Julia Hübner, Katja Politt u. Marc Schutzeichel

Überlegungen zur Reliabilität der Buchstabenschreibung in frühneuhochdeutschen Handschriften

<https://doi.org/10.1515/bgsl-2018-1001>

Abstract: Reliability issues emerge when investigating initials in transcriptions of ENHG manuscripts. Our goal is to demonstrate how a methodological framework of reliability classes affects and facilitates the evaluation of annotated corpus data based on existing text editions. We hypothesize that deviant measurements in the capitalization of sentence-internal word tokens occur when reliability classes and the representativeness of allographs are accounted for. Our data, derived from SIGS project annotations, shows that the capitalization mapped per part of speech can be represented by the size of a set of allographs, thus pointing out the importance of the letter as discrete factor for each text of the corpus.

I. Problemstellung

Die vorliegende Untersuchung geht der Frage nach, inwieweit der Einfluss der in der Regel als gegeben aufgefassten editorischen Transkription auf weiterreichende linguistische Fragestellungen modelliert und überprüft werden kann. Ausgangspunkt der Überlegungen ist die große Heterogenität historischer Handschriften, zu der bekanntlich jede an der Manuskriptverfertigung beteiligte Hand in Form von spezifischen aber auch für sie typischen Idiosynkrasien beiträgt (vgl. Maas [u. a.] 1989, S. 147). Ein anderer bedeutsamer Aspekt bei der Betrachtung von Diversität in diesem Bereich ist der Grad der Bearbeitung: Es kann sich, etwa bei juristischen Texten, um eine Mitschrift, Reinschrift oder eine Abschrift

Julia Hübner: FU Berlin, SFB 980 ›Episteme in Bewegung‹, Schwendenerstraße 8, D-14195 Berlin, E-Mail: julia.huebner@fu-berlin.de

Katja Politt: Leibniz Universität Hannover, Deutsches Seminar, Königsworther Platz 1, D-30167 Hannover, E-Mail: katja.politt@germanistik.uni-hannover.de

Dr. Marc Schutzeichel: Universitäts- und Landesbibliothek Münster, Wissenschaftliche Informationssysteme, Krummer Timpen 3–5, D-48143 Münster, E-Mail: m.schutzeichel@uni-muenster.de

handeln (vgl. Macha [u. a.] 2005, S. XVII). Die für eine erfolgreiche Transkription unerlässliche Bestimmung der Buchstaben orientiert sich neben der genauen Beobachtung der graphischen Gegebenheiten (Form, Größe, Ähnlichkeitsbeziehungen) auch – und bei schwierigen Fällen ganz wesentlich – an der nächstgrößeren Basiseinheit, nämlich dem Wortkörper, den die Graphe konstituieren. Eine bereits erkannte Wortform, deren Bedeutung im umgebenden Syntagma passend eingebracht werden kann, liefert demnach eine Bestätigung für die Identifikation einer unsicheren Lesung einzelner Graphe. Zahlreiche Editionen, denen pro Texteinheit nur ein einziges Manuskript bzw. nur eine Version zugrunde liegen, beweisen die Zuverlässigkeit dieses Prinzips. Auf die Bedeutung des Wortkörpers kann sich jedoch in der Regel der Transkribierende bei der Entscheidung über vorliegende Groß- oder Kleinschreibung nicht stützen. Wortartenzugehörigkeiten, etwa bei Konversionen, oder der Ausschluss gewisser Homonyme mittels der Großschreibung bilden allenfalls ein schwaches Indiz, wenn es um historische Sprachstufen geht. Im Folgenden möchten wir das hierin liegende Problem in den Vordergrund stellen und anhand der Frage nach der Großschreibung Verfahrensweisen aufzeigen, die einen effizienten, kritischen Umgang mit bereits geleisteten Transkriptionen in historischen Korpora, denen Handschriften zugrunde liegen, ermöglichen.

II. Zielsetzung

Das Ziel der vorliegenden Untersuchung ist die Entwicklung eines verallgemeinerbaren Verfahrens zur Messung der graphematischen Reliabilität in historischen Handschriften. Der Terminus Reliabilität wird in empirischen Untersuchungen gewöhnlich im Kontext der Reproduzierbarkeit von Ergebnissen verwendet. Im Folgenden ist das Begriffspaar reliabel gegenüber nicht-reliabel als ein Maß für die Verlässlichkeit der Messgröße (Allo-)Graph als unmittelbare Konstituente der satzinternen Großschreibung aufzufassen. Das zu beschreibende Phänomen soll mittels korpuslinguistischer Methoden erschlossen werden. Die Klassifizierung der Schriftzeichen in den Handschriften erfolgt manuell und setzt eine gründliche Sichtung aller Kontexte voraus. Bisher wurde der Untersuchung von Handschriften in Bezug auf vergleichbare Fragestellungen wenig Beachtung geschenkt, da die Arbeit mit uneinheitlichen Manuskripten eine Vielzahl an Problemen mit sich bringt. Allerdings erscheint sie für Fragestellungen, die die Graphematik betreffen, besonders lohnenswert. In diesem Fall darf in der durch verschiedene Hände erzeugten Variabilität ein wichtiger Anhaltspunkt für die Interpretation und Interpretierbarkeit von Schriftzeichen vermutet werden.

Ein weiterer nützlicher Anhaltspunkt ist die Buchstabenverteilung im Korpus. Sie ist ein separater Faktor, der je nach Untersuchungsgegenstand Zählungen und damit Ergebnisse beeinflussen kann. Indem das mögliche und das tatsächlich vorgefundene Spektrum der Buchstaben festgestellt werden, ergibt sich ein Profil des Korpus, ein graphematischer Fingerabdruck. Dieser hängt in erster Linie von der Verteilung und den Häufigkeiten der Lexeme im Korpus ab. Häufige Wörter erhöhen zwangsläufig auch die Vorkommenshäufigkeit der Buchstaben dieser Wörter. Somit lässt sich die Variabilität quantifizieren. Während durch die Formseite des schriftsprachlichen Zeichens seine Qualität vor der Schablone eines bekannten Alphabets in der Regel eindeutig zu bestimmen ist (<a>, , <c>, ...), lässt sich dem Schriftzeichen auf der graphetischen Ebene eine besondere Form der Quantität beimessen, die den Terminus im besten Wortsinn gebraucht (lat. *quantitās* ›Größe‹). In den nachfolgenden Überlegungen wird die erläuterte Problemstellung anhand der Thematik der Groß- und Kleinschreibung illustriert. Die Quantität eines Schriftzeichens erlaubt eine Aussage über die tatsächliche Zeichengröße zwischen den Polen Groß- und Kleinschreibung auf dieser Skala. Mit der Einteilung in Majuskel einerseits und Minuskel andererseits sind zwei Quantitätsbereiche abgesteckt, die jedoch nicht immer eine klare Abgrenzung aufweisen. Zuweilen findet sich keine solche Opposition für ein Initial. Die qualitativen Unterschiede der Buchstabenformen untereinander sind in den meisten Fällen erheblich größer als die quantitativen Unterschiede in den Instanzen eines Graphems. Die Verteilung der Majuskeln lässt sich also keineswegs allein über die Häufigkeit der Wörter, deren Initial gezählt wird, exakt vorhersagen. Dennoch sind die häufigsten Initiale in erster Linie abhängig von besonders hochfrequenten Wortformen, etwa das <d> beim satzintern je nach Textsorte häufig kleingeschriebenen Definitartikel sowie im Weiteren das <s> im Pronominalbereich, das <v> bei dem häufigen Wort *vnd* (>und<) oder das <g> im Partizip II.

Grundlegend für die Messbarkeit der Initiale und damit der Großschreibung ist das Konzept des graphematischen Worts. Ein Graphem wird nur dann in die Berechnung einbezogen, wenn es als Initial eines graphematischen Worts als eine »ununterbrochene Graphemkette« (Fuhrhop 2008, S. 194) realisiert wird. Ein graphematisches Wort wird dabei im Idealfall durch seine Abgrenzung durch zwei Spatien bestimmt (vgl. Jacobs 2005, S. 22 [mit Anm. 19]). Bei der Feststellung des Initials ist in der Regel zunächst zwischen Majuskel- und Minuskel-schreibung zu differenzieren. Jede der beiden Varianten wird als eigenes Zeichen gewertet. Es wird jedoch nicht zwischen einfachen und komplexen Graphemen (Digraphe wie <ch> oder Trigraphe wie <sch>) unterschieden. Als Initial gilt uns bei Polygraphen nur das erste Graph. Die Segmentierbarkeit der Polygraphie ist den Schreibern in unserem Korpus offenbar bewusst, denn bei der wortinitialen Großschreibung wird lediglich das erste Graph eines Polygraphs großgeschrie-

ben.¹ Der hier verwendete Graphembegriff entspricht dem ›Grundgraphem‹ von Gallmann (1985, S. 10).

III. Großschreibung

Aus umgekehrter Perspektive lässt sich argumentieren, dass bei allen graphematischen Untersuchungen der Umgang mit handschriftlichen Dokumenten eine besondere Herausforderung darstellt. Bei historischen Untersuchungen liegt hier neben der nicht-standardisierten Sprache aufgrund der handschriftlichen Variation zusätzlich nicht-standardisierte Schrift vor. Auch aus diesem Grund sind Handschriften bisher in nur geringem Ausmaß überhaupt auf graphetischer Ebene betrachtet worden (vgl. Fuhrhop/Peters 2013, S. 184).

Darüber hinaus betonen Maas [u. a.] (1989, S. 148), dass immer auch das Weltwissen und das Erwartete die Interpretation der Grapheme beeinflussen. Wenn in einem bestimmten Kontext von dem Transkribierenden ein Allograph von erwartet wird, so wird die Qualität des Graphs zunächst als b vorinterpretiert. Schließlich erfolgt die Bestimmung der Größe. Obwohl schriftgeschichtlich für die lateinische Schrift und ihre Vorgänger der Kleinbuchstabe als jüngere Innovation gelten kann, muss spätestens ab frühneuhochdeutscher Zeit die Minuskel wegen ihres deutlich häufigeren Vorkommens in den Texten als Normalfall angesehen werden, der durch den Spezialfall der Majuskel modifiziert wird. Die Bestimmung der Groß- und Kleinschreibung unterliegt also im Gegensatz zur Bestimmung der Buchstabenqualität einem besonderen Bias. Dieser Effekt verstärkt sich, wenn bei der Bestimmung des Schriftzeichens und der Schriftzeichenfolge die Wiedergabe der Lautung des ganzen Wortkörpers und eben nicht die Großschreibung des Initials im Vordergrund steht. Mit anderen Worten kann eine editorische Situation entstehen, bei der die exakte Bestimmung der Groß- bzw. Kleinschreibung der Bestimmung der Buchstabenqualität untergeordnet wird. Die Groß- oder Kleinschreibung wird somit beim Lesen und Transkribieren zu einem sekundären Aspekt des Schriftzeichens.

Generell kann die Großschreibung eines Buchstabens zunächst als binäre Größe erfasst werden. Der Parameter kann für die Frage nach der Eigenschaft,

¹ Ausnahmen von dieser Regel finden sich beispielsweise im Lexem *HErr* und gehen auf eine bewusste, systematische Entscheidung zurück. Schon Martin Luther gibt in der Vorrede der Erstausgabe seiner Übersetzung des Alten Testaments (1523) bekannt, dass er die Doppelmajuskel bei *HErr* nutzt, um das hebräische Adonai (*ādōnāy*) vom Tetragrammaton (wiedergegeben als *HERR*) und dem gewöhnlichen Lemma *Herr* zu unterscheiden (vgl. Weber [1958], S. 33).

groß geschrieben zu sein, die Werte ›wahr‹ oder ›falsch‹ annehmen. Eine solche Konzeptualisierung wird den tatsächlichen graphischen Gegebenheiten in den Handschriften allerdings nur scheinbar gerecht. Es bietet sich stattdessen eine Differenzierung nach Formdistinktheit und Skalendistinktheit an. Binarität findet sich nur in formdistinkten Graphpaaren. Dagegen ist ein skalendistinktes Minuskel-Majuskel-Paar formmäßig identisch, jedoch anhand eines graduellen Größenfaktors, der Skalierung, im Gebrauch zu unterscheiden. Es gilt, die Elemente des je vorgefundenen Grapheminventars den beiden Klassen zuzuweisen. Auch bei formdistinkten Varianten kann zusätzlich die Skalierung genutzt werden (und umgekehrt), um auffallend große Graphen darzustellen, die etwa am Zeilen- oder Seitenanfang platziert werden (siehe im Folgenden). Für eine Einteilung in Graphklassen ist jedoch das Vorhandensein einer Formdistinktion eines Minuskel-Majuskel-Paars in einem Text entscheidend. Die Formdistinktheit allein garantiert schon eine eindeutige Identifikation der Allographe, während Skalendistinktheit immer mit der relativen Größe der Schriftzeichen in der Umgebung, insbesondere mit der aktuellen Zeile und Vorzeile, in Beziehung steht und somit erheblich mehr Spielraum für Interpretationen lässt (vgl. Maas [u. a.] 1989, S. 149).

Es bleibt zu bedenken, dass auch innerhalb der Majuskeln keineswegs Uniformität besteht. Grund für die unterschiedlichen Realisierungen sind die oben skizzierten Parameter, die die Größe und Formgestalt eines Graphs bestimmen, sowie die unterschiedlichen Hände, die Anteil an der Verfertigung eines Textes haben. Innerhalb des Textverlaufs ist von einer formalen Variation eines jeden Schriftzeichens auszugehen. Die Zugehörigkeit zu einem spezifischen Graphem lässt sich allein anhand einer Ähnlichkeitsbeziehung festmachen. Alle realisierten Varianten eines Graphems bilden ein Similaritätskontinuum, das vergleichbar ist mit den realisierten Varianten eines Phonems.

Abb. 1 zeigt kontextbezogene unterschiedliche Realisierungen der Majuskel <D>. Die Beispiele stammen aus der im Folgenden einzuführenden Datenbasis. Die Kontexte sind in diesem Fall verschiedene Positionen innerhalb des Textverlaufs bzw. in unterschiedlichen Segmenten des Seitenaufbaus. Die auffällige Ausgestaltung der Buchstaben am Zeilen- und vor allem am Abschnittsanfang veranschaulicht, dass teils mehr als zwei distinkte Formen zur Auswahl stehen, wenngleich die Realisierungen der Majuskeln große Ähnlichkeiten zueinander aufweisen. Es ist davon auszugehen, dass die Größe des jeweiligen Buchstabens immer in Verbindung mit dem Schriftsystem (Druck oder Schreibschrift, Struktur des Textes) steht (vgl. Nottbusch 2008, S. 35). Die Kombination aus Schreibschrift und freier Textstruktur lässt die größte Variabilität zu. Während Studien zu den graphischen Realisierungen eines einzelnen spezifischen Schreibers (vgl. zu den Briefen Luthers Moulin 1990) über mehrere Texte hinweg durch die schiere Zahl an beobachteten Wiederholungen leichter zu Entscheidungen hinsichtlich der

Großschreibung gelangen, wirft eine Sammlung kurzer Texte individueller, nicht identifizierter Schreiberhände Fragen auf. Abweichende Realisierungen in der Schreibung sind mitunter nicht von der regelhaften Realisierung zu unterscheiden, wenn die Gesamtanzahl gering ist.

Realisierung	Position
Textanfang
Seitenanfang
Absatzanfang
Zeilenanfang
satzintern, wortinitial
–	Binnenmajuskel ²
satzintern, wortintern

Abb. 1: Relative Größe: Die Majuskelvarianten von <D> in unterschiedlichen Positionen, nebst der Minuskelrealisierung (Jever 1592). Konstanter Maßstab, Vergrößerungsfaktor des Scans³ durchgängig 100 %

² Nur bei der sogenannten Binnenmajuskel finden sich groß geschriebene Graphemrealisierungen auch im Wortinnern. Diese ist aber im Editions-korpus äußerst selten. <D> ist an keiner Stelle im Gesamtkorpus der von Macha [u. a.] (2005) edierten Protokolle als Binnenmajuskel vorzufinden.

³ Wir danken den Mitarbeiterinnen und Mitarbeitern am ehemaligen Lehrstuhl von Prof. Dr. Jürgen Macha für den Zugang zu den im DFG-Projekt ›Kanzleisprache des 17. Jahrhunderts. Untersuchungen zu Sprache und Kommunikation in Hexenverhörprotokollen‹ verarbeiteten Faksimiles.

Vielfach sind es Varianten einzelner Graphe, die Probleme für die Transkription aufwerfen. Im Protokoll⁴ ›Georgenthal 1597‹ erscheint das im Korpus zumeist zuverlässig zu deutende Initial <O> in zwei Versionen. In beiden Fällen ist der Kontext vergleichbar. Es handelt sich jeweils um einen Seitenanfang sowie um das erste Wort. Darüber hinaus handelt es sich um Reklamanten, beide Schriftzeichen werden schon am unteren Rand der vorausgehenden Seite homoform dargestellt.

S. 4: Transkribiert als <Ob>

S. 5: Transkribiert als <Ob>

Z. 128

Z. 168

Abb. 2: ›Georgenthal 1597‹, Varianten des Wortes <Ob>

In beiden Fällen erfolgte die Transkription mit einer Majuskel. Das Beispiel von S. 4 ist formdistinkt zur Minuskel, während dasjenige von S. 5 skalendistinkt ist. Es stellt lediglich eine größer skalierte Variante der Minuskelform da. Der Kontext der Wiederholung durch Reklamanten belegt, dass es sich bei diesem Beispiel nicht um eine zufällige, einzigartige Fluktuation handelt. Vermutet werden kann am Beispiel von S. 4 eine zusätzliche Markierung der Recto-Seite durch eine Formdistinktion nach erfolgter Skalierung der Größe des Schriftzeichens. In diesem Zusammenhang eröffnen sich Kontexte, in denen trotz einer formalen Opposition zwischen Majuskel und Minuskel die Transkription in Zweifel gezogen werden kann.

Um eine wortinitiale Distinktion zwischen Majuskel- und Minuskelschreibung zu ermöglichen, ist ein erneuter Blick in die Handschriften unerlässlich. Erst im direkten Vergleich der Graphe lässt sich zuordnen, ob es sich um ein formdistinktes Paar handelt. Beispielhaft wird dies in der folgenden Übersicht dargestellt, die das handschriftliche Graphinventar für die Protokolle ›Georgenthal 1597‹ und ›Jever 1592‹ zeigt. Ausgewählt wurden repräsentative Beispiele, bevorzugt aus der

⁴ Es sei angemerkt, dass die Bezeichnungen ›Hexenverhörprotokoll‹ oder ›Hexereiprotokoll‹ irreführend sein können. Bei den verhörten Frauen und Männern handelt es sich bekanntlich um durch Verleumdung und Folter unschuldig Verurteilte. Im Folgenden möchten wir daher die Ausdrücke ›Protokoll‹ bzw. ›Verhörprotokoll‹ bevorzugen. Für die Beschreibung des Korpus siehe Abschnitt ›Datenbasis‹.

Zeilenmitte. Ein unmittelbar ins Auge fallendes Ergebnis ist die identische Realisierung unterschiedlich transkribierter Grapheme innerhalb eines Protokolls, wie <I> und <J> in ›Georgenthal 1597‹. Jede Ausprägung eines Graphems weist unterschiedlich starke Variabilität auf. Dies führt dazu, dass einige Realisierungen konstant gleich sind, während andere zu einem gewissen Grad variieren.

Tab. 1: Minuskeln und Majuskeln am Beispiel der Protokolle ›Georgenthal 1597‹ und ›Jever 1592‹. Die Initiale <q>, <x>, <y>, <ä>, <ö> und <ü> sind nicht mit aufgeführt, da sie nicht vorkommen

	a	b	c	d	e	f
Georgenthal Min.			-
Georgenthal Maj.
Jever Min.			-
Jever Maj.			-		-
	g	h	i	j	k	l
Georgenthal Min.				-
Georgenthal Maj.
Jever Min.				-

Tabelle fortgesetzt

Jever Maj.		-
	m	n	o	p	r	s
Georgenthal Min.
Georgenthal Maj.				-
Jever Min.
Jever Maj.
	t	u	v	w	z	
Georgenthal Min.
Georgenthal Maj.		-
Jever Min.		-
Jever Maj.		-			-	

Die Tabelle verdeutlicht, dass nicht in allen Fällen eine starre Zuordnung von handschriftlich realisierten Graphen zu Graphemen erfolgen kann und dass nicht alle Graphe homoform realisiert werden. Für eine Auswahl möglichst repräsentativer Graphrealisierungen gilt es zu beachten, dass eine Verallgemeinerbarkeit für die ausgewählten Formen für alle Absätze eines Protokolls oder für mehrere Protokolle nicht gegeben ist. Aus diesem Grund kann bei der Analyse von Handschriften nur ein Abgleich mit der prototypischen Form des zugrundeliegenden Graphems erfolgen. Auf dieses Problem weisen auch Maas [u. a.] (1989, S. 146) hin. Demnach kann eine Edition, die auf der Grundlage von Handschriften entstanden ist, nicht als transparent wiedergegebene Schriftform gelten, da die Editoren gezwungenermaßen in die Gestalt des Textes, etwa die Platzierung von Absätzen oder Zeilenumbrüchen, eingreifen. Bei der Transkription steht die Entscheidung im Vordergrund, welche Graphe jeweils in der Handschrift realisiert wurden. Macha [u. a.] (2005, S. XX) definieren zwar in Bezug auf die Transkription drei Grade der Verlässlichkeit. Diese Abstufung hängt jedoch lediglich davon ab, ob das handschriftliche Protokoll vorliegt oder ob stattdessen eine andere Edition Grundlage der Transkription war. Mit anderen Worten: Die in diesem Fall erreichte Verlässlichkeit ist nicht mit der Reliabilität der Schriftzeichen im eigentlichen Sinn gleichzusetzen. In ambigen Fällen wurde von den Editoren zugunsten der Leserfreundlichkeit entschieden. Für die Groß- und Kleinschreibung bedeutet dies eine Entscheidung für die Kleinschreibung. Wenn für Majuskel und Minuskel formidentische Zeichen verwendet wurden, beruht die Transkription auf der relativen Größe des Zeichens (vgl. Macha [u. a.] 2005, S. XXIII).

Der Fokus der Erforschung der satzinternen Großschreibung liegt bisher auf den Drucken (vgl. Weber 1958; Bergmann/Nerius 1997). Gerade der Blick auf handschriftliche Urkunden, Prozessakten oder Briefe erscheint aber sinnvoll, da diese stark produzentenorientiert sind und somit wertvolle Einblicke in die sprachliche Kompetenz des Schreibenden sowie die dem Schreibvorgang zugrundeliegenden kognitiven Prozesse gewähren. Dagegen können Drucke durchaus auch von illiteraten Personen erstellt werden.

Entscheidend für die erfolgreiche Bestimmung der Einflussfaktoren auf die satzinterne Großschreibung sind somit vorrangig Frequenzzählungen, d. h. die Gegenüberstellung der Auftretenshäufigkeiten wortinitialer Großschreibung versus wortinitialer Kleinschreibung. Die Gesamtanzahl und -verteilung einzelner Buchstaben⁵ unterliegt jedoch ebenfalls einer textspezifischen Variabilität, die es zu berücksichtigen gilt.

⁵ Die hier zugrundeliegenden Konzepte und Termini variieren je nach Autor und Forschungsrichtung (vgl. Maas 1989, S. 149). Diese Diskussion soll an dieser Stelle nicht fortgeführt werden, stattdessen schließen wir uns den Definitionen von Rezec (2009, S. 10) an.

IV. Datenbasis

Das für diese Untersuchung herangezogene Korpus basiert auf 56 Verhörprotokollen, die in der Edition von Macha [u. a.] (2005) veröffentlicht wurden. Es handelt sich hierbei um Transkripte von Handschriften, die einen Zeitraum von 1570 bis 1653 abdecken und über ganz Deutschland verteilt sind, wodurch sich sechs Regionen ergeben: »Westliches und Östliches Norddeutschland«, »Westliches und Östliches Mitteldeutschland« und »Westliches und Östliches Süddeutschland«⁶. Der vorliegende Sprachstand ist eine späte Phase des Frühneuhochdeutschen, die als Übergangszeit zum Neuhochdeutschen ab der Mitte des 16. Jahrhunderts angesehen wird (vgl. Reichmann/Wegera 1993, S. 5). Zwar weisen alle Texte einen frühneuhochdeutschen Sprachstand auf, dieser ist jedoch durch die regionale Verteilung stark von den jeweiligen Varietäten beeinflusst. Der Schwerpunkt der Editoren bei der Zusammenstellung der Texte lag vor allem auf einer möglichst breiten Sammlung und einer Aufbereitung der handschriftlichen Daten für eine maschinelle Weiterverarbeitung (vgl. Macha [u. a.] 2005, S. XIV), mit besonderem Fokus auf eine »ausgewogene Ortspunktverteilung« (ebd., S. XVII). Für die Bestimmung der Buchstabenreliabilität ist dies nicht ausreichend. An dieser Stelle setzt die vorliegende Studie an.

Die Zufälligkeit der Überlieferung sowie die Rücklaufquote auf Archivanfragen nach Verhörprotokollen durch das Editionsprojekt von Macha [u. a.] (2005) führen dazu, dass nicht alle Räume in gleichem Maß in der Edition repräsentiert sind. So sind beispielsweise für das Gebiet des Östlichen Mitteldeutschlands vier Verhörprotokolle aus einer Zeitspanne von 1597 bis 1640 vorhanden, während für das mit Blick auf Hexenprozesse historisch aktivere Gebiet Westliches Norddeutschland 16 Protokolle mit einem zeitlichen Rahmen von 1570 bis 1649 enthalten sind. Um einen gleichmäßigen und in Bezug auf das Editions-korpus repräsentativen Ausschnitt untersuchen zu können, wurden insgesamt 18 Protokolle ausgewählt, die auf der Basis der räumlichen und zeitlichen Distribution sowie in Bezug auf den Textumfang ein ausgewogenes Kernkorpus⁷ darstellen. Die regionale Einteilung von Macha [u. a.] (2005) wurde beibehalten und ist somit durch jeweils drei Protokolle repräsentiert um einen möglichst vergleichbar gestaffelten Zeitraum abzudecken. Auf diese Weise wird eine Zeitspanne von 1588 bis 1630 abgedeckt. Zudem wurde durch die Auswahl sichergestellt, dass die Textlänge nur in geringem Maß variiert. Die durchschnittliche Länge der Protokolle beträgt etwa 1400 Tokens. Das so ausgewählte Korpus wurde mit einer Mehrebenen-

⁶ Vgl. Macha [u. a.] (2005), S. XVIII f., für eine Übersicht über die räumliche Verteilung des Editions-korpus.

⁷ Stand der annotierten Daten, die für diese Auswertung benutzt wurden, ist der 02.04.2015.

annotation versehen, die im Rahmen des SIGS-Projekts (›Die Entwicklung der satzinternen Großschreibung im Deutschen‹)⁸ erstellt wurde. Dem Projekt liegt die Annahme zugrunde, dass sich die Herausbildung der satzinternen Großschreibung entlang morphosyntaktischer (Wortartenzugehörigkeit, Numerus) und kognitiv-semantischer (Referentialität, Belebtheit) Prinzipien vollzieht. Damit kann die Entwicklung der satzinternen Großschreibung innerhalb eines systematischen, linguistisch beschreibbaren Prozesses verortet werden, dessen Endpunkt in einer rein syntaktisch motivierten Großschreibung vermutet werden darf (vgl. Schutzeichel/Szczepaniak 2015).

Für die Aufbereitung des Kernkorpus lassen sich folgende Arbeitsphasen einteilen, von denen nur der Schritt von der dritten zur vierten Phase im SIGS-Projekt unternommen wurde:

Tab. 2: Arbeitsphasen, die zur Erstellung des SIGS-Kernkorpus führen

Phase 1	Phase 2	Phase 3	Phase 4
Manuskripte/Faksimiles →	Transkription →	Digitale Edition →	Annotation →
<i>Urgicht</i> -Teil ⁹ der handschriftlichen Verhörprotokolle	Edition (Macha [u. a.] 2005)	Grundlage des SIGS-Korpus	Erstellung des Kernkorpus

V. Reliabilität

Auf der Basis der vorangegangenen Ausführungen zum Graphinventar in den Verhörprotokollen sowie zu den Besonderheiten der Textsorte kann die Quantifizierung der Reliabilität der einzelnen Grapheme folgendermaßen erfolgen: Das vorhandene Initialinventar wird in zwei Klassen eingeteilt. Die Grundklasse, in der alle Grapheme enthalten sind, unterscheidet dabei noch nicht zwischen reliablen und nicht-reliablen Graphemen. Es wurde jedoch festgestellt, dass sich die einzelnen Grapheme hinsichtlich der Distinktivität von Majuskel und Minuskel unterscheiden. Diese Einteilung erzeugt einen weiteren Parameter für die

⁸ Das SIGS-Projekt unter der Leitung von Klaus-Michael Köpcke (Universität Münster) und Renata Szczepaniak (Universität Hamburg) wird finanziell unterstützt durch die Deutsche Forschungsgemeinschaft (DFG), SZ 280/ 2-1 und KO 909/ 12-1, unter der Mitarbeit von Fabian Barteld und Marc Schutzeichel. An der Annotation der Texte beteiligt waren zudem Annemarie Bischoff, Lisa Dücker, Julia Hübner, Johanna Legrum, Katja Politt, Nicolai Pudimat, Eleonore Schmitt, Annika Vieregge und Nicholas Wieling.

⁹ Hierbei handelt es sich um die Aussage bzw. das Hauptbekenntnis der/des Angeklagten, auf der die spätere Verurteilung beruht.

Analyse. Es ist nun möglich, den Einfluss der Graphe der Handschriften auf die Großschreibung zu quantifizieren, denn durch das Auslassen oder Hinzuzählen bestimmter Initiale verändert sich auch die Menge der berücksichtigten Tokens.

Die Vielschichtigkeit der anzusetzenden Unterschiede erlaubt den Umgang mit mehreren Reliabilitätsklassen, die um die größtmögliche Aussagekraft konkurrieren. Die Reliabilitätsklassen dieser Studie erfordern eine erneute Sichtung der Handschriften nach den oben erläuterten Kriterien. Die Menge der reliablen Elemente einer Klasse umfasst alle Graphe, die nach den unten dargestellten Parametern eindeutig zwischen Majuskel und Minuskel unterscheiden. Nicht beurteilbare Elemente sind Graphe, die nur in einer Variante auftreten, entweder nur als Majuskel oder nur als Minuskel realisiert sind und über deren Formdistinktion folglich nicht entschieden werden kann.

Tab. 3: ›Georgenthal 1597‹ – Reliabilitätsklassen, die für das Beispiel anzusetzen sind

	Basis	Parameter	Reliable Elemente ¹⁰	Nicht-reliable Elemente	Nicht-beurteilbar ¹¹
Reliabilitätsklasse 0	Gesamtes Initialinventar des Protokolls	keine	alle	keine	keine
Reliabilitätsklasse 1	Editionsnotizen Macha [u. a.] (2005, S. XXIII)	Verschiedenartigkeit und relative Größe »aus heutiger Sicht« (ebd.)	a, ä, b, c, d, e, f, g, i, j, k, l, m, n, o, ö, p, q, r, s, t, u, ü, w	h, v, z	keine
Reliabilitätsklasse 2	Erneute Sichtung und Bewertung aller Zeichen in den Handschriften (im Bsp.) / pro Protokoll	formdistinkt und/oder skalendis-tinkt	a, b, d, e, f, g, h, i, k, m, n, o, r, s, t, v, w, z	keine	c, j, p, u
Reliabilitätsklasse 3	Erneute Sichtung und Bewertung aller Zeichen in den Handschriften (im Bsp.) / pro Protokoll	ausschließlich formdis-tinkt	a, b, d, e, f, g, h, i, m, n, o, r, s, t	k, v, w, z	c, j, p, u

¹⁰ Die reliablen Elemente variieren wie zu erwarten je nach untersuchtem Protokoll.

¹¹ Nichtbeurteilbarkeit liegt dann vor, wenn ein Graphem nur in Minuskelform oder nur in Majuskelform vorkommt.

Da diese Reliabilitätsklassen nur geringe Unterschiede zueinander aufweisen, liegt es nahe, das Unterscheidungssystem auf zwei Klassen zu vereinfachen. Reliabilitätsklasse (0) berücksichtigt das gesamte verfügbare Initialinventar. Aus den Klassen (1) bis (3) ergeben sich für das Beispielprotokoll¹² ›Georgenthal 1597‹ die nicht-reliablen Elemente ›h‹, ›k‹, ›v‹, ›w‹, ›z‹.

VI. Texthomogenität

Obwohl alle Texte des SIGS-Kernkorpus der Textsorte Verhörprotokoll zuzuordnen sind, weisen sie deutliche Unterschiede auf und erfordern eine protokollspezifische Herangehensweise. Zunächst ist festzustellen, dass nicht die gesamten Protokolle in der Edition von Macha [u. a.] (2005) enthalten sind, sondern nur die Urgicht. Abweichungen vom prototypischen Urgicht-Teil eines Verhörprotokolls sind durch die weit gesteckten Rahmenvorgaben begründet: Ein zeitliches Spektrum von 1588 bis 1630 und eine große räumliche Divergenz resultieren in vielen weiteren – außersprachlichen – Faktoren, die Heterogenität beisteuern. So geht mit der Nähe zu Schreibzentren und juristischen Zentren ein für die Gestaltung des Manuskripts wesentlicher Aspekt einher, während sich der textuelle Aufbau nach dem tatsächlichen Verfahren bzw. Verhör richtet. Die Anzahl der Schreiber ist nicht konstant, zuweilen ist von mehr als einer Hand auszugehen (z. B. ›München 1600‹ und ›Rosenburg 1618‹). Dieser Punkt wird teils durch unbekannte Provenienz der Handschriften sowie Korrekturen und Abschriften weiter erschwert. Die räumliche und zeitliche Verortung wirkt sich zudem maßgeblich auf die dialektalen und schreibsoziologischen Eigenschaften der vorgefundenen sprachlichen Verhältnisse aus. Trotz dieser Unterschiede zwingt der Umstand, dass es sich um das Protokoll eines juristischen oder juristisch relevanten Ereignisses handelt, die Texte in eine weitgehend vorgegebene Rahmenstruktur, die maßgeblich durch die Wiedergabe von Frage- und Antwortsequenzen der Verhör-situationen charakterisiert ist. Vergleichbare inhaltliche Tendenzen, etwa wiederkehrende Themen des Motiv-Kanons der Hexenverfolgung (vgl. Topalović 2003, S. 8–35, am Beispiel der Stadt Osnabrück) machen das Korpus insgesamt zu einer Menge vergleichbarer textueller Einheiten.

Ausschlaggebend für die Homogenität des Faktors ›Buchstabe‹ ist der Umstand, dass alle Texte des Korpus auf der Edition von Handschriften basieren.

¹² Neben dem exemplarisch detailliert besprochenen ›Georgenthal 1597‹ wurden auch die Protokolle ›Jever 1592‹, ›Günzburg 1613‹, ›Rosenburg 1618‹, ›Schweinfurt 1616‹ und ›Stralsund 1630‹ als Fallstudien herangezogen.

Innerhalb eines Protokolls liegen Schriftzeugnisse nur eines Schreibers, in Ausnahmefällen auch von mehr als einem Schreiber vor. In der graphematischen Forschung werden zwei Formen der Variabilität von handschriftlichen Produktionen unterschieden: Interindividualität und Intraindividualität (vgl. Nottbusch 2008, S. 46 f.). Die folgende Tabelle demonstriert die formale Unterschiedlichkeit in der Schreibung der zwei vermuteten Hände im Protokoll Rosenberg 1618.

Tab. 4: »Rosenburg 1618« – Unterschiede in der Realisierung der Grapheme durch zwei Hände

	a		d		o	
	Min.	Maj.	Min	Maj	Min	Maj
Hand 1
Hand 2

Der Herausforderung der unterschiedlichen formseitigen Realisierungen der Grapheme im Korpus sind die Editoren der Verhörprotokolle mittels einer gründlichen Methodik und mit konsequent eingehaltenen Editionsgrundsätzen (vgl. Macha [u. a.] 2005, S. XXII–XXIV) begegnet. Es gilt zu berücksichtigen, dass dabei die Problematik der Groß- und Kleinschreibung zu einem hohen Maß reflektiert wurde, obgleich sie keineswegs im editorischen Hauptinteresse lag. Es bleibt insofern offen, ob die für eine solche Fragestellung hinreichende Erfassung der Großschreibung erfüllt ist. Die Edition liefert somit die erste Klassifizierungsmöglichkeit der Initiale des Korpus in eine reliable Gruppe einerseits und eine nicht-reliable Gruppe andererseits. Unter diesem Aspekt der Reliabilität treten solche Grapheme in den Vordergrund, deren Majuskel- und Minuskelformen sich lediglich anhand der Skalierung diskriminieren lassen. Reliabel hinsichtlich der Transkription ist nach diesen Regeln, was eindeutig zu identifizieren ist, daher stellen die explizit von den Editoren aufgeführten Problemkandidaten <z>, <v>, <h> (vgl. Macha [u. a.] 2005, S. XXIII) die erste Klasse der nicht-reliablen Initiale. Sie wurden für die Edition im Zweifelsfall, insbesondere allerdings im – hier irrelevanten – Wortinnern mit Kleinschreibung transkribiert (vgl. ebd.). Die relative Größe der Schriftzeichen liefert den Anhaltspunkt für die Entwicklung zweier grundlegender Parameter: Für die im Folgenden zu erstellende Gruppe nicht-

reliabler Graphen sollen die Eigenschaften Formdistinktion und Skalendistinktion grundlegend sein. Die Merkmale formdistinkt und skalendistinkt zeichnen dabei eine Opposition aus. Die Handschriften zeigen auch Majuskel-Minuskel-Paare, die formdistinkt und gleichzeitig skalendistinkt sind.

Zwangsläufig ist eine zeit-, raum- und handübergreifende Klassifikation der Majuskelvarianten hinsichtlich ihrer Reliabilität nicht zu leisten. Vielmehr muss die Klassifikation idealerweise je Text erfolgen. Die unterschiedliche graphische Gestaltung der Majuskel im Vergleich zu der zugehörigen Minuskel mit demselben Lautwert unterliegt demnach einer Anzahl an Parametern. Wie oben gezeigt, ist dabei der verfügbare Platz im Manuskript ein wesentlicher Faktor. Dies zeigt sich immer wieder an der Länge der Hasten sowie an Stauchungen im Oberband und vor allem an den Anfängen von Zeilen, Absätzen und Seiten. Aufbauend auf diesen Überlegungen können die folgenden Einflussfaktoren für Gestalt und Größe des Schriftzeichens isoliert werden: 1.) Das Schriftzeichen selbst, 2.) Intention und Regelbewusstsein des Schreibers, 3.) Position im Text, 4.) Layout des Manuskripts sowie 5.) umgebende Schriftzeichen.

Es stellt sich die Frage, wie sich die Transkription von Initialen in Handschriften als eigenständiger Faktor für die Messung der Großschreibung auswirkt und modellieren lässt. Hierzu soll mit den Konzepten der Reliabilität und Repräsentativität in Bezug auf die transkribierten Schriftzeichen vor dem Hintergrund des im Folgenden dargestellten Korpus operiert werden. Anhand von Fallstudien wird gezeigt, dass bestimmte Grapheme bei isolierter Betrachtung wiederkehrend als nicht-reliabel in Bezug auf die Wiedergabe der Großschreibung identifiziert werden können. Die Berücksichtigung ausschließlich reliabler Belege führt zu einer Reduktion der verwendbaren Datensätze und schlägt sich somit unmittelbar in den auszuwertenden Anteilszählungen, d. h. den Frequenzen der großgeschriebenen Tokens nieder. Da gleichzeitig eine erneute Transkription mit dem Schwerpunkt der Unterscheidung von Groß- und Kleinbuchstaben bei bereits bestehenden Editionen nicht ökonomisch erscheint, ist es das Ziel der vorliegenden Studie die Schwankungen in der Graphreliabilität abzubilden und in die mittels deskriptiv-statistischer Methoden erstellten Anteilszählungen einzuberechnen.

VII. Hypothesen

Die bisher angestellten Überlegungen orientieren sich an der Frage, welche Buchstaben bei der Identifizierung der Majuskeln in den Handschriften als uneindeutig gelten können. Dabei kann angenommen werden, dass eine auf diesen Erkenntnissen basierende Wertung in Form von unterschiedlichen Klassifika-

tionsversuchen der Buchstaben in eine reliable und eine nicht-reliable Gruppe eine genauere Einschätzung des Verhältnisses zwischen großgeschrieben und kleingeschriebenen Tokens erlaubt. Der formbasierten Klassifikation (Reliabilitätsklasse [3]) ist dabei gegenüber der skalenbasierten Klassifikation der Majuskeln der Vorzug zu geben, da sie in erster Linie nicht von der relativen Größe der umgebenden Schriftzeichen abhängig ist, sondern ihre Diskriminationsfähigkeit auch an einem isolierten Schriftzeichen leisten kann.

Bei der Durchsetzung der satzinternen Großschreibung werden nicht alle Wortarten in gleichem Maße erfasst (vgl. Kaempfert 1980; Moulin 2004; Nerius 2003). Erst bei der allmählichen Festlegung der Großschreibung auf die Nomina findet ein Ausgleich bei den übrigen Wortarten statt. Besonders relevant für die Überprüfung der Projekthypothesen ist somit die Entwicklung der Großschreibung bei Nomina und Eigennamen, die als Kerne von Nominalphrasen das Ergebnis einer syntaktisch gesteuerten satzinternen Großschreibung darstellen. Es ist also zunächst von zwei unterschiedlichen Modellierungen auszugehen. In Modell A (Reliabilitätsklasse [0]) wird die Buchstabenreliabilität als eigener Faktor ignoriert. Dagegen wird in Modell B die Buchstabenreliabilität als eigener Faktor angesehen und explizit mit Blick auf die Großschreibung analysiert, indem in der direkten Arbeit mit den zugrundeliegenden Faksimiles der Manuskripte eine Klassifizierung in reliable und nicht-reliable nach den oben entworfenen Prinzipien vorgenommen wird.

Es wird angenommen, dass Modell B insofern genauere Aussagen über die Entwicklung der satzinternen Großschreibung ermöglicht, als hier protokollspezifische (Schrift-)Variation Berücksichtigung findet, während in Modell A dieser Aspekt nicht von den übrigen Parametern abgelöst werden kann. Darüber hinaus ist davon auszugehen, dass sich bei einem direkten Vergleich der Messwerte der beiden Modelle sichtbare Abweichungen ergeben. Dies muss dabei nicht unbedingt zu einer Widerlegung der vorherigen Ergebnisse führen, sondern ist vielmehr als eine Präzisierung der Messwerte anzusehen. Die Aussagekraft der Reliabilitätsklassen untereinander ist unterschiedlich groß und variiert zusätzlich von Hand zu Hand, sodass die Klassen für jeden neuen Text erneut bestimmt werden müssten. Liegen in einem Text Passagen unterschiedlicher Hände vor, so müssen diese separat als zwei unterschiedliche Texte behandelt werden. Unter der Voraussetzung eines edierten Korpus kann das hier angewendete Verfahren allerdings verallgemeinert werden.

VIII. Repräsentativität

Gerade bei der Einteilung der satzintern großgeschriebenen Tokens nach Wortartenzugehörigkeit zeigt sich ein weiterer Aspekt aus dem Kernbereich korpuslinguistischer Prämissen: ein Maß an Repräsentativität. Bezogen auf die Großschreibung besagt Repräsentativität, in welchem Ausmaß der Anteil der Majuskeln in einem Text von einzelnen Graphemen getragen wird. Anders ausgedrückt lässt sich die Großschreibung als Phänomen eines Alphabets, genauer gesagt, eines Grapheminventars erfassen, dem eine bestimmte Menge an graphischen Zeichen angehört. Die einzelnen Realisierungen der Grapheme kommen im Korpus sowie pro Protokoll nicht gleichermaßen häufig vor. Dieser Umstand begründet sich durch die begrenzte Kombinierbarkeit und Frequenz der sprachlichen Zeichen auf der Wort- und Satzebene. Es ist davon auszugehen, dass die Großschreibung einer geschlossenen Wortklasse wie die der Artikelwörter, die sich zwangsläufig aus erheblich weniger verschiedenen Initialen als die einer offenen Klassen wie die der Nomina konstituiert. Einzelne Grapheme, beispielsweise das <E>, schlagen sich hier demnach prozentual nicht im selben Maße nieder wie in der offenen Klasse.

Werden Aussagen über die Entwicklung der satzinternen Großschreibung gemacht, kann nur durch die Entwicklung einer quantifizierbaren Repräsentativität ermittelt werden, in welchem Ausmaß die einzelnen Texte des Korpus sowie im Einzelnen die jeweiligen Allographe die Entwicklungstendenzen in der Schreibung der Initiale anteilmäßig repräsentieren. Stützen sich die Ergebnisse auf Veränderungen in einzelnen, wenigen Allographen, sinkt der Wert der Repräsentativität. Für jeden auftretenden Buchstaben des Alphabets, und somit für jedes Graphem, das im Korpus durch Allographe repräsentiert wird, lässt sich ein Wert bestimmen, der anzeigt, in welcher Relation sein Vorkommen im Vergleich zu Einzeltext und Korpus steht. Auf diese Weise entsteht ein die Aussagekraft der Antworten auf die weiteren Untersuchungsfragen einschätzbar machender Parameter. Die Einzigartigkeit der Protokolle erlaubt es dabei, für jeden Text eine Art graphematischen Fingerabdruck zu erstellen, vor dessen Hintergrund eine Gewichtung der Belege großgeschriebener Tokens pro Kategorie möglich wird.

Im idealen Fall sind alle Buchstaben des Grapheminventars in gleicher Zahl als Initiale in einem Protokoll vertreten. Da dies aus linguistischer Perspektive unrealistisch erscheint, kann die Messung der Repräsentativität erfolgen, indem ein Vergleichswert aus den tatsächlichen Okkurrenzen und dem Ideal einer gleichmäßigen Verteilung gebildet wird. Der Grad der Approximation an das Repräsentativitätsideal bestimmt den Messwert. Berechnet man für alle Grapheme eines Protokolls das durchschnittliche Vorkommen, so ist ein Protokoll hinsichtlich seines Grapheminventars in dem Maß repräsentativ, wie sein graphemati-

scher Fingerabdruck (siehe Abb. 3) einer gleichmäßigen Verteilung ähnlich ist. Der so berechnete Wert ist also eine Eigenschaft und Messgröße des jeweiligen Graphs.

Der Repräsentativitätsgrad besitzt wenig Aussagekraft, wenn ein einzelner Text in seiner gesamten Tokenzahl betrachtet wird. Für die Frage nach der satz-internen Großschreibung ist aber zwangsläufig die Klassifizierung der Tokens nach verschiedenen Ausprägungen von Variabilität durchzuführen. Dies kann in Form von Wortartenzugehörigkeit, Belebtheitsklassen und unter zahlreichen anderen Gesichtspunkten geschehen. Bildet man die Auftretenshäufigkeit der Initiale pro Text ab, so ergibt sich folgender Ansatz: Wie wahrscheinlich ist es, bezogen auf die Initiale des Protokolls, dass bei einer zufälligen Stichprobe eines Tokens aus dem Protokoll das jeweilige Initial vorliegt? Zur Beantwortung dieser Frage sind folgende Berechnungsschritte geeignet:

1. Ermittlung aller Tokens eines Protokolls (Tokenisierung)
2. Feststellen und Auszählen der vorkommenden Initiale der Tokens
3. Berechnung der Wahrscheinlichkeit $p = \text{Vorkommen} / \text{Tokenzahl}$

Damit ergibt sich für das Protokoll ›Georgenthal 1597‹ die folgende Verteilung:

Abb. 3: Graphematisches Profil der Initiale in ›Georgenthal 1597‹, satzintern. Der schwarze Ring nahe des Zentrums zeigt die Durchschnittshäufigkeit

Am Beispiel der Wortartenzugehörigkeit soll im Folgenden demonstriert werden, welche Abhängigkeiten sich zwischen den Ergebnissen der anteilmäßigen Großschreibung und einzelnen Buchstabengruppen ergeben. Mithilfe der quantifizierten Faktoren Reliabilität und Repräsentativität wird dabei eine Validierung des Phänomens satzinterne Großschreibung möglich.

IX. Auswertung

Die nicht-reliablen Elemente aus Modell B (siehe Abb. 4 am Beispiel von ›Georgenthal 1597‹) sind ›k‹, ›v‹, ›w‹, ›z‹. Für die Frequenzzählung der satzintern großgeschriebenen Tokens heißt dies, dass Tokens, deren Initial eines dieser Elemente ist, in Modell B nicht mitgezählt werden, während sie in die Zählung von Modell A eingehen¹³. Dadurch ergibt sich eine Diskrepanz, die sich am Beispiel der wortartenspezifischen satzinternen Großschreibungen demonstrieren lässt. Durch den Wegfall von Tokens mit nicht-reliablem Initial ergeben sich Reduzierungen in der Anzahl sowohl von großgeschriebenen, als auch von kleingeschriebenen Tokens. Daher ist ein Anstieg wie auch ein Abfall der anteilmäßigen satzinternen Großschreibung denkbar. Im folgenden Beispiel zeigt Modell B durchgängig größere Anteilswerte. Das heißt, der Wegfall der Tokens mit nicht-reliablen Initialen führt insgesamt zu einer anteilmäßig höheren Großschreibung, wie die Abbildung veranschaulicht.

Es zeigt sich, dass Modell B für das Protokoll ›Georgenthal 1597‹ durchgängig mehr satzinterne Großschreibung nachweist. Der Unterschied ist insbesondere bei den Berufsbezeichnungen und Eigennamen deutlich zu erkennen. Eigennamen und Berufsbezeichnungen erreichen beinahe den Wert durchgängiger Großschreibung. Dieses Ergebnis verdeutlicht, dass Modell B die Durchsetzung der Großschreibung exakter abbildet. Als Konsequenz ist davon auszugehen, dass die Durchsetzung der satzinternen Großschreibung im Beispieltext schon erheblich weiter fortgeschritten ist als die unkorrigierte Zählweise (Modell A) suggeriert. Auch für die Präpositionen kann die satzinterne Großschreibung erheblich nach oben korrigiert werden. Eine mögliche Erklärung ist die Auszeichnung der Präposition als Kopf einer Phrase. Deren Annotation steht im Untersuchungskorpus zum gegenwärtigen Zeitpunkt aber noch aus.

¹³ Für Modell A kann nur Reliabilitätsklasse (0) angesetzt werden, da bei der Sichtung der Handschriften bereits in Fallstudien gezeigt werden kann, dass es immer nicht-reliable Elemente gibt.

Abb. 4: Wortarten nach Reliabilitätsklassen am Beispiel ›Georgenthal 1597‹.

ADJ = Adjektive; APPR = Präposition; ART = Artikel; NB = Nomina, die Berufsbezeichnungen sind; NE = Nomina Propria; NN = Nomina Appellativa; PRO = Pronomina; V = Verben

Bei der Betrachtung je nach Wortart wird deutlich, weshalb die Auswahl der zu berücksichtigenden Initiale bei der Zählung der satzinternen Großschreibung unterschiedliche Auswirkungen hat. Mit Blick auf die Repräsentativität der Buchstaben lässt sich feststellen, dass innerhalb der Wortarten nicht jedes Initial gleichermaßen repräsentiert ist. Anders gesagt: Jede Wortart weist ein unterschiedliches Initialprofil auf. Die folgende Grafik schlüsselt die gemessenen Anteile der satzintern großgeschriebenen Tokens nach ihrer Wortart und gleichzeitig nach ihrem jeweiligen Initial auf:

Abb. 5: Repräsentation der Wortarten in ›Georgenthal 1597‹ nach Modell A durch verschiedene Initialprofile, gemessen an der Auftretenswahrscheinlichkeit der Initialen pro Wortart.

(ADJ = Adjektive; APPR = Präpositionen, ART = Artikel, NB = Berufsbezeichnungen, NE = Nomina Propria, NN = Nomina Appellativa. PRO = Pronomina, V = Verben)

Wie der Abbildung zu entnehmen ist, gestaltet sich die Verteilung der Initialen innerhalb der einzelnen Wortarten sehr unterschiedlich. Je nachdem, welche Initialen in die nicht-reliable Klasse gerechnet werden, sind die Wortarten daher auch mehr oder weniger stark betroffen. Gilt beispielsweise das <d> als nicht-reliabel, hat dies große Auswirkungen auf die Artikelwörter, da sie von nur zwei Initialen, <d> für bestimmte und <e> für unbestimmte Artikel, repräsentiert werden. Vergleichbares lässt sich für die häufigen Initialen der Präpositionen (<i>, <m>, <v>, <z>) beobachten. Die Klasse der Pronomina ist hingegen weniger stark betroffen, da sie in dem Protokoll ›Georgenthal 1597‹ aus elf Initialen besteht. Nicht betroffen wären beispielsweise die Verben, da hier <d> nicht wortinitial auftritt. Generell lässt sich daher folgern: Je geringer die Diversität im Initialprofil einer Wortart, desto stärker wirkt sich die Klasseneinteilung in reliable und nicht-reliable Initialen auf diese Wortart aus. Die Auswirkungen der für ›Georgenthal 1597‹ als nicht-reliabel eingestuft Initialen <k>, <v>, <w>, <z> auf die verschiedenen Wort-

arten lassen sich somit anhand der Initialprofile nachvollziehen. Die Klasse der Eigennamen wird durch 20 verschiedene Initiale konstituiert. Alle Elemente der nicht-reliablen Klasse finden sich auch im Initialprofil der Eigennamen, dessen Anzahl sich nach Modell B auf 15 verringert. Fast alle bisher als klein geschrieben transkribierten Eigennamen fallen damit in die nicht-reliable Klasse, es bleibt nur <l> mit einem Beleg. Der Anteil der Großschreibung steigt damit nahezu auf 100 Prozent (98,2 %) an, was in dieser zeitlichen Phase der Großschreibung den Erwartungen entspricht. Die starke Abweichung zu Modell A verdeutlicht, dass es nicht möglich ist, auf dessen Basis die Wirklichkeit der sprachlichen Verschriftlichungsstrategien abzubilden.

Des Weiteren ist zu berücksichtigen, dass die wiederkehrend als nicht-reliabel gedeuteten Grapheme mit dieser Methode auf Kontexte verallgemeinert werden, in denen sie nicht überprüft wurden. Dies geschieht auf Basis der Annahme, dass die Elemente der nicht-reliablen Klasse aufgrund ihrer formalen Eigenschaften uneindeutig sind. Bei dieser Verallgemeinerung gilt ein Buchstabe als nicht-reliabel, wenn er in den Fallstudien häufiger als der Mittelwert als nicht-reliabel identifiziert wurde. Für diese Messung möchten wir ein Kontinuum vorschlagen.

In den Fallstudien wurden die ›Problemkandidaten‹ und nicht ›entscheidbaren Fälle‹ wie folgt gezählt:

Tab. 5: Problemkandidaten

k	z	o	d h l v	c n p w
6	4	3	2	1

Tab. 6: Nicht entscheidbare Fälle

j	c	p	z	h l ö	e i k o u
7	5	4	3	2	1

Die beiden Skalen reichen in Leserichtung von maximal häufig bis minimal häufig. Da auch die nicht entscheidbaren Fälle nicht zu den reliablen Elementen gezählt werden können, ergibt sich aus der Synthese der beiden Skalen das Gesamtbild.

sichtigen, dass die Präzisierung durch genauere Sichtung der Handschriften eine solche Verallgemeinerung unnötig machen würde.

Die Berücksichtigung der Reliabilität bei der Auswertung der anteilmäßigen Großschreibung hat weitreichende Folgen, wenn das gesamte Kernkorpus nach Modell B ausgezählt wird:

Tab. 8: Anteilmäßige satzinterne Großschreibung im SIGS-Kernkorpus unter Berücksichtigung der in den Fallstudien als nicht-reliabel festgestellten Grapheme

	Ort	Jahr	Modell A	Modell B	Differenz
1.	Alme	1630	0.34	0.34	0.00
2.	Baden-Baden	1628	0.42	0.40	-0.02
3.	Bamberg	1628	0.42	0.36	-0.06
4.	Gaugrehweiler	1610	0.22	0.31	0.09
5.	Georgenthal	1597	0.49	0.55	0.06
6.	Günzburg	1613	0.25	0.29	0.04
7.	Güstrow	1615	0.29	0.22	-0.07
8.	Hamm	1592	0.23	0.24	0.01
9.	Jever	1592	0.32	0.33	0.01
10.	Lemberg	1630	0.40	0.40	0.00
11.	Meldorf	1618	0.29	0.29	0.00
12.	München1	1600	0.19	0.26	0.07
13.	München2	1600	0.57	0.48	-0.09
14.	Ostrau	1628	0.74	0.84	0.10
15.	Perleberg	1588	0.07	0.05	-0.02
16.	Riedlingen	1596	0.49	0.53	0.04
17.	Rosenburg1	1618	0.51	0.56	0.05
18.	Rosenburg2	1618	0.33	0.34	0.01
19.	Schweinfurt	1616	0.52	0.42	-0.10
20.	Stralsund	1630	0.32	0.32	0.00

Tab. 8 zeigt, dass die Variation der satzinternen Großschreibung im Korpus groß ist. Der Anteil der satzintern großgeschriebenen Initiale je nach Protokoll schwankt, was auf das Erscheinungsjahr (von 1592 bis 1630) und die geographische Lage des Erscheinungsortes zurückzuführen ist. Jedoch lassen sich die Unterschiede nicht allein durch zeitliche oder räumliche Distribution erklären.

Auf das ganze Untersuchungskorpus ausgeweitet zeigen sich Differenzen im Anteil der Großschreibung der Nomina und Eigennamen, die von null bis zehn Prozent reichen. In Modell B kommen Abweichungen sowohl durch den Wegfall kleingeschriebener als auch großgeschriebener Tokens zustande, wenn deren Initial zur Menge der nicht-reliablen Elemente gehört. Da sich die nach Modell B berechneten Anteile ausschließlich auf reliable Elemente stützen, kann die ansonsten für Handschriften problematische Zählung der Großbuchstaben präzisiert werden. Ohne die Einteilung in Reliabilitätsklassen anhand des Kriteriums Formendistinktion ergeben sich teils Unterschiede im zweistelligen Prozentbereich, die gravierende Konsequenzen für die Interpretation der Ergebnisse mit sich bringen. Aus den Ergebnissen wird somit deutlich, dass die vorherigen Zählungen nicht grundlegend falsch waren, dass es jedoch möglich ist, sie durch die hier vorgeschlagene Methode zu präzisieren. So ist eine genauere Auswertung auf der Basis der Handschriften und der Transkription möglich.

Die Daten in der Tabelle verdeutlichen, dass eine Unterscheidung zwischen reliablen und nicht-reliablen Initialen für graphematische Untersuchungen zwingend erforderlich ist. Im Ergebnis liegen schließlich zwei Modelle vor, die auf der Reliabilität der Transkription der Initiale beruhen. Modell A geht davon aus, dass alle Buchstaben als reliabel zu transkribieren sind. Allerdings zeigt ein Blick in die Handschrift, dass dies aufgrund der fünf Einflussfaktoren auf Gestalt und Größe des Schriftzeichens nicht möglich ist. Anhand eindeutiger Kriterien lassen sich die Grapheme in Reliabilitätsklassen einteilen. Obwohl sich Effekte möglicher Transkriptionsfehler und -uneindeutigkeiten nie komplett aufheben lassen, können sie auf diese Weise minimiert werden.

X. Fazit

Jeder handgeschriebene Text ist in graphetischer Hinsicht einzigartig. Er kann nicht unabhängig von den in dieser Untersuchung ermittelten Faktoren analysiert werden. Einen konsequenten Zugang zum graphematischen Profil eines Textkorpus bietet daher nur die Sichtung und Auswertung aller zugrundeliegenden Manuskripte. Hier sind sicherlich, aus ökonomischen Gründen, Grenzen vorhanden. Die in dieser Studie entworfene Methode kann jedoch zielgerichtet angewendet und ausgeweitet werden und dient somit als Schablone für weitere Untersuchungen an Manuskripten, die sich auf bereits bestehende Editionen stützen. Für Forschungsfragen, die direkten oder indirekten Bezug zur Großschreibung aufweisen, gilt es, insbesondere den Fokus auf jene Grapheme zu richten, die sich hier protokollübergreifend als problematisch erwiesen haben,

da diese durch den unternommenen Vergleich bereits einen regionalen, idiolektalen und zeitlichen Durchschnittswert implizieren. Zudem bleibt zu betonen, dass handschriftliche Texte nicht aufgrund ihrer spezifischen Eigenschaften bei linguistischen Untersuchungen unberücksichtigt bleiben dürfen. Der Umgang mit ihnen erfordert aber eine besondere Herangehensweise, die der Ambiguität der handschriftlichen Zeichen Rechnung trägt. Eine solche Herangehensweise bei der Transkription mit besonderem Fokus auf die Groß- und Kleinschreibung stellt sich wie folgt dar:

Zunächst sind Buchstabengrenzen in den Text einzuziehen. Die dadurch entstehenden Zeichen werden, falls möglich, ungeleitet vom Wortkörper transkribiert. Der Vorteil dieser Strategie besteht in der geringen Beeinflussung der Transkribierenden durch die Bedeutung der einzelnen Wörter und Sätze. Ist die qualitative Identifikation des Schriftzeichens gelungen, mag sich die Entscheidung zwischen Minuskel oder Majuskel, gerade bei nicht-reliablen Graphen, unbewusst an der Semantik, der Wortart oder dem syntaktischen Umfeld orientieren. Der Einfluss dieser Fehlerquelle wird durch die Segmentierung in singuläre Graphe unterbunden. Die Transkription findet in mindestens zwei Durchgängen statt und wird von mindestens zwei Transkribierenden geleistet. Im direkten Vergleich der Transkripte werden Übereinstimmungen und Unterschiede sichtbar. Für dieses Szenario erscheint der Einsatz eines unterstützenden bzw. automatischen EDV-Systems (Feststellen des *Inter Annotator Agreement*) sinnvoll. Erst bei den im Vergleich gefundenen Abweichungen bzw. nicht entschlüsselbaren Zeichen wird in einem zweiten Schritt der ganze Wortkörper zur Transkription herangezogen. Besonders bei der Anzeige, der Verteilung und dem Vergleich der einzelnen Segmente sollten so weit wie möglich Computerprogramme eingesetzt werden. Darüber hinaus können die Ergebnisse mit gegenwärtigen Werkzeugen bei der automatischen Texterkennung von historischen Handschriften verglichen werden (vgl. Worch/Gottfried 2014). Von der digitalen Paläographie sind in diesem Kontext im Zuge des Aufschwungs der *Digital Humanities* wichtige Impulse für die Erforschung der Entwicklung der Großschreibung zu erwarten. Wenngleich sich bisherige Projekte zu dem Themenfeld ›*Optical Character Recognition*‹ zuerst auf historische Drucke konzentrieren (etwa das DFG-Projekt OCR-D¹⁴), sind in der Verbindung von Datenbanken und Annotationswerkzeugen aus dem Bereich Bild und Text Schlüsselemente für die Bewältigung des hier skizzierten Problems zu sehen. Andere Software, beispielsweise das Werkzeug ›Transkribus‹¹⁵, eröffnen die Perspektive auf die Erforschung von Handschriften mittels virtueller Forschungsumgebungen. Die Berücksichtigung der Majuskel-

14 URL: <http://www.ocr-d.de/> (Aufrufdatum: 12.03.2018)

15 URL: <https://transkribus.eu/Transkribus/> (Aufrufdatum: 12.03.2018)

Reliabilität erfährt unter Zuhilfenahme digitaler Verfahrensweisen eine dem Zweck angemessene Skalierbarkeit.

Die diesem Beitrag zugrunde gelegten Texte verweisen in eine Zeit grammatischer wie auch graphischer Emergenz und sind nicht isoliert zu verstehen. Aus den besprochenen Ergebnissen lassen sich Konsequenzen für die Arbeit mit annotierten Transkripten handschriftlicher Texte, historischer sowie neuzeitlicher, ableiten. Dabei ist zunächst festzuhalten, dass eine Sichtung der Handschriften und Notizen in der Basisedition bzw. der tatsächlichen handgeschriebenen Texte unerlässlich ist. Nur so kann eine angemessene Klasseneinteilung zur Messung der Reliabilität geschehen. Das Ausmaß der Reliabilität der Transkription ist bei der Arbeit auf der Basis einer Edition bereits notwendigerweise festgelegt, da beim Transkriptionsprozess selbst eine Entscheidung getroffen wurde. Erst die Sichtung der Handschriften und die Einteilung in Buchstabenklassen erlaubt die Bestimmung der Reliabilität anhand der formdistinkten Merkmale der Initiale. Das Ausmaß der Repräsentativität hingegen ist erst in Abhängigkeit der Parameter der jeweiligen Fragestellung zu bestimmen. Wir konnten am Beispiel von handgeschriebenen Texten zeigen, dass die Grapheme als eigenständiger Faktor eine weitere Form von Repräsentativität beisteuern, die über die Repräsentativität der Textsorte, einer vergleichbaren Satzlänge und dergleichen hinausgeht. Diese verdeutlicht, dass supraphemische Phänomene, wie die Großschreibung, von einzelnen Buchstaben unterschiedlich stark getragen werden. Der Unterschied zeigt sich, sobald für die Auswertung Klassen, wie etwa auf Basis der Wortartenzugehörigkeit oder je nach Untersuchungsgegenstand beispielsweise für verschiedene Wortfelder, gebildet werden.

Bei kleineren Klassen (z. B. ›Wortart: Artikel‹) kann es vorkommen, dass der Effekt der Großschreibung unter Umständen nur von einem einzigen reliablen Buchstaben getragen wird. In einem solchen Fall werden keine Aussagen über das Phänomen der Großschreibung gemacht, sondern im Extremfall über die Großschreibung eines einzelnen Schriftzeichens. Diese Diskrepanz wird jedoch ohne die Messung der Repräsentativität verschleiert. Die Belege stützen sich somit auf weniger Evidenzmaterial und sind damit nur schwer verallgemeinerbar. Protokollspezifische Entscheidungen über die Reliabilität müssen im Idealfall für jeden einzelnen Buchstaben in jedem Text getroffen werden. Vorstellbar ist eine Datenbank, die wiederkehrende Vergleichswerte für Buchstaben mit den Parametern ›Zeit‹, ›Ort‹, ›Text‹, ›Position im Text‹, ›Textsorte‹, ›Alphabet¹⁶ und ›Schreiber/in‹ in Verbindung bringt, um die Reliabilität hinsichtlich der Groß-

¹⁶ Das Konzept ist auf alle Schriftsysteme übertragbar, die über Majuskeln und Minuskeln verfügen. Zu beachten bleibt aber, dass das Merkmal der Formdistinktheit in den historischen Schreibschriften anderer Buchstabenalphabete womöglich weniger stark zum Tragen kommt.

schreibung in größerem Maßstab evaluieren zu können. Aus den Gemeinsamkeiten von problematischen Schriftzeichen kann versucht werden, übergeordnete formal-gestalterische Tendenzen abzuleiten, die dazu führen, dass eine Schriftzeichenform als nicht-reliabel eingestuft wird.

Es hat sich gezeigt, dass die für korpuslinguistische Untersuchungen relevanten Faktoren Reliabilität und Repräsentativität je nach Untersuchungsgegenstand neu definiert werden müssen. Bei der Fragestellung der satzinternen Großschreibung in handschriftlichen Texten stellt sich nicht nur die Frage nach der Repräsentativität der Textsorte oder der Einheitlichkeit der Satzlänge sondern auch nach der Vergleichbarkeit der Initialprofile der einzelnen Texte. Beim Vergleich der Großschreibung innerhalb eines oder mehrerer Texte unterschiedlicher Schreiber dürfen nur dieselben reliablen Initiale mit in die Berechnung einbezogen werden. Während bei gedruckten Texten von der Einheitlichkeit der Schreibungsvarianten ausgegangen werden kann, ist der Blick in den Originaltext bei Handschriften unerlässlich. Die graphische Realisierung des Schriftzeichens selbst muss kritisch hinterfragt werden, besonders wenn der Zugang zu den Originaldaten bzw. den Anmerkungen der Transkribierenden nicht mehr gewährleistet ist. Durch die vorgeschlagene Methode verringert sich dabei zwar die Anzahl der zur Untersuchung auswählbaren Wörter, die Aussagekraft der Ergebnisse wird jedoch deutlich erhöht und erlaubt eine präzisere Auswertung der historischen Daten. Das hier vorgestellte Verfahren zeichnet sich durch seine Erweiterbarkeit aus: Die Grenzwerte, die bestimmen, welche graphischen Elemente als nicht-reliabel markiert werden und somit aus der Zählung herausfallen, können je nach Fragestellung festgelegt werden. Das am Beispiel des SIGS-Kernkorpus dargelegte Vorgehen bietet lediglich eine Empfehlung in Form eines Kontinuums, das die Identifikation wiederkehrender unsicherer Elemente erlaubt.

Literatur

- Bergmann, Rolf u. Dieter Nerijs 1997: Die Entwicklung der Großschreibung im Deutschen von 1500 bis 1700, 2 Bde., Heidelberg (Germanische Bibliothek, Neue Folge, Reihe 3, Untersuchungen 29).
- Fuhrhop, Nanna 2008: Das graphematische Wort (im Deutschen). Eine erste Annäherung, in: *Zeitschrift für Sprachwissenschaft* 27, S. 189–228.
- Fuhrhop, Nanna u. Jörg Peters 2013: Einführung in die Phonologie und Graphematik, Stuttgart.
- Gallmann, Peter 1985: Graphische Elemente der geschriebenen Sprache. Grundlage für eine Reform der Orthographie, Tübingen (Germanistische Linguistik 69).
- Günther, Hartmut 1988: Schriftliche Sprache. Strukturen geschriebener Wörter und ihre Verarbeitung beim Lesen, Tübingen (Konzepte der Sprach- und Literaturwissenschaft 40).

- Jacobs, Joachim 2005: *Spatien. Zum System der Getrennt- und Zusammenschreibung im heutigen Deutsch*, Berlin u. New York (Linguistik. Impulse & Tendenzen 8).
- Kaempfert, Manfred 1980: Motive der Substantiv-Groß-Schreibung. Beobachtungen an Drucken des 16. Jahrhunderts, in: *Zeitschrift für Philologie* 99, S. 72–98.
- Maas, Utz [u. a.] 1989: Überlegungen zur graphischen Analyse. Am Beispiel frühneuzeitlicher Texte, in: *Driemaandelijke Bladen voor taal en volksleven in het oosten van Nederland* 41, S. 149–186.
- Macha, Jürgen [u. a.] (Hgg.) 2005: *Deutsche Kanzleisprache in Hexenverhörprotokollen der Frühen Neuzeit*, 2 Bde., Berlin u. New York.
- Moulin, Claudine 1990: Der Majuskelgebrauch in Luthers deutschen Briefen (1517–1546), Heidelberg (Germanische Bibliothek, Neue Folge, Reihe 3, Untersuchungen und Einzeldarstellungen 13).
- Moulin, Claudine 2004: Das morphematische Prinzip bei den Grammatikern des 16. und 17. Jahrhunderts, in: *Sprachwissenschaft* 29, S. 33–73.
- Nerius, Dieter 2003: Graphematische Entwicklungstendenzen in der Geschichte des Deutschen, in: Werner Besch [u. a.] (Hgg.): *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*, Bd. 3, Berlin (Handbücher zur Sprach- und Kommunikationswissenschaft), S. 2461–2472.
- Nottbusch, Guido 2008: *Handschriftliche Sprachproduktion. Sprachstrukturelle und ontogenetische Aspekte*, Tübingen.
- Rezec, Oliver 2009: *Zur Struktur des deutschen Schriftsystems*, Diss. München.
- Reichmann, Oskar u. Klaus-Peter Wegera (Hgg.) 1993: *Frühneuhochdeutsche Grammatik*, Tübingen (Sammlung kurzer Grammatiken germanischer Dialekte A, Hauptreihe 12).
- Schutzzeichel, Marc u. Renata Szczepaniak 2015: Die Durchsetzung der satzinternen Großschreibung in Norddeutschland am Beispiel der Hexenverhörprotokolle, in: Markus Hundt u. Alexander Lasch (Hgg.): *Deutsch im Norden. Varietäten des norddeutschen Raumes*, Berlin u. Boston (Jahrbuch für Germanistische Sprachgeschichte 6), S. 151–167.
- Topalović, Elvira 2003: *Sprachwahl – Textsorte – Dialogstruktur. Zu Verhörprotokollen aus Hexenprozessen des 17. Jahrhunderts*, Trier.
- Weber, Walter Rudolf 1958: *Das Aufkommen der Substantivgrösschreibung im Deutschen. Ein historisch-kritischer Versuch*, München.
- Worch, Jan-Hendrik u. Björn Gottfried 2014: Choosing Shape Features by means of Genetic Algorithms for Glyph-Clustering of Historical Documents, in: *International Journal of Computer Applications* 102/3, S. 1–6.