lassen sich zu allen Hypothesen Wackernells Zweifel und Gegenargumente anführen, was die Forschung der jüngeren Vergangenheit zu einem radikalen Rückzug veranlasste: Einige gingen sogar so weit, für die drei fraglichen Zonen A, B und C mangels ausreichender Beweise überhaupt nur mehr einen einzigen Schreiber anzunehmen (Spechtler; Werner; Welker).

Auch an der methodischen Vorgehensweise Wackernells gibt es aus Sicht der modernen Linguistik einiges zu kritisieren: Pauschalaussagen über paläographische Detailbeobachtungen haben höchstens den Status persönlicher Eindrücke, wenn sie nicht objektiv belegt werden können²⁵ – man bedenke, dass die zeitgenössischen Leser seiner detaillierten Abhandlung über die Schreiber des cpg 329 noch nicht einmal ein Faksimile zur Verfügung hatten, um die Aussagen wenigstens visuell zu prüfen. Eine Möglichkeit, hier Abhilfe zu schaffen, bestünde in der (partiellen) Vertiefung der Transliteration, wie sie am Beispiel der <h>-Varianten demonstriert werden konnte. Um auch die *mikroskopischen* Feinheiten der Graphvarianz zu erfassen, wurde von uns eine *Komponentenanalyse* versucht, die den Graph in seine morphologischen Bauelemente (Schleife, Schafffuß und Bogen) zerlegt und diese Variablen jeweils getrennt beschreibt (Hofmeister-Winter 2007 97–100, bes. Tab. 2, und 114, Abb. 9). Diese Untersuchung, die sich wegen des nicht unerheblichen Deskriptionsaufwandes auf die ›neuralgischen‹ Zonen des Codex (rund um die von Wackernell konstatierten Schrebergrenzen) beschränkte, lieferte tatsächlich signifikante Ergebnisse hinsichtlich der von A, B und C bevorzugten Kombinationsformen, zeigte aber auch deutlich das Problem der *stilistischen Schwankungsbreite* von Schreiberhänden auf, das von der Forschung bisher noch nicht eingehend untersucht worden ist.

Problematisch an Wackernells Vorgehen ist weiters die Beziehung von Phänomenen mit sehr geringer Belegzahl oder gar ›Hapaxlegomena‹: Es stimmt zwar, dass im cpg 329 die Graphvariante Ligatur-r in der angenommenen Zone A überhaupt nicht vorkommt, aber da sie in Zone B und C nur 1% aller <r>-Belege stellt, ist die Signifikanz dieses Phänomens als äußerst schwach einzustufen (Hofmeister-Winter 2007 100–1). Wackernell findet sich zwar in Einklang mit der traditionellen Paläographie, die bei der Beurteilung von *Schreiberindividualität* seit jeher den Blick schwerpunktmäßig den Abweichungen, dem Außergewöhnlichen widmet, das eben oftmals nur vereinzelt auftritt. Jedoch er-

²⁴ Im Fall der Texte Hugos von Montfort ist damit zu rechnen, dass diese ursprünglich der vorarlbergischen Herkunft des Dichters entsprechend mehr oder weniger stark alemannisch gefärbt waren; die Eintragung in den von Hugo selbst in Auftrag gegebenen Codex erfolgte jedoch in der Zeit, die er aus privaten wie beruflichen Gründen in der Steiermark verbrachte (ab ca. 1414), vermutlich durch Schreiber aus der Region, welche die Texte womöglich eigenständig in unterschiedlichem Maße bairisch überformten (wobei ungeklärt bleiben muss, wie weit nicht schon die Kopiervorlagen der über mehrere Jahrzehnte entstandenen Dichtung dialektal schwankten). Vgl. Hugo von Montfort 2005 XXIV–XXV.

²⁵ Vgl. z. B. Wackernell über die Schaffform von Lang-s bei den verschiedenen Händen (Hugo von Montfort 1881 CXVII).

schweren selten vertretene Merkmale die Bestimmung von Schreibergrenzen insofern, als keine flächendeckende Markierung der Schreiberbereiche gegeben ist. Schon aus diesem Grund sollte das Augenmerk – nicht nur, aber mit gleich viel Akribie – auf *hochfrequente* Phänomene gelegt werden, die eine möglichst gleichmäßige Streuung über den ganzen Text aufweisen.

Sowohl Layout- und schriftstilistische als auch Graphieunterschiede dürfen nicht vorbehaltlos als Ausdruck von Schreiberindividualität betrachtet werden. Sie stehen unter dem Einfluss vielfältiger Faktoren, die oft außerhalb des Schreiberindividuums liegen und daher nicht von diesem gesteuert werden; entsprechend umsichtig gilt es, sie im Rahmen von statistischen Auswertungen zu berücksichtigen:

- Das *Layout* kann z. B. vom Auftraggeber veranlasst sein, der sich ein bestimmtes Design wünscht; im Fall von Arbeitsteilung müssen sich alle Schreiber dem Gesamtkonzept beugen, dies gilt vor allem ab dem zweiten Schreiber. Dennoch sind Umstände denkbar, die dazu führen, dass die Kontinuität des Layouts durchbrochen wird: Ein Wechsel der Schreiberhand ist nur eine Möglichkeit unter vielen.
- Die abwechselnde/gleichzeitige Verwendung verschiedener *Formvarianten* eines Schriftsymbols ist noch kein Beweismittel für die Authentifizierung einer Schreiberhand, denn es ist nicht unüblich, dass ein (zumal geübter) Schreiber mehrere stilistische Varianten in seinem Repertoire hat, ja sogar mehrere Duktus (sog. Anlasshände), die er entweder willkürlich, kontextabhängig oder aber bewusst für verschiedene Zwecke/stilistische Ansprüche einsetzt.²⁶
- Unterschiede der *Graphie* schließlich, also Abweichungen hinsichtlich der Umsetzung gesprochener Laute in Schrift, können, müssen aber nicht auf den individuellen Schreibusus der beteiligten Personen zurückgehen, sodass diese den Vorlagentext sozusagen bewusst ihren eigenen Graphiegewohnheiten und damit indirekt ihrem Dialekt angeglichen hätten. Es muss immer auch mit inhomogenen Vorlagen gerechnet werden.

Die Phänomene, in denen Wackernell (und mit bzw. nach ihm die traditionelle Paläographie) auch anhand statistischer Gesichtspunkte nach Schriftindividualität gesucht hat, hängen großteils von äußeren Einflüssen ab. Für die Fahndung nach Schriftindividualität sind jedoch besser solche *Schriftmerkmale* aufzuspüren und auszuwerten, die möglichst ›unbelastet‹ von Vorbildern und Normen sind und am besten beim Schreibvorgang nicht im Zentrum der Aufmerksamkeit des Schreibers liegen und daher nicht ›bewusst‹ ausgeführt werden. Derartige Merkmale sind eher im peripheren Bereich des Schriftsystems zu entdecken und in der Regel dadurch charakterisiert, dass sie nicht un-

²⁶ So beobachtet bei der Augsburger Berufsschreiberin Clara Hätzlerin, die u. a. zwei völlig verschiedene <d>-Formen gebraucht, die sie teils (fast) ausschließlich, teils willkürlich vermischt einsetzt. Vgl. Hofmeister und Hofmeister-Winter 2008 104–7 und 114–5 (Abb. 4–6).

mittelbar sprachsystemrelevant sind, sondern redundant oder als Zierelemente ›überflüssig‹ und daher in gewisser Weise entbehrlich (Hofmeister-Winter 2007 106–7).

Zwei Elemente der Schrift, auf die diese Eigenschaften in hohem Maß und dabei mit statistisch signifikanter Frequenzstreuung zutreffen, scheinen der *i-Punkt* und haarfeine *Zierstriche* an bestimmten Graphen zu sein: Diese unscheinbarsten Elemente des Schriftsystems, deren Existenz sich oft nur am Original verifizieren lässt, werden aus schreibpsychologischer Sicht kaum jemals bewusst gesetzt, sondern verdanken sich als periphere Merkmale der betreffenden Schriftsymbole eher einem *graphomotorischen Reflex*, ausgelöst durch den einmal erworbenen Usus. Die *i-Punkt-Setzung* scheint – zumindest in der Zeit von ihrem ersten Auftreten ab 1320 (Schneider 1999 49) bis zur Normierung des *i-Punkts* im Schriftsystem als ›Diakritikum‹ – deshalb noch in starkem Maße ›unbelastet‹, weil sie dem Belieben des Individuums unterstellt und daher weitgehend vorlagenunabhängig ist. Für gegenwartssprachliche handschriftliche Texte liegen noch keine Vergleichszahlen vor, jedoch sollte sich hier die Frequenz aufgrund der Normiertheit des *i-Punktes* der 100 %-Marke annähern. Dennoch ist denkbar, dass die durchschnittliche Zahl der ›vergessenen‹ *i-Punkte* auch in der Gegenwart Schreiberindividuen charakterisiert. Solche Schrift Elemente können als *psychometrische Merkmale* bezeichnet werden, da sie den Schriftzügen quasi wie eine DNA eingeschrieben sind und in Summe die ›Grammatik der Schreiberhände‹ prägen. In der Basistransliteration des cpg 329 wurden einige dieser Schreibereigenheiten mittels der oben beschriebenen Codierungsmethode festgehalten (vgl. Abschnitt 2.1).

Zur statistischen Beobachtung von charakteristischen Zügen einer Hand im Bereich der schriftsystemrelevanten Elemente eignen sich ganz besonders auch *Abkürzungen*. Wieder gilt es zunächst das Repertoire zu erheben und in der Basistransliteration differenziert zu codieren, sodann Frequenz und Streuung der verschiedenen Formtypen zu ermitteln und schließlich mit Hilfe einer graphetischen Analyse die oft individuell geprägte Verwendungsweise für bestimmte Graphsequenzen oder formelhafte Wörter zu untersuchen. Zusätzlich kann die Ermittlung des generellen ›*Kürzungsindikators*‹ lohnend sein, also der Frequenz des Einsatzes von Kürzungssymbolen im Allgemeinen (am besten zeilenbezogen) – dies aber nur unter bestimmten Voraussetzungen, nämlich wenn ein Codex strikte Layoutvorgaben mit Randausgleich (in der Typographie als ›Blocksatz‹ bezeichnet) aufweist, der die Schreiber zur ökonomischen Befüllung des zur Verfügung stehenden Schreibraumes zwingt.²⁷ Das ist im cpg 329 allerdings nicht der Fall, weshalb keine Notwendigkeit zu Kürzungen bestand, so dass Schrift Elemente dieser Kategorie eher sporadisch auftreten und ihre Frequenzwerte geringe Aussagekraft besitzen.

²⁷ Vgl. Hofmeister 2001 94. Das Ziel dieser Studie war allerdings etwas anders gelagert: Hier ging es darum, im cpg 848 verschiedene Schreibschichten nachzuweisen, nämlich Grundschrift und Nachträge, für die von den Grundschriftschreibern ein exakt bemessener Freiraum ausgespart worden war, der von den Nachtragsschreibern möglichst ohne erkennbare Bruchlinie ausgefüllt werden musste.

Die statistische Auswertung des ›mikro-codierten‹ Materials kann beispielsweise folgende Fragestellungen behandeln:

- Ob z. B. i-Punkte gesetzt sind bzw. in welcher *durchschnittlichen Frequenz*: Schon diese Werte können von Schreiberindividuum zu Schreiberindividuum signifikant differieren. Für eine objektive Feststellung des Durchschnitts ist es allerdings erforderlich, von den durch Wackernell prädisponierten Zonen abzusehen und das Material in einer Weise zu analysieren, die Frequenzschwankungen unvoreingenommen anzeigt. Wir haben zu diesem Zweck den Wert für jede Spalte gesondert ermittelt, und zwar die Belegzahl pro Zeile. Im cpg 329 kristallisierten sich auf diese Weise drei deutlich voneinander abgesetzte Zonen heraus: A, B/C und D (Hofmeister-Winter 2007 107 und 116, Abb. 14). Im Vergleichskorpus von Veit Feichter (ausgewertet wurden alle drei bekannten Codizes des Schreibers) zeigte die durchschnittliche Frequenz des i-Punktes innerhalb von gut zehn Jahren lediglich minimale Abweichungen.²⁸
- Ob der i-Punkt-Gebrauch bestimmten *Normen* folgt: Hier ergaben sich im Fall der Analyse der Schriften der Clara Hätzlerin, die dem Projekt DA*m*als als ›Eichwerkzeug‹ dient (vgl. Abschnitt 1.3), erstaunliche Einsichten, nämlich dass diese routinierte Schreiberin offenbar für sich klare Regeln entwickelte, die sie noch dazu peinlich genau einhielt. Die genauere Untersuchung dieser Regeln ergab, dass die Schreiberin äußerst ökonomisch verfuhr, indem sie i-Punkte nur dort setzte, wo sie zur Differenzierung des Basisgraphen <i> unbedingt nötig sind bzw. den Les-/Verstehensprozess unmittelbar fördern, also in direkter Nachbarschaft mit solchen Buchstaben, die aufgrund ihrer Bauweise (Kurzschäfte) zu Missverständnissen führen können; das sind im Fall des kursiven Schreibsystems der Hätzlerin u, m, n (Hofmeister und Hofmeister-Winter 107–8).

Die Eruierung dieser schreiberindividuellen Normen liefert viel exaktere Befunde als die bloße Ermittlung der durchschnittlichen i-Punkt-Frequenz innerhalb eines Textes. Da eine verbindliche überindividuelle Schreibnorm im 15. Jh. noch nicht ausgeprägt war, kann die Erstellung eines *Schreiberprofils* wertvolle Indikatoren für die Authentifizierung von Schreiberhänden liefern. Die Zuverlässigkeit dieser Methode muss allerdings durch die serienmäßige Analyse anderer Schreiber weiter abgesichert werden.

Durchgeführt wurden die beschriebenen Untersuchungen mit Hilfe von Beleglisten, die zu jedem Beleg auch den genauen Fundort angeben, so dass jeder einzelne Repräsentant eines bestimmten Phänomens bis in die Handschrift zurückverfolgt werden kann. Dies ist zur Kontrolle der Belege unerlässlich, denn da Transliterationen in der Regel nicht von allem Anfang an die Disambiguierung aller einzelnen Wortformen durch Annotationen vorsehen, müssen Homographen auf diese Weise identifiziert und ggf.

²⁸ Die Abweichung vom Gesamtdurchschnitt (98,8 %) beträgt weniger als 1 %. Vgl. Hofmeister-Winter 2007 108–9.

ausgesondert werden. Zugleich bietet die *exakte Beleg-Verortung* maximale Transparenz

A		B		C		D	
 001ra12W04	 011ra04W03	 013ra14W03	 045rb11W03	 047ra02W01	 047rb07W03	 049rb05W03	 050rb07W03
 001va17W04	 011ra18W03	 013ra14W05	 045va12W05	 047rb06W02	 047rb12W05	 049vb01W05	 051va03W02
 001vb12W03	 011ra22W05	 013rb07W04	 045vb07W05	 047rb07W03	 047rb16W03	 049vb03W04	 051va05W03
 002va11W03	 011ra29W03	 013va03W04	 046ra02W01	 047rb12W05	 047va01W02	 049vb07W04	 051vb06W04
 002va16W02	 011va01W07	 013vb05W03	 046rb05W03	 047rb16W03	 047va04W04	 049vb21W05	 052ra16W06
 002va25W05	 011va03W05	 013vb11W04	 046va03W03	 047va01W02	 047va10W04	 049vb30W06	 052va20W03
 002vb02W03	 012rb02W05	 014ra01W04	 046va10W03	 047va04W04	 047va12W04	 050ra01W05	 052va20W07
 002vb15W06	 012va07W03	 014ra07W03	 046va11W03	 047va10W04	 047va18W04	 050ra16W02	
 002vb26W03	 012va14W05	 014ra16W05	 046va23W03	 047va12W04	 047va27W04	 050ra22W03	
 003ra06W03	 012vb02W03	 014ra17W04	 046va28W02	 047va18W04	 047va28W01	 050ra34W04	

Abbildung 4. Liste von <die>-Belegen mit exakten Positionsangaben.

und erleichtert die Nachvollziehbarkeit der Untersuchungen (vgl. Abb. 4; die Stellenangabe umfasst Blattzahl, Spalte, Zeile sowie – mit der Sigle »W« – die Wortnummer des Belegs innerhalb der Zeile).

3.2 Technische Umsetzung der statistischen Auswertung

Ausgehend von der im Abschnitt 2.2 dargestellten detaillierten Erfassung aller Schriftsymbole und Zeichenpositionen ergibt sich eine Vielzahl von Möglichkeiten, diese Grundinformationen statistisch auszuwerten. So kann man z. B. automatisch Zeichen- und Wortindizes – ergänzt durch entsprechende Belege aus dem Digitalisat – erzeugen oder Zeichen- und Worthäufigkeiten für den gesamten Codex, je nach Fragestellung auch pro Seite, Spalte oder Zeile, evtl. zusätzlich hinsichtlich Buchstabenvarianten detailliert, berechnen. Für weitergehende Auswertungen und die Erzeugung entsprechender Schaubilder ist der Export über den DAMALS-Annotator in standardisiertem Format (csv-Dateien) vorgesehen: Damit lassen sich dann (auch via Excel-Dateien) *Visualisie-*

rungen z. B. der Frequenzverteilung anhand besonders aussagekräftiger graphischer Diagramme generieren.²⁹

4 ›Musterorientierte Schriftbild-Erfassung‹ (dritte Authentifizierungs-Säule)

4.1 Allgemeines zum musterorientierten Ansatz

Jegliche mündliche Sprachreproduktion bedient sich eingeübter Muster, indem nicht Einzellaute artikuliert werden, sondern miteinander zu Klangbildern verschmelzende Lautfolgen von Silben oder Wörtern; als individuell eingeübte Schemata steigern sie die Ökonomie und Persönlichkeit unserer Kommunikation. Wenn Sprache *schriftlich* ausgedrückt wird und auch diese ›sekundäre‹ Form der Sprachreproduktion flüssig (also nicht Buchstaben für Buchstaben malend, sondern etwa in routinierter Kursive) erfolgt, ist mit demselben Phänomen einer ›Amalgamierung‹ zu rechnen:³⁰ Dies mag man sich dadurch erklären, dass die Art, wie die Buchstaben innerhalb einzelner Silben oder Wörter zusammengefügt werden, ebenfalls einer Art *mentaler Matrix* gehorcht, durch deren eintrainiertes ›Lexikon‹ die Feinmotorik einer schreibenden Hand gesteuert wird; ähnlich wie auf der Laut-Ebene ist dann auch in den Schriftzügen mit charakteristischen Verschleifungen und ensembleartigen *Musterbildungen* zu rechnen. Dies darf zum einen für komplexer geformte Buchstaben und Ligaturen angenommen werden, zum andern aber auch für ganze Silben bzw. Morpheme und Wörter: Bei deren Verschriftung kann es daher zu *stempelartigen* Ausprägungen kommen.

Die ›klassische‹ Paläographie hat ihre bisherigen Befundungen von Schriftzügen fast ausschließlich auf die Beobachtung von Einzelbuchstaben gestützt, allenfalls ergänzt um Ligaturen. Damit hat sie unbestreitbar ein beachtliches Register an Kriterien entwickelt, das außer für die *zeitliche* und *räumliche* Einordnung von Schriftproben auch für das Erkennen von *individuellen* Schrift-Charakteristika und damit für das Unterscheiden einander recht ähnlicher Handschriften äußerst hilfreich ist (Schneider 1999). Dennoch blieben – wie schon am Beginn dieses Beitrags festgestellt – zahlreiche Zweifelsfälle bislang ungeklärt. Um nun sowohl auf dieser graphetischen Ebene als auch auf der (hier vergrößert so genannten) ›Wort-Ebene eine Verbesserung unserer Seh- und Trennschärfe zu erreichen, bedient sich DAMALS aller modernen Möglichkeiten der *digitalen Bildverarbeitung*, indem elektronische Abbildungen von Handschriften mit einbezogen werden.

²⁹ Diese Funktionalitäten werden vom DAMALS-Annotator derzeit nur teilweise angeboten, sollen jedoch im Fortsetzungsprojekt MOSES komplettiert werden.

³⁰ Zur gestalthaften Abbildung lexikalischen Sprachmaterials vgl. Aitchison. Über allgemeine graphematische Zusammenhänge reflektiert anschaulich Grabowski.

Voraussetzung für eine graph- und wortgenaue Befundung von Schriftzügen ist die Verknüpfung aller Transliterations-Elemente mit den dazugehörigen Bildinhalten. Sie wird (vom unten näher beschriebenen und eigens für DAMaIS entwickelten) *Annotator-Tool* geleistet: In zwar sehr mühsamer, am Ende jedoch ebenso lohnender Arbeit werden mit seiner Hilfe alle Graphe und Graph-Elemente durch eine exakte Positionsangabe mit dem Bild verbunden. Erst durch diese individuelle *Adressierung* gelingt es, von Beleg zu Beleg Vergleichsmessungen durchzuführen und entsprechende Divergenz- oder Konvergenzwerte zu gewinnen, die in die XML-Beschreibung einfließen können.

Um nun etwa für diverse *Leitbuchstaben/Kenngraphien* vergleichbare Messwerte zu erhalten, werden entweder ganze Buchstaben oder auch nur auffallend variante Teile von ihnen (wie Unterlängen, Schlaufenformen oder Diakritika) herausgegriffen und miteinander verglichen. Die bislang überzeugendsten Erfolge können dabei (wie schon oben erwähnt) durch die gesonderte Erfassung der Positionierung des i-Punkts erzielt werden, weil dieses Auszeichnungselement zu jenen gehört, die sich allem Anschein nach einer bewussten Steuerung entziehen und entsprechend viel über die individuelle *psychomotorische Routine* einer Schreiberhand verraten.

Darüber hinaus wird der Blick aber auch auf die bislang fast vollständig ignorierte *Wort-/Silben-Ebene* erweitert; auch dafür bietet die elektronische Bildverarbeitung bislang ungeahnte Möglichkeiten: So können alle dafür in Frage kommenden Graphsequenzen als ganze vermessen, ausgeschnitten und in diverse Listen exportiert werden, wo sie dann – nach Ähnlichkeit oder Position dicht an dicht gereiht – in Form eines ›virtuellen Morphings‹ neu beurteilbar werden. Am Beispiel des Wortes ›die‹ scheint dies durch DAMaIS bereits geglückt, und zwar sowohl für die zu authentifizierenden Hände im cpg 329 (vgl. Abb. 4) als auch für die identifizierte Hand der Clara Hätzlerin (Hofmeister und Hofmeister-Winter).

Die soeben erwähnte Auswahl des bestimmten Artikels ›die‹ als ›Stempelwort‹ führt zur Frage nach ihrem methodischen Hintergrund und einer Ausschau nach weiteren tauglichen Wörtern für eine musterorientierte Untersuchung. Diese Selektion gehorcht, kurz gesagt, folgenden Kriterien: Als Stempelwort eignen sich lexikalisch-morphologische oder silbische Einheiten, sofern sie *hochfrequent* sind und dazu – in Summe – *alle paläographisch wichtigen Kenngraphien* abdecken. Moderne Wortfrequenzlisten helfen zwar bei der Einengung dieser Suche in historischen Texten,³¹ da eben Grundformen aus dem Bereich der Pronomina, Präpositionen oder Konjunktionen seit jeher häufig auftreten, aber eine darauf fußende Liste wird nicht für alle An-

³¹ Vgl. die Statistik des Projekts ›*Wortschatz*‹ des Instituts für Informatik/Abteilung für Sprachverarbeitung an der Universität Leipzig [zuletzt eingesehen am 7.2.2009]. Demnach sind die 25 häufigsten Wortformen der deutschen Gegenwartssprache (in absteigender Reihenfolge): der, die, und, in, den, von, zu, das, mit, sich, des, auf, für, ist, im, dem, nicht, ein, Die, eine, als, auch, es, an, werden. Eine ähnliche Statistik lässt sich auch für diverse frühneuhochdeutsche Textcorpora ermitteln (vgl. die folgende Anm.).

wendungsfälle geeignet sein,³² denn da gilt es auf textsortenspezifische Bedingungen Rücksicht zu nehmen: Durch derartige Grundperspektiven von Texten kann z. B. das Auftreten des an sich häufigen Personalpronomens »ich« ausgeschlossen oder stark eingeschränkt sein. Somit empfiehlt es sich, ein *Register* von präsumtiven Stempelwörtern/musterbildenden Graphsequenzen anzulegen, aus dem dann für jeden Untersuchungsfall ein *maßgeschneidertes Bündel* (mit vielen Leitgraphien) ausgewählt wird.

In Summe können alle erwähnten Prozeduren für die schon mehrfach erwähnte *Brillenfunktion* von *DAmals* genützt werden: Durch Visualisierung aller Messdaten, die sich (entlang der Aufzeichnungsrichtung) als Abweichungen von bis dahin etablierten Stempelwort-Grundmustern ergeben haben, kann man jede signifikante Abweichung direkt im digitalen Faksimile markieren. Wie sich diese »*Rückprojektion*« technisch umsetzen lässt, wird nun in Abschnitt 4.2 erläutert.

4.2 Technische Umsetzung der Mustererkennung

Um das Ziel von *DAmals* – eine intersubjektiv besser argumentierbare Entscheidung hinsichtlich des Auftretens eines Schreiberwechsels – zu unterstützen, wurden Bildverarbeitungsmethoden auf der *Makroebene* (Einbeziehung des gesamten Codex bzw. zumindest mehrerer Seiten oder Spalten daraus) und der *Mikroebene* (Untersuchung und Vergleich von Zeichen, Buchstaben oder Wörtern) umgesetzt. Das Ergebnis der automatisch ablaufenden Methoden der Makroebene ist die unmittelbare – mit Unsicherheit behaftete – Einschätzung, ob ein Schreiberwechsel aufgetreten ist. Die Mikroebene ergibt weitere objektive Informationen, die den subjektiven Entscheidungsprozess unterstützen.

Der Entscheidungsprozess auf der *Makroebene* läuft wie folgt ab:³³ In einem ersten Schritt werden in den Digitalisaten jene Bereiche identifiziert, die Text enthalten, da die nachfolgenden Analysen durch graphische Elemente (Fleuronnée etc.) gestört werden würden. Danach werden diese Textblöcke in Segmente (die ein oder mehrere Zeilen enthalten können) getrennt und diese ihrerseits in Regionen zerlegt, die grob Buchstaben gleichzusetzen sind. Diese Regionen werden nun weiter analysiert, wobei einerseits Maße errechnet werden, die von der Kontur der Region abgeleitet werden, andererseits solche, die auf der Strichstärke in bzw. Abständen zwischen Regionen beruhen.

Basierend auf der Kontur werden drei Maße berechnet: der Winkel eines Kontursegmentes gegenüber der Horizontalen, der Winkel, den zwei aufeinanderfolgende Kon-

³² Ein hilfreiches Tool zur Frequenzbestimmung von Wörtern auch in historischen Texten bietet Wordle. (Es ist hier allerdings zu beachten, dass zwar die vom Programm ermittelten Zahlenangaben stimmen, aber bei der graphischen Ausgabe dieser Werte – die keinen sprachwissenschaftlichen, sondern einen ästhetischen Zweck verfolgt – gerade die häufigsten Wörter ausgeblendet werden.)

³³ Die dargestellte Methode wurde von Severin Stampfer in seiner Diplomarbeit »Discrimination of Scribes in Medieval Manuscripts« (Graz 2009) unter anderem basierend auf Arbeiten von Bulacu und Schomaker sowie Srihari et al. entwickelt.

tursegmente einschließen, sowie die Winkel zweier Kontursegmentpaare, die einander horizontal oder vertikal unmittelbar gegenüberstehen.

Hinsichtlich der Strichstärke in bzw. Abständen zwischen Regionen werden sog. Lauflängen berechnet, das sind die Längen horizontaler bzw. vertikaler Segmente zwischen Konturpixeln entweder der Schrift selbst oder des Hintergrundes. Die Abstände innerhalb von Hintergrundregionen umfassen dabei die Ausmaße eingeschlossener Regionen sowie die Abstände zwischen Buchstaben, innerhalb von Vordergrundregionen werden Strichstärken und Buchstabenproportionen erfasst.

Alle diese Einzelmaße werden nun für die einzelnen Regionen eines Segmentes berechnet und in *normalisierten Histogrammen* je Segment zusammengefasst. Die normalisierten Histogramme aller Maße bilden in Summe einen sog. *Deskriptor*, der die charakteristischen Eigenschaften dieser Region beschreibt.

Dieser Deskriptor ist nun der Ausgangspunkt, um die eigentliche Detektion von Schreiberwechseln durchzuführen. Der Vorgang läuft wie folgt ab: Unter der Annahme, dass ein Schreiberwechsel am Beginn eines Abschnittes unwahrscheinlich ist, werden für eine vordefinierte Anzahl von teilweise überlappenden Segmenten am Beginn eines Abschnittes die einzelnen Deskriptoren berechnet und zu einem Gesamtmodell zusammengefasst. Dieses Gesamtmodell soll die charakteristischen Eigenschaften des präsumtiven Schreibers repräsentieren. Nun werden für die folgenden Segmente wiederum die Deskriptoren berechnet und über die *Ähnlichkeit* zwischen Gesamtmodell und Deskriptor für neue Segmente die Wahrscheinlichkeit der Zugehörigkeit berechnet. Ein Wechsel wird als erkannt angenommen, wenn genügend Deskriptoren vom aktuellen Modell stark abweichen, wobei der notwendige Grad der Abweichung ein Parameter der Methode ist.

Wurde ein Wechsel erkannt, so wird ein neues Gesamtmodell für den nächsten Schreiber aufgrund einer vorgegebenen Anzahl von Deskriptoren ab dem Schreiberwechsel erstellt und der Vorgang wiederholt sich. Wurde kein Wechsel erkannt, wird das aktuelle Gesamtmodell mit den für die neu hinzugekommenen Segmente berechneten Deskriptoren verfeinert. Alle diese Schritte laufen vollständig *automatisch* ab, also ohne die Notwendigkeit von Benutzereingaben.

Die Methode wurde mit unterschiedlicher Parametrisierung auf Datensätzen mit bis zu 16 verschiedenen (bekannten) Schreibern, getestet. Dabei konnten Erkennungsraten zwischen 80 und 91 % erzielt werden, wobei im besten Fall 92 % der vom System gemeldeten Wechsel korrekt waren.

Für die *Mikroebene* werden einzelne Buchstaben oder kurze Buchstabenfolgen (maximal im Ausmaß einzelner Wörter, nämlich der hochfrequenten Stempelwörter, vgl. Abschnitt 4.1) herangezogen und dafür folgende Deskriptoren berechnet:

- Abstand zwischen Basisgraph und diakritischem Zeichen
- Winkel zwischen Basisgraph und diakritischem Zeichen

Als Ausgangspunkt für die Bestimmung dieser beiden Deskriptoren wird der Punkt oben in der Mitte des Basisgraphen genommen und davon ausgehend zur Mitte der Region des diakritischen Zeichens gemessen.

- Strichstärke
- Eingeschlossene Fläche
Diese beiden Deskriptoren werden mit dem oben beschriebenen Verfahren zur Berechnung von Lauflängen ermittelt. Für die Ermittlung des Maßes für die Strichstärke wird der Mittelwert eines Teiles der kurzen Lauflängen herangezogen.
- Schriftneigung
Hierzu wird das Histogramm der Winkel zwischen Kontursegmenten und der Horizontalen in einer Region berechnet; der Schriftneigungsgrad ergibt sich als das Maximum im Histogramm dieser Werte.
- Visuelle Ähnlichkeit
Für die Berechnung der visuellen Ähnlichkeit zwischen Buchstabenregionen gibt es eine Reihe von Methoden, basierend z. B. auf der Farbverteilung, den dominanten Farben oder der Textur des Bildes. Für das Projekt DAMaLS hat sich die Verwendung so genannter *Kovarianz-Matrizen*, in denen statistische Informationen zur Farb- bzw. Grauwertverteilung innerhalb einer Region gesammelt werden, als am geeignetsten herauskristallisiert bzw. der Einsatz des SIFT-Deskriptors (Lowe), der die Häufigkeit von Richtungen in einer Region beschreibt.

Um von der Auflösung der Digitalisate unabhängig zu sein und somit codex-übergreifende Analysen durchführen zu können, werden alle längen- bzw. flächenbezogenen Deskriptoren auf die Fläche der Region normalisiert.

Zur weiteren Analyse können entweder der Verlauf all dieser Deskriptoren einzeln oder die Unterschiede im Verlauf herangezogen werden. Zur Berechnung der Unterschiede wird im einfachen Fall von Einzelwerten (wie z. B. für Schriftneigung oder Winkel) der Absolutbetrag der Differenz dieser Einzelwerte, für komplexere Maße (wie z.B. beim visuellen Vergleich) speziell abgestimmte Unterschiedsmaße (deren Erklärung den Rahmen dieses Beitrages sprengen würde) herangezogen. Für die *Visualisierung* des Verlaufes können einfache Liniengraphiken verwendet werden. Im Projekt sind jedoch spezielle Methoden entwickelt worden, die den Verlauf eines Maßes in Form eines in der Transparenz modulierten Bandes darstellen, unter dem die Digitalisate durchscheinen und somit – wie schon am Ende von Abschnitt 4.1 dargestellt – der Verlauf gleichsam in die Bilder zurückprojiziert wird. Mit dieser Methode können auch mehrere Maße gleichzeitig dargestellt werden, wobei die Bänder auch überlappend angeordnet werden können und sich so der Effekt insgesamt verstärkt. Bei gleichzeitiger unterschiedlicher Färbung der Bänder bleibt die individuelle Interpretation der Einzelverläufe weiterhin möglich; siehe dazu Abb. 5 (Darstellung der Unterschiede zwischen den einzelnen <die>-Belegen anhand der Strichstärke des <d> im linearen Verlauf des Dokuments als

transparenzmoduliertes Band über den Digitalisaten) und Abb. 6 (Darstellung der Unterschiede zwischen den einzelnen <die>-Belegen anhand der Strichstärke des <d>, der inneren Volumina des <d> sowie der Strichstärke des <i> als transparenzmodulierte Bänder über den Digitalisaten).

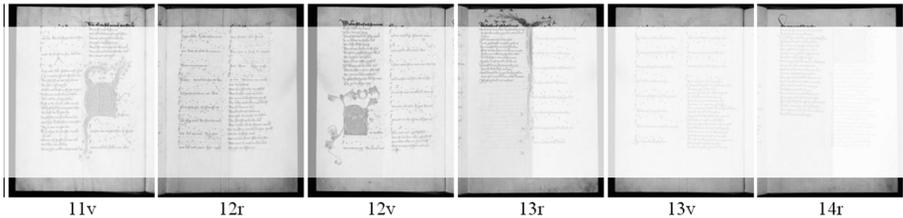


Abbildung 5. Darstellung eines Deskriptors im linearen Verlauf des Dokuments als transparenzmoduliertes Band über den Digitalisaten.



Abbildung 6. Darstellung dreier Deskriptoren im linearen Verlauf des Dokuments als transparenzmodulierte, teilweise überlappende Bänder über den Digitalisaten.

Für die komplexeren Maße oder Kombinationen von einfachen Maßen kann in einem weiteren Schritt unter der Annahme, dass die Maße aussagekräftig in Hinsicht auf die Schreiberhand sind, versucht werden, diese zu *gruppieren* (zu »clustern«). Hierzu werden die Distanzen zwischen den Maßen herangezogen, um Belege mit kleinen Distanzen zu Gruppen zusammenzufassen. Es können mehrere Gruppierungsvarianten mit einer unterschiedlichen Gruppenanzahl als Vorgabe berechnet werden. Stellt man nun die Zuordnung der einzelnen Vorkommen einer Buchstabenfolge über deren Position im Dokument als Verlauf dar, geben die wechselnden Zuordnungen über die Stabilität eines Clusters Auskunft. In Abb. 4 ist dies beispielhaft für das Wort »die« dargestellt.³⁴

³⁴ Die Tabelle listet die ersten und letzten 10 Belege der Form <die> jedes der von Wackernell festgelegten Schreiberbereiche A, B, C, D. In der direkten Gegenüberstellung der Bildausschnitte treten mikroskopische Duktus-Unterschiede zwischen A und B deutlich zutage, während B und C in markanten Details (z. B.

Abb. 7 zeigt eine Anordnung der einzelnen <die>-Vorkommen basierend auf dem Unterschiedsmaß, Abb. 8 zeigt den Verlauf der Zuordnung zu Clustern bei der Vorgabe »drei Cluster« und Abb. 9 bei der Vorgabe »vier Cluster« zu bilden. Die Interpretation der Unterschiede in der Zuordnung zu Clustern zwischen Abb. 8 und Abb. 9 stärkt nun

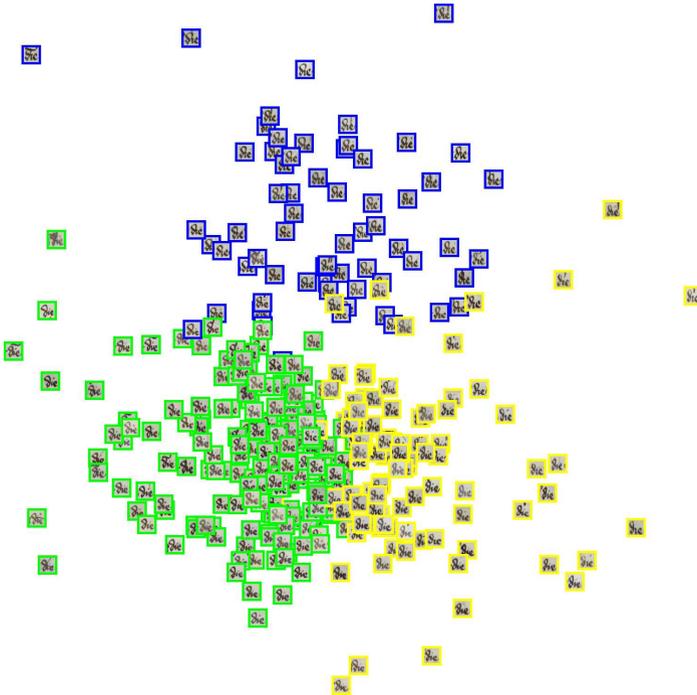


Abbildung 7. Kumulierende Zuordnung aller <die>-Belege im cpg 329 zu Clustern bei der Vorgabe »drei Cluster«.

Form und Position des i-Punktes) Übereinstimmung zeigen. Vgl. ausführlich Hofmeister und Hofmeister-Winter.

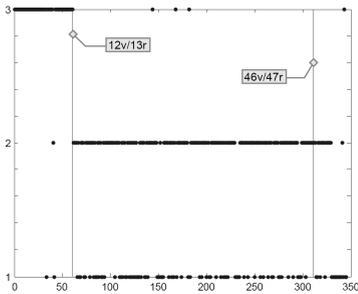


Abbildung 8. Lineare Zuordnung aller <die>-Belege im cpg 329 zu drei Clustern.

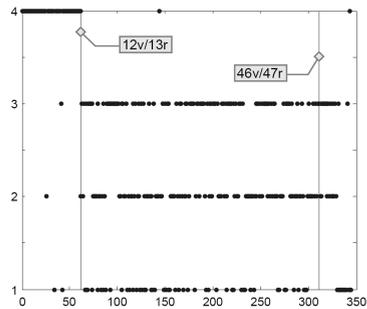


Abbildung 9. Lineare Zuordnung aller <die>-Belege im cpg 329 zu vier Clustern.

die schon von Wackernell gebildete Hypothese (vgl. Abschnitt 2.1), dass ein Schreiberhandwechsel in cpg 329 zwischen fol. 12v und 13r signifikant hervortritt, wogegen seine These eines weiteren Wechsels zwischen fol. 46v und 47r nicht erhärtet werden kann.

5 Ausblick: Weiterentwicklung von DAMaIS mit Integrierung in MOSES

Das methodisch anpassungsfähige Projektdesign von DAMaIS – rund um seine (bislang einzigartige) *mikrographische* Bild-Text-Erfassung – sollte es erlauben, zahlreiche weitere Schriftdokumente auf allfällige Schreiberhandwechsel hin zu untersuchen, und zwar prinzipiell auch für andere Sprachen. Damit könnte eine Art *Fahndungskartei* heranwachsen, welche wiederum die Grundlage für die Zeichnung einer historischen ›Wanderkarte‹ von Schriftdokumenten quer durch Europa bilden mag.

Das *Limit* für die Authentifizierung von Schreiberhänden wird von DAMaIS erst dort erreicht, wo sich in den Schriftzügen (z. B. wegen dominanter kalligraphischer Bestrebungen oder materieller Einschränkungen) keine ausreichenden Spuren von psychomotorischen Einprägungen nachweisen lassen. Aus dieser Einsicht kann per Umkehrschluss und unter Zugrundelegung einer diachronen Betrachtungsweise für eine Schriftkultur zugleich der Beginn ihrer *Individualisierung* eruiert werden, also der Zeitpunkt, ab wann innerhalb einer Sprachgemeinschaft deren Verschriftungstechnik so etwas wie einen schreiberspezifischen ›Fingerabdruck‹ ausgebildet hatte.

Bestärkt durch das große (sogar europaweite) Medien-Echo, das die soeben geschilderten Entwicklungsperspektiven von DAMaIS mittlerweile erfahren haben,³⁵ wurde

³⁵ Vgl. die auswahlhafte Dokumentation der Print- und Filmberichte über DAMaIS seit einer Presseaussendung im März 2008 auf der Projekt-Homepage.

vom Projektleiter das ›Meta-Projekt‹ MOSES (›Musterorientiertes System zur Erfassung von Schriftindividualität‹) konzipiert. Es soll DAMaLS integrieren, aber darüber hinaus auch für *gegenwartsbezogene* Problemlösungen im Zusammenhang mit der Verifizierung von individuellen Schriftzügen offen sein. Durch entsprechende Verfeinerungen und sprachspezifische Anpassungen der Authentifizierungstechnik mag es möglich werden, z. B. auch der *forensischen Schriftforschung* zu dienen; praxisnahe Vorstudien finden dazu bereits statt.³⁶

Eine erfreulich konkrete Aussicht auf die Umsetzung von MOSES hat sich Anfang 2009 durch seine Einbindung in eine Initiative der Ludwig Boltzmann-Gesellschaft ergeben. Diese strebt die Entwicklung einer (weltweit vernetzten) ›Archiv-, Text- und Editionstheorie‹ an. Sollte dieser Antrag – trotz einer zur Zeit (global wie national) sehr angespannten Finanz- und Förderungslage – erfolgreich sein, ließe sich 2010 in Graz ein eigenes MOSES-Forschungsinstitut gründen; durch offizielle ›Letters of Intent‹ an die LBG sowohl von der Karl-Franzens-Universität Graz als auch von der Forschungsgesellschaft Joanneum Research wurden dafür jedenfalls die ersten Schritte gesetzt.

Bibliographie

- Aitchison, Jean. *Wörter im Kopf. Eine Einführung in das mentale Lexikon*. Aus dem Englischen von Martina Wiese. Tübingen: Niemeyer, 1997.
Arbeitsgemeinschaft für germanistische Edition.
 <<http://www.ag-edition.org/html/archiv.html>>.
- Bibliotheca Palatina – digital*. UB Heidelberg.
 <<http://www.ub.uni-heidelberg.de/helios/digi/palatina-digital.html>>.
- Bromm, Gudrun. »Neue Vorschläge zur paläographischen Schriftbeschreibung.« *Methoden der Schriftbeschreibung*. Hrsg. Peter Rück. Stuttgart: Thorbecke, 1999. 21–43.
- Bulacu, Marius and Lambert Schomaker. »Text-independent writer identification and verification using textural and allographic features.« *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.4 (2007): 701–17.
- Cappelli, Adriano. *Lexicon abbreviatarum*. 2., verb. Aufl. (Nachdr. d. Ausg. Leipzig 1928.) München: Hieronimus, 1988.
 [cpg 329, *Online-Präsentation*]. <<http://diglit.ub.uni-heidelberg.de/diglit/cpg329>>.
- DAMaLS, *Projekthomepage*. <<http://www.uni-graz.at/wernfried.hofmeister/damals>>.
- Das Brixner Dommessnerbuch. Mit elektronischer Rohtextversion und digitalem Voll-*

³⁶ Das gesamte Autorenteam dieses Beitrags ist eingeladen, im Juni 2009 auf dem Kriminologenkongress der Gesellschaft für Forensische Schriftuntersuchung (siehe [Tagungsprogramm](#)) die bislang entwickelte Leistungsfähigkeit von DAMaLS anhand der Befundung strittiger handschriftlicher Dokumente unter Beweis zu stellen, und hofft seinerseits, dabei neue Entwicklungsimpulse zu erhalten.

- faksimile auf CD-ROM*. Hrsg. Andrea Hofmeister-Winter. Innsbruck: Institut für Germanistik, 2001.
- Gesellschaft für Forensische Schriftuntersuchung*. <<http://www.gfs2000.de>>; [*Informationen zum Tagungsprogramm*]. <http://www.gfs2000.de/main_d1.htm>.
- Grabowski, Joachim. »Bedingungen und Prozesse der schriftlichen Sprachproduktion.« *Psycholinguistik/Psycholinguistics. Ein internationales Handbuch*. Hrsg. Gert Rickheit, Theo Herrmann und Werner Deutsch. Berlin: de Gruyter 2003. 355–368.
- Grun, Paul A. *Schlüssel zu alten und neuen Abkürzungen*. (Reprint d. Ausg. 1966.) Limburg/Lahn: Starke, 2002.
- Hofmeister, Wernfried. »Die Edition als ›offenes Buch‹: Chancen und Risiken einer Transponierungs-Synopse, exemplarisch dargestellt an der Dichtung *Von des todes gehugede* des sog. Heinrich von Melk.« *Produktion und Kontext. Beiträge der Internationalen Fachtagung der Arbeitsgemeinschaft für germanistische Edition im Constantijn Huygens Instituut, Den Haag, 4. bis 7. März 1998*. Hrsg. H. T. M. van Vliet. Tübingen: Niemeyer 1999. 23–39.
- Hofmeister, Wernfried. »Der Mut zur Lücke: Auf den Spuren von Textnachträgen in der Manessischen Liederhandschrift. Ein Beitrag zu einer ›Überlieferungs-Philologie‹ des Mittelalters.« *Entstehung und Typen mittelalterlicher Lyrikhandschriften. Akten des Grazer Symposiums, 13.–17. Oktober 1999*. Hrsg. Anton Schwob und András Viskelety unter Mitarbeit von Andrea Hofmeister-Winter. Bern: Lang, 2001. 79–106.
- Hofmeister, Wernfried. »Ein Autor ›outet‹ sich: Hugo von Montfort (1357–1423) im rezeptionellen Spannungsfeld.« *Autor – Autorisation – Authentizität. Beiträge der internationalen Fachtagung der Arbeitsgemeinschaft für Germanistische Edition in Verbindung mit der Arbeitsgemeinschaft Philosophischer Editionen und der Fachgruppe Freie Forschungsinstitute in der Gesellschaft für Musikforschung, Aachen, 20. bis 23. Februar 2002*. Hrsg. Thomas Bein, Rüdiger Nutt-Kofoth und Bodo Plachta. Tübingen: Niemeyer, 2004. 165–72.
- Hofmeister, Wernfried. »Perspektiven und Auswirkungen des Edierens am Beispiel der neuen Hugo von Montfort-Ausgabe.« *Aller weishait anevang / ist ze brüfen an dem ausgang. Akten des Symposiums zum 650. Geburtstag Hugos von Montfort, Dornbirn, 19.–22. September 2007*. Hrsg. Klaus Amann und Elisabeth De Felip-Jaud (im Druck).
- Hofmeister, Wernfried und Andrea Hofmeister-Winter. »Schriftzüge unter der High-Tech-Lupe. Theoretische Grundlagen und erste praktische Ergebnisse des Grazer Forschungsprojekts DAMaLS (›Datenbank zur Authentifizierung mittelalterlicher Schreiberhände‹).« *editio. Internationales Jahrbuch für Editionswissenschaft* 22 (2008): 90–117.
- Hofmeister-Winter, Andrea. *Das Konzept einer »Dynamischen Edition« dargestellt an*

- der Erstausgabe des »Brixner Domesnerbuches« von Veit Feichter (Mitte 16. Jh.). Theorie und praktische Umsetzung.* Göppingen: Kümmerle, 2003.
- Hofmeister-Winter, Andrea. »Textkritik als Erkenntnisprozeß: sehen – verstehen – deuten.« *editio. Internationales Jahrbuch für Editionswissenschaft* 19 (2005): 1–9.
- Hofmeister-Winter, Andrea. »Die Grammatik der Schreiberhände. Versuch einer Klärung der Schreiberfrage anhand der mehrstufig-dynamischen Neuausgabe der Werke Hugos von Montfort.« *Edition und Sprachgeschichte. Baseler Fachtagung 2.–4. März 2005.* Hrsg. Michael Stolz in Verbindung mit Robert Schöller und Gabriel Viehhauser. Tübingen: Niemeyer, 2007. 89–116.
- Hofmeister-Winter, Andrea. [*Forschungshomepage*].
<<http://www.uni-graz.at/~hofmeisa>>.
- Hugo von Montfort. Mit Abhandlungen zur Geschichte der deutschen Literatur, Sprache und Metrik im XIV. und XV. Jahrhundert.* Hrsg. J[osef] E[duard] Wackernell. Innsbruck: Wagner, 1881.
- Hugo von Montfort. Das poetische Werk.* Hrsg. Wernfried Hofmeister. Mit einem Melodie-Anhang von Agnes Grond. Berlin: de Gruyter, 2005.
- Hugo von Montfort, [Editions-Homepage].*
<<http://www.gewi.uni-graz.at/montfort-edition>>.
- Institut für Informatik/Abteilung für Sprachverarbeitung an der Universität Leipzig. [*Projekt »Wortschatz«*]. <<http://wortschatz.uni-leipzig.de>>.
- Krämer, Sigrid. *Scriptores codicum medii aevi. Datenbank von Schreibern mittelalterlicher Handschriften.* CD-Rom und Beiheft. Augsburg: Rauner, 2003.
- Lowe, David G. »Distinctive image features from scale-invariant keypoints.« *International Journal of Computer Vision* 60.2 (2004): 91–110.
- Parzival-Projekt.* <<http://www.parzival.unibe.ch>>.
- Produktion und Kontext. Beiträge der Internationalen Fachtagung der Arbeitsgemeinschaft für Germanistische Edition im Constantijn Huygens Instituut, Den Haag, 4. bis 7. März 1998.* Hrsg. H. T. M. van Vliet. Tübingen: Niemeyer, 1999.
- Schieb, Gabriele. »Editionsprobleme altdeutscher Texte.« *PBB [Halle/S.]* 89 (1967): 404–30.
- Schlögl, Waldemar. Rez. von *L'expertise des écritures médiévales. Recherche d'une méthode avec application à un manuscrit du XIe siècle: Le lectionnaire de Lobbes, Codex Bruxellensis 18018, Gand 1973 [...]*, von Léon Gilissen. *Deutsches Archiv für Erforschung des Mittelalters* 33 (1977): 264–65.
- Schneider, Karin. *Paläographie und Handschriftenkunde für Germanisten. Eine Einführung.* Tübingen: Niemeyer, 1999.
- Schneider, Karin. »Akzentuierung in mittelalterlichen deutschsprachigen Handschriften.« *Edition und Sprachgeschichte. Baseler Fachtagung 2.–4. März 2005.* Hrsg. Michael Stolz in Verbindung mit Robert Schöller und Gabriel Viehhauser. Tübingen: Niemeyer, 2007. 17–24.

- Schubert, Martin. »Sprechende Leere. Lücke, Loch und Freiraum in der Großen Heidelberger Liederhandschrift.« *editio. Internationales Jahrbuch für Editionswissenschaft* 22 (2008): 118–38.
- Spechtler, Franz V. *Die Heidelberger Handschrift cpg 329 und die gesamte Streuüberlieferung*. In Abbildung hrsg. von Eugen Thurnher, Franz V. Spechtler und Ulrich Müller. Göttingen: Kümmerle, 1978. 12–20.
- Srihari, Sargur N., Sung-Hyuk Cha, Hina Arora, and Sangjik Lee. *Handwriting identification: Research to study validity of individuality of handwriting and develop computer-assisted procedures for comparing handwriting*. Buffalo (NY): University at Buffalo, 2001.
- Wege zum Text. Überlegungen zur Verfügbarkeit mediävistischer Editionen im 21. Jahrhundert*, Grazer Kolloquium, 17.-19. September 2008. Hrsg. Wernfried Hofmeister und Andrea Hofmeister-Winter. Tübingen: Niemeyer, 2009 (im Druck).
- »Wege zum Text«, [Tagungshomepage].
<<http://www.uni-graz.at/wernfried.hofmeister/wegezumtext>>.
- Welker, Lorenz. »Die Melodien des Burkhard Mangold.« *Hugo von Montfort. Einführung zum Faksimile des Codex Palatinus Germanicus der Universitätsbibliothek Heidelberg*. Mit Beiträgen von Franz Viktor Spechtler u. a. Wiesbaden: Reichert, 1988. 47–60.
- Werner, Wilfried. »Die Handschrift und ihre Geschichte.« *Hugo von Montfort. Einführung zum Faksimile des Codex Palatinus Germanicus der Universitätsbibliothek Heidelberg*. Mit Beiträgen von Franz Viktor Spechtler u. a. Wiesbaden: Reichert, 1988. 7–11.
- Wordle (Programm). <<http://www.wordle.net>>.
- Wortfrequenzlisten der Universität Leipzig.
<<http://wortschatz.uni-leipzig.de/Papers/top10000de.txt>>.
- Zentrum für Informationsmodellierung in den Geisteswissenschaften (ZIMig) an der Karl-Franzens-Universität Graz. <<http://www.uni-graz.at/zim>>.