

Algorithmisch unterstützte Literarkritik

Eine grammatikalische Analyse zur Bestimmung von Schreibstilen

Michael Tschuggnall, Günther Specht, Christian Riepl
Innsbruck, München

Einleitung

Am Beginn der literaturwissenschaftlichen Analyse eines alttestamentlichen Textes steht in der von Wolfgang Richter begründeten Methodologie der Aspekt der „Einheit oder Zusammengesetztheit des Textes“¹, der im methodischen Schritt der Literarkritik zu behandeln ist. Sie dient dazu, Texte aus ihrem weiteren Kontext nach außen abzugrenzen und nach innen deren Einheitlichkeit bzw. Uneinheitlichkeit nachzuweisen. Als Kriterien werden dazu Doppelungen, Spannungen, Wiederholungen und weitere Beobachtungen herangezogen, die sich sowohl auf die Ausdrucks- wie auch die Inhaltsseite der Texte beziehen können. Der Fortschritt in der Grammatikforschung² ermöglicht eine zunehmend verfeinerte sprachwissenschaftliche Beschreibung der Ausdrucksseite von Texten. Grammatikalische Analysen gelten als eine wichtige Grundlage der Literarkritik³ und erlauben exaktere Beobachtungen, differenziertere Bewertungen der literarkritischen Kriterien und begründetere Urteile bei der Isolierung von sog. „kleinen Einheiten“ und deren Verhältnisbestimmung, worauf wiederum die weiteren literaturwissenschaftlichen Analysemethoden aufbauen.⁴

¹ RICHTER 1971, 49-72.

² Zum ebenenspezifischen Grammatikmodell siehe RICHTER 1978, 1979, 1980. Die dort eingeführten Ebenen Wort, Wortfügung und Satz wurden um die Ebene Satzfügung ergänzt durch die Forschungen zu den Großsatzformen u.a. bei IRSIGLER 1991, 1993.

³ Siehe dazu z.B. VANONI 1984, der in der literarkritischen Untersuchung u.a. umfangreiche Valenzanalysen durchführt und RIEPL 1993, der die sprachwissenschaftliche Beschreibung aller sprachlichen Ebenen (Wort, Wortfügung, Satz, Satzfügung) als Voraussetzung nicht nur für die Literarkritik, sondern für alle methodischen Schritte einer literaturwissenschaftlichen Analyse (Form-, Gattungs-, Traditions- und Redaktionskritik samt darauf aufbauender Inhaltsanalyse) sieht.

⁴ Einen Überblick zur literaturwissenschaftlichen Forschungsrichtung am Alten Testament von 1971 bis 1989 gibt SEIDL 1989. Zum Modell einer rechnergestützten Anwendung dieser Methoden siehe RIEPL 1999.

Im Bereich der Informatik wurde bereits gezeigt, dass es möglich ist, Autoren moderner Texte allein aufgrund der von ihnen verwendeten Grammatik zu identifizieren. Durch das automatisierte Finden von Stilbrüchen ist es so beispielsweise möglich, Plagiate zu erkennen⁵, Schriftstücke Autoren zuzuordnen⁶ oder auch jeweilige Einzelbeiträge in gemeinschaftlich verfassten Texten zu isolieren⁷. Die Hauptidee besteht dabei darin, die aus der Analyse natürlicher Sprache resultierenden Syntaxbäume einzelner Sätze miteinander zu vergleichen, um anschließend strukturelle Eigenheiten eines Autors zu quantifizieren. Beispielsweise bietet die Grammatik die Freiheit, einen Satz syntaktisch auf viele verschiedene Arten zu formulieren, ohne dabei den semantischen Gehalt zu ändern. Genau diese Freiheit der Formulierung wird oft intuitiv von Autoren verwendet. Sie schafft einen individuellen sprachlichen Ausdruck, der zur Identifikation des Verfassers herangezogen werden kann.

Durch die von *AMOS*⁸ erzeugten, auf *SALOMO*⁹ basierenden und in der MultiBHT¹⁰- bzw. der aktuellen Biblia Hebraica transcripta – Forschungsdatenbank¹¹ verwalteten Grammatikstrukturen ist es nun möglich, die entwickelten Algorithmen auch auf sämtliche alttestamentlichen Texte anzuwenden, mit dem Ziel, kleine kongruente Einheiten bzw. unterschiedliche Verfasser in einem Text aufgrund des Schreibstils satz- bzw. teilsatzgenau identifizieren zu können. Es wurden zwar bereits computerunterstützte Analysen von Bibeltexten durchgeführt, jedoch basieren diese meist ausschließlich auf Statistiken über vorher definierte Wortgruppen oder Zeichenketten. Zudem wurde dabei in vielen Fällen auf eine englische Übersetzung der Bibel zurückgegriffen¹². In dieser Arbeit werden erstmals vollautomatisiert Stilbrüche und damit Autorenezuordnungen untersucht, wobei die auf den Grammatikforschungen von Wolfgang Richter basierenden sprachwissenschaftlichen Analysen der objektsprachlichen Transkriptionen der alttestamentlichen Texte zugrunde gelegt werden.

⁵ TSCHUGGNALL / SPECHT, 2013.

⁶ TSCHUGGNALL / SPECHT, 2014 (1).

⁷ TSCHUGGNALL / SPECHT, 2014 (2).

⁸ SPECHT, 1990.

⁹ ECKARDT, 1987.

¹⁰ SPECHT, 1995.

¹¹ RIEPL, 2016. Die Datenbank ist unter dem URL <http://www.bht.gwi.uni-muenchen.de> erreichbar.

¹² Siehe beispielsweise KOPPEL et al., 2011.

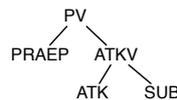
Im Folgenden wird der Ansatz näher beschrieben, der durch eine anschließende Anwendung auf Gen 1,1-2,24 und Jon 1-4 exemplarisch diskutiert wird.

Ansatz

Die Grundlage für alle vorgestellten Analysen bilden die von *AMOS* erzeugten Bäume morphosyntaktischer Einheiten, die über die *Biblia Hebraica transcripta* – Forschungsdatenbank zugänglich sind¹³. Diese werden in folgender Form gespeichert:

$$(PV (PRAEP l') (ATKV (ATK [h]a) (SUB raqī')))$$

In dieser Arbeit wird ausschließlich die morphosyntaktische Struktur betrachtet, von den konkreten Wörtern wird abstrahiert. Eine Visualisierung des Baumes sieht schließlich wie folgt aus:



Da die von uns entwickelten Algorithmen¹⁴ auf der Satzebene operieren, hier aber morphosyntaktische Strukturen die Datenbasis bilden, wurden drei verschiedene Varianten von Algorithmen getestet:

- a) *WORTVERBINDUNG*: Jede Wortverbindung¹⁵ wird als „Satz“ betrachtet. Abbildung 1 zeigt alle Wortverbindungen aus Gen 1,12a.aR.b. In dieser Variante wurden zusätzlich weitere Konfigurationen getestet, die Wortverbindungen auch als Teil eines größeren Satzes interpretieren und verarbeiten.

¹³ *AMOS* hat dabei meist mehrere Varianten von Grammatikbäumen für ein und dieselbe Wortverbindung berechnet, die sich aus der Anwendung der grammatischen Regeln ergeben und zu Mehrdeutigkeiten bzw. verschiedenen Auslegungsmöglichkeiten führen. Richter hat letztlich alle Bäume durchgesehen und sich für die aktuellen in der Datenbank gespeicherten Lösungsmöglichkeiten entschieden.

¹⁴ Siehe TSHUGGNALL / SPECHT, 2013-2015.

¹⁵ Gemeint sind morphosyntaktisch begründete Wortverbindungen, die zugleich satztheoretisch ein Satzglied repräsentieren, im Gegensatz zu einzelnen morphosyntaktisch begründeten Wortverbindungen, die Teil einer rekursiven Wortverbindung sein können.

- b) *SATZTYP*: Alle Wortverbindungen mit demselben Satztyp (z.B. 12a) werden zusammengefügt und als „Satz“ betrachtet. Abbildung 2 illustriert diese Variante am vorangehenden Beispiel. Diese Methode erfasst auch alle satzhafte Erweiterungen von Satzgliedern.

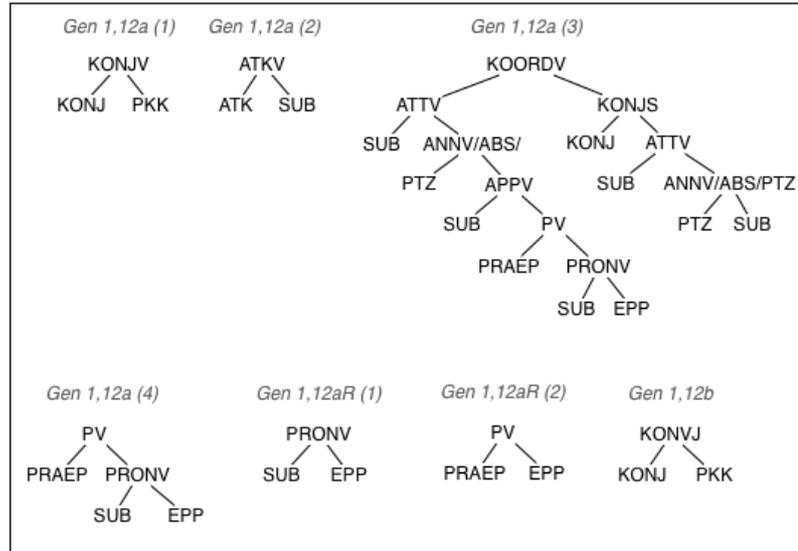


Abbildung 1: Grammatikbäume in der Variante *WORTVERBINDUNG*.

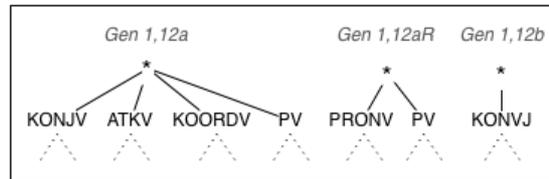
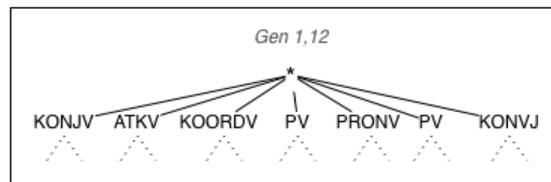


Abbildung 2: Grammatikbäume in der Variante *SATZTYP*.

Abbildung 3: Grammatikbaum in der Variante *VERS*.

- c) *VERS*: alle Wortverbindungen eines Verses werden zusammengefügt und als „Satz“ betrachtet. Abbildung 3 zeigt den resultierenden Grammatikbaum in Gen 1,12. Die Zusammenfassung nach Vers mag zwar vorerst irreführend klingen, da ein Vers keiner grammatikalischen Einheit entspricht, dennoch können durch die anschließende Extraktion syntaktischer Teilstrukturen (wie später beschrieben) durchaus interessante Informationen gewonnen werden.

Die durch die jeweilige Variante erzeugten Grammatikbäume bilden im nächsten Schritt die Grundlage zur Analyse des Schreibstils, der auf zwei verschiedene Arten bestimmt wurde.

Basierend auf den Algorithmen zur intrinsischen Plagiatserkennung¹⁶ werden in der ersten Untersuchung für alle Bäume sog. pq-Gramme¹⁷ berechnet und mittels einer Distanzmetrik miteinander verglichen. Ein pq-Gramm repräsentiert einen Teilausschnitt der Höhe p und der Breite q aus einem gegebenen Syntaxbaum. Beispielsweise können unter Verwendung der Höhe $p=2$ und der Breite $q=2$ aus dem Baum von Gen1,12a (3) aus Abbildung 1 u.a. folgende pq-Gramme extrahiert werden:

KOORDV – ATTV – SUB – ANNV/ABS/
 KOORDV – KONJS – KONJ – ATTV
 ...

¹⁶ „intrinsisch“ bedeutet dabei, dass Plagiate ausschließlich innerhalb eines Dokuments über Stilbrüche erkannt und nicht über Vergleiche mit möglichen externen Quellen identifiziert werden.

¹⁷ Siehe AUGSTEN et al., 2010.

Das Konzept der pq-Gramme berücksichtigt zusätzlich Baumgrenzen, indem es nicht existierende Randknoten mittels * markiert. Alle möglichen pq-Gramme aus Gen 1,12a (1) wären somit:

* – KONJV – * – KONJ
 * – KONJV – KONJ – PKK
 * – KONJV – PKK – *
 KONJV – KONJ – * – *
 KONJV – PKK – * – *

Der Vergleich von Grammatikbäumen wird dann mittels der pq-Gramm-Distanz durchgeführt, die auf den Mengen aller möglichen pq-Gramme der jeweiligen Bäume basiert und ein Maß dafür ist, wie unterschiedlich bzw. ähnlich sich die Bäume sind. Mit der berechneten Distanz wird anschließend versucht, syntaktisch signifikant abweichende Sätze zu identifizieren. Als auffällige Sätze werden dabei solche betrachtet, deren durchschnittliche Distanz zu allen anderen Sätzen ein vordefiniertes Maß überschreitet. Der Ansatz liefert also Hinweise auf grammatisch auffällige Strukturen im Text, die zunächst isoliert und unter literarkritischem Aspekt als einzelne „Zusätze“ oder als literarkritische Kriterien zur Bestimmung von „kleinen Einheiten“ gewertet werden können.

Anwendung

Abbildung 4 zeigt eine exemplarische Analyse von Gen 1,1-2,4a mit der Variante *WORTVERBINDUNG*. Die durchgezogene Linie zeigt dabei die Grenze zwischen unauffällig und auffällig¹⁸. Da die Methode *WORTVERBINDUNG* einzelne Wortverbindungen betrachtet, fallen vor allem solche Verbindungen auf, die sich in ihrer Länge sehr unterscheiden. Im Beispiel sind dies etwa Gen 1,16x oder Gen 1,25a, die durch die Verwendung von *KOORDV*-Verbindungen¹⁹ sehr lang sind.

¹⁸ Die Trennung basiert auf den statistischen Größen Mittelwert μ und Standardabweichung σ und liegt je nach Anwendungsfall zwischen $\mu + 2\sigma$ und $\mu + 3\sigma$.

¹⁹ Als *KOORDV* werden auf Wortfügungsebene durch die Konjunktion *w* gleichgeordnete Wortverbindungen bezeichnet; eine *KOORDV* ist definiert durch die Regel: *KOORDV* := WA | WV + KONJS [+ KONJS ...]; wobei gilt: WA := SUB | ADJ | IMP | NUM | VBNOM | EN | PRON | ADV | DEIKT; und WV := CSV | APPV | ATTV | ANNV | ATKV | PV | ABSV | ADVV | PRONV | NEGV | KONJS; ferner gilt: KONJS := KONJ + WV. Ein Wort (WA) oder eine Wortverbindung (WV) kann durch beliebig viele Konjunktionssyndesen (KONJS) erweitert werden.

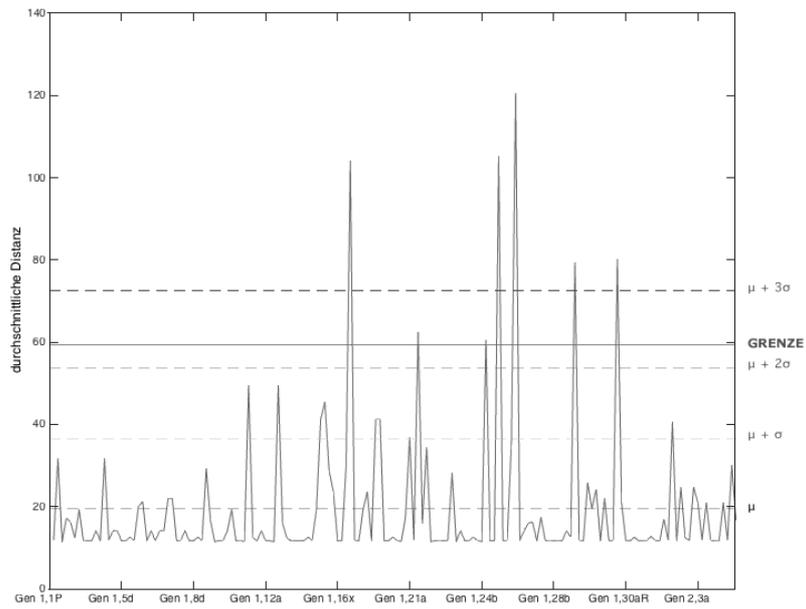


Abbildung 4: Intrinsischer Vergleich von Gen 1,1-2,4a
(Variante: *WORTVERBINDUNG*).

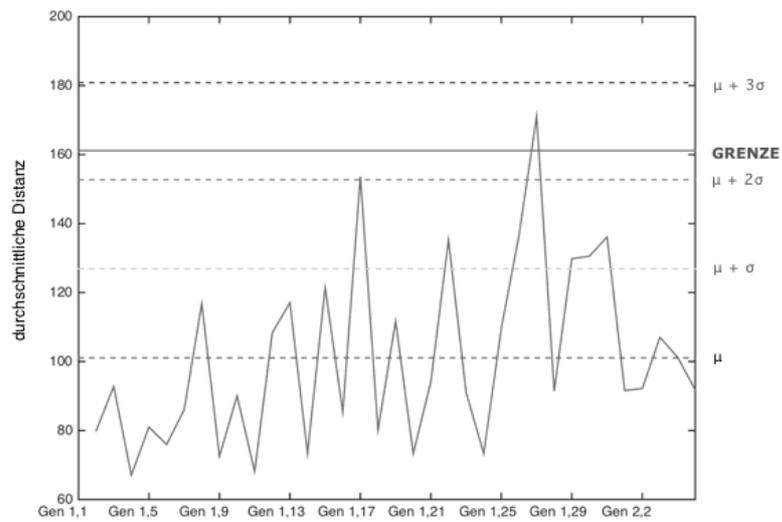


Abbildung 5: Intrinsischer Vergleich von Gen 1,1-2,4a
(Variante: *VERS*).

Betrachtet man hingegen dieselbe Stelle mit der gewissermaßen gröber auflösenden Variante *VERS*, so werden solche Ausreißer eingeebnet, wie in Abbildung 5 ersichtlich ist. Nach dem Zusammenfassen zu Versen ist also nur mehr Gen 1,26 auffällig²⁰.

In einer weiteren Untersuchung, die auf Algorithmen zur automatischen Erkennung von Einzelautoranteilen eines Mehrautorenwerkes beruht²¹, wurde versucht, Wortverbindungen nach deren Autorenschaft zu gruppieren. Die Vorgehensweise unterscheidet sich dabei etwas von der vorherigen intrinsischen Analyse. Während zuvor einzelne Bäume miteinander verglichen wurden, werden hier sog. Profile mit den am häufigsten vorkommenden Baumstrukturen bzw. pq-Grammen erstellt. Beispielsweise mag ein Verfasser die Verbindung (PRONV (SUB EPP)) vermehrt einsetzen, während ein anderer diese selten bis gar nicht einsetzt. Ein konkretes Profil (mit $p=2$ und $q=3$), berechnet aus den Vorkommen der Strukturen in Gen 1,1-3, sieht wie folgt aus:

<i>pq-Gramm</i>	<i>Häufigkeit [%]</i>
PV – PRAEP – * – * – *	5,05
CSV – SUB – * – * – *	5,05
ATKV – ATK – * – * – *	4,04
...	...
KOORDV – KONJS – * – KONJ – SUB	1,01
* – ATKV – * – ATK – SUB	1,01

So wird für alle analysierten Einheiten²² ein Profil erstellt, um diese anschließend nach Ähnlichkeiten gruppieren zu können. Diese Gruppierung erfolgt dabei mithilfe von sog. *Clustering-Algorithm*²³, die automatisiert

²⁰ Innerhalb des Verses ist im Satz Gen 1,26c das Präpositionalobjekt auf Grund der ihm zugrundeliegenden KOORDV, bestehend aus einer PV und 4 weiteren KONJS, auffällig. Solche Reihungen fallen zunächst wegen ihrer Länge auf. Ähnliche Beobachtungen lassen sich auch über größere Textabschnitte leicht sammeln. Sie liefern Kriterien für die literarkritische Untersuchung, die diese hinsichtlich primärer und sekundärer Einheiten zu bewerten hat.

²¹ TSCHUGGNALL / SPECHT, 2014.

²² Je nach Variante wieder jeweils einzelne Wortverbindungen (*WORTVERBINDUNG*) oder auch zusammengefasste Bäume (*SATZTYP*, *VERS*).

²³ Für das Clustering wurde das WEKA-Framework (HALL et al., 2009) mit verschiedenen Clusteralgorithmen evaluiert. Anschließend an die bereits erwähnten Baumvarianten wurden auch noch zusätzliche Zusammenfassungen getestet. Hier wurden beispiels-

signifikante und unwichtige Eigenschaften unterscheiden und die Profile nach den ersteren Kriterien schließlich gruppieren („*clustern*“).

Abbildung 6 zeigt eine Clustering-Analyse über den Textabschnitt Gen 1-2 und markiert die von den Algorithmen identifizierten zusammenhängenden Einheiten mit der jeweils selben Farbe. Es ist deutlich ersichtlich, dass einerseits Textteile von Gen 1 überwiegend einem Autor (hell) zugewiesen werden, während andererseits Textteile von Gen 2 überwiegend einem anderen Verfasser (dunkel) zugeordnet werden.²⁴ Bedenkt man, dass die verwendeten Methoden ausschließlich mit den morphosyntaktischen Analyseergebnissen arbeiten, erzielt das

Gen 1,1P-Gen 1,4a	Gen 1,4c-Gen 1,6b	Gen 1,6c-Gen 1,8b	Gen 1,8c-Gen 1,10b	Gen 1,10c-Gen 1,12b
Gen 1,13a-Gen 1,15a	Gen 1,15aI-Gen 1,18vI2	Gen 1,19a-Gen 1,21aR	Gen 1,21b-Gen 1,23c	Gen 1,24a
Gen 1,24c-Gen 1,27b	Gen 1,27c-Gen 1,29b	Gen 1,29bR1-Gen 1,31b	Gen 1,31c-Gen 2,3b	Gen 2,3c-Gen 2,5c
Gen 2,5dI-Gen 2,8b	Gen 2,9a-Gen 2,11b	Gen 2,11bR-Gen 2,14cP	Gen 2,15a-Gen 2,17a	Gen 2,17aP-Gen 2,19bI
Gen 2,19c-Gen 2,21c	Gen 2,21d-Gen 2,23d	Gen 2,24a-Gen 2,25a		

Abbildung 6: Clustering von Gen 1-2 nach Autoren
(Variante: *SATZTYP*).

weise je 2-10 *WORTVERBINDUNG*, *SATZTYP* oder auch *VERS*-Einheiten zusammengefasst und in einem einzigen Profil verarbeitet.

²⁴ Im Diagramm nicht berücksichtigt sind Sätze, die keine Wortverbindungen enthalten, wie z.B. Gen 1,4b.10d.12c.18b.21c.25c. Eine Einbeziehung der Satzebene in die oben vorgestellten Algorithmen und Methoden würde zusätzliche, für die Literarkritik wichtige Beobachtungen zu textlichen Gliederungsmerkmalen liefern.

Verfahren schon einen relativ guten Annäherungswert an das in der alttestamentlichen Exegese herrschende Urteil²⁵ zumindest bezüglich einer äußeren literarkritischen Abgrenzung der „kleinen Einheiten“ Gen 1,1-2,4a und Gen 2,4b-25.

Um so erstaunlicher ist allerdings das Ergebnis, das die Methoden am Buch Jona liefern. Wie aus Abbildung 7 ersichtlich ist, isoliert die Clustering-Analyse den (poetischen) Textabschnitt Jon 2,3-10²⁶ eindeutig aus dem umgebenden (prosaischen) Text. Offensichtlich reichen dazu Berechnungen anhand der morphosyntaktischen Analysen aus. Interessant wäre nun, diese Methoden auch auf andere Textabschnitte (z.B. Ex 14-15) oder Werke des Alten Testaments anzuwenden, um zu sehen, ob sich dort ebenfalls solche eindeutigen Zuordnungen allein auf der Ausdrucksseite der Wortfügungsebene ergeben.

Jon 1,1a-Jon 1,3a	Jon 1,3aI-Jon 1,3eI	Jon 1,4a-Jon 1,5b	Jon 1,5c-Jon 1,5f	Jon 1,6a-Jon 1,7a	Jon 1,7e-Jon 1,8c
Jon 1,8d-Jon 1,9c	Jon 1,9cR-Jon 1,10e	Jon 1,10f-Jon 1,12a	Jon 1,12b-Jon 1,13aI	Jon 1,13c-Jon 1,14d	Jon 1,14f-Jon 1,16b
Jon 1,16c-Jon 2,1b	Jon 2,3a-Jon 2,4a	Jon 2,4b-Jon 2,6b	Jon 2,6c-Jon 2,8a	Jon 2,8aI-Jon 2,10d	Jon 2,11b-Jon 3,2cR
Jon 3,3a-Jon 3,4b	Jon 3,4c-Jon 3,6a	Jon 3,6b-Jon 3,7b	Jon 3,7bI-Jon 3,8b	Jon 3,8c-Jon 3,10b	Jon 3,10c-Jon 4,2b
Jon 4,2c-Jon 4,3a	Jon 4,3b-Jon 4,5c	Jon 4,5d-Jon 4,6bI1	Jon 4,6bI2-Jon 4,7c	Jon 4,8a-Jon 4,8e	Jon 4,8eI-Jon 4,9c
Jon 4,9d-Jon 4,10bR3	Jon 4,10bR4-Jon 4,11aRR				

Abbildung 7: Clustering von Jon 1-4 nach Autoren
(Variante: *SATZTYP*).

²⁵ Demnach werden in Gen 1-2 grob zwei Textabschnitte, nämlich Gen 1,1-2,4a und Gen 2,4b-25 (mit einem sekundären Einschub in Gen 2,10-14), vermutet; diese werden jeweils unterschiedlichen Quellenschichten zugeordnet, die wiederum von verschiedenen Redaktionen bearbeitet worden sein könnten. Vgl. SMEND 1984, 47 und 86.

²⁶ In der alttestamentlichen Exegese ist das literarkritische Urteil bezüglich Jon 2,3-10 unumstritten. Siehe dazu SMEND 1984, 177: Das Gebet, das Jona im Bauch des Fisches spricht, wird übereinstimmend als sekundärer Einschub eines zweiten Autors betrachtet.

Ausblick

In diesem Beitrag konnten wir zeigen, dass informatische Methoden die literaturwissenschaftliche Analyse alttestamentlicher Texte, insbesondere den methodischen Schritt der Literarkritik unterstützen können und dabei neue, spannende Sichtweisen bei der Sammlung und Bewertung literarkritischer Kriterien eröffnen. Zum ersten Mal konnten detaillierte Evaluierungen auf rein morphosyntaktischer Ebene durchgeführt werden, ohne dabei auf semantische oder lexikalische Kriterien zurückzugreifen. Die exemplarischen Auswertungen zeigen äußerst vielversprechende Ergebnisse auf deutlich höherem Niveau als gemeinhin vermutet.

Künftige Arbeiten sollen deshalb nicht nur Berechnungen und Analysen der Ausdrucksseite auf Wortfügungsebene, sondern auch der Wort-, Satz- und Satzfügungsebene untersuchen.²⁷ Zudem sollen die Untersuchungen auf das gesamte alttestamentliche Textkorpus ausgedehnt werden. Weiter lassen sich die hier vorgestellten Methoden nicht nur im Rahmen der Literarkritik für die Isolierung von „kleinen Einheiten“ bzw. die Autorenerkennung einsetzen. Auch andere Methoden der Literaturwissenschaft, wie z.B. die Formkritik, können die informatischen Methoden anwenden, um automatisiert textliche Strukturmerkmale²⁸, die sich etwa auf die Intention von Texten auswirken können, zu erkennen.

Literatur

- ARGENTON, Hans, Indexierung und Retrieval von Feature-Bäumen am Beispiel der linguistischen Analyse von Textkorpora, DISDBIS 40, St. Augustin 1998.
- AUGSTEN, Nikolaus / BÖHLEN, Michael / GAMPER, Johann, The pq-gram Distance Between Ordered Labeled Trees, in: ACM Transactions on Database Systems, 2010.
- ECKARDT, Walter, Computergestützte Analyse althebräischer Texte, ATSAT 29, St. Ottilien 1987.
- HALL, Mark et al., The WEKA Data Mining Software: an Update, in: ACM SIGKDD Explorations Newsletter 11/1 2009, 10-18.

²⁷ Das Datenmaterial der grammatischen und semantischen Analysen ist in der „Biblia Hebraica transcripta - Forschungsdatenbank“ zugänglich.

²⁸ In einem der Tests der Clustering-Analyse konnten mit verfeinerten Einstellungen in Gen 1 strukturelle Wiederholungen wie z.B. die Tageszählung bzw. Formeln („Es wurde Abend und es wurde Morgen“) isoliert werden.

- IRSIGLER, Hubert, Thronbesteigung in Ps 93? Der Textverlauf als Prozeß syntaktischer und semantischer Interpretation, in: Groß, W. / Irsigler, H. / Seidl, T. (Hg.): Text, Methode und Grammatik, FS W. Richter, St. Ottilien 1991, 155-190.
- , Großsatzformen im Althebräischen und die syntaktische Struktur der Inschrift des Königs Mescha von Moab, in: Irsigler, H. (Hg.), Syntax und Text. Beiträge zur 22. Internationalen Ökumenischen Hebräisch-Dozenten-Konferenz 1993 in Bamberg, ATSAT 40, St. Ottilien 1993, 81-121.
- KOPPEL, Moshe et al., Unsupervised Decomposition of a Document Into Authorial Components, in: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, 2011, 1356-1364.
- RECHENMACHER, Hans / VAN DER MERWE, Christo H. J., The Contribution of Wolfgang Richter to Current Developments in the Study of Biblical Hebrew, in: Journal of Semitic Studies 50.1 2005, 59-82.
- RICHTER, Wolfgang, Exegese als Literaturwissenschaft. Entwurf einer alttestamentlichen Literaturtheorie und Methodologie, Göttingen 1971.
- , Grundlagen einer althebräischen Grammatik. A. Grundfragen einer sprachwissenschaftlichen Grammatik. B. Die Beschreibungsebenen: I. Das Wort (Morphologie), ATSAT 8, St. Ottilien 1978.
- , Grundlagen einer althebräischen Grammatik. B. Die Beschreibungsebenen: II. Die Wortfügung (Morphosyntax), ATSAT 10, St. Ottilien 1979.
- , Grundlagen einer althebräischen Grammatik. B. Die Beschreibungsebenen: III. Der Satz (Satztheorie), ATSAT 13, St. Ottilien 1980.
- RIEPL, Christian, Sind David und Saul berechenbar? Von der sprachlichen Analyse zur literarischen Struktur von 1 Sam 21 und 22, ATSAT 39, St. Ottilien 1993.
- , Wie wird Literatur berechenbar? Ein Modell zur rechnergestützten Analyse althebräischer Texte, in: Jahrbuch für Computerphilologie 1 (1999), hg. v. Deubel V. / Eibl K. / Jannidis F., Paderborn 1999, 107-134.
- , Biblia Hebraica transcripta – Das digitale Erbe, in: Rechenmacher, H. (Hg.), In Memoriam Wolfgang Richter, ATSAT 100, St. Ottilien 2016.

- SEIDL, Theodor, Die literaturwissenschaftliche Methode in der alttestamentlichen Exegese. Erträge – Erfahrungen – Projekte. Ein Überblick, in: Münchener Theologische Zeitschrift 40,1 1989, 27-38.
- SMEND, Rudolf, Die Entstehung des Alten Testaments, 3. durchges. Aufl., Stuttgart 1984.
- SPECHT, Günther, Wissensbasierte Analyse althebräischer Morphosyntax. Das Expertensystem Amos, ATSAT 35, St. Ottilien 1990.
- , MultiBHT – Ein multimediales Datenbanksystem zur Sprachanalyse; vorgestellt am 29. März 1995 an der LMU München vor der Fakultät für Altertumskunde und Kulturwissenschaften. Technische Universität München, Institut für Informatik, München 1995.
- SPECHT, Günther / ZIRKEL, Martin, MultiMAP/2. Netzzugang und Netzbetrieb für das multimediale Datenbanksystem MultiMAP: TUM-19920, Technische Universität München, Institut für Informatik, München 1999.
- TSCHUGGNALL, Michael / SPECHT, Günther, Detecting Plagiarism in Text Documents Through Grammar-Analysis of Authors, in: Proceedings of the 15th Conference on Datenbanksysteme für Business, Technologie und Web (BTW), 2013, 241-259.
- , Automatic Decomposition of Multi-Author Documents Using Grammar Analysis, in: Proceedings of the 26th GI-Workshop on Grundlagen von Datenbanken, 2014.
- , Enhancing Author-ship Attribution By Utilizing Syntax Tree Profiles, in: Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2014, 195-199.
- VANONI, Gottfried, Literarkritik und Grammatik. Untersuchung der Wiederholungen und Spannungen in 1 Kön 11-12, ATSAT 21, St. Ottilien 1984.

Abstracts

In diesem Artikel wird ein Ansatz vorgestellt, der es ermöglicht, kleine Texteinheiten von alttestamentlichen Texten über eine automatisierte Analyse des Grammatikstils zu erkennen und unterschiedlichen Verfassern zuzuordnen. Exemplarische Auswertungen von Gen 1,1-2,25 und Jon 1-4 zeigen vielversprechende Ergebnisse.

This paper proposes an approach to automatically detect small units of scripts from the Old Testament and to attribute authors by analyzing the grammatical style. Exemplary evaluations of Gen 1,1-2,25 and Jon 1-4 reveal promising results.