

**Wortstatistik – ein immer beliebter werdendes exegetisches Handwerkszeug
auf dem (mathematischen) Prüfstand¹**

Scholastika Deck – Freiburg

In einer Zeit, in der die biblische Exegese immer mehr auf Methoden anderer Wissenschaften zurückgreift, um die Texte noch gründlicher zu erschließen, der Exeget also sich mit Sprach- und Literaturwissenschaft, Geschichte, Wirtschaftswissenschaften, Soziologie, Psychologie etc. und eben auch mit Statistik beschäftigen muß, wird es immer schwieriger, die Wirkung und Bedeutung der in solchen Wissenschaften verwendeten Methoden abzuschätzen und damit die Gewichtung der Ergebnisse und die Tragfähigkeit der Argumentation zu beurteilen.

In einem in VF 31 (1986) erschienenen Überblick zur Jesajaforschung zieht C Hardmeier gegen eine sich immer mehr verbreitende exegetische Arbeitsweise zu Felde, die er als "Neoliterarkritik" bezeichnet. Er wettet gegen eine positivistische Überschätzung von Sprachbefunden, weil sie sich auf stilkritische und vokabelstatistische Einzelbeobachtungen stütze, die sich bei näherem Zusehen als von geringer Signifikanz erwiesen.² Eines seiner Gegenargumente liegt in dem Hinweis, daß die hebräische Bibel nur ein kleiner, möglicherweise einseitiger Ausschnitt aus dem Althebräischen darstelle, und die von den kritisierten Autoren³ aufgeführten Beobachtungen für ein statistisches Urteil quantitativ unzureichend seien. In der Tat ist die Jesajaexegese ein reiches Beobachtungsfeld für verschiedene Varianten wortstatistischer Argumentation in der Exegese. Man findet beachtenswerte Indizien neben Argumenten, die einem Statistiker "mathematische Gänsehaut" verursachen. Ein anschauliches Beispiel dafür ist J. Vermeylens Analyse des Vokabulars von Jes 1,18ff in seiner zweibändigen Studie zum Jesajabuch.⁴

1)Es handelt bei diesem Artikel um einen für die Veröffentlichung als "biblische Notiz" leicht bearbeiteten Exkurs aus meiner in FzB 67 erschienenen Dissertation "Die Gerichtsbotschaft Jesajas: Charakter und Begründung", Würzburg 1991.

2)Vgl.C.Hardmeier, Jesajaforschung im Umbruch, 13.

3)Er bezieht seine Kritik vor allem auf O.Kaiser, W.Werner und R.Kilian.

4)Du Prophète Isaïe à l'apoclyptique. Miroir d'un demi-millénaire d'expérience religieuse en Israël, 2 Bde, Paris 1977/78, 59. Auf dieser Seite sind alle Tugenden und Gefahren der wortstatistischen Methode auf engem Raum vereinigt.

Er stellt z.B. fest, daß Jesaja שמע wenig benutze,⁵ daß כירכ und שלנו sonst bei Jesaja nicht mehr vorkämen und versucht, aus dem gestreuten Vorkommen (oft in späteren Texten) der sehr speziellen Verben לכך und ארם oder des Substantivs ררלע irgendwelche Schlüsse zu ziehen.⁶ Wie wenig diese Argumentation trägt, macht man sich schnell klar: Man stelle sich vor, man ziehe aus gehörten Ansprachen eines Predigers über einen Zeitraum von 35 Jahren eine Stichprobe von etwa 40–50 Sätzen und prüfe dann, wie oft wohl das Wort "Schnee" darin vorkommt.⁷ Wenn es sich nicht gerade um einen Gebirgsdorfpfarrer handelt, würde es doch sehr verwundern, wenn dies mehr als einmal der Fall wäre. Der nur einmalige Gebrauch des Wortes "Schnee" sagt mit Sicherheit *überhaupt nichts* über die Jesajanität aus, auch nicht mit irgendeiner Wahrscheinlichkeit.

Aber dies ist ein allgemeines Problem bei der Argumentation mit Wahrscheinlichkeiten, das man bei Arbeiten von Anhängern der Wortstatistik immer wieder beachten muß: Manche Beobachtung erlaubt den Schluß, daß ein Indiz für oder gegen Authentizität vorliegt, bei einer anderen ist ein solcher Schluß nicht vertretbar. Darüber scheinen sich jedoch die Verfechter dieser Methode ganz offensichtlich selten Rechenschaft abzulegen. Es kann hier nicht darum gehen, Möglichkeiten und Grenzen statistischer Tests zu erörtern, aber ich möchte einige grundsätzliche Überlegungen über erlaubte und unerlaubte Schlüsse in der Wortstatistik vorlegen, die sich ergeben, wenn man die (quantitativen) Anfragen eines Biblikers an seine Texte mit den mathematisch-statistischen Gesetzen konfrontiert.⁸

Die Situation, in der sich ein Statistiker befindet, der statistische Verfahren in der Bibelarbeit anwenden will, stellt sich wie folgt dar: Die hebräische Bibel (im Folgenden B) ist eine echte Teilmenge aus der Grundgesamtheit (im Folgenden G), nämlich der gesprochenen und geschriebenen hebräischen Sprache in allen Lebensbereichen über viele Jahrhunderte hinweg. Dieser Ausschnitt ist *nicht* repräsentativ, da die einzelnen Jahrhunderte und die verschiedenen Lebensbereiche unterschiedlich stark vertreten

5) "On le trouve plus de 1100 fois dans l'A.T. Isaïe le connaît mais l'emploie peu ... En revanche, le verbe est très fréquent dans le livre de Jérémie ... et dans la littérature deutéronomienne ...", ebd.

6) Vgl. J. Vermeylen, Du prophète Isaïe à l'apocalyptique, 59.

7) Das dürfte quantitativ den etwa 100 Versen entsprechen, die Vermeylen auf den Propheten Jesaja zurückführt.

8) Für eine gründliche Information zu Fragen der Statistik und insbesondere zur Testtheorie sei verwiesen auf: R. Hafner, Wahrscheinlichkeitsrechnung und Statistik, Wien/New York 1989, (besonders 368–441); für den mathematischen Laien vielleicht leichter zugänglich sind: A. Linder/W. Berthold, Elementare statistische Methoden I, (UTB 796) Basel/Stuttgart 1979 (besonders 157–185) und M.R. Spiegel, Statistik, Düsseldorf u.a. 1976 (besonders 167–200).

sind.⁹ Man mag auch vermuten, daß in der Bibel eher gehobene als Umgangssprache vorliegt. In dieser Teilmenge (B) gibt es nun weitere Untermengen (im Folgenden U_1, U_2 , etc.) von Wörtern, die dem gleichen Sprecher/Schreiber, z.B. Jesaja, zuzuordnen sind. Die Auswahl dieser Worte aus der tatsächlichen Gesamtpredigt (beispielsweise Jesajas) in der hebräischen Bibel ist wiederum mit Sicherheit nicht repräsentativ, so daß ein weiterer Unsicherheitsfaktor für statistische Überlegungen hinzukommt.¹⁰ Die Frage, ob ein kleines Teilstück der Bibel (= eine Perikope; im folgenden P) einer solchen Untermenge zugehört oder eben einer anderen, etwa der Untermenge deuteronomistischer Autoren (U_b), ist das Problem, das mit statistischen Mitteln einer Lösung zugeführt werden soll.¹¹ Um die Situation korrekt zu beschreiben, muß man sich klar machen, daß U eine unbekannte Größe ist, d.h., daß man nur eine Teilmenge U' kennt – das sind die Perikopen, die mit ziemlicher Sicherheit einem Autor (beispielsweise Jesaja) zuzuordnen sind (also ist U' eine Teilmenge von U: $U' \subset U$). Die Frage, ob eine Perikope einer bestimmten Schicht angehört oder nicht ($P \subset U' ?$), kann nur positiv entschieden werden, wenn irgendwelche Beziehungen (i.d.R. quantitativer Art) zwischen P und U' hergestellt werden. Dabei kommen folgende Grundkonstellationen mit ihren zugehörigen Risiken und Fehlschlüssen vor:

1) Um den ersten Fehlschluß – aus der Tatsache, daß ein Wort (W) aus einer Perikope in U' sonst nicht mehr vorkommt, zu einem Argument zu machen – als solchen zu erkennen, bedarf es eigentlich keiner besonders tiefeschürfenden Statistikkennntnisse. Nach dem oben Gesagten leuchtet ein, daß dies reiner Zufall sein kann. Dazu ist U' meist viel zu klein und eben nicht repräsentativ.¹² Auch der Fall des Hapaxlegomenon für die hebräische Bibel für ein Wort $W \in P$ (\in := Element) läßt allein keinerlei Schluß für oder wider $P \subset U'$ zu. Das einmalige Vorkommen kann an der – bezüglich der oben genannten Grundgesamtheit 'hebräische Sprache' – nicht proportionalen und inhomogenen Teil-

9) Beispielsweise ist das achte Jahrhundert mit viel weniger Texten vertreten als z.B. die exilische oder frühnachexilische Zeit, oder ist der kaufmännische Bereich (Verträge) im Vergleich zu seinem realen Vorkommen verständlicherweise unterrepräsentiert.

10) Die von einem Propheten überlieferten Texte sind uns nur gemäß dem Interesse der Tradenten ausgewählt zugänglich, sind also auch unter quantitativem Aspekt durch mehrere "Filter" gegangen (Hörer, Schüler, Situation und Intention von Redaktoren). Eine entsprechend der Themen *proportionsgetreue* Auswahl der realen Prophetenpredigt ist unrealistisch. Von daher sind Argumente, die aus der Tatsache gewonnen sind, daß ein Thema (oder bestimmte Wörter) bei einem Propheten nur einmal oder selten vorkommen (Solche Argumentation kommt vor!), praktisch ohne Gewicht.

11) Also: $\{U_1, U_2, \dots\} \subset B \subset G$. (c := Teilmenge).

12) Das kann z.B. daran liegen, daß das in P behandelte Thema in U' nur einmal vorkommt. Bei Jesaja wäre hier die Kultkritik zu nennen.

menge 'hebräische Bibel' liegen. Man macht sich auch sofort klar, daß aus der Tatsache, daß irgendein Wort *nicht* in dieser Perikope benützt wird, das man hier vielleicht erwartet hätte, ebenso keinerlei Schluß gezogen werden kann. Dazu ist das Teilstück (P) viel zu klein, ebenso die Auswahl U aus der tatsächlichen Gesamtpredigt (z.B. des Jesaja), erst recht U'.

2) Anders gelagert ist die folgende Frage: Wenn in einer Untermenge (z.B. der jesajanischen Texte: U_j) ein Wort mehrmals vorkommt, das in der ganzen sonstigen Bibel gar nicht oder offensichtlich signifikant selten vorkommt, d.h.: der Quotient

$$\frac{\text{Vorkommen eines bestimmten Wortes bei Jesaja}}{\text{Anzahl der Wörter jesajanischer Texte}}$$

muß wesentlich größer sein als der Quotient

$$\frac{\text{Vorkommen dieses Wortes in der übrigen Bibel}}{\text{Anzahl der Wörter in der übrigen Bibel}}$$

was in Anbetracht des im zweiten Quotienten riesigen Nenners bei den genannten Voraussetzungen ohne statistischen Test sofort als gegeben erkannt werden kann.¹³ Das klassische Beispiel bei Jesaja ist der קִרְשׁ יִשְׂרָאֵל.¹⁴ In einem solchen Fall kann das Vorkommen in einem Teilstück (Perikope) als *ein* Indiz für die Zugehörigkeit zur Untermenge, in der dieses Wort signifikant häufig ist, gewertet werden, wenn nicht besondere inhaltliche Gründe das häufigere Vorkommen plausibel machen, d.h. wenn z.B. Themen verhandelt werden, die sonst in der Bibel selten oder gar nicht mehr vorkommen. Bei Wörtern, die auch in der ganzen Bibel häufig vorkommen, ist ein statistischer Test erforderlich (der hier wegen der fehlenden Proportionalität zwischen U, B, G mit großem Fehlerisiko verbunden ist), um zu klären, ob die Häufigkeit des Wortes in der Untermenge nicht doch noch im *Zufallsbereich* des Ereignisses 'Vorkommens in der Perikope' liegt. Allerweltswörter sind daher für eine wortstatistische Argumentation sehr unergiebig, wenn nicht besondere Redewendungen oder typische Kombinationen hinzukommen.

3) Der umgekehrte Fall – ein Wort, das in der hebräischen Bibel häufig vorkommt, ist in einer Untermenge relativ selten – macht viel mehr Schwierigkeiten, vor allem, wenn dann als Argument für Nichtzugehörigkeit zur Untermenge herhalten soll, daß dieses Wort in einem Teilstück häufiger vorkomme. Wenn man die unter 2) definierten Quotienten heranzieht und vergleicht, wird einem rasch klar, daß für das Urteil 'mit Wahrscheinlichkeit signifikant abweichend' ein objektives Kriterium erforderlich ist. Schließlich wird hier eine kleine Stichprobe mit einer sehr großen verglichen, wobei man

¹³) Um im Folgenden vereinfachen zu können, führe ich die Abkürzungen W (=Wort) und |W| (=Vorkommen eines Wortes) ein. Zusätzlich gebrauche ich das mathematische Symbol # für Anzahl.

¹⁴) Da eben der Anteil jesajanischer Stücke im Jesajabuch nicht geklärt ist und dieser Ausdruck auch von späteren Autoren, die die jesajanische Tradition aufgegriffen oder verarbeitet haben, verwendet wurde, kann in diesem Fall nicht ausgeschlossen werden, daß diese den Ausdruck an für sie wichtigen Stellen plaziert haben. Der oben geführte Schluß gilt nur für die ganze Schicht, in der das Wort signifikant häufig vorkommt, und das ist in diesem Fall eben das ganze Jesajabuch.

erst noch klären muß, wie sich das auf die jeweiligen Quotienten auswirkt (kleiner Nenner und kleiner Zähler bei der Untermenge und im Vergleich zum riesigen Nenner doch kleiner Zähler beim Quotienten, der sich auf den Rest der Bibel bezieht). Was heißt hier 'relativ selten'? Es gibt statistische Testverfahren, bei denen man die Irrtumswahrscheinlichkeit vorgeben kann (in der Regel 2,5 % oder 5 %) und dann ein Kriterium erhält, ob ein Quotient aus einer Stichprobe signifikant vom entsprechenden Quotienten einer anderen Stichprobe oder dem der Grundgesamtheit abweicht. Die Intuition allein führt hier leicht in die Irre, wie jeder Statistiker bestätigen wird.¹⁵ Hier wäre beim Beispiel Jesaja zu vergleichen:

$$\frac{|W| \text{ bei Jesaja}}{\# \{ \text{Wörter in } \bar{U} \}} \quad \frac{|W| \text{ im Rest von B (bzw. ganz B)}}{\# \{ \text{Wörter im Rest von B (bzw. ganz B)} \}}$$

und dann ein Quotiententest (oder Anteilstest) durchzuführen. Vermeylens Argumentation zum Verb עָשָׂה gehört beispielsweise in diese Schlußkategorie. Um dann bei ein- oder mehrfachem Vorkommen dieses in der Untermenge nachgewiesen signifikant seltenes Wortes in einer Perikope zu schließen, diese Perikope gehöre einer anderen Untermenge zu (in unserem Beispiel: ist nicht jesajanisch), muß eine weitere statistische Überlegung hinzukommen: Wenn der Quotient

$$\frac{|W| \text{ in } P}{\# \{ \text{Wörter in } P \}}$$

signifikant verschieden ist vom obigen Quotienten

$$\frac{|W| \text{ in } U'}{\# \{ \text{Wörter in } U' \}}$$

kann eine ähnliche Situation wie im Fall 2) vorliegen, doch je kleiner der Anteil sicher authentischer jesajanischer Texte ist, desto eher kommt man in den in der dortigen Schlußanmerkung dargelegten Grenzfall. Außerdem können auch inhaltliche Gründe die Ursache sein, daß ein bestimmtes Wort in einer Perikope häufiger benützt wird. Hier kommt die Einseitigkeit der Auswahl jesajanischer Worte in der hebräischen Bibel als Schwierigkeit noch dazu, so daß selbst die mit einer korrekten statistischen Methode errechnete Signifikanz noch einmal relativiert werden muß.

4) Am häufigsten wird in der Redaktionskritik bei wortstatistischen Überlegungen mit Lieblingswörtern oder bevorzugten Redewendungen bestimmter Redaktionsschichten gearbeitet (z.B. das sogenannte 'deuteronomistische Vokabular'). Die Sachlage ist ähnlich gelagert wie in 3). Zu vergleichen sind:

$$a) \frac{|W| \text{ in der Redaktionsschicht } U_1 \text{ (z.B. } U_1)}{\# \{ \text{Wörter dieser Schicht} \}} \quad c) \frac{|W| \text{ in } U_1 \text{ (z.B. Jesaja)}}{\# \{ \text{Wörter in } U_1 \}}$$

15) Jeder, der mit wortstatistischen Methoden arbeiten will, mache sich einmal die Mühe, in Lehrbüchern bei Aufgaben zu Quotiententests erst einmal zu raten, ob eine Abweichung signifikant ist oder nicht und sie dann erst durchzurechnen. Es wird ihn für die Zukunft von der Neigung zu vorschnellen intuitiven Vermutungen heilen.

$$b) \frac{|W| \text{ in der restlichen Bibel}}{\# \text{ W\u00f6rter in der restlichen Bibel}}$$

$$d) \frac{|W| \text{ in P}}{\# \text{ W\u00f6rter von P}}$$

Dazu sollte als Kontrolle noch der Quotient

$$e) \frac{|W| \text{ in } B \setminus \{U_1, U_2\}^{16}}{\# \text{ W\u00f6rter in } B \setminus \{U_1, U_2\}}$$

herangezogen werden. Die fragliche Wendung ist durch Vergleich der Quotienten a) und b) als signifikant f\u00fcr diese Schicht zu erweisen. Die M\u00f6glichkeiten, Grenzen und Risiken sind unter 2) genannt. Am sichersten ist es, die Quotienten a) und e) zu ben\u00fctzen. Dann ist auch ein Quotiententest \u00fcber das signifikant seltene Vorkommen in der einen gegen\u00fcber dem signifikant h\u00e4ufigen in der anderen Untermenge sinnvoll, d.h. Quotienten a) und c) sind zu vergleichen. Die entscheidende Frage ist aber, ob tats\u00e4chlich Quotient c) und d) signifikant voneinander abweichen, d.h. das Wort kommt in der Perikope gegen\u00fcber dem sonstigen Vorkommen in der Untermenge (U_i) signifikant h\u00e4ufig vor, also die unter 3) beschriebene Situation mit all ihren Unw\u00e4gbarkeiten und Risiken. Nur wenn all' das positiv beantwortet wird und anderweitige - etwa inhaltliche - Erkl\u00e4rungen in jedem Fall ausscheiden, kann man mit einigerma\u00dfen guten Gr\u00fcnden zumindest einen Teil der Perikope der anderen Untermenge (U_i) zuweisen. F\u00fcr die Praxis erhebt sich hier, abgesehen von den umstrittenen Datierungen und Zuordnungen der Redaktionsschichten das Problem, da\u00df es f\u00fcr Redaktionsschichten keine Konkordanz gibt und auch die Wortanzahlen jeweils erst festgestellt werden m\u00fcssen.

Da durch die nichtrepr\u00e4sentativen Auswahlen aus Grund- bzw. Teilmenge selbst die korrekt ermittelten statistischen Ergebnisse mit gr\u00f6\u00dferer Unsicherheit belastet und somit nur mit einer geringeren Wahrscheinlichkeit g\u00fcltig sind, erscheinen mir die ohne statistische Fundierung vorgenommenen Schl\u00fc\u00dfe nach 3) und 4), wie man sie heutzutage fast landl\u00e4ufig antrifft, wenn nicht gewichtige zus\u00e4tzliche Beobachtungen hinzukommen, abenteuerlich. Schl\u00fc\u00dfe nach 1) sind peinlich.

Ob diese \u00dcberlegungen f\u00fcr einen Nichtmathematiker hilfreich sein k\u00f6nnen, wei\u00df ich nicht. Es w\u00e4re aber schon viel gewonnen, wenn der eine oder andere Exeget daraus die Motivation mitn\u00e4hme, das Handwerkszeug "Wortstatistik" ganz vorsichtig und behutsam zu benutzen, weil sonst Fehlschl\u00fc\u00dfe geradezu vorprogrammiert sind. Hilfreicher als Beobachtungen zum rein quantitativen Gebrauch eines Wortes, sind sicher inhaltliche und am jeweiligen Kontext orientierte Beobachtungen zum Sprachgebrauch eines Wortes in verschiedenen Zeiten oder literarischen Schichten.

¹⁶Vorkommen des Wortes in der Bibel au\u00dfer der beiden fraglichen Untermengen; '\ ' ist mathematisches Symbol f\u00fcr 'au\u00dfer'.