

**Thomas Bliesener, Lena Deyerling,
Arne Dreißigacker, Isabel Henningsmeier,
Merten Neumann, Jonas Schemmel,
Carl Philipp Schröder und Laura Treskow (Hrsg.)**

Kriminalität und Kriminologie im Zeitalter der Digitalisierung

Neue Kriminologische Schriftenreihe

Forum Verlag Godesberg

Thomas Bliesener, Lena Deyerling, Arne Dreißigacker, Isabel Henningsmeier, Merten Neumann, Jonas Schemmel, Carl Philipp Schröder und Laura Treskow (Hrsg.)

Kriminalität und Kriminologie im Zeitalter der Digitalisierung

Neue Kriminologische Schriftenreihe

Der Kriminologischen Gesellschaft e.V. Band 119

Herausgeber:

Kriminologische Gesellschaft (KrimG, vormals NKG)

Wissenschaftliche Vereinigung deutscher, österreichischer und schweizerischer Kriminologen e.V.

Kriminalität und Kriminologie im Zeitalter der Digitalisierung

Herausgegeben von:

Thomas Bliesener, Lena Deyerling, Arne Dreißigacker, Isabel Henningsmeier, Merten Neumann, Jonas Schemmel, Carl Philipp Schröder und Laura Treskow

Mit Beiträgen von:

Dirk Baier, Neele Bäker, Doris Bender, Barbara Bergmann, Lorenz Biberstein, Aaron Bielejewski, Nicole Bögelein, Patricia Bruns, Felix Butz, Lena Carl, Tamara Dangelmaier, Axel Dessecker, Joachim Dietzenschmidt, Rebecca Endtricht, Diego Farren, Lena Fecher, Jannik M. K. Fischer, Stephanie Fleischer, Katja Girschik, Friederike Gödecke, Moritz Golombek, Thomas Görgen, Daniel Haake, Alisa Hastedt, Maria-Anna Hirth, Katrin Höffler, Sabine Hohmann-Fricke, Amina Hoppe, Stefanie Horn, Edith Huber, Benedikt Iberl, Irina Jugl, Johannes Kaspar, Peter Kieseberg, Sonja King, Jörg Kinzig, Janosch Kleinschnittger, Jonas Knäble, Dominic Kudlacek, Katharina Leimbach, Eva Link, Friedrich Lösel, Paulina Lutz, Nora Markwalder, Paula Emily Matthies, Till Neuhaus, Stella Nüschen, Figen Özsöz, Bettina Pospisil, Tamina Preuß, Katharina Reisch, Josefine Röder, Michael Roth, Henri Ruotsalainen, Thilo Sauter, Jonas Schemmel, Sarah Schreier, Jessica Schütz-Wilke, Walter Seböck, Jens Struck, Stefan Suhling, Isabelle Thormann, Laura Treskow, Albert Treytl, Thaya Vester, Catharina Vogt, Daniel Wagner, Anna Katharina Weiterer, Rüdiger Wulf, Chiara Wüller, Ulrike Zähringer, Sabine Zeymer

Forum Verlag Godesberg GmbH
Mönchengladbach 2023

Bibliografische Informationen 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.

© 2023 Forum Verlag Godesberg GmbH, Mönchengladbach
Herstellung: BoD – Books on Demand, Norderstedt
Printed in Germany

ISSN: 2509-5056

ISBN: 978-3-96410-038-2 (Printausgabe)

ISBN: 978-3-96410-039-9 (Online-Ausgabe/PDF-Dokument)

Aus dem Inhaltsverzeichnis:

Kriminologische Methodologie

Reproduziert kriminologische (qualitative) Forschung Ungleichheitsstrukturen? Überlegungen zur Praxis der empirischen Forschung
Nicole Bögelein, Katharina Leimbach..... 605

Datenschutz, Forschungsethik und gute wissenschaftliche Praxis in der Kriminologie. Eine Initiative am Institut für Kriminologie der Universität Tübingen
Rüdiger Wulf..... 620

Forensische Kriminologie und Cognitive Science – mögliche Synergien idiographischer Einzelfallkriminologie und situativer Ansätze der Kognitionswissenschaft
Michael Roth..... 642

Forensische Linguistik: auf der Suche nach dem Idiolekt, u. a. digital
Isabelle Thormann..... 660

Forensische Linguistik: auf der Suche nach dem Idiolekt, u. a. digital

Isabelle Thormann

Gliederung

- | | |
|---|--|
| 1. Einleitung | 6.1 Typografie und Textstruktur, Textsorte/-funktion, Stil, Register |
| 2. Die Teildisziplinen der forensischen Linguistik | 6.2 Orthografie und Interpunktion |
| 3. Das Zauberwort „Idiolekt“ | 6.3 Syntax |
| 4. Voraussetzungen für erfolgreichen Textvergleich und belastbare Befunde | 6.4 Morphologie und Morphosyntax |
| 4.1 Authentizität der Vergleichstexte | 6.5 Lexik und Semantik |
| 4.2 Minimum an Textmenge | 7. Qualitative und quantitative Verfahren der Textanalyse |
| 4.3 Zeitnähe der Entstehung der zu vergleichenden Texte | 7.1 Es wird digital: quantitative Verfahren |
| 4.4 Fehler bzw. Normabweichungen | 7.2 Qualitative Verfahren |
| 5. Achtung, dezeptive Strategien | 7.2.1 Satzglieder |
| 6. Die linguistischen Untersuchungsbe-
reiche | 7.2.2 Hypotaxe und die Thormann'schen
Treppenstufen |
| | 7.3 Digitale Verfahren in der forensi-
schen Textanalyse |

1. Einleitung

Einer der wichtigen Bereiche in der forensischen Linguistik ist die Autorschaftserkennung. Hier werden speziell in Fällen anonymer Bedrohung aller Art, Cyber-Mobbing, bei Erpressungsversuchen, Ehrverletzungsdelikten usw. quantitative und qualitative Verfahren und deren Kombination eingesetzt, um den in einem inkriminierten Text und in hoffentlich vorliegenden Vergleichstexten enthaltenen Idiolekt eines Textverfassers¹, also den individuellen spezifischen Sprachgebrauch einschließlich eines bestimmten Wortschatzes und speziellen Ausdrucksweisen, herauszuarbeiten. Dabei sind etwaige dezeptive

¹ In diesem Text wird aus Gründen der besseren Lesbarkeit nur die männliche Form der personenbezogenen Substantive, Personal- und Possessivpronomen verwendet (generisches Maskulinum). Selbstverständlich sind alle Geschlechter gleichermaßen gemeint.

Strategien der Textverfasser zu berücksichtigen und zu erkennen. Digitales spielt in der forensischen Linguistik in zwei Hinsichten eine Rolle: Erstens wird Software eingesetzt, um Texte im Rahmen der quantitativen Analyseverfahren (auch „Stilometrie“) zu sezieren, zu beschreiben und zu vergleichen; zweitens liegen oft Texte vor, die „in der digitalen Welt für die digitale Welt“ erstellt wurden, also zu untersuchende Texte, die in den sozialen Medien, auf Bewertungsportalen, im Darknet – oder gar in Form von Social Bots – erstellt und dort gelesen werden.

In diesem Beitrag geht es um die Methoden mit typischen Beispielen, wobei gezeigt wird, dass im Bereich der Syntax, also des Satzbaus, viel Idiolektales zu finden ist, da sich Textverfasser hier besonders schlecht verstellen können. Außerdem gibt es einen Einblick in die Binnenstrukturierung, die Möglichkeiten und Methoden der forensischen Linguistik und u. a. die Unterschiede zwischen quantitativen und qualitativen Verfahren bei der forensischen Textanalyse.

2. Die Teildisziplinen der forensischen Linguistik

Eine mögliche und mir als sehr sinnvoll erscheinende Unterteilung der Teildisziplinen der forensischen Linguistik ist diese:

Sprache (primär Texte) als Streit- und Untersuchungsgegenstand

- Urheberschaftsfragen (Autorschaftserkennung, -bestimmung [ca. „Sprachprofiling“], Identifizieren des Idiolekts)
- Texte aller Art als Beweismittel vor Gericht, Authentizität (mutmaßlich gefälschte Verfügungen von Todes wegen/Testamente, [kodierte] Absprachen im Darknet, Vortäuschen einer Straftat, Love Scamming, Texte auf Beurteilungsplattformen, Kundenbewertungsportale)
- Sprache (primär Texte) als Medium kriminellen Handelns (Sprechakttheorie) wie Erpressung, Bedrohung, Ehrverletzungsdelikte (Beleidigung, üble Nachrede, Verleumdung), Bestechung(sversuch), Mobbing, Nötigung(sversuch), Hetz-/Schmutzkampagnen, Volksverhetzung, Anstiftung usw.
- Fälschungserkennung (u. a. Urkunden, Zeugnisse; Abschiedsbrief bei Suizid evtl. erzwungen, echtes Bekennerschreiben vs. Trittbrettfahrertexte)

- Erkennung dezeptiver Strategien (Verstellung, z. B. Deutsch als Muttersprache vs. Deutsch als Fremdsprache; Tarnung von Verfassern speziell bei Erpressung, Nötigung und anonymer Bedrohung)
- Plagiat-Erkennung (Literatur, Wissenschaft, akademische Arbeiten)
- Hand- und Maschinschrift-Untersuchung (kriminaltechnische Untersuchungen [Schreibsysteme, MIC/ [Machine Identification Code], „tracking dots“ bei Farbaudrucken), Forensische Handschriftenanalyse (ca. Graphologie)
- Glaubhaftigkeitsbeurteilung (speziell von Zeugenaussagen [Undeutsch-Hypothese]; merkmalsorientierte Inhaltsanalyse, international CBCA [criteria-based content analysis], Lügnererkennung, Erkennen von Befangenheit), Glaubwürdigkeitsbeurteilung (Kriminalpsychologie, Psycholinguistik, linguistische Psychologie)
- Psycholinguistik (forensische Psychologie, kognitive Linguistik, Sprach-/Textpsychologie, Psyketing [Psychologie u. Marketing]) für die Ursachen bestimmter Ausdrucksweisen
- Texte im Marken- und Patentrecht
- Phonetik (Stimmerkennung, -Vergleich, Sprecherprofile, Transkription)

Verständlichkeitsbeurteilung, Auslegung

- Vereinbarungen, Verträge, Definitionen, Versicherungskonditionen, Anweisungen, Verbraucherinformationen, Instruktionen aller Art, Gebrauchsanweisung, Beipackzettel [evtl. irreführend, missverständlich]
- Beleidigende [u. a. Anspielungen enthaltende] veröffentlichte Texte [auch Rapsong-Texte], Plakate
- Behördenanschriften (z. B. Bescheid), Belehrung, Gesetzestexte, Protokolle (Vernehmungsprotokolle: Beschuldigten-, Zeugenvernehmung), Transkriptionen (z. B. von Abhörprotokollen)
- Beschilderung (evtl. irreführend, missverständlich)
- Kurzmitteilungen aller Art (mit besonderer Berücksichtigung von Medien-Textsorten/Social Media wie Tweets, Facebook-, Instagram-, WhatsApp-, Tiktok-Nachrichten, Chatroom-Inhalten, Memes), Erkennen von Social Bots

Sprache (inkl. Texte) und ihre Wirkung, sprachliches Verhalten bei der Polizei und im Gerichtssaal

- Sprechhandlungen durch die Exekutive (Polizei), Judikative (Gericht) und der Organe der Rechtspflege; von Sachverständigen (Darstellung, Schlüssigkeit, Beweiskraft von Aussagen), Befehl, Anordnung, Beweisbeschluss
- Behördenanschriften (z. B. Bescheid), Gefährderansprache (und -ansprechen), Belehrung, Vernehmung, Geständnis, Zeugenaussage, Rechtsbeugung, Einschüchterung, "Sanktionsschere" (Rechtswidrigkeit von Verständigung im Strafverfahren), innere Schlüssigkeit und möglicher Widerspruch zum Akteninhalt, Transparenz- und Mitteilungsgebote, das "letzte Wort" des Angeklagten vor der Urteilsfindung
- Besonderheiten bestimmter Sprecher/Schreiber (traumatisierte Personen/Opfer, Kinder, Menschen mit kognitiven Einschränkungen im weitesten Sinne (nach Schlaganfall usw.), Notwendigkeit „Leichter (oder einfacher) Sprache“
- Verdolmetschung, Übersetzung

Sprache als Beschreibungs- und Schulungsgegenstand

- Schulung von Nicht-Juristen (Sachverständigen, Dolmetschern und Übersetzern); „Rechtssprache“ (Umfang, Didaktik usw.)

In dem Bereich, um den es hier primär geht, der Autorschaftserkennung, werden quantitative und qualitative Verfahren eingesetzt, um einen inkriminierten Text mit vorliegenden Vergleichstexten auf Gemeinsamkeiten zu untersuchen, wobei auch etwaige dezepive Strategien der Textverfasser zu berücksichtigen sind.

3. Das Zauberwort „Idiolekt“

Bei dem bei Urheberschaftsfragen zu identifizierenden Idiolekt eines Textverfassers geht es um den einem Menschen eigenen Sprachgebrauch mit einem bestimmten Wortschatz und spezifischen, typisch erkennbaren Ausdrucksweisen, welche von Sozialisation, Schul- und Ausbildung, Freundeskreis, Beruf, Fachwissen, Wohnort[en] und Regiolekt[en] beeinflusst worden sind. Dabei ist wichtig:

„Nicht, wie ‚ungewöhnlich‘ ein sprachlicher Ausdruck schlechthin ist, sondern wie konsistent und typisch er für den Sprachgebrauch eines Sprechers als (allgemein) ungewöhnliche sprachliche Handhabung einer Situation auftritt, ist von Interesse.“²

Bei der Wahl jedes einzelnen Wortes, jeder Einleitung, Anrede, Höflichkeitsform, Tempusform, Redewendung, jedes Satzzeichens, jeder Groß- und Kleinschreibung, jeder Aktiv- oder Passivkonstruktion, jeder Abfolge von Haupt- und Nebensätzen (Hypotaxe) usw. usf. trifft ein Verfasser (unbewusst) Entscheidungen.³

Bei allen festgestellten sprachlichen Phänomenen, die auf solchen Entscheidungen beruhen, wird dann gefragt:

„Ist diese Ausdrucksweise idiolektal?“

In anderen Worten: Ist sie kennzeichnend für den spezifischen Sprachgebrauch eines (mutmaßlichen) Verfassers?

Wird sie von ihm unbewusst verwendet, was bedeutet, dass er sie nicht vorstellen kann (konnte)?

4. Voraussetzungen für erfolgreichen Textvergleich und belastbare Befunde

Es müssen primär fünf Bedingungen erfüllt sein, damit die erarbeiteten Ergebnisse eines Textvergleichs für die Autorschaftserkennung und -bestimmung belastbar sind. Kurzgefasst sind es diese Bedingungen:

1. Die Authentizität der Vergleichstexte muss gesichert sein.
2. Es muss ein Minimum an Textmenge vorliegen.
3. Zeitnähe der Entstehung der zu vergleichenden Texte muss gegeben sein.
4. Textsortenkompatibilität muss gegeben sein.
5. Fehler bzw. Normabweichungen (Konvertierungsfehler, Kompetenz- und Performanzfehler) sind zu differenzieren.

² Kniffka (1981), S. 591 f.

³ Siehe dazu auch Rottler/Martin (2020), S. 38.

4.1 Authentizität der Vergleichstexte

Die sprachlichen Besonderheiten von jeweils zwei Texten werden verglichen, um festzustellen, ob sie von derselben Person verfasst wurden. Es liegt auf der Hand, dass ganz sicher sein muss, dass der/die Vergleichstext(e) von der Person stammt bzw. stammen, die verdächtigt wird, den inkriminierten Text verfasst zu haben.

Es muss klar sein, welche Art von Aufgabe bzw. Frage vorliegt:

- Es liegen ein inkriminierter Text und ein Vergleichstext vor, wobei bei dem Vergleichstext gesichert ist, dass er von der Person verfasst wurde, die verdächtigt wird, den inkriminierten Text verfasst zu haben. Es ist festzustellen, ob beide Texte von derselben Person verfasst wurden.
Das Ergebnis kann dann lauten: „Der inkriminierte Text wurde mit hoher Wahrscheinlichkeit von dem Verfasser des Vergleichstextes verfasst.“
- Es liegen ein inkriminierter Text und ein anderer Text (bzw. andere Texte) vor. Es ist festzustellen, ob beide Texte von derselben Person verfasst wurden.
Das Ergebnis kann dann lauten: „Der inkriminierte Text wurde mit hoher Wahrscheinlichkeit von dem Verfasser des anderen Textes verfasst.“
Da jedoch nicht sicher bestimmt werden kann, wer den anderen Text verfasst hat, betrifft diese Ungewissheit auch die Person, die den inkriminierten Text verfasst hat.
- Es liegen mehrere Texte – z. B. mehrere inkriminierte und/oder mehrere Vergleichstexte vor. Es ist festzustellen, ob alle Texte von einer Person oder von mehreren verschiedenen Personen verfasst wurden.
Das Ergebnis kann dann lauten: „Die Texte wurden mit hoher Wahrscheinlichkeit von derselben Person verfasst“ oder „Die Texte wurden mit hoher Wahrscheinlichkeit von mehreren verschiedenen Personen verfasst“.

Es kommt oft vor, dass linguistische Gutachten angefordert werden, sich die Parteien jedoch darum streiten, welche Texte als Vergleichstexte heranzuziehen sind.

Ob gesichert ist, dass ein bestimmter Vergleichstext von einer bestimmten Person verfasst wurde, hängt von vielen Faktoren ab. Bei der Beschaffung solcher Texte kann auch „geschummelt“ worden sein, ein solcher Text kann von jemandem verändert worden sein, der etwa ein Interesse daran hat, dass keine Übereinstimmung des Idiolekts des inkriminierten Textes mit dem des Vergleichstextes (bzw. mehrerer Vergleichstexte) festgestellt wird und

folglich dieselbe Person als Verfasser gemutmaßt wird. Als sicher sind solche Texte anzusehen, die tatsächlich verschickt wurden und beim Empfänger ankamen und bei denen der Verfasser außerdem vor der Untersuchung seine Autorschaft bestätigt.

Es ist wegen des Zeit- und Geldaufwands ausgesprochen ärgerlich, wenn einem Linguisten gefälschte Texte zum Vergleich zur Verfügung gestellt wurden und die Fälschung später festgestellt wird. Falls eine solche Fälschung vorliegt und sie für ein linguistisches Gutachten herangezogen wurde, die Fälschung jedoch nie aufgedeckt wird, kann das gravierende Folgen haben.

Ein authentischer Text ist also ein Text, der von dem tatsächlichen Verfasser stammt und nicht gefälscht oder verändert wurde und wenn eine entsprechende Bestätigung des tatsächlichen Verfassers vorliegt.

Es gibt auch schwierige Fälle, bei denen eine Person, die beispielsweise nicht schreiben kann (weil sie verletzt oder behindert ist), einer anderen Person den Text diktiert. In einem solchen Fall ist zu beachten, dass sich bei dem Text die Orthografiekompetenz der schreibenden und nicht der diktierenden Person zeigt. Kritisch sind auch Texte, die von einer anderen Person als dem Verfasser „Korrektur gelesen“ oder in anderer Weise bearbeitet wurden. Es gibt auch Texte, die von Verfassergruppen erstellt werden. Solche Texte sind meist als Vergleichstexte unbrauchbar.

4.2 Minimum an Textmenge

Da nur solche Auffälligkeiten gewertet werden dürfen, die mehrmals bzw. oft bzw. im besten Fall „systematisch“ (d. h. sehr oft) vorkommen (also nicht einzelne Auffälligkeiten bzw. solche, die nicht hochfrequent sind), ist es sehr wichtig, dass die zu untersuchenden Texte einen gewissen Minimalumfang haben, denn in kurzen Texten finden sich keine bzw. nicht genügend wiederholende sprachliche Besonderheiten. Wenn ein Verfasser z. B. ein substantiviertes (und folglich großzuschreibendes) Verb zweimal klein schreibt und/oder „das“ statt „dass“ schreibt und/oder ein bestimmtes Adjektiv, das normalerweise nicht häufig vorkommt, verwendet, ist das zwar interessant; es dürfen jedoch nur solche Auffälligkeiten gewertet werden, die mehrmals bzw. besser oft bzw. „hochfrequent“ bzw. „systematisch“ vorkommen.

Mehr als ein Drittel der Tatschriften (35 %), die im BKA ausgewertet und archiviert werden, ist weniger als 100 Wörter lang, weitere 29 % sind kürzer als 200 Wörter.⁴

4.3 Zeitnähe der Entstehung der zu vergleichenden Texte

Der Mensch verändert sich im Laufe seines Lebens, und auch seine Ausdrucksweise verändert sich. Das wird besonders stark von Ausbildung aller Art und von den Menschen bestimmt, mit denen er kommuniziert. Daher ist es wichtig, dass die Texte eine zeitliche Nähe haben. Bei einem 50-Jährigen ist es nicht gravierend einschränkend, wenn ein großer Zeitraum von beispielsweise zehn Jahren zwischen der Entstehungszeit der zu vergleichenden Texte liegt, bei einem 30-Jährigen wäre das allerdings bedenklich.

4.4 Textsortenkompatibilität

Bei dem Vergleich von Texten in der forensischen Linguistik ist es wichtig, dass die Texte möglichst derselben Textsorte angehören, dass also nicht ein Drohschreiben mit einem Kochrezept verglichen wird. In der Textsortenforschung gibt es verschiedene Theorien, die jedoch alle gemein haben, dass nach Form und Gebrauch bzw. Absicht des Textes klassifiziert wird. Ich finde folgende, von mir selbst verwendete Klassifizierung sinnvoll.

Textsorten

- Informativ-neutrale Texte
 - Konstative Texte
 - Reportative Texte
 - Prädiktative Texte
 - Deklarative Texte
 - Deskriptive Texte
 - Expositive Texte
 - Introduktive Texte
 - Textstrukturierende Texte
 - Konklusive Texte
 - Referentielle Texte
 - Komparative Texte
- Interaktiv-kommunikative Texte
 - Konventionell-konfirmative, affirmative und resentative Texte

⁴ Fobbe (2013), S. 9.

- Kommunikativ-reaktive Texte
- Konventionell-behaviorale Texte
- Sentativ-kommentative Texte
- Promissive Texte
- Kontentative Texte
- Judikative Texte
- Initiativ-appellative Texte
 - Deklative Texte
 - Laudative und improbativ Texte
 - Konsultative Texte
 - Revokative Texte
 - Assertative Texte
 - Suggestive Texte
 - Suggestiv-konventionelle Texte
 - Suggestiv-persuasive Texte
 - Direktive Texte
 - Mandative Texte
 - Demandative Texte
 - Exerzitive Texte

Quelle: eigene Darstellung

Für die Unterteilung der Textsorten in der forensischen Linguistik und speziell für Urheberfragen ist das Kriterium der Intention, also der Absicht des Verfassers, bzw. auch der Intension, also dem inneren Bestreben, einen Text herzustellen, wichtig.⁵

Jemand drückt sich anders aus, wenn er einen offiziellen Bericht für einen Vorgesetzten schreibt, als wenn er eine private Mitteilung an seine Nachbarin verfasst. Er schreibt auch anders, wenn er Vorwürfe macht, als wenn er einen Dankesbrief verfasst. Daher ist zunächst u. a. darauf zu achten, was miteinander verglichen wird, dass also die Textsorte, das Register bzw. der Grad der Förmlichkeit, der Stil und die Absicht bzw. Funktion des Textes übereinstimmen.

Texte unterscheiden sich auch je nachdem, ob der Verfasser nervös oder in Eile, betrunken, ängstlich, gelassen oder verliebt ist. Umstände dieser Art sind jedoch oft nicht bekannt. Literarische Textsorten sind in der forensischen

⁵ Intention: Absicht, Vorhaben, Plan; damit verwandt ist das Wort „Tendenz“ (etwas auf etwas hin Gerichtetes); der Fokus ist ein nach außen gerichtetes Ziel.

Intension: Eifer, innerer Antrieb, Bestreben, Anspannung, etwa sogar „Berufung“; Gegenteil von „Extension“. Das lateinische Wort *intensio* bedeutet „Anspannung“; das sieht man auch an dem englischen Wort „tension“ ([An-]Spannung) und dem Adjektiv „intensiv“.

Linguistik lediglich bei Plagiats-Untersuchungen von Interesse, um die es jedoch hier nicht geht.

4.5 Fehler bzw. Normabweichungen

Bei den für eine Fehleranalyse zu betrachtenden Fehlern ist zunächst zu bemerken, dass die Bezeichnung „Normabweichung“ sinnvoller und weniger hart ist, das Wort „Fehler“ ist u. a. in dem Kompositum „Fehleranalyse“ leichter handhabbar und daher gebräuchlich.

Bei den für die forensische Textanalyse betrachteten Fehlern sind folgende Begriffe zu unterscheiden:

- Konvertierungs- und Transkriptionsfehler
- Kompetenz- und Performanzfehler

4.5.1 Konvertierungs- und Transkriptionsfehler

In den meisten Fällen liegen Texte im PDF-Format als Ergebnis eines Scans vor, oder sie wurden, sofern es sich um handschriftlich verfasste Texte handelt, abgetippt (transkribiert). In beiden Fällen ist darauf zu achten, dass keine Fehler bzw. Normabweichungen korrigiert wurden, sondern dass das Transkriptionsergebnis Normabweichungen aller Art wie auch alle Leerzeichen enthält, was später bei der Betrachtung der Getrennt- und Zusammenschreibung wichtig sein kann. Es kommt auffällig oft vor, dass Hilfskräfte in Sekretariaten beauftragt werden, einen handschriftlich verfassten Text „abzutippen“ und nicht auf Feinheiten wie auf Abstände und unklare Buchstabengrößen (wegen Unsicherheit des Verfassers bzgl. der Groß- und Kleinschreibung) achten. Es kommt auch vor, dass ein Text nach einer Umwandlung vom PDF- in das Word-Format Konvertierungsfehler enthält, weil z. B. die „Vorlage“ unscharf war. Dann kann es beim Textvergleich zu falschen Ergebnissen kommen.

4.5.2 Kompetenz- und Performanzfehler

Die Möglichkeit von Fehlinterpretationen, die sich aus Konvertierungsfehlern und aus einer Fehlinterpretation von Kompetenz- bzw. Performanzfehlern ergeben können, ist zu minimieren.

Kompetenzfehler sind solche Fehler, bei denen die Norm nicht gelernt wurde und nicht beherrscht wird. Sie ergeben sich systematisch aus dem

zugrundeliegenden Wissen eines Textverfassers. Sie werden von ihm in vergleichbaren Situationen voraussagbar immer wieder produziert.

Performanzfehler hingegen sind Flüchtigkeitsfehler, die dem Verfasser unterlaufen, weil er einen Bestandteil eines Satzes nach dem Aufschreiben einer ersten Formulierung geändert und bestimmte Kasus-/Genus-/Numerus-Änderungen vorzunehmen vergessen hat, oder die ihm als Folge mangelnder Konzentration unterlaufen, oft bedingt durch Eile, Stress, Nervosität, Übermüdung usw., also „zufällige“ Fehler.

Typische Beispiele für Performanzfehler sind „in der Nachbar“ (statt „in der Nachbarn“), „nicht fünf, sondern sechs Stunden“ (statt „sondern“). Oft sind solche Tippfehler erkennbar, weil der falsche Buchstabe auf der Tastatur nahe bei dem richtigen liegt, oder es handelt sich um ausgelassene oder zusätzliche Buchstaben oder um Buchstabendreher.

Typische Kompetenzfehler sind „Russland ist eine Förderung.“, „Die Miene von dem Stift ist leer.“, „Er hat eine eidesstaatliche Versicherung abgegeben.“

5. Achtung, dezeptive Strategien

Es ist damit zu rechnen, dass der Verfasser von sich ablenken will und Verstellungsversuche unternimmt. Hier stellen sich die Fragen:

- Sind dezeptive Strategien erkennbar?
 - z. B. Versucht der Verfasser zu schreiben, als hätte er Deutsch als Fremdsprache gelernt („DaF“/Deutsch als Fremdsprache)?
 - z. B. Versucht der Verfasser, den Eindruck zu machen, als gehöre er einer anderen sozialen Schicht/Bildungsschicht an?
 - z. B. Wurde der Text systematisch manipuliert?
- Wurde abgeschrieben (z. B. von vorhandenen bzw. öffentlich zugänglichen Texten)?⁶

Bei den Verstellungsstrategien ist zunächst grundsätzlich zu unterscheiden:

1. (Lediglich) Verschleierung bzw. Verheimlichung der eigenen Identität (kein Absender, keine Unterschrift)

⁶ Fobbe (2011), S. 52.

2. Imitation anderer Personen bzw. Personengruppen (fingierte Autorschaft, oft mit erfundenem Absendernamen, „auktoriale Inszenierung“)⁷

Zu den dezeptiven Strategien gehört auch das Thema „Plagiat“, um das es hier nicht geht.

Was wird bei den dezeptiven Strategien verändert und was nicht?

Bei dem, was vom Verfasser nicht verändert werden kann, handelt es sich um „Idiolektales“, und das ist das Interessante.

Ein sprachlich ungeschulter Verfasser kann bewusst orthografische Fehler einbringen, er kann Flexionsendungen ändern, verstümmeln oder tilgen, er kann für das Verständnis verzichtbare Funktionswörter (z. B. Präpositionen, Artikel) ändern oder entfernen (z. B. sämtliche Artikel in „die“ ändern). Veränderungen grammatischer Strukturen hingegen, insbesondere des Satzbaus, kann er nicht verändern bzw. wird dies derart stümperhaft tun, dass die Manipulation leicht erkannt werden kann. Außerdem ist die konsistente Veränderung des Sprachstils in einem längeren Text sehr anstrengend und wird selten durchgehalten. Erfahrungsgemäß werden Eingangspassagen eines Textes massiv manipuliert, im weiteren Textverlauf oder aber in Folgebriefen einer Briefserie nimmt der Verstellungsgrad dann ab, und der Verfasser verfällt in seine gewohnten sprachlichen Routinen.

Eine erste Orientierung in der Bewertung eines fraglichen Textes hinsichtlich der Frage, ob ein Verstellungsversuch vorliegt oder nicht, bietet die Beantwortung der folgenden Fragen:

1. Beschränkt sich die Fehlerhaftigkeit weitgehend auf die Ebenen der Orthografie und der Flexionsmorphologie (Wortformenbildung)?
2. Treten korrekte Vorkommen neben inkorrekten Vorkommen einer sprachlichen Kategorie auf?
3. Nimmt die Fehlerhaftigkeit im Verlauf des Textes ab?
4. Werden idiomatische Ausdrücke korrekt eingesetzt?
5. Sind Wortverwendung/Beherrschung der Idiomatik und Fehlerhaftigkeit auf oberflächenstruktureller Ebene vereinbar?
6. Gibt es inkorrekte Wortverwendungen?
7. Wird das Verständnis des Textes gefährdet?⁸

⁷ Ausführliche Erläuterungen zu Verstellungsstrategien finden sich in Hessler (2023), speziell in den Kapiteln 13, 15, 16 und 17.

⁸ Vgl. Dorn (2009), S. 81 f.

6. Die linguistischen Untersuchungsbereiche

Bei einem Textvergleich sollten möglichst alle der im Folgenden aufgeführten und beschriebenen fünf Untersuchungsbereiche betrachtet werden:

1. Textstruktur und Typografie
2. Orthografie und Interpunktion
3. Lexik und Semantik
4. Morphologie u. Morphosyntax
5. Syntax

Speziell im Bereich der Syntax, also des Satzbaus, ist viel Idiolektales zu finden, da sich Textverfasser hier besonders schlecht verstellen können. Bei Kurznachrichten aller Art (Tweets, SMS, bei Texten auf Bewertungsportalen, im Darknet usw.) ist nicht viel Syntax vorhanden, jedoch Kollokationen, Kookkurrenzen und idiolektale Ausdrücke, für deren Untersuchung sich quantitative Verfahren eignen.

6.1 Typografie und Textstruktur, Textsorte/-funktion, Stil, Register

Dazu gehören das Erscheinungsbild (auch „Phänotyp“), Textgenre, Inhalt des Textes und Absicht des Verfassers (z. B. *initiativ-appellativ* [Drohbrief, Erpressungsversuch, Koch-/Arzt-Rezept], *informativ-neutral* [z. B. Testament, Reise-, Sport-/Wetter-Bericht, Zeugnis], *interaktiv-kommunikativ* [z. B. Anfrage, Privatbrief [Verbalstil], Behörden-, Verwaltungssprachstil (Nominalstil, vs. Verbalstil in Privatbriefen); die Satzgrenze überschreitende („transphrastische“), textlinguistische, textkonstituierende Trennungs- bzw. Zusammenhangskennzeichnung von semantischen und syntaktischen Beziehungen (Kohärenzen, Kohäsionen);

Textkonstitution (-struktur, -aufbau): Brief, E-Mail, SMS, Länge/Zeichenanzahl (Text, Satz, Wort), Wort- u. Satzanzahl, Satzlängen (Wort- u. Zeichenanzahl pro Satz);

Typographisches: Besonderheiten: Formatierung, Absätze, Ränder, Satz (Block-, Flattersatz), Abstände, Einrückungen, Leerzeilen, Leerzeichen, Tabs, Font(s) u. Schriftgröße(n), Silbentrennung, Hervorhebungen, Grußformeln, typografische Besonderheiten bei Datumsangaben u. Abkürzungen (nicht lexikalisch), Aufzählungen, Satzzeichen;

Inhaltliches: Grundaussage und -absicht (Information, Drohung, ...), Gesamtmitteilung des Textes, inhaltliche Gliederung des Textes, Einzelaussagen der Absätze, Thema-Rhema-Gliederung (TRG), Brüche im Textaufbau,

Unklares/Unverständliches, Unlogisches, mangelnde Konsistenz, Unstimmigkeiten aller Art;

Hinweise auf den/die Verfasser: Einzelperson vs. Gruppe (evtl. inkonsistente Schreibung, Interpunktion, Normabweichungen aller Art), evtl. Hinweise auf Alter und/oder soziale/ethnische Zugehörigkeit (primär lexikal. Besonderheiten, evtl. Verstellung), auf Geschlecht (evtl. „Genderlekt“), Fachwissen (Fachausdrücke und Fakten), Bildungsstand, Sprachstand (Deutsch als Fremdsprache oder Zweitsprache gelernt, pidgin-sprachliche Ausdrucksweisen, z. B. Ich nix Arbeit statt Ich bin arbeitslos), Ghetto-slang bzw. Kiezdeutsch (lexikalisch u. [morpho-]syntaktisch, z. B. Musstu gucken), Einflüsse (z. B. Ethnolekt, Dialekt, Regiolekt, Funktiolekt), interlinguale Einflüsse (oder Interferenzen), als Textbausteine verwendete Phrasen, Register (der Funktion des Textes angepasste Ausdrucksweisen)

6.2 Orthografie und Interpunktion (ohne „dass“/„das“ und Komma)

Dazu gehören: GKS (Groß-/Kleinschreibung [u. a. bedeutungsdifferenzierende Höflichkeitsform, sogen. „Honorifikum“]), GZS (Getrennt-/Zusammenschreibung, u. a. Komposita), alte vs. neue Rechtschreibung, Phonem-Graphem-Zuordnung (u. a. Dehnungs-h [Strähne], ß/ss/s, Schreibung von Fremdwörtern und Affixen [prä/pre-]), Zeichensetzung bzw. Verwendung von Satzzeichen (zur Gliederung von Sinneinheiten im Satz und Text): Punkt, Fragezeichen, Ausrufezeichen, Anführungszeichen, Semikolon, Doppelpunkt, Leerzeichen; Wortzeichen (Apostroph/Auslassungszeichen, Bindestrich/Trennstrich [u. a. beim Zeilenumbruch], Auslassungspunkte), Kennzeichnung von Einschüben („Parenthesen“ [z. B. mit Gedankenstrichen oder Klammern]); evtl. mangelnde Kohärenz mit der Aussage.

6.3 Syntax (einschließlich „dass“/„das“ und Komma)

Hier geht es um Satzteile, Satzbau (Hypo-/Parataxe, Nebensatzarten), Satzanfänge (mit welchem Satzteil), Kommata, Besetzung der Satzgliedpositionen (Adverbial als Präpositionalphrase oder als Nebensatz), Realisierung von Satzteilen (z. B. Objekt als Nominalphrase oder als Nebensatz), Art und Anzahl der Unterordnungs- bzw. Verschachtelungsebenen (Haupt- und Nebensätze und Kommata oder Linksattributionen), Verb-/Satzklammer (evtl. häufige Ausklammerung bzw. „Nachfeldbesetzung“), Topikalisierung (auch „Vorfeldbesetzung“), freier Dativ, Nominal- vs. Verbalstil.

6.4 Morphologie und Morphosyntax

Dazu gehören: Konjugation (Person, Numerus, Tempus, Modus, Genus verbi); Deklination (Kasus, Numerus, Genus); Kongruenz (Subjekt-Prädikat- und Kasus-Numerus-Genus-Kongruenz) und Kohärenz in Sätzen (Pronomina, Attribute)

6.5 Lexik und Semantik

Es geht um Wortarten (eventuell auffällig oft verwendete Wortart, z. B. Adverbien), Attribute (Arten: eher Einzelwörter in Form von Adjektiven oder Partizipien oder komplexe Linksattributionen [siehe Syntax]), Verwendung von Idiomen [für ein Individuum typische Ausdrucksweise], Metaphern [bildliche Ausdrucksweise, z. B. der Zahn der Zeit], Phraseologismen [zu fester Form verwachsene Folge von Lexemen, z. B. etwas auf dem Schirm haben], Floskeln [nichtssagende Redensarten, z. B. quasi], Partikeln, Modewörtern; Schimpfwörtern; FVG/Funktionsverbgefügen [z. B. zur Anwendung kommen], Abstrakta [z. B. Hass], Tautologien [Doppel-Aussagen, z. B. nie und nimmer], Pleonasmen [mehrfach gleiche, überflüssige Aussagen ohne Informationsgewinn, z. B. Fußpedal], Verb-Adverb-Gefügen [z. B. beiseite schaffen], Gendergerechte Sprache bzw. Darstellungsmethoden.

7. Quantitative und qualitative Verfahren der Textanalyse

Die in der forensischen Textanalyse hauptsächlich verwendeten Verfahren sind die Stilanalyse und die Fehleranalyse. Die Abweichung von einer oder mehreren Normen zu beschreiben, fällt vorrangig in den Bereich der Fehleranalyse, die daher nicht Teil der Stilanalyse ist. Bei der Stilanalyse ist wichtig: Es gibt Stil, der keine Normabweichung darstellt.⁹

7.1 Es wird digital: quantitative Verfahren (auch computergestützte Verfahren, Stilometrie)

Ein wertvolles Hilfsmittel bei den quantitativen Verfahren ist Konkordanz-Software. Sie zerlegt Texte und Sätze in Worthäufigkeitslisten, die

⁹ Siehe auch Fobbe (2011), S. 107 f.

beispielsweise so aussehen (IT steht für „inkriminierten Text“, VT steht für „Vergleichstext“):

IT	VT1	VT2	VT3	VT4	Neutrales Vergleichskorpus
und	und	und	und	und	und
ist	ich	ich	ich	ich	die
auch	die	zu	die	ist	ich
sie	von	die	sehr	das	der
die	wir	sehr	der	sehr	sie
nicht	der	das	ist	die	in
man	es	es	es	es	das
als	ist	dass	zu	mit	zu

und zeigt Kookkurrenzen auf, also Wörter, die auffällig häufig mit einem anderen Wort (oder mehreren anderen Wörtern) gemeinsam auftreten, und zwar als unmittelbarer linker Nachbar, als unmittelbarer rechter Nachbar oder gemeinsam in einem Satz. Das ergibt Auflistungen bestimmter Wörter mit ihren „Nachbarn“ wie hier beim Beispiel „Es/es“:

gt und **betrügt!** **Es** gibt so **viele** Feinde, dass u
 ositiv über **ihn!** **Es** gibt unzähligen **Ärger** mit E
 wir **denken** **es** ist an **der** Zeit, dass Sie als
 ihm **gekündigt.** **Es** ist nicht **die** erste Geliebte d
 ritte **eingeleitet.** **Es** wurde alles **über** einen länd

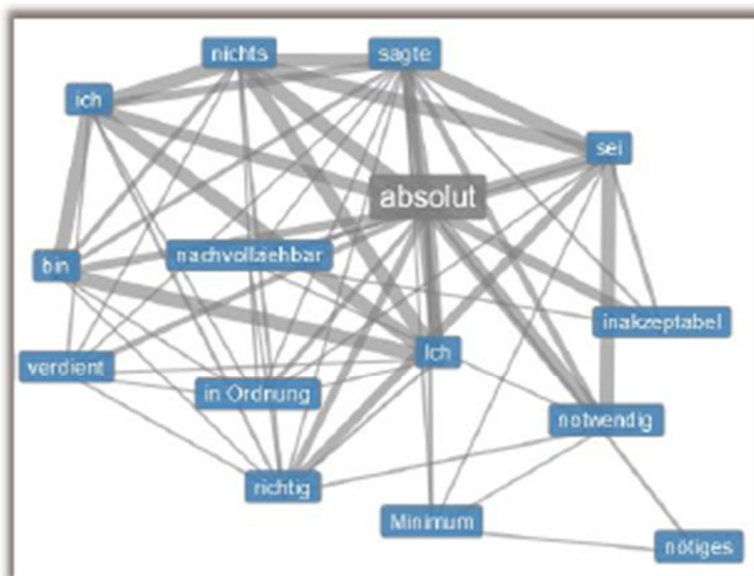
oder „das/dass“:

1 da sie als naiv bekannt **sind** und **das** **Leben** ohne Arbeit genie\xDFe
 2 : dass was man will! **Auch** wenn **das** **Leben** so sehr bequem ist! Aber
 3 ung begrabscht hat und **die** **sich** **das** **nicht** gefallen lassen hat. Nur p
 4 wuerde! Wir gehen **davon** **aus,** **dass** **Sie** von der Kuendigung von
 5 wir denken es ist an **der** **Zeit,** **dass** **Sie** als Ehefrau und Mutter de
 6 ie Mitarbeiter dabei **sind!** **Auch** **das** **Thema** Schwarzarbeit werden
 7 betruengt! Es gibt so **viele** **Feinde,** **dass** **uns** an seiner und auch ihrer
 8 ins Bett geht, ist **sicherlich** **nicht** **dass** **was** man will! Auch wenn das

Hier zum Lexem „gibt“, und es fällt auf, dass der linke Nachbar hier immer „Es/es“ lautet:

81	Natürlich wird nicht alles verfolgt werden, aber es gibt unzählig viele Möglichkeiten u ich mö	WhatsApp
82	nd daraus lassen sich keine Einzelfälle ableiten! Es gibt einen sehr guten Onkologen in Regen:	WhatsApp
83	u uns bestmöglich bzgl Bestrahlung beraten! Denn es gibt etliche Varianten der Bestrahlung u da	WhatsApp
84	♥ [07.10.18, 20:10:21] TA : Liebe Familie, es gibt schöne News...gestern u heute haben	WhatsApp
85	für ihn bedeuten, was auch gut ist...es gibt keine schlimmere Lage als Tumor im	WhatsApp
86	ähnlich wie beim letzten Befund im Jan, es gibt eine kleine Auffälligkeit...hierzu kann h	WhatsApp
87	Wetterbericht ja bescheiden mit Regen u kühl...es gibt ein Brunch beim Käfer...jedoch keine	WhatsApp
88	Rehkitz zum Mittagessen in der Sonne reserviert! Es gibt nur noch einen Tisch für 5...alles	WhatsApp
89	dazu spielt d Gewebe eine große Rolle...es gibt kaum Immunologen in MUC für Gliob	WhatsApp

Kookkurrenzen lassen sich auch mit Hilfe der sehr hilfreichen Seite <https://corpora.uni-leipzig.de> der Universität Leipzig feststellen, die u. a. Wortgraphen erstellt. Hier das Beispiel der „Nachbarn“ des Wortes „absolut“:



Wörter, die mit „absolut“ gemeinsam in einem Satz vorkommen, werden von der Universität Leipzig so aufgeführt:

Wörter gemeinsam im Satz
ist (4.913), notwendig (3.109), inakzeptabel (3.106), nichts (2.069), sei (1.854), ich (1.713), nötiges (1.671), richtig (1.616), sagte (1.604), bin (1.538), Minimum (1.525), ich (1.520), verdient (1.480), nachvollziehbar (1.432), in Ordnung (1.352), dass (1.171), unverständlich (1.155), Ordnung (1.043), überzeugt (968), b.2 (840), sicher (821), Stimmrech- (796), notwendige (789), Es (762), Kontakte (747), hatte (739), i.S.d. (726), finde (708), nicht (707), Ausübungszeit- (693), kein (663), richtige (639), nachvollziehen (638), Wertpapierleihe (637), N/A (612), sind (597), mich (579), wir (579), reduzieren (559), falsch (548), raum (546), Fälligkeit (526), WpHG (523), legitim (507), wäre (484), Stimm- (476), aber (475), Instruments (468), Entscheidung (461), sinnvoll (450), bezeichnete (432), unverantwortlich (432), gerechtfertigt (431), Nr. 1 (417), Verhalten (415), kontraproduktiv (399), keinen (398), uns (397), notwendiges (396), Sinn (393)

Die signifikanten Nachbarn von „absolut“ werden nach „Links“ und „Rechts“ so unterschieden:

Links	Rechts
ist (6.893), bin (1.651), sind (1.254), ich (959), es (786), war (741), das (725), ein (723), mir (573), sei (563), aber (550), mich (509), te (347), wäre (319), uns (311), natürlich (261), eine (261), als (232), wir (202), Es war (182), wirklich (181), hier (165), Sicht (161), für (154), Laufzeit (154), Es gibt (144), seien (132), also (128), dies (124), hat (119), daher (99), auch (92), stimmt (91), Situation (87), derzeit (78), Vorgehensweise (77), U-Trainern (76), Fördersystem (76), Zeitpunkt (75), waren (72), Budgetgrenze (71), m.E. (67), bisher (66), nicht (63), da (63), menschlich (62), Hygieneboards (61), ja (57), Moment (56), habe (55), Aufklärungsgremium (54), sich (53), Lage (52), doch (49), sowohl (48), Irina (47), somit (46), politisch (46), zwar (46), momentan (46)	

Links	Rechts
notwendig (5.922), inakzeptabel (5.076), nichts (4.888), nicht (4.479), richtig (3.626), verdient (2.996), nötiges (2.440), in Ordnung (2.407), kein (2.227), sicher (2.152), keine (2.080), unverständlich (1.882), nachvollziehbar (1.692), notwendige (1.663), überzeugt (1.544), keinen (1.286), richtige (1.131), sinnvoll (997), legitim (821), nachvollziehen (815), unverantwortlich (801), zufrieden (790), verständlich (771), falsch (771), kontraproduktiv (765), gerechtfertigt (711), notwendiges (687), berechtigt (675), okay (623), unzureichend (618), korrekt (605), falsche (593), Notwendige (573), Sinn (569), recht (546), sehenswert (518), begeistert (516), notwendigen (506), tabu (494), keine Ahnung (456), gar (448), lächerlich (435), unnötig (431), vertretbar (418), entscheidend (417), kein Problem (412), wichtig (405), top (389), unmöglich (375), verrückt (366), vorbildlich (353), ausreichend (341), ernst (338), fahruntüchtig (335), widerlich (333), realistisch (330), fantastisch (328), vermeidbar (325), beeindruckend (325), im Vordergrund (304)	

7.2 Qualitative Verfahren

Im Rahmen der qualitativen Verfahren werden Untersuchungen zum Erkennen sprachlicher Phänomene durchgeführt, die nicht mit Software erkannt werden können, speziell die Abfolge von Satzteilen bzw. Satzelementen und Hypotaxe, also die Anzahl und Abfolge der Subordinationsstufen bei Nebensätzen.

7.2.1 Satzelemente

Es werden – der Einfachheit halber – die fünf Satzteile Prädikat, Subjekt, Adverbial, Objekt, Prädikativ unterschieden. Außerdem gibt es das Attribut, welches Zusatzinformationen zu einem Wort im Satz liefert.

Textbeispiel:

Sie sind ein verdammter teurer Scheis-Rechtsanwalt und ein kleiner Arbeitnehmer der keine Chance gegen seinen Arbeitgeber hat soll fettichgemacht werden und Sie helfen ganich, weil Sie gehören auch zu denen Kapitalistenschweinen.

Die einfache Satzteil-Analyse würde so aussehen:

<i>Sie sind ein verdammter teurer Scheis-Rechtsanwalt und ein kleiner Arbeitnehmer der keine Chance gegen seinen Arbeitgeber hat soll fettichgemacht werden und Sie helfen ganich, weil Sie gehören auch zu denen Kapitalistenschweinen!</i>	S-P-Pk- UND-S-P- UND-S-P-A
--	----------------------------------

Die genauere, komplexere, detailliertere Analyse sieht so aus:

S(2.)-P(Kop)-Pk(mit2AttrLinks)-
UND-S(3.mit2AttrLinksUndRechts
RelSatz)-P1(Mod)-P2-P3(Passiv)
UND-S(2.)-P1-P2(Neg)-A(SatzAdvKaus)

Dann wird ein Text weiter untersucht, z. B. daraufhin, mit welchem Satzteil die Sätze beginnen. Dabei handelt es sich um eine unbewusste Entscheidung des Verfassers. Wird ein ganzer inkriminierter Text daraufhin untersucht und mit beispielsweise zwei anderen Texten verglichen, kann das Ergebnis so aussehen (IT steht für „inkriminierten Text“, VT steht für „Vergleichstext“):

Satzbeginn in allen Texten	IT		VT1		VT2	
Sätze insgesamt	22	%	13	%	37	%
Satzbeginn nicht mit dem S	3	13,6	8	61,5	8	21,6
Satz elliptisch, unübliche Satzgliedfolge, Satzkonstruktionsbruch, Aufforderung, Frage	3	13,6	0	0	2	5,4
Satzbeginn mit S	16	73	5	38,5	27	73
Satzbeginn mit A	3	14	7	54	5	13,5
Satzbeginn mit O	1	5	1	7,7	4	12
Satzbeginn mit Pk	0	0	0	0	1	2,7
Satzbeginn mit „Es“	3	13,6	0	0	1	2,7
Satzbeginn mit „Man“	0	0	1	7,7	0	0
Satzbeginn mit „Ich“	1	5	0	0	1	2,7

7.2.2 Hypotaxe und die Thormann'schen Treppenstufen

Dann wird der Text daraufhin untersucht, wie viele und welche Arten von Nebensätzen er enthält, und es werden Aussagen dieser Art gemacht:

- „x % aller Sätze des Textes enthalten keinen Nebensatz,
- x % enthalten einen Nebensatz,
- x % enthalten zwei Nebensätze und
- x % enthalten einen komplexen Nebensatz.“

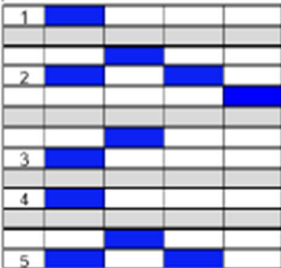
Was die Struktur von Haupt- und Nebensätzen betrifft, kann die Darstellung der hypotaktischen Elemente und Stufen so aussehen (wobei es hier nicht um das Lesen des Textes und die Bedeutung der Sätze geht, sondern lediglich um die Darstellung der Hypotaxe, also der Über- und Unterordnung der Sinneinheiten):

680

Isabelle Thormann

1	Wenn du dich nicht mehr eintrinst		
2	Ich habe dich mehrmals genannt in den letzten Monaten		
		Dass Du mir das Apartment überschreiben sollst	
3	Breit würdest du es besser		
4	Das habe ich auch dazu gesagt		
5	Ich bin keine Terroristin oder Altklerikerin		
6	Ich bin nur eine sehr entsäuerliche, sehr sehr selbstbewusste und bald geschiedene Ehefrau		
		Du nicht bereit ist	Ich auszugehen
7	Wenn du nicht bis zum 10. Mai die Vollmacht für überschreiben hast	Musst du dich nicht wundern	Wenn es dort brennt oder wenn deine geliebte Lina polizeimäßige Schutzmaßnahmen bekommt oder wenn deine Mutter stirbt und sich einen Ötenschenkelbruch zuzieht
8	Nimm das nicht auf die leichte Schulter		
9	Das rede ich dir		
10	Du weißt selber	Dass ich sehr konsequent sein kann	
11	Ich an deiner Stelle würde das Ernst nehmen		
12	Du brauchst auch gar nicht erst zu versuchen	Die Angelegenheit von Herrn Koch regeln zu lassen	
13	Der würde dir nicht helfen		
14	Das kannst du mir glauben		

Wenn der Text entfernt wird und nur noch die Subordination in Form von blauen Kästchen zu sehen ist, entstehen die „Thormann'schen Treppenstufen“, die z. B. so aussehen:

Inkrimierter Text (IT)	Vergleichstext 1 (VT1)	Vergleichstext 2 (VT2)
		
3 Sätze (60%) beginnen NICHT mit dem Hauptsatz. Sätze 1 und 4 haben keine Hypotaxe, sie bestehen aus nur einem Satz; einem Hauptsatz. Alle hypotaktisch konstruierten Sätze beginnen mit dem Nebensatz (!). Satz 2 besteht aus vier horizontalen Elementen, Satz 5 aus drei, Satz 3 aus zwei horizontalen Elementen. Bei Satz 5 ist das horizontale vierte Element entweder die Fortsetzung des Hauptsatzes oder ein neuer Hauptsatz. Ein Satz hat vier horizontale Elemente, ein Satz drei, die anderen zwei.	Alle Sätze beginnen mit dem Hauptsatz. Sätze 1 und 3 haben keine Hypotaxe, sie bestehen aus nur einem Satz; einem Hauptsatz. Satz 2 weist zunächst zwei Hauptsätze nacheinander auf und danach einen vom zweiten Hauptsatz abhängigen Nebensatz. Sätze 4 und 5 haben zwei Elemente; es folgt jeweils auf den Hauptsatz ein abhängiger Nebensatz. Ein Satz hat drei horizontale Elemente, alle anderen zwei.	Alle Sätze beginnen mit dem Hauptsatz. Sätze 1, 2 und 4 haben keine Hypotaxe, sie bestehen aus nur einem Satz; einem Hauptsatz. Nur Sätze 3 und 5 haben zwei Elemente, und zwar folgt jeweils auf den Hauptsatz ein abhängiger Nebensatz. Kein Satz hat mehr als zwei horizontale Elemente.
Der inkriminierte Text enthält mehrere Sätze, die mit einem Nebensatz beginnen, wohingegen die beiden Vergleichstexte keine solche Inzidenz enthalten.		

7.3 Digitale Verfahren in der forensischen Textanalyse

Es gibt vornehmlich zwei Fragen zu beantworten:

1. Gibt es auch für geringe inkriminierte Textmengen und/oder für solche Texte, die aus kurzen bzw. elliptischen Sätzen bestehen und die bei SMS und anderen Kurzmitteilungen in den sozialen Medien typisch sind, quantitative Methoden und statistische Auswertungen („forensische Stilometrie“), die quantifizierbare Merkmale verlässlich bestimmen können?
2. Wie weit kann man in Anbetracht der großen zu analysierenden Textmengen bei der forensischen Textanalyse auf den zeitaufwändigen Einsatz hochqualifizierter Linguisten verzichten?

Antwort (auf beide Fragen):

Das Untersuchen geringer Textmengen mit quantitativen Verfahren ist eine große Herausforderung, durch die *Jan Lars Svartvik*, ein schwedischer Linguistik-Professor, im Jahr 1967 bekannt wurde. Er unterzog (relativ kurze) Polizeiprotokolle einer forensischen Textanalyse und untersuchte – in dem speziellen Fall – primär die Wortwahl (also etwas Lexikalisches) und die Form der verwendeten Nebensätze (also etwas Grammatisch-Syntaktisches) und konnte inkonsistente Merkmale herausarbeiten, und sein Ergebnis war gerichtlich belastbar. Es führte dazu, dass der Verurteilte postum freigesprochen wurde.¹⁰ Für die von *Svartvik* verwendeten Verfahren wie auch für die in diesem Beitrag dargestellten Verfahren kann Software eingesetzt werden, allerdings immer nach entsprechender Auswahl der sprachlichen Merkmale durch einen hochqualifizierten Linguisten.

Für jede Analyse werden zunächst (von einem Linguisten, der die qualitativen Analyseverfahren beherrscht) Stilmerkmale bestimmt, auf die hin der Text dann (von auf quantitative Analyseverfahren und entsprechende Software spezialisierten Linguisten) analysiert wird. Es wird also erst in diesem zweiten Schritt digital, und es können große Textmengen mithilfe von Software untersucht werden.

¹⁰ Details dazu u. a. in *Fobbe* (2011), S. 110 ff.

Literatur

- Coulthard, M./Johnson, A. (eds.) (2010): *The Routledge Handbook of Forensic Linguistics*. London, New York: Routledge.
- Dern, C. (2009): *Autorenerkennung. Theorie und Praxis der linguistischen Tatschreibenanalyse*. Stuttgart: Boorberg.
- Ehrhardt, S. (2016): Texte als Straftat und im Straftatkontext. In: Felder, E./Vogel, F. (Hg.): *Handbuch Sprache im Recht*. Berlin: de Gruyter, S. 547-566.
- Ehrhardt, S. (2018): Authorship Attribution Analysis. In: Visconti, J.: *Handbook of Communication in the Legal Sphere*. Handbooks of Applied Sciences (HAL) Bd. 14. Berlin: de Gruyter, S. 169-200.
- Fobbe, E. (2011): *Forensische Linguistik. Eine Einführung*. Tübingen: Narr.
- Fobbe, E. (2013): Aufgaben und Möglichkeiten der forensischen Linguistik. *SIAG-Journal Zeitschrift für Polizeiwissenschaft und polizeiliche Praxis* (4), S. 4-15.
- Fobbe, E. (2017): Forensische Linguistik. In: Felder, E./Vogel, F. (Hg.): *Handbuch Sprache im Recht*. Berlin: de Gruyter, S. 271-290.
- Hessler, S. (2023): *Autorschaftserkennung und Verstellungsstrategien*. Tübingen: Narr (Reihe TBL/Tübinger Beiträge zur Linguistik).
- Kämper, H. (1996): Nachweis der Autorschaft. Methodische Überlegungen zur linguistischen Textidentifizierung und Täterermittlung. *Kriminalistik*, 50 (8-9), S. 561-566.
- Kniffka, H. (1981): Der Linguist als Gutachter bei Gericht. Überlegungen und Materialien zu einer „Angewandten Soziolinguistik“. In: Peuer, G./Winter, S. (Hg.): *Angewandte Sprachwissenschaft. Grundfragen – Bereiche – Methoden*. Bonn: Bouvier, S. 584-634.
- Kniffka, H. (2007): *Working in Language and Law. A German Perspective*. Basingstoke: Palgrave Macmillan.
- Knobloch, J. (1981): Forensische Linguistik: Eine Anwendung sprachwissenschaftlicher Methoden auf Rechtsfragen. In: Peuer, G./Winter, S. (Hg.): *Angewandte Sprachwissenschaft. Grundfragen – Bereiche – Methoden*. Bonn: Bouvier, S. 635-637.
- Rathert, M. (2006): *Sprache und Recht*. Heidelberg: Universitätsverlag Winter GmbH.
- Rottler, P./Martin, L. (2020): *Die geheimen Muster der Sprache*. München, Redline-Verlag.
- Stamatatos, E. (2009): A survey of modern authorship attribution methods. *American Society for Information, Science and Technology*, 60 (3), S. 538-556.
- Thormann, I./Hausbrandt, J. (2016): *Rechtssprache*. Berlin: BDÜ-Fachverlag.
- Thormann, I. (1994): *Wirtschaftsenglisch. Kategorisierung wirtschaftsenglischer Äußerungen*. Egelsbach/Köln/Washington: Verlag Hansel-Hohenhausen.
- Thormann, I. (2018): Forensische Linguistik. In: Lang-Groth, I./Neef, M. (Hg.): *Facetten der deutschen Sprache*, Frankfurt am Main: Peter Lang, S. 117-129.
- Träger, F. (2012): *Aufgaben und Methodik der Autorenerkennung im Rahmen der Forensischen Linguistik*. GRIN.