

Bibliotheca Teubneriana Latina (BTL-2). CD-ROM mit viersprachigem Benutzerhandbuch (D, E, I, F), hrsg. von P. TOMBEUR (Centre Traditio Litterarum Occidentalium). K.G. Saur Verlag, München/Leipzig und Brepols Publishers, Turnhout 2002.

Die lateinischen Bände der Bibliotheca Teubneriana sind 1999 zum ersten Mal in elektronischer Form veröffentlicht worden; die u.a. auch in dieser Zeitschrift besprochene CD-ROM¹ ist nun in zweiter Auflage erschienen. Bereits die erste Ausgabe hat ein starkes, ganz überwiegend positives Echo gefunden, und die BTL wird zu Recht als unverzichtbar für die wissenschaftliche Arbeit bezeichnet.² Hier soll es nun darum gehen, ob dieses Potential in der erweiterten Fassung noch ausgebaut werden konnte, d.h. welche konkreten Änderungen gegenüber der ersten Auflage vorgenommen wurden. Den Hintergrund hierfür bildet eine grundsätzliche Reflexion über die Arbeit mit lateinischen Textdatenbanken: Was können klassische Philologen von einer solchen Datenbank und der entsprechenden Suchsoftware erwarten? Im Anschluß an diese prinzipiellen Überlegungen wird im Einzelnen geprüft, inwieweit die BTL-2 (gerade auch im Vergleich mit anderen Datenbanken bzw. Programmen) diesen Anforderungen bzw. Wünschen gerecht wird.³

Gegenstand der wissenschaftlichen Arbeit sind die Texte, deren sorgfältige Erfassung und möglichst umfassende Kodierung daher vorrangiges Ziel bei der Konzeption einer Textdatenbank sein muß. Die Texte sollten dem aktuellen

¹ H. LÜHKEN, GFA 2 (1999) 1137-1146; <http://www.gfa.d-r.de/2-99/luehken.pdf>); weitere im folgenden herangezogene Rezensionen: U. SCHMITZER, Gymnasium 108 (2001) 279-283 (mit Bildschirmfotos unter <http://telemachos.phil.uni-erlangen.de/database/libri/btl-rez.html>), zitiert als SCHMITZER (Rez. BTL-1); C. SCHULZE, Bochumer philosophisches Jahrbuch für Antike und Mittelalter 5 (2000) 254-259. Die BTL-2 ist ebenfalls von SCHMITZER besprochen worden, nur online zugänglich unter <http://telemachos.phil.uni-erlangen.de/database/libri/btl2-rez.html>, zitiert als SCHMITZER (Rez. BTL-2), außerdem von R. KIRCHNER für H-Soz-u-Kult (<http://hsozkult.geschichte.hu-berlin.de/rezensionen/id=1713&type=rezcdrom>).

² SCHMITZER (Rez. BTL-1) 283.

³ Hierbei werden die Datenbanken und Programme einbezogen, die dem Rez. zur Verfügung standen: Poesis-1, ed. P. MASTANDREA/L. TESSAROLO, Zanichelli Editore, Bologna 1995; Poesis-2, ed. P. MASTANDREA/L. TESSAROLO, Zanichelli Editore, Bologna 1999. Auf dieser Software basiert auch (offensichtlich mit leichten Verbesserungen) die CD-ROM Poetria Nova, von der mir bisher nur ein Prospekt und eine Autorenliste vorliegt (zu allgemeinen Informationen vgl. <http://www.sismel.it>). An Rechercheprogrammen für PHI #5.3 (hrsg. vom Packard Humanities Institute, Los Altos [California] 1991) wird herangezogen: Pandora für Apple Macintosh, Lector für DOS, Diogenes (Perl-basiert); für Informationen hierzu vgl. z.B. <http://www.gwdg.de/~hluehke/athena/CD-ROM-Datenbanken.html>.

Forschungsstand entsprechen; die Zählung sollte einheitlich nach dem Indexband des Thesaurus erfolgen (auch dann, wenn eine andere als die dort genannte Ausgabe verwendet wurde), alternative Zählungen können zur Orientierung zusätzlich beigegeben sein.⁴ Autorennamen, Pseudonyme und Werktitel sollten sowohl voll ausgeschrieben als auch mit der ThLL-Abkürzung abgespeichert sein; als weitere Angabe ist eine (zumindest ungefähre) Datierung jedes Autors bzw. Werkes unabdingbar. Eine Zuordnung zu Gattungen wäre außerdem sehr hilfreich, auch eine Grobunterteilung in Prosa und Dichtung ist für bestimmte Zwecke nützlich.⁵ Mindestens das Satzende sollte markiert sein (ggf. auch schwächere Interpunktion), bei Versen das Versende.⁶ Für das computerunterstützte Recherchieren ist außer der Satzdefinition auch die Festlegung wichtig, was genau als ein Wort angesehen wird. Hier sollten Enklitika wie selbständige Wörter behandelt werden, damit beim Suchen nicht unabsichtlich Wörter unberücksichtigt bleiben, an die ein *-que*, *-ve* o.ä. angehängt ist. Ideal für das computergestützte Arbeiten mit Texten wäre natürlich eine vollständige Lemmatisierung des Wortbestands, d.h. die Zuordnung aller Wortformen zu ihrer Grundform. Mit einer solchen Datenbank könnte man dann beispielsweise direkt nach *ferre* in allen Flexionsformen suchen, ohne sich Gedanken um die einzelnen Formen zu machen. Für eine umfassende Datenbank dürfte dies wohl außerhalb des Möglichen liegen, doch ist eine solche Arbeit immerhin für ein so großes Textcorpus wie die Werke des Augustin bereits in einem sehr fortgeschrittenen Stadium.⁷

⁴ Dies empfiehlt sich vor allem bei Autoren, die im Indexband des ThLL nicht besonders benutzerfreundlich gezählt werden, wie beispielsweise die Werke des Ausonius oder die Hymnen des Ambrosius. Daß die verwendete Ausgabe bibliographisch exakt vermerkt sein sollte, versteht sich von selbst; diese Angaben sollten möglichst einfach abgerufen und auch beim Drucken und Exportieren auf Wunsch mit ausgegeben werden können.

⁵ Weitere Angaben wie etwa die geographische Zuordnung, die beim TLG noch zusätzlich erfaßt wird, sind für die lateinische Literatur weniger wichtig. Hier könnten, gerade in der Spätantike, eher Faktoren wie die Haltung zum Christentum eine Rolle spielen, die sich aber nur schlecht in starre Schemata einordnen lassen.

⁶ Eine noch weitergehende Kodierung, die bei jedem Wort den Sitz im Vers erfaßt, würde wohl den Rahmen sprengen. Der dem Rez. vorliegende Prospekt zur CD-ROM *Poetria Nova* (vgl. oben Anm. 3) deutet darauf hin, daß hier immerhin Hexameter und Pentameter markiert wurden, was sehr hilfreich sein kann und mit vertretbarem Aufwand realisierbar wäre.

⁷ Das *Corpus Augustinianum Gissense* auf CD-ROM, von dem in Kürze eine erweiterte Fassung zu erwarten ist, bietet außer dem vollständigen Augustinustext und einem umfassenden Wortindex auch eine Rohlemmatisierung des gesamten Wortbestandes (mit nur geringen Einschränkungen); vgl. die Website des Projekts (<http://www.augustinus.de>). Eine einfacher zu verwirklichende Alternative sind intelligente Suchalgorithmen bzw. die Verwendung von Indizes bei der Suche.

Für die so aufbereiteten Texte muß dem Benutzer nun ein Programm zur Verfügung gestellt werden, das eine Vielzahl von unterschiedlichen Aufgaben zu erfüllen hat. Die Software sollte über einen möglichst langen Zeitraum für einen möglichst großen Benutzerkreis verfügbar sein. Die Beschränkung auf ein einziges Betriebssystem scheint dabei angesichts der rasanten technischen Entwicklung von Hard- und Software ebensowenig ratsam wie die Verbreitung in nur einem einzigen Medium. Die Datenbank und das sie erschließende Programm sollten daher idealerweise betriebssystemunabhängig sein oder zumindest für mehrere Systeme parallel entwickelt werden. Auch die Möglichkeit, sie in Computernetzwerken zu benutzen, sollte unbedingt implementiert sein. Daß die Arbeit mit dem Programm nicht nur einfach und komfortabel, sondern auch schnell vonstatten gehen sollte, liegt auf der Hand.

Die wohl einfachste Form des Zugriffs auf eine solche Volltextdatenbank ist das gezielte Aufschlagen einer bestimmten Stelle. Hier sollte man eine einfache Möglichkeit vorsehen, ein Werk „aufzuschlagen“, indem man eine bestimmte Stelle (z.B. definiert durch Buch, Kapitel, Paragraph; Buch und Vers; ggf. auch Seite und Zeile) angibt. Eine solche Stelle sollte auch (mit angemessenem Kontext) exportiert und gedruckt werden können.⁸

Für sprachliche Untersuchungen wichtiger sind die Recherchemöglichkeiten, die eine Volltextdatenbank bietet. Für die Suche nach Wörtern oder Wortteilen sind umfassende Möglichkeiten erforderlich, sogenannte wildcards zu setzen, d.h. Platzhalter für einen oder mehrere Buchstaben. Diese sollten an beliebiger Stelle plaziert werden können (auch mehrfach in einem Wort); insbesondere die Trunkierung von Wörtern sollte links, rechts oder an beiden Stellen möglich sein. In bezug auf verschiedene Schreibvarianten wäre es gut, die strenge Logik etwas zu lockern, so daß beispielsweise *u* und *v* als äquivalent angesehen werden können (ebenso *i* und *j*, Akkusativendungen auf *-es* und *-is*) und Wörter mit Assimilationen wie *-mm-* auch dann gefunden werden, wenn im Text *-nm-* geschrieben ist etc. Wünschenswert wäre auch die Möglichkeit, gezielt nach Wortwiederholungen, Alliterationen etc. zu suchen.

Häufiger als nach Einzelwörtern wird man nach Kombinationen von Wortformen suchen, etwa wenn es um Belege für eine bestimmte Junktur geht. Dafür muß ein variabler Abstand angegeben werden können, innerhalb dessen die Wörter bzw. Wortbruchteile im Text stehen. Hierfür ist es aus philologischer Sicht vernünftig, eine bestimmte Anzahl von Wörtern als Abstand anzu-

⁸ Dies ist auch deshalb wichtig, weil derartige Volltextdatenbanken schon jetzt immer stärker den Charakter einer elektronischen Bibliothek annehmen, in der schwer zugängliche oder gar nicht mehr aufgelegte Werke zur Verfügung stehen.

geben oder festzulegen, daß die Suchwörter innerhalb eines Satzes (ggf. auch Teilsatzes, wenn Kommata o.ä. berücksichtigt sind) stehen sollen; in der Dichtung kann auch die Eingrenzung des Suchkontexts auf einen oder mehrere Verse sinnvoll sein. Nicht immer wird man im gesamten Textcorpus recherchieren wollen. Eine Auswahl der zu durchsuchenden Autoren und Werke sollte auf mehrere Arten möglich sein. Sinnvoll sind hierbei die (bei der Erstellung der Textgrundlage bereits zu berücksichtigenden) Kriterien Chronologie und Gattungszugehörigkeit. Zusätzlich sollte man manuell eine Liste nach eigenen Kriterien zusammenstellen und auch für spätere, wiederholte Suchen abspeichern können.

Die Ergebnisse, die eine solche Suche erbracht hat, müssen nun in übersichtlicher Form präsentiert und zur Weiterverarbeitung in eine Datei exportiert oder gedruckt werden können. Nur dann kann man etwa mit einer Liste von Suchergebnissen die einzelnen Fundstellen genauer überprüfen bzw. die Ergebnisse für die eigene Arbeit nutzbar machen. Für die Anzeige am Bildschirm ist es zunächst hilfreich, zwischen einer Kurz- und einer Langform wechseln zu können. So gewinnt man anhand einer einfachen Liste, die z.B. nur die Fundstellen (idealerweise mit nach ThLL abgekürzten Stellenangaben) mit jeweils nur sehr kurzem Kontext (etwa 1 Zeile) enthält, leichter einen Überblick darüber, ob man die Suche ggf. weiter einschränken muß, um nicht eine zu große Zahl von Ergebnissen durchsehen zu müssen. Von dieser Kurzform aus sollte man in der Lage sein, entweder schnell in die ausführlichere Liste zu wechseln (die dann einen größeren Kontext enthält) oder den Kontext einer einzelnen Stelle anzuzeigen. Für das Drucken und Exportieren der eventuell weiter manuell selektierten Ergebnisse sollte dann ebenfalls zwischen Kurz- und Langform gewählt werden können, möglichst mit Abstufungen (Größe des Kontexts, bibliographische Angaben ja/nein, Autoren und Werke abgekürzt oder ausgeschrieben). Die Reihenfolge der Ausgabe sollte chronologisch, alphabetisch oder nach Maßgabe der manuell erstellten Autorenliste möglich sein, damit man z.B. die Ergebnisse direkt sprachgeschichtlich untersuchen kann. Beim Druck sollten zumindest elementare Formatierungsmöglichkeiten (Zeichensatz, Punktgröße, Zeilenabstand, Ränder; ggf. auch Seitennumerierungen, eigene Kopfzeilen etc.) gegeben sein; beim Exportieren sollte zwischen verschiedenen Dateiformaten (ein einfaches Textformat ohne weitere Auszeichnungen, ein Austauschformat wie RTF etc.) gewählt werden können.

In elektronischer Form vorliegende Daten können relativ leicht auch statistisch weiter ausgewertet werden. Außer einer Exportfunktion, die eine solche Weiterverarbeitung ermöglicht, wäre es daher zu begrüßen, wenn man zumindest

rudimentäre Daten auch direkt aus dem Suchprogramm heraus abfragen könnte. Angaben wie etwa die relative und/oder absolute Häufigkeit von Worten oder Wortverbindungen bei einzelnen Autoren oder in bestimmten Gruppen von Autoren oder Werken sind ohne großen Aufwand zu ermitteln. Mit Hilfe einer möglichst anschaulichen Präsentation kann auch der Laie solche Ergebnisse leicht interpretieren. Eine weitere Anwendungsmöglichkeit ist das Erstellen von Indizes und Konkordanzen. Auch dies ist programmieretechnisch relativ leicht zu bewerkstelligen. Mit einer „idealen“ lateinischen Textdatenbank sollte man daher auch auf einfache Wortindizes und Konkordanzen (z.B. im Format keyword-in-context) zugreifen können. Sinnvoll wäre dabei eine Verknüpfung mit statistischen Funktionen und dem Volltext: So sollte man die Konkordanz an einer bestimmten Stelle „aufschlagen“ können, statistische Informationen zum entsprechenden Wort vorfinden, die Belege einfach durchsehen und auf den Volltext an der entsprechenden Stelle zurückgreifen können.

Inwieweit erfüllt nun die Bibliotheca Teubneriana Latina diese Anforderungen an eine „ideale“ lateinische Volltextdatenbank?⁹ Daß sich für das Lateinische keine dem TLG vergleichbare Institution gefunden bzw. gegründet hat, ist bereits verschiedentlich bedauert worden.¹⁰ Auch wenn es dem Verlag B.G. Teubner in Zusammenarbeit mit Brepols gelungen ist, die für die abgesteckte Epoche umfangreichste lateinische Textdatenbank zu begründen (eine Entwicklung, die dank des Engagements des Saur Verlags erfreulicherweise noch intensiviert werden soll), erfolgt somit die Auswahl der verwendeten Textausgaben nicht nach philologischen Kriterien, sondern nach verlegerischen Gesichtspunkten. Daß dadurch auch bei wichtigen Autoren wie Vergil oder Terenz teilweise deutlich veraltete Ausgaben zur Anwendung kommen, ist bereits verschiedentlich kritisiert worden.¹¹ Hieran wird sich jedoch nicht viel ändern lassen, es sei denn, einige weitere Verlage könnten für eine Kooperation gewonnen werden, was jedoch angesichts der harten Konkurrenzsituation in einem solch kleinen Markt kaum vorstellbar erscheint. Hier bleibt nur festzuhalten, daß die jetzt vorliegende Datenbank das beste ist, was unter den

⁹ Hier ist natürlich zu unterscheiden zwischen solchen Punkten, die einfach durch Geschichte und Anlage des Projekts bedingt und somit relativ unabänderlich sind und solchen, die durch Modifikationen an Textbestand und Software verändert, d.h. verbessert werden können.

¹⁰ Vgl. z.B. SCHMITZER (Rez. BTL-1) 279.

¹¹ Als Konsequenz daraus hat man in Einzelfällen einen vom gegenwärtigen Standard stark abweichenden Text auf der CD-ROM, wie SCHMITZER (Rez. BTL-2) anschaulich vorführt. Bei dem von ihm gewählten Beispiel Vergil fällt außerdem auf, daß RIBBECK z.B. Aen. 2,114 und 11,89 *ecus* statt *equus* schreibt – wer denkt wohl an solche Schreibvarianten, wenn er eine Junktur mit *equus* auf der CD-ROM recherchiert?

gegebenen Umständen zu erwarten bzw. zu realisieren ist. In der neuen Auflage ist die Textgrundlage gegenüber BTL-1 um 190 neue Werke erweitert worden, so daß der Gesamtbestand an enthaltenen Wörtern um etwa ein Viertel auf nun ca. 8,45 Millionen angewachsen ist.¹² Unter den Neuaufnahmen besonders hervorzuheben ist die Eingliederung der kompletten Grammatici Latini in die Datenbank – damit ist die BTL-2 auf dem Gebiet der heidnischen Spätantike konkurrenzlos.

Doch gibt es im Bereich der Textkonstitution auch vermeidbare Einschränkungen: So fehlen entgegen der Ankündigung bisher u.a. die Scholien zu Horaz, Juvenal, Lucan und Persius.¹³ Auch sonst vermißt man auf der CD-ROM einige Autoren und Werke, von denen Teubner-Ausgaben vorliegen, so z.B. die Panegyrici Latini (ed. BAEHRENS, Leipzig 1874/1911) oder die Fragmente von Sueton (ed. REIFFERSCHIED, Leipzig 1860; auf PHI #5.3 enthalten).¹⁴ Auch sind nicht, wie ursprünglich vorgesehen, alle Neuerscheinungen der Jahre 1998 bis 2000 enthalten: Die neue Ausgabe von Boëthius' *Consolatio philosophiae* (ed. MORESCHINI, München/Leipzig 2000) kam wohl zu spät, um noch einge-

¹² Eine Liste der enthaltenen Autoren und anonymen Werktitel ist im Prospekt zur CD-ROM enthalten, der auch unter http://www.saur.de/pdf/40503_d.pdf im Internet verfügbar ist. Die Zahlenangaben sind diesem Prospekt entnommen. Nur zum Vergleich: Die im Jahr 2000 erschienene Version E des TLG enthält laut Auskunft der Website (<http://www.tlg.uci.edu>) etwa 72 Millionen Wörter; der enorme Unterschied ist nicht nur durch die etwas andere Zielsetzung bedingt – der TLG soll die griechische Literatur möglichst bis in die Gegenwart erfassen, während die BTL u.a. die christliche Literatur weitgehend ausklammert –, sondern auch dadurch zu erklären, daß die Arbeit am TLG bereits 1972 begann, während die BTL ein relativ junges Projekt ist.

¹³ Für diesen auf der BTL-2 nun so gut ausgebauten Bereich der grammatischen Texte aus der Spätantike würde man sich außerdem auch den Terenzkommentar des Donat (ed. WESSNER, Leipzig 1902-1908) wünschen, ebenso den Kommentar des Lactantius Placidus zur Thebais des Statius (ed. SWEENEY, Stuttgart/Leipzig 1997). Auch die Interpretationes Vergilianae des Tiberius Claudius Donatus (ed. GEORGII, Leipzig 1905) wären trotz ihrer vergleichsweise geringen Bedeutung keineswegs überflüssig.

¹⁴ Weitere Autoren und Werke, die Eingang hätten finden können, nennt z.B. KIRCHNER zu Beginn seiner Rez. (allerdings sind die von ihm in Fußnote 8 genannten weiteren Ausgaben z.T. noch gar nicht erschienen oder so neu, daß sie ohnehin erst für BTL-3 zu erwarten sind). Diesen Auslassungen steht eine zwar nicht störende, aber doch auffallende Ergänzung gegenüber: Wie kommt Iulianus Toletanus (ausgehendes 7. Jh.) auf die BTL-2, wo doch die Ausgabe (M.A.H. MAESTRE YENES, Toledo 1973, Publicaciones del Instituto Provincial de Investigaciones y Estudios, serie secunda, vol. 5) nicht zur Bibliotheca Teubneriana gehört? Daß im kleinen Rahmen Ergänzungen möglich sind, zeigt auch die bereits auf der BTL-1 verwendete Columella-Ausgabe (ed. V. LUNDSTRÖM, A. JOSEPHSON, S. HEDBERG, Uppsala 1897-1968). Hier wäre noch eine ganze Reihe von Kooperationen wünschenswert.

arbeitet werden zu können.¹⁵ Doch ist z. B. Valerius Maximus nur in der alten Edition von KEMPF (1888) enthalten, obwohl der neue Text von BRISCOE 1998 erschienen ist.¹⁶ Für die bereits geplante BTL-3 sollte also außer den angekündigten zusätzlichen Werken aus Mittelalter und Neuzeit sowie BT-Neuerscheinungen auch eine ganze Reihe weiterer Aufnahmen erfolgen, um diese Lücken zu schließen. Schließlich ist für sprachliche Untersuchungen ein möglichst vollständiges Textcorpus von großer Wichtigkeit.

Abgesehen von diesen Lücken sollten jedoch auch einige vorhandene Texte manuell überarbeitet werden: Die Epitoma des Florus wird z.B. nach der Zweiteilung zitiert, dabei sind jedoch nur das Buch sowie Seite und Zeile angegeben, nicht aber das jeweilige Kapitel. Die ebenfalls überlieferte Einteilung in vier Bücher, die sowohl im ThLL-Index als Referenz verzeichnet ist als auch z.B. von K. SALLMANN befürwortet wird (HLL 4, München 1997, § 462.B.1, S. 330), wird als Zusatz in Klammern mit dem Vermerk „olim“ beigegeben. Schlimmer ist es bei den Briefen Cic. ad Brut., die nur nach der Zählung von SHACKLETON BAILEY angegeben werden – im Thesaurusindex wird hier die OCT-Ausgabe von WATT (1958) zugrunde gelegt. Auch die Verwunderung über die Entscheidung, Horaz nicht in der ebenfalls bei Teubner erschienenen Ausgabe von BORZSÁK zu bieten, kann hier nur erneut zum Ausdruck ge-

¹⁵ Calvetus, De rebus Indicis (ed. MARTOS, München/Leipzig 1998) wurde dagegen vermutlich bewußt noch nicht aufgenommen und dürfte mit weiteren Werken des Mittelalters und der Neuzeit für die BTL-3 erwartet werden.

¹⁶ Auch die Einarbeitung neuerer, korrigierter Auflagen ist nicht vollständig erfolgt: Für Manilius wird im Memento nur die erste Auflage 1985 angegeben, die Änderungen in der zweiten Auflage 1998 (vgl. dazu R. MAYER, CR N.S. 49 [1999] 267) sind jedoch eingearbeitet; allerdings ist die Einfügung des *hic* durch BENTLEY in 2,193 auf der CD-ROM nicht als solche gekennzeichnet. Bei Tibull weicht die zweite Auflage 1998 an etlichen Stellen von der ersten ab (vgl. die Rezension von J.L. BUTRICA im Bryn Mawr Classical Review, <http://ccat.sas.upenn.edu/bmcr/1999/1999-03-19.html>), hier bietet die BTL-2 nur die erste Auflage. Bei Velleius Paterculus dagegen steht im Memento lediglich die Jahreszahl 1998; daß es sich um eine zweite Auflage handelt, ist nicht ersichtlich (die Änderungen gegenüber der Erstausgabe von 1988 bei R. MAYER, CR N.S. 49 [1999] 270; offensichtlich ist tatsächlich die Neuauflage verwendet worden). Auch bei Lucan ist die zweite Auflage der Edition von SHACKLETON BAILEY (Stuttgart/Leipzig 1997) nicht eingearbeitet worden (vgl. die Praefatio und Lucan. 6,663), obwohl der Druckfehler in 3,26 (*et nullam* auf der CD-ROM gegenüber *et nullum* in beiden Auflagen der gedruckten Ausgabe) berichtigt ist. Unberichtigt geblieben ist dagegen der Fehler in Lucan. 9,681 (dort ist ein *-que* zu viel). Auf das grundsätzliche Problem, das diese Unterschiede zwischen gedruckten und elektronischen Fassungen aufwerfen, hat SCHMITZER (Rez. BTL-2) zu Recht hingewiesen. Seine Anregung, Listen solcher Corrigenda leicht zugänglich zu veröffentlichen (am besten im WWW und auf den CD-ROMs im jeweiligen Memento), sollte dringend aufgenommen werden.

bracht werden.¹⁷ Die eigenartige Groß- und Kleinschreibung in Caesars *Bellum civile* erhellt ein Blick in die verwendete Ausgabe: Alle Worte, die am Zeilenanfang stehen, sind in der CD-ROM-Fassung groß geschrieben worden.¹⁸ Außerdem verdient die Zeichensetzung mehr Beachtung: Zwar ist es prinzipiell wünschenswert, im Zweifelsfall lieber einen größeren Textabschnitt als einen Satz zu definieren, doch hat die auf der BTL-1 und BTL-2 angewandte Methode zum Teil extreme Konsequenzen: So führt die Zeichensetzung mancher Herausgeber zu „Sätzen“ von über 100 Versen Länge wie Ov. Ib. 255-418 oder Ov. Ib. 419-538.¹⁹ Die Definition der einzelnen Wortformen dagegen ist vorbildlich: Enklitika werden wie einzelne Wörter behandelt, was manche Arbeit beim Trunkieren erspart.²⁰ Die Gruppierung der Texte nach Gattungen, die auf der BTL-1 begonnen wurde, ist nicht weitergeführt worden; somit sind diese Möglichkeiten einer Eingrenzung des zu durchsuchenden Textmaterials, die ohnehin in der jetzigen Form keineswegs zufriedenstellen, weiterhin auf Cicero und Seneca beschränkt.²¹

Die Software ist ein Faktor, der sehr viel leichter und umfassender beeinflusst werden kann als der zur Verfügung stehende Textbestand. Ein sehr großer Teil der Anforderungen, die aus philologischer Sicht an die Datenbank gestellt werden, kann unabhängig von Fragen der Textkonstitution von dem Programm erfüllt werden; auch sind Erweiterungen der Software unabhängig von der Datenbank möglich.²² Dies gilt indes nicht für die grundsätzlichen Fragen nach Betriebssystem und Netzwerkfähigkeit; hier bleibt nur zu hoffen,

¹⁷ Vgl. bereits SCHMITZER (Rez. BTL-1) 282f.; SCHULZE 257. Gerade eine konservative Ausgabe wie die von BORZSÁK hätte sich für eine solche Datenbank eher angeboten. Vielleicht wäre es sogar möglich gewesen, beide Ausgaben aufzunehmen und dem Anwender selbst die Entscheidung zu überlassen, welche er bevorzugt. Gleiches gilt außer für Catull (vgl. LÜHKEN 1138 Anm. 7) übrigens auch für Valerius Flaccus, von dem außer der für die CD-ROM verwendeten Ausgabe von EHLERS (Stuttgart 1980) auch eine Leipziger Teubneriana von COURTNEY aus dem Jahr 1970 vorliegt (vgl. allerdings die Kritik von EHLERS an dieser Ausgabe *Gnomon* 48 [1976] 255-260). Mindestens wäre hier Nachbearbeitung nötig gewesen: Die von A.J. KLEYWEGT in seiner Rezension (*Mnemosyne* Ser. IV 41 [1988] 202-206, hier 206) aufgeführten Druckfehler sind nicht bereinigt (z.B. 1,277 *testitudine* statt *testudine*; Auslassung von *terras* in 1,567 etc.).

¹⁸ Ähnlich entstellt erscheint der Text der Appendix Sallustiana, wo (fast?) alle Kommata offensichtlich bei der Aufnahme in die Datenbank ausgefallen sind. Dies hat zwar keine Auswirkungen auf das Durchsuchen des Texts, macht diesen aber schwer handhabbar.

¹⁹ Ähnlich schon SCHMITZER (Rez. BTL-2).

²⁰ Dies hebt die BTL wohlthuend sowohl von PHI #5.3 als auch von der Poesis-CD-ROM ab.

²¹ Vgl. LÜHKEN 1143 Anm. 22.

²² Vorbildlich ist in dieser Hinsicht beispielsweise die Entwicklung der Software für die „Digitale Bibliothek“ von Directmedia Publishing, Berlin: Die aktuelle Version ist stets im Internet unter <http://www.digitale-bibliothek.de> abrufbar, so daß jeder Anwender von der fortwährenden Weiterentwicklung profitieren kann.

daß es bei den derzeitigen Beschränkungen nicht bleibt bzw. auf Veränderungen der Gegebenheiten schnell genug reagiert werden kann.²³

Größtes Manko bleibt nach wie vor das Fehlen eines einfach zu handhabenden „Browsers“, mit dem die Texte analog zu einem gedruckten Buch erschlossen werden können. Hier ist eine Nachbesserung unabdingbar. Diese hätte eigentlich bereits für die vorliegende Fassung vorgenommen werden müssen und sollte spätestens mit der BTL-3 erfolgen.²⁴

Für die konkreten Recherchemöglichkeiten gilt, daß alles, was bisher möglich ist, zuverlässig und vor allem sehr schnell vonstatten geht.²⁵ Jedoch sind insbesondere hinsichtlich der Komplexität von Suchanfragen oft allzu enge Grenzen gesetzt: Gerade angesichts der enormen Suchgeschwindigkeit²⁶ ist unverständlich, warum nicht mehr Verschachtelungen möglich sind.²⁷ Rechtsschreibvarianten können so bei den Suchen nur eingeschränkt berücksichtigt werden. Als vorbildlich in diesem Bereich kann Lector hervorgehoben werden, das für verschiedene Schreibungen vier abgestufte Toleranzgrade anbie-

²³ Wenn man den zahlreichen Beiträgen in Newsgroups Glauben schenken kann, scheint es recht schwierig zu sein, die BTL-2 in einem Netz zu betreiben, hierzu kann der Rez. aber mangels eigener Erfahrung keine Stellung nehmen. Die Beschränkung auf Windows wurde bereits früher moniert; weitsichtig sind hier nur TLG und PHI, die mit ihrem offenen Konzept im Prinzip von jedem Betriebssystem genutzt werden können. Angesichts der Entwicklung scheint es ratsam, derartige Datenbanken auch über eine WWW-Schnittstelle verfügbar zu machen, wie es beispielsweise beim TLG schon realisiert worden ist (<http://www.tlg.uci.edu>). Auch der Saur Verlag bietet einige Datenbanken über das Internet an, mit der Technik sind also dort schon Erfahrungen vorhanden.

²⁴ Hierfür müßte lediglich auf der Registerkarte *textus* eine entsprechende Navigation verfügbar gemacht werden.

²⁵ Ein kleiner Fehler ist allerdings beim Suchen von links trunkierten Wortformen zu beobachten: Steht direkt nach dem gesuchten Wort ein Fragezeichen im Text, wird dieses bei den *sententiae* nicht farbig hervorgehoben (man suche z. B. nach **erri* bei Lucan, das erste Ergebnis ist Lucan. 1,8 *quae tanta licentia ferri*, das Wort *ferri* ist nicht hervorgehoben).

²⁶ Nur wenn Worte links trunkiert werden, benötigt die Software etwas länger für eine Suche, alle anderen Rechercheaufträge bewegen sich im Sekundenbereich. Hinsichtlich der Geschwindigkeit ist eine vergleichbare Software für PHI #5.3 dem Rez. noch nicht bekannt, die Poesis-2 arbeitet, wie bereits die Vorgängerversion, ähnlich schnell.

²⁷ Schon eine so einfache Anfrage wie *in %3 v?ltu*, womit man nach den im Abstand von drei Worten in dieser Reihenfolge aufeinanderfolgenden Worten *in* und *vultu* oder *voltu* suchen könnte, wird als zu komplex abgelehnt. Auch die Beschränkung, die Operatoren *%* und */* nicht in direktem Kontakt mit Klammern verwenden zu können, sollte dringend aufgehoben werden (vgl. bereits LÜHKEN 1141f.). Die Möglichkeiten von Diogenes in diesem Bereich sind dagegen äußerst vielfältig, doch ist die Suchsyntax und die Handhabung dieses Programms wenig anwenderfreundlich.

tet. Auch die Suche nach Wiederholungen von Wörtern²⁸ sollte erleichtert werden.²⁹ Daß als Suchkontext eine bestimmte Anzahl Wörter oder ein ganzer Satz angegeben werden kann,³⁰ ist vernünftig, jedoch vermißt man für Suchen in der Dichtung die Möglichkeit, auch nach Junktoren innerhalb eines oder mehrerer Verse suchen zu können.³¹

Für die Präsentation der Ergebnisse steht lediglich eine Langform zur Verfügung; hier wäre zusätzlich eine Kurzform wünschenswert, die gut als eine weitere Registerkarte in das Programm integriert werden könnte.³² Ebenfalls sollte unbedingt die Angabe der Fundstellen verbessert werden: Immer noch

²⁸ Hier setzt ebenfalls Lector Maßstäbe: Außer der Möglichkeit, nach Wiederholungen von beliebigen Wörtern suchen zu können (und nicht nur nach der konkreten Wiederholung eines bestimmten!), kann hier sehr komfortabel nach beliebigen Alliterationen (die Anzahl gleicher Buchstaben ist ebenso frei wählbar wie die der Wiederholungen) und sogar Häufungen von Buchstaben recherchiert werden.

²⁹ Das von mir in der Besprechung der BTL-1 herausgegriffenen Beispiel, nach der Gemination *fuit, fuit* bei Cicero zu suchen, hat SCHULZE in seiner Rezension aufgenommen und die vorhandene Möglichkeit als ausreichend bezeichnet (S. 256f. Anm. 8). Dies erscheint mir aus zweierlei Gründen unangemessen: Erstens ist es meiner Ansicht nach der falsche Ansatz, derlei stupide und fehlerträchtige Arbeiten selbst erledigen zu wollen. Genau für solche Aufgaben ist der Computer bzw. das Suchprogramm zuständig und nicht der Benutzer; von der Software kann und muß verlangt werden, daß solche Wiederholungen zuverlässig herausgefiltert werden. Zweitens mag man in diesem konkreten Fall mit zwei Minuten lästigem Blättern (und, im Falle der BTL, handschriftlichem Notieren der ausgesonderten Stellen!) auskommen. Doch abgesehen davon, daß man seine Zeit sehr viel sinnvoller verbringen kann (die Software würde für diesen Vorgang wohl deutlich unter einer Sekunde benötigen), lassen sich leicht Beispiele finden, in denen die Durchsicht der Ergebnisse wesentlich länger dauern würde.

³⁰ Die Problematik zu lang definierter Sätze gehört in den Bereich der Textkonstitution, vgl. oben S. 1144.

³¹ Diese bietet z.B. Lector, das es darüber hinaus auch noch erlaubt, gezielt nach Wörtern am Anfang oder Ende von Sätzen bzw. Versen (= Zeilen) zu suchen.

³² Analog zu den anderen Registerkarten könnte diese etwa *loci* genannt werden. Vorbildlich ist hier die Poesis-2, die in Weiterentwicklung der ersten Fassung nun als Abkürzungen nicht mehr die gewöhnungsbedürftigen Siglen aus P. MASTANDREA, *De fine versus* (Hildesheim etc. 1993) verwendet, sondern Autoren und Werke nach ThLL abkürzt (mit leichten Abweichungen u.a. bei der Appendix Vergiliana und der Alcestis von Barcelona). Als Kontext wäre hier ein Vers bzw. bei Prosawerken eine Druckzeile sinnvoll; für das Wechseln zwischen *loci*, *sententiae* und *textus* sollten eingängige Tastenkombinationen definiert werden (wie bereits jetzt für das Wechseln zwischen *sententiae* und *textus*). Hier ist bei der Poesis-2 ein gewisser Rückschritt gegenüber der alten DOS-Version zu konstatieren: Das dort noch sehr praktische Wechseln zwischen Kurz- und Langform mittels Tastenkombinationen ist in der neuen Version weggefallen; das ständige Öffnen neuer Fenster mittels Mausclick ist kein Fortschritt. Bei der BTL dagegen ist die Handhabung mit den einzelnen Registerkarten praktisch und sollte beibehalten werden. Auch Pandora bietet für das Suchen auf der PHI-CD-ROM eine einfache Möglichkeit, zwischen Kurzanzeige und Kontext zu wechseln.

verweist das Programm den Anwender auf den entsprechenden Satzanfang, was mühevoll Auszählen des genauen Verses zur Folge hat, wenn sich das gesuchte Wort in einem Satz befindet, der mehrere Verse umfaßt.³³ Diese Art der Präsentation von Ergebnissen hat außerdem eine weitere unglückliche Konsequenz: Kommt ein gesuchtes Wort oder eine recherchierte Junktur gleich mehrfach innerhalb eines Satzes vor, so werden diese Belege als ein einziges Ergebnis angezeigt.³⁴ Auch die Reihenfolge der Ergebnisse³⁵ kann nicht befriedigen; ferner vermißt man die Möglichkeit, sie umzusortieren bzw. irrelevante Ergebnisse vor der Weiterverarbeitung der Daten manuell auszusortieren. Die allzu rigide Einschränkung hinsichtlich der Druck- und Exportmöglichkeiten ist ebenfalls nicht gelockert worden – auch hier sollten die Möglichkeiten spätestens bei der BTL-3 erweitert werden. Schließlich bleiben elementare statistische Funktionen ein Desiderat. Mustergültig ist auch hier die Poesis-2 mit einer komfortablen Möglichkeit, einfache Übersichten zu erstellen. Gleiches gilt für die dort verfügbaren Wortindizes zu den einzelnen Autoren.³⁶ Dieser letzte Erweiterungsvorschlag ist jedoch zweitrangig, zunächst sollte sich die Weiterentwicklung auf wichtigere Dinge konzentrieren.

Um nicht einen falschen Eindruck zu erwecken: Die BTL-2 kann mit Fug und Recht als ein Meilenstein in der Entwicklung lateinischer Volltextdatenbanken bezeichnet werden. Die bisher verfügbaren Funktionen sind grundsätzlich begrüßenswert, sollten aber weiter ausgebaut werden. Die Textbasis ist im großen und ganzen gut, zumal der Umfang beträchtlich gewachsen ist. Dieser Fortschritt sollte beibehalten bzw. weiter ausgebaut werden. Es warten noch viele Teubnerausgaben auf ihre Eingabe in die Datenbank. Die vorhandenen Texte müssen in manchen Fällen noch nachbearbeitet werden, insbesondere die Dokumentation hinsichtlich der Abweichungen von den gedruckten Ausgaben sollte genauer ausgeführt werden. Für die Software müssen – abgesehen von der dringend einzufügenden Funktion, gezielt Stellen anzusteuern – vor allem einige Beschränkungen aufgehoben werden (Komplexität der Suchanfragen, Druck und Export). Insgesamt stellt die BTL-2 faszinierende Möglichkeiten zur Verfügung; man darf auf die nächste Auflage gespannt sein. Und noch ein Lob ist angebracht: Angesichts der enormen Erweiterung

³³ Sucht man beispielsweise nach der Formulierung *altae moenia Romae*, wird als Fundstelle bei Vergil Aen. 1,1 angegeben. Derlei Angaben schränken den Nutzen der Ergebnislisten stark ein, hier ist eine Nachbesserung dringend geboten.

³⁴ Dieser Fehler ist auch bei der Poesis-2 festzustellen.

³⁵ Vgl. bereits LÜHKEN 1143.

³⁶ Was elektronische Konkordanzen betrifft, sei hier nur kurz auf das kostenlos verfügbare Programm conc für Apple Macintosh verwiesen, an dem sich die Konkordanzfunktion innerhalb einer Textdatenbank orientieren sollte (vgl. <http://www.sil.org/computing/conc>).

ist die Verzögerung gegenüber dem Zeitplan nur minimal. Möge es der BTL-3 und eventuell weiteren Nachfolgern ähnlich ergehen.

Henning Lühken
Seminar für Klassische Philologie
Universität Göttingen
Humboldtallee 19
D-37073 Göttingen
e-mail: hluehke@gwdg.de