

Jonas Richter

Varianz im *Mittelhochdeutschen Wörterbuch*. Herausforderungen und Möglichkeiten der Darstellung

Abstract: Diatopic and diachronic variation in Middle High German is rarely addressed explicitly in the new *Mittelhochdeutsches Wörterbuch* (MWB). The need to base the dictionary on edited texts makes it difficult for the lexicographical team to cover varietal specifics of MHG with certainty. This paper analyzes how variation is mentioned in the MWB and, based on examples of unpublished entries, discusses the problems and limits of covering MHG variation in the dictionary. A final section describes the public presentation of the digital archive of lemmatized references that is available at MWB Online, and outlines a project to enrich the list of primary sources with dialectal and temporal details about the texts. Although neither of the two are at the heart of the dictionary project, both promise to mitigate the scarcity of information on variation in the MWB.

Schlagwörter: Lexikographie, Mittelhochdeutsch, Variation, Varianz, Symptomwertangaben

1 Einleitung

Das Mittelhochdeutsche ist eine variantenreiche Sprache. Unterschiede, beispielsweise zwischen regionalen Ausprägungen, sozialen Milieus oder im zeitlichen Wandel, gibt es auf allen sprachlichen Ebenen (phonologisch, grammatisch, lexikalisch usw.). Diese Variation zu erfassen und zu beschreiben ist eine der Aufgaben germanistischer Mediävistik. Die Erwartung liegt nahe, dass auch ein Wörterbuch sich an dieser Arbeit beteiligt. In welcher Weise dies im *Mittelhochdeutschen Wörterbuch* (MWB), einem gemeinsamen Projekt der Göttinger und der Mainzer Akademien, geschieht und welche Schwierigkeiten dabei auftreten, ist Thema dieses Werkstattberichts.

Es wird meist differenziert nach diatopischer (regionaler), diachronischer (zeitlicher), diastratischer (sozialer) und diaphasischer (kommunikativ-situativer) Variation. Durch die Überlieferungssituation des Mittelhochdeutschen sind natürlich längst nicht alle Arten von Varianz erfassbar. Die erhaltenen Quellen decken nur

Dr. Jonas Richter: *Mittelhochdeutsches Wörterbuch*, Akademie der Wissenschaften zu Göttingen, Geiststraße 10, 37073 Göttingen, jrichte5@gwdg.de

einen kleinen Ausschnitt von Texttypen bzw. kommunikativen Anlässen und natürlich auch nur bestimmte soziale Gruppen der Sprachgemeinschaft ab. Diastatische Unterschiede beispielsweise werden bestenfalls in literarisch pointierter Form sichtbar, z.B. im Flämeln des Protagonisten im „Helmbrecht“ Wernhers des Gärtners. Das Streben des jungen Helmbrecht nach sozialem Aufstieg findet hier Ausdruck durch Verwendung fremdsprachlicher Versatzstücke. Aber textkritisch sind die entsprechenden Zeilen nicht so klar, dass daraus eindeutige Aussagen gewonnen werden könnten (vgl. Voorwinden 2003).

Angesichts der Überlieferungslage lassen sich am Quellenmaterial also am ehesten diatopische und diachrone Varianz beobachten. Die diatopische Differenzierung des Mittelhochdeutschen verändert sich im vom MWB bearbeiteten Zeitraum. Einerseits erlaubt die mit der Zeit zunehmende Überlieferungsdichte eine detailliertere Untergliederung, andererseits spielt beispielsweise die militärische Expansion und Siedlungspolitik des Deutschen Ordens eine Rolle. Wegera (2000: 1307–1308) fasst die Entwicklungen und Schwierigkeiten knapp zusammen.

Die diatopische und diachronische Einordnung der im MWB verwendeten Belege steht allerdings vor Problemen, die über die entsprechende Bestimmung der Handschriften hinausgeht. Um die Projektziele des MWB innerhalb des ambitionierten Zeitrahmens zu erreichen, wurde bei der Konzeption entschieden, die Quellen nicht aus den Handschriften zu zitieren, sondern aus Textausgaben. Die Entscheidung hat pragmatische Gründe: Der Zugriff auf Belege über Ausgabenglossare und Wörterbücher, die Editionen zitieren, ist wesentlich weniger zeitaufwändig, wenn die Belege nicht noch auf die den Editionen zugrundeliegenden Handschriften zurückverfolgt werden müssen – auch bei der zunehmenden digitalen Verfügbarkeit der Handschriften. Die Bearbeitung des mittelhochdeutschen Wortschatzes ist im vorgesehenen Projektrahmen nur unter der Nutzung von Editionstexten vorstellbar.

Diese Entscheidung, das Belegmaterial aus Editionen zu gewinnen, steht aber auch im Zusammenhang mit der Quellenauswahl anhand der Entstehungs- anstelle der Überlieferungsdatierung. Zwischen der Entstehung mittelalterlicher Texte und ihren ältesten erhaltenen Überlieferungszeugen liegt in vielen Fällen eine große zeitliche Distanz (vgl. Gärtner 2000a: 106). Die Zusammenstellung unseres Quellenkorpus stützt sich auf die Entstehungsdatierungen der Texte, selbst wenn diese nur in jüngeren Handschriften überliefert sind. Bei einer Handschrift, die unikal einen gegenüber dem Überlieferungszeugen deutlich älteren Text überliefert, stellt sich daher die Frage, wie zuverlässig sie den älteren Sprachstand bewahrt, und mithin wie Belege aus der Edition dieser Quelle in Bezug auf diachronische Varianz zu beurteilen sind.

Gärtner (2000b: 21) betont in seiner Darstellung des MWB-Corpus, dass moderne, handschriftennahe Editionen bevorzugt werden. Dennoch bringt es die Geschichte der altgermanistischen Editionsphilologie mit sich, dass das Belegmaterial des MWB auf Ausgaben basiert, die in einem Zeitraum von über 200 Jahren entstanden und sehr unterschiedlichen editorischen Prinzipien folgen. So orientieren sich

einige der Zitierausgaben, um das offensichtlichste Problem anzusprechen, an Vorstellungen einer überregionalen normalmittelhochdeutschen Dichtersprache. Natürlich haben sich die mediävistischen Vorstellungen vom Normalmittelhochdeutschen an bestimmten Texten und Überlieferungsträgern entwickelt, sie lassen sich aber nicht ohne weiteres auf beliebige andere Texte übertragen (vgl. zu diesem Themenfeld Solms 2014 und Kragl 2015 mit unterschiedlichen Stoßrichtungen). Dennoch: Als Lesehilfen erfreuen sich normalisierende Eingriffe bei Editionen einer gewissen Beliebtheit. Als Basis variationslinguistischer Aussagen jedoch können die so hergestellten Textausgaben kaum dienen.

Angesichts dieser Umstände hält das Vorwort des MWB fest, dass „die Ermittlung von raum-, zeit- und textsortenspezifischen Symptomwerten nur in sehr begrenztem Maße möglich“ ist (MWB 1, VIII f.). Auch Aussagen über regionale Verteilungen können, so das Vorwort, nur „eingeschränkte Gültigkeit“ haben (MWB 1, IX). Bereits Grubmüller (2000: 15) geht in aller Kürze auf diese Unsicherheiten ein.

Im Folgenden wird es zunächst anhand von einigen Beispielen darum gehen, welche Angaben zu diatopischer und diachronischer Varianz im MWB gemacht werden. Dabei beziehe ich mich nur auf Fälle, in denen diatopische oder diachronische Merkmale explizit genannt werden. Außen vor gelassen sind also etwa morphologische Varianten, die zum Ansatz als Doppellemma führen, oder Schreibformen, die im Formteil aufgelistet werden. Anschließend diskutiere ich an Fällen aus der laufenden Artikelarbeit exemplarisch die Schwierigkeiten, mit denen das Redaktionsteam sich konfrontiert sieht. Der Beitrag endet mit Überlegungen, wie die Erschließung von Varianz im Mittelhochdeutschen dennoch auf Basis unseres Projektmaterials gefördert werden kann.

2 Angaben zu diatopischen und diachronischen Merkmalen

2.1 Diachronische Angaben

Im Artikel *âkust* stF., der in mehrere Bedeutungspositionen untergliedert ist, steht zur ersten Bedeutung ‚Laster, sündhafte Begierde‘ die Angabe: „häufig in frühmhd. geistl. Texten“; sie ist eingeschoben zwischen die Bedeutungsangabe und einen Hinweis auf eine Belegsammlung Leitzmanns, die natürlich eingesehen, aber nicht vollständig im Artikel wiedergegeben wurde. Die diachronische Information ist hier durch eine nicht-terminologische Frequenzangabe („häufig“) und den Bezug auf eine bestimmte Gruppe von Texten („geistl.“) eingeschränkt.

1 MWB 1, 137, 58 f. bzw. <http://www.mhdwb-online.de/wb/3519000>.

âkust *stF.* 1 'Laster, sündhafte Begierde', häufig in frühmhd. geistl. Texten (Belegsammlung Leitzmann, *Lexik. Probl.*, S. 33): die sibēn achust [*Todsünden*], / die gotes gaist hat vertriben / mit sinen geben sibēn SIEBENZ 14; so nevehtent in den brusten / die tugende mit den achusten VATERUNS 94; giriskheit. unt hûrlust. und ander werltliche

Abb. 1: Artikel *âkust* (MWB 1, 137 f.)

Bei *bestôzen* *stV.* ist ebenfalls eine einzelne Bedeutungsposition diachronisch näher bestimmt als „nur frühmhd.“, auch diese Information folgt der semantischen Angabe.²

bestôzen *stV.* 1 'jmdn. aus etw. verstoßen' (nur frühmhd.), mit *Akk.d.P.* und *Gen.d.S.*: duo er si dere wunnône [*Paradies*] bestiez GEN 500. 1427; einen erben [...], / den nieman sines riches bestiezze HOCHZ 213. 142; SÜKL V 804; LITAN 549 (*Fassung G*). – dannen der tievel / durch sin ubermuot grozen / in die helle wart pestozen ARNOLTSIEB 48,18 2 'auf jmdn. stoßen, zu jmdm. gelangen',

Abb. 2: Artikel *bestôzen* (MWB 1, 683 f.)

Im Gegensatz dazu folgt im dritten Beispiel *brouchen* *swV.* die diachronische Information „überwiegend frühmhd. belegt“ direkt dem Lemmaansatz und der grammatischen Angabe, sie geht also den einzelnen Gliederungspositionen voraus und bezieht sich daher auf das Lemma, nicht auf eine bestimmte Verwendungsweise.³

brouchen *swV.* überwiegend frühmhd. belegt. 1 'etw. biegen, formen' golt ketenne [...] gebrôht in wurmes wîs TRUDHL 23,23; WILL 18, 3; alsô der tuot der ûz wahsse ein bilede machet, / alsô brouchet er den leim GEN 110; der engel unde diu sêle [...] sint ûf dich [*Gott*] gebrouchet [*nach dir gebildet*], daz sie dîne forme in sich enpfâhen DVASCHR 368,26. 323,34; in stæines weise [...] den nie-

Abb. 3: Artikel *brouchen* (MWB 1, 1033 f.)

² MWB 1, 683, 40 f. bzw. <http://www.mhdwb-online.de/wb/17676000>.

³ MWB 1, 1033, 58 bzw. <http://www.mhdwb-online.de/wb/25743000>.

2.2 Diatopische Angaben

Ähnliches lässt sich bei diatopischen Angaben beobachten. So kann zu *anegülte* stF., das lediglich in zwei Urkunden belegt ist, festgehalten werden: „nur in Freiburg i.Br.“.⁴ Beim schwachen Verb *²dingen* wird zu Unterposition 1.5 angemerkt, Belege für diese Verwendung stammten „überwiegend aus dem österr.-bair. Sprachraum“.⁵ Wie schon bei den diachronischen Informationen begegnen uns auch hier Aussagen zur Frequenz.

anegülte stF. in der Wendung *ze ~ geben* ‘Pfand, Sicherheit’ (nur Freiburg i.Br.): vnd gilt er in den sehs wochen nýt, der schvltheize vnd zwene der viervndzweinzigen svln mit dem kleger ze dem hvs gan vnd svln ime ez ze angvltte geben URK-CORP (WMU) 1797AB,17. 248B,28. 248B,7

Abb. 4: Artikel *anegülte* (MWB 1, 242)

czv seynem rechten: holung vnd gesprech, als offt ym des not wer SCHÖFFIGLAU 122; URKCORP (WMU) 1358,30; RBMAGDEB 269,24 **1.5** ‘*sich an ein (höheres) Gericht wenden, Berufung einlegen*’ oft mit *präp. Erg.* (überwiegend aus dem österr.-bair. Sprachraum): so mach der chlagær oder der antwurtær denne wol an vns dingen URKCORP (WMU) 2345,20; swer an den keiser dingete ê er verteilet würde in swelicheme lande iz were man müstin in ime senden PRLPZ (L) 84,11. 84,10; [*der*

Abb. 5: Artikel *²dingen* (MWB 1, 1310)

Im Artikel zum Pron. *¹ez* wird die Formenvielfalt knapp beschrieben, dabei heißt es zum Sg. Nom. und Akk. *ez*: „daneben frühmhd. noch (md. auch später) *iz*, *hez* mit prothetischem *h*- (ÄPhys 9,5), mfrk. *it*, *id*“.⁶ Hier sind diachronische und diatopische Angaben teilweise verschränkt und mit einer Belegreferenz kombiniert und, anders als in den vorigen Beispielen, auf graphemische und phonetische Varianz bezogen.

4 MWB 1, 242, 22 bzw. <http://www.mhdwb-online.de/wb/246541600>.

5 MWB 1, 1310, 50 f. bzw. <http://www.mhdwb-online.de/wb/30783000>.

6 MWB 2, 21, 24–26 bzw. <http://www.mhdwb-online.de/wb/45969000>.

¹ëz *Pron. der 3. Pers. Neutr. In diesem Art. wird nur der Sg. behandelt, vgl. ër. 1 Deklination und Form (vgl. ²⁵Mhd. Gr. § M 41): – Sg. Nom. und Akk. ez, daneben frühmhd. noch (md. auch später) iz, hez mit prothetischem h- (ÄPHYS 9,5), mfrk. it, id; Gen. neben es auch is, für es tritt aber auch das urspr. refl. sîn ein (z.B. GRUD B 4; HERB 15880); zur Vermischung des Nom. und Akk. Sg. Neutr. ez mit Gen. Sg. Neutr. (Mask.) es als Folge des Zusammenfalls von s und 3 vgl. ²⁵Mhd.*

Abb. 6: Artikel ¹ëz (MWB 2, 21)

Im letzten Beispiel, beim Adv. *gevach*, findet sich wiederum eine Spezifizierung der Texte, in denen das Lemma vorkommt: „nur belegt im Omd. und in Deutschordensliteratur“.⁷ Das Vorwort hält übrigens fest, dass neben ausdrücklichen Nennungen von Textsorten auch implizite Angaben möglich sind: „Falls es keine explizite Angabe zur Textsorte gibt, kann gelegentlich auch die Quellenangabe einen Hinweis auf fachsprachliche Besonderheiten geben.“ (MWB 1, IX)

gevach *Adv. nur belegt im Omd. und in Deutschordensliteratur; auch gevache. In der NVJER-Ausgabe in Anlehnung an gevêch mit â angesetzt (vgl. dazu DWB 4,1,1,2059). 1 'häufig, wiederholt; in vielerlei Hinsicht' Darius sust sprach: / 'din got, dem du hie gevach / hast gedient, irlose dich!' DANIEL 5390; dar nach ich [Daniel] sach / ein anders [Tier], daz was gevach / gestalt als ein grimmic ber ebd. 5744. 4668; sô wurdin sî sô oft*

Abb. 7: Artikel *gevach* (MWB 2, 653)

2.3 Zusammenfassung – Was sehen wir an diesen Beispielen?

Wenn entsprechende Angaben gemacht werden, können sie bei Bezug auf das gesamte Belegmaterial zum Lemma im Artikelkopf („Formteil“) stehen, andernfalls erscheinen sie bei den einzelnen Artikelpositionen. Es gibt keine einheitlich vorgeschriebene Position für Varianzinformationen. Die Aussagen bleiben auf das jeweilige Lemma bezogen; es gibt also zum Beispiel in der Regel keine Verweise auf andere, „komplementäre“ Ausdrücke, die in anderen Schreibsprachen eine entsprechende Bedeutung bzw. Funktion tragen.

⁷ MWB 2, 653, 25 f. bzw. <http://www.mhdwb-online.de/wb/58788000>.

Die Angaben sind nicht verpflichtend. Sie sind außerdem selten: In vielen Fällen bilden unsere Artikel zwar Varianten ab, klassifizieren sie aber nicht nach diachronischen oder diatopischen Gesichtspunkten. Zu den Gründen hierfür zählen, dass das in vielen Fällen schmale Belegmaterial keine sicheren Aussagen zulässt, aber leider auch, dass angesichts des straffen Zeitrahmens die Erfassung der Varianz eine niedrigere Priorität hat als die Beschreibung von Bedeutung und Gebrauch der Stichwörter.

An den diachronischen Informationen zu den Verben *bestôzen* und *brouchen* lässt sich bemerken, dass die Angaben nicht streng formalisiert sind. So steht in einem Fall das Partizip „belegt“ in der Formulierung, im anderen Fall fehlt es: „nur frühmhd.“ – „überwiegend frühmhd. belegt“. Da varianzbezogene Aussagen nicht zum Hauptaufgabenbereich des MWB zählen, gibt es weniger formale Festlegungen zu den Angaben.

Nachfolgend demonstriere ich an zwei Beispielen, welche Fragen sich uns im Hinblick auf Varianz stellen und vor welchen Schwierigkeiten das MWB bei seiner Bearbeitung steht.

3 Schwierigkeiten bei der Beurteilung von Varianz

3.1 *larrûn*, *latrôn*

Der Artikel zum Substantiv *larrûn*, *latrôn* (‚Straßenräuber, Strolch‘) ist bisher unveröffentlicht. Es handelt sich um ein Lehnwort mit drei Belegen in deutlich unterschiedlicher Schreibung:⁸

- Renew 34066 = Ulrich von Türheim, Rennewart (Hübner 1938): *da han ich erslagen/ vil starker manne nuene,/ und waren daz **larruene** [Lesarten: larivne, karrune, garruone]:/ sie wolten mich beraubet han*
- UvEtzAlex 14823 = Ulrich von Eschenbach [Etzenbach], Alexander. (Toischer 1888): *daz [Gottes Barmherzigkeit] ist uns, hêrre, bewîset wol/ an dem **latrône**,/ der an dem criuze frône,/ got hêrre, an dich umb riuwe warp*
- EbvErf 2538 = Ebernand von Erfurt, Heinrich und Kunegunde (Bechstein 1860): *torste ich in [Bischof Brûn] heizen **loterûn** [Hs. lod-]/ den selben ungetrûwen wiht,/ durch in selben liez ichz niht*

Der Rennewart stammt aus dem Oberdeutschen der Mitte des 13. Jahrhunderts. Ulrichs von Etzenbach Alexander entstand Ende des 13. Jahrhunderts in Böhmen;

⁸ Zitiert ist jeweils nach Versnummer. Die Hervorhebung des Stichworts stammt von mir. Übergeschriebene Buchstaben (e bzw. o über u) sind hier hinter den Buchstaben, über dem sie in der Ausgabe stehen, auf die Grundlinie gesetzt.

seine Überlieferung teilt sich in drei Redaktionen, ist aber weniger zahlreich als beim Rennewart. Ebernand von Erfurt schrieb seine Kaiserlegende von Heinrich und Kunigunde in thüringischer Schreibsprache um 1220. Der Text ist demnach von seiner Entstehung her etwas älter als die anderen beiden, überliefert ist er jedoch nur unikal in einer Handschrift aus dem 15. Jahrhundert. In diesem Beleg ist die Diskrepanz zwischen der konjizierten und der handschriftlichen Form (laut Apparat) angegeben.

Hinter den beiden anderen Belegen stehen 41 (Rennewart) bzw. 12 (Alexander) Überlieferungsträger, auf deren vielfach divergierende Textformen uns die Editionen jedoch nur teilweise Zugriff erlauben. Im Rennewart-Beleg sind mehrere Lesarten aufgeführt. Der Herausgeber von Ulrichs Alexanderroman, Wendelin Toischer, rekonstruiert laut Verfasserlexikon (VL 9, 1258) ein Dichteroriginal gegen die Überlieferung. Der Text ist zwar in zwölf Handschriften überliefert, es gibt in Toischers Ausgabe jedoch keine Lesartenangabe zu dieser Stelle. Es existiert bisher keine Edition dieses Alexanderromans, die die handschriftliche Überlieferung besser wiedergibt.⁹

Genügt also vor diesem Hintergrund die Beleglage, um nach der Artikelbearbeitung im Wörterbuch festzuhalten, dass *larrûn* bzw. *latrôn* nur oberdeutsch und ostmitteldeutsch, aber nicht westmitteldeutsch gebräuchlich war? Dafür müssten eigentlich alle Überlieferungsträger geprüft werden. Alle drei Texte stammen aus dem 13. Jahrhundert; wir kennen keine Belegstellen aus Texten, die nach 1300 entstanden sind. Aber natürlich stammen Überlieferungsträger der drei zitierten Texte aus dem 14. und 15. Jahrhundert. Weder das FWB noch das DWB behandeln das Wort – es war also in den jüngeren Sprachperioden zumindest sehr ungebräuchlich, vielleicht auch gar nicht mehr bekannt und wurde nur noch in der Überlieferung der zitierten Texte fortgeschrieben.

Können wir andererseits die Formenvielfalt sowie die überraschenden Lesarten wie *karrune*, *garruone* (beide aus der ersten Hälfte des 14. Jahrhunderts, schwäbisch bzw. thüringisch) dadurch erklären, dass das Wort bereits in mittelhochdeutscher Zeit nie wirklich gebräuchlich war, und unsere Belege vereinzelte ad-hoc-Entlehnungen zeigen? Solche Fragen liegen nahe; sie zu beantworten ist kaum möglich.

Wie geht der Artikelentwurf mit dieser Beleglage um? Angesetzt ist das Stichwort als Doppellemma *larrôn*, *latrûn*; die beiden Ansätze entsprechen ungefähr den Schreibformen in Rennew und UvEtzAlex; die dritte (*loderun*) ist im Formteil erwähnt. Die Schreibung mit d in dieser dritten Schreibform könnte aus dem Altfranzösischen übernommen sein (auf den entsprechenden Eintrag im *Altfranzösischen Wörterbuch* von Tobler/Lommatzsch wird im Artikel hingewiesen). Bechstein änder-

⁹ Im Editionsbericht des Handschriftencensus ist allerdings ein entsprechendes Vorhaben von Christophe Thierry vermerkt, s. <https://editionsbericht.de/?n=Thierry1> (abgerufen am 3. 8. 2020).

te als Herausgeber des Textes die Schreibung zu *loterûn*, weil er einen Anschluss an mhd. *loter* vermutete. Der Vokal o in der ersten Silbe ist ungeklärt. Es ist denkbar, dass Bechstein weder die beiden anderen Belege noch die altfranzösische Entsprechung kannte.

Im Artikelentwurf steht als diachronische Angabe: „ab dem 13. Jh. belegt“. Da wir im MWB aber nicht nach der Überlieferung datieren, sondern nach der Entstehungszeit, wäre „nur im 13. Jh. belegt“ eigentlich präziser. Wäre diese scheinbar präzisere Formulierung für die Nutzerinnen und Nutzer aber auch wirklich hilfreicher, oder nicht vielleicht sogar irreführend? Denn die Überlieferung reicht ja bis ins 15. Jahrhundert.

larrûn, latrôn *stM.* zu *afrz.* larron (auch ladrón, ladrun, vgl. *Tobler/Lommatzsch* 5,196–200), zu *lat.* latro, s. *Suolahti* 1,150; *loderun* evtl. unter Einfluss von mhd. *loter*. Ab dem 13. Jh. belegt. 'Straßenräuber, Strolch, Verbrecher' da han ich erslagen / vil starker manne nûne, / und waren daz larrûne [*Laa. larivne, karrune, garrûne*]; / sie wolten mich beraubet han RENNEW 34066; daz [*Gottes Barmherzigkeit*] ist uns, hêrre, bewîset wol / an dem latrône, / der an dem criuze frône, / got hêrre, an dich umb riuwe warp UVETZALEX 14823; torste ich in [*Bischof Brûn*] heizen loterûn [*Hs. lod-*] / den selben ungetrûwen wiht, / durch in selben liez ichz niht EBVER 2538 Richter

Abb. 8: Artikel *larrûn* (MWB 3, in Vorbereitung)

3.2 *last*

Im Gegensatz zu *larrûn* ist *last* ein sehr gut belegtes Stichwort mit etwa 130 Belegen aus 50 unterschiedlichen Quellen. Es kommt sowohl als Maskulinum als auch als Femininum vor. Um diese Genusvarianz soll es im Anschluss gehen.

Die älteren mittelhochdeutschen Wörterbücher geben für das Femininum nur einen Beleg an (Herborts von Fritzlar Liet von Troye). Während bei BMZ (BMZ 1, 926b, Z. 36) einfach nur „stf. Herb. 3749“ angegeben ist, macht Lexer (Lexer 1, 1836) daraus eine diatopische Aussage: „md. auch fem. (Herb.)“.

Das Belegmaterial des MWB legt nahe, dass Lexers offene Formulierung, dass „mitteldeutsch auch“ das feminine Genus vorkomme, Bestand hat. Anlässlich dieses Beitrags habe ich das Belegmaterial zu *last* nach Schreibsprachen ausgezählt. (Normalerweise nehmen wir bei der Bearbeitung der Belege zu einem Stichwort keine vergleichbaren Auszählungen vor.) In unserem Material zu *last* finden sich etwa siebzehn Texte, die sich als mitteldeutsch klassifizieren lassen; in sieben von diesen siebzehn finden sich Belege für das Femininum. Die Zahlen beruhen auf den

Editionstexten, die ich nicht systematisch hinterfragt bzw. an den Handschriften überprüft habe.

In ihrer Ausgabe der mitteldeutschen Verserzählung *Rittertreue* ist Meier-Branecke der Lesart der Erfurter Fragmente (E) gefolgt, die feminines „last“ bezeugt. Sie begründet dies in einer Anmerkung unter Verweis auf Lexer: „Mit E lese ich last als Femininum, da dies die im Md. herrschende Form ist (Lexer I, 2, Sp. 1836).“¹⁰ Die editorische Entscheidung gründet sich also auf die Annahme, das Fem. solle „die im Md. herrschende Form“ sein. Das lässt sich aus der Angabe „md. auch fem.“ in Lexers Handwörterbuch – zumal mit nur einem angegebenen Beleg – nicht herauslesen. Da das MWB seine Belege aus Editionen zieht, ist für uns interessant, dass die Entscheidungen bei der Textherstellung teilweise auf Angaben in Lexers Wörterbuch – und zukünftig vermutlich auch auf denen im MWB – fußen, und dass in Einzelfällen diese Aussagen auch Gefahr laufen, überinterpretiert zu werden.

Stützen aber nicht nur die Editionstexte, sondern auch die Handschriften die diatopische Zuschreibung „md. auch fem.“? Ist das Femininum vielleicht in anderen Schreibsprachen ebenso belegt, in den Editionen aber von den Herausgebern „bereinigt“ worden? Es dürfte deutlich werden, dass eine genaue Überprüfung der gesamten Überlieferung der etwa 130 Textstellen, die das Stichwort enthalten, die Arbeit am einzelnen Artikel und, zum Prinzip erhoben, die Arbeit am Wörterbuch insgesamt im Hinblick auf die zu gewinnenden Erkenntnisse unvertretbar in die Länge zöge. Es ist außerdem anzunehmen, dass diese Prüfung nur selten wirklich etwas Neues brächte – die Editionen sind überwiegend verlässlich; auch an der Entscheidung Meier-Braneckes ist im Ergebnis nichts auszusetzen.

Das Femininum ist ab etwa 1200 belegt, das Maskulinum bereits ein paar Jahrzehnte früher. Frühmittelhochdeutsche Belege finden sich in unserem Material nicht. Wäre es also sinnvoll, anzugeben, dass *last* in der zweiten Hälfte des 12. Jahrhunderts aufkommt, zunächst als Maskulinum und erst später auch als Femininum belegt ist? Nein, denn *last* ist bereits althochdeutsch belegt! Im AWB (AWB 5, Sp. 641) ist das entsprechende althochdeutsche Stichwort als Femininum geführt. Das Maskulinum ist demnach die spätere Entwicklung – anders, als es die mittelhochdeutsche Beleglage für sich allein genommen andeutet. Eine Erklärung für die frühmittelhochdeutsche Lücke in der Belegchronologie habe ich nicht, abgesehen von der insgesamt dünnen Überlieferung aus dieser Zeit. An diesem Beispiel kann man sehen, dass auch aus einer vergleichsweise großen Belegmenge für ein mittelhochdeutsches Lemma nicht ohne weiteres chronologische Entwicklungen abgelesen werden können.

¹⁰ Meier-Branecke (1969).

last *stM.* *md.* auch *stF.* (z.B. ATHIS E 18; DANIEL 4673; HERB 3749; MARCOPOLLO 69,24. 74,11; MARLB-RH 101,10; HESLAPK 4050; RITTERTREUE 840).
1 '(Trag-)Last, Gewicht' **1.1** *allg.*: wie swærer
 last sich neiget zer erden zenter KLD:KZL 2: 11,1;

Abb. 9: Artikel *last* (MWB 3, in Vorbereitung)

In den beiden vorausgegangenen Beispielen habe ich zu zeigen versucht, welche Schwierigkeiten sich im Belegmaterial bei dem Versuch präsentieren, Varianzaspekte verlässlich zu beschreiben. Die Beispiele helfen hoffentlich auch zu verstehen, warum wir in vielen Fällen lieber zurückhaltend mit entsprechenden Beschreibungen sind, und warum das Vorwort unseres Wörterbuchs den Aussagen zu diesen Sachverhalten lediglich „eingeschränkte Gültigkeit“ zuschreibt.

4 Ausblick

Zum Schluss möchte ich einen Ausblick auf andere Möglichkeiten geben. Dabei wird es um zwei Aspekte gehen, zunächst das Belegarchiv, das auf MWB-Online öffentlich abrufbar ist; anschließend um die Annotation des digitalen Quellenverzeichnisses mit Symptomwertangaben und die Erwartung, diese Informationen in ein zukünftiges lexikographisches Informationssystem einzubinden.

Sowohl das elektronische Belegarchiv als auch das Quellenverzeichnis sind zunächst einmal projektinterne Werkzeuge, die der Erarbeitung des *Mittelhochdeutschen Wörterbuchs* dienen. Sie sind Teil eines Artikelredaktionssystems, das eigens für das Projekt entwickelt wurde (vgl. Queens/Recker-Hamm 2005; Plate/Recker-Hamm 2001). Leider wurde im damaligen Projektantrag nicht berücksichtigt, dass die EDV-Grundlage eines Langzeitprojektes der Wartung und Anpassung bedarf und im Verlauf des Projektes eine Erneuerung oder Umstellung auf ein anderes System empfehlenswert oder gar nötig sein könnte. In einem solchen Umbruch befindet sich das MWB, und dank zusätzlicher finanzieller Mittel unserer Trägerakademien kann gegenwärtig eine neue Redaktionsumgebung vorbereitet werden, die das bisherige System ersetzen wird. Auch wenn gegenwärtig noch offen ist, wieviel Bedürfnisse das neue System wird befriedigen können, haben wir natürlich als eines der Ziele die nachhaltige Nutzbarkeit unserer Daten – auch für externe Anfragen – als wünschenswert im Blick.

4.1 Belegarchiv

Das Belegarchiv, ursprünglich lediglich ein internes Hilfsmittel, besteht aus lemmatisierten Wortformen aus etwa 200 elektronisch aufbereiteten Editionstexten. Diese elektronischen Texte bilden nur einen Ausschnitt aus dem Corpus des MWB.

Wie das Quellencorpus des MWB zusammengestellt und hinsichtlich der Erschließungsweisen untergliedert ist, und welche Teile digital zur Auswertung vorbereitet wurden, beschreibt Gärtner (2000b). Ein digitales Grundcorpus wurde dreigeteilt und die drei Gruppen in unterschiedlicher Tiefe lemmatisiert, bevor mit der Artikelredaktion, also der Erarbeitung des Wörterbuchs im engeren Sinne, begonnen wurde. Neben diesen digitalisierten Texten werden natürlich auch nicht-digitalisierte Editionen ausgewertet – einerseits anhand von Ausgabenglossaren, andererseits anhand von Wörterbüchern, Konkordanzen und lexikologischen Studien. Diese Auswertung geschieht jeweils nur lemmabezogen während der laufenden Arbeit. Nach der vorbereitenden Lemmatisierung der Texte des Grundcorpus sind inzwischen rund 60 weitere Quellentexte in das elektronische Textarchiv eingespeist worden. Sie werden ebenfalls lemmabezogen ausgewertet; im Gegensatz zu den nicht elektronisch vorliegenden Quellen können hier jedoch die herangezogenen Wortformen lemmatisiert und damit dauerhaft in das digitale Belegarchiv aufgenommen werden. Diese Lemmatisierung bei der schrittweisen Bearbeitung einzelner Stichwörter zielt nicht auf thesaurusartige Vollständigkeit des herangezogenen Materials, sondern orientiert sich am Bedarf bei der Redaktion jedes einzelnen Artikels. Dadurch wächst das Belegarchiv langsam, aber stetig an. Aktuell umfasst das elektronische Textarchiv etwa 7,9 Mio. Wortformen, davon sind etwa 1,48 Mio. Wortformen lemmatisiert, also im Belegarchiv abrufbar.

Wir sind froh, dass wir das Belegarchiv seit einigen Jahren über unsere Online-Präsenz www.mhdwb-online.de öffentlich anbieten können. Finanzielle Mittel für die weitere Betreuung der Daten und ihrer Präsentation im WWW gibt es aktuell leider nicht. Das Belegarchiv bietet eine KWIC-Konkordanz (keyword in context) aller zu dem aufgerufenen Stichwort lemmatisierten Textstellen in unserem elektronischen Textcorpus. Interessierte sollten bei der Benutzung im Hinterkopf behalten, dass die Lemmatisierung in seltenen Fällen fehlerhaft sein kann. Es handelt sich bei dem Belegarchiv um ein Arbeitsinstrument, nicht um ein abschließendes Ergebnis unseres Projekts.

Dank des öffentlichen Belegarchivs ist es möglich, Belegmaterial zu sehen, das über die im Artikel gezeigten Belegstellen hinaus geht und somit potenziell ein breiteres Spektrum von Varianz bietet. Natürlich kann das elektronische Belegarchiv die Zurückhaltung in den Wörterbuchartikeln (in Bezug auf Varianzbeschreibung) nur bedingt kompensieren. Zunächst einmal kann nicht davon ausgegangen werden, dass alle Vorkommen des Stichworts in den elektronisch vorliegenden Texten erfasst sind.

Aus den elektronischen Texten stammt außerdem nur ein Teil des Belegmaterials, das zur Erarbeitung der Artikel durchgesehen wird bzw. je nach Arbeitsstand schon gesichtet wurde. Da das MWB kein Thesaurus ist, sind die in den Artikeln gezeigten Belege, ebenso wie die lemmatisierten Textstellen im Belegarchiv jeweils nur eine Teilmenge des ausgewerteten Materials (mit Überschneidungen). So kann das Belegarchiv zwar dabei helfen, über die in den Artikeln gebotenen Belege hinaus Beispiele für Varianz zu finden – allerdings ohne Anspruch auf Vollständigkeit oder Repräsentativität.

Weitere Einschränkungen ergeben sich z.B. daraus, dass

- die elektronischen Texte leider nicht die Apparate der Ausgaben umfassen (das ist beim Umgang mit den Textstellen zu bedenken);
- es im öffentlichen Zugriff auf das Belegarchiv bisher nicht möglich ist, nur eine Liste der Wortformen zu einem Lemma auszuwerfen oder die Belege nach unterschiedlichen Kriterien zu sortieren; besonders bei großen Belegmengen kann dies unbefriedigend sein.

Zum jetzigen Zeitpunkt ist es unklar, ob im MWB-Projekt oder anderweitig zukünftig Möglichkeiten gefunden werden, die Nutzungsmöglichkeiten des Belegarchivs zu erweitern.

4.2 Annotation des Quellenverzeichnisses mit Symptomwerten

Mit dem Quellenverzeichnis, von dem ich hier spreche, ist nicht das Verzeichnis gemeint, das dem gedruckten Wörterbuch vorangestellt ist (und das mit jeder Lieferung ergänzt wird), sondern das darüber hinausgehende Quellenverzeichnis, das in unser internes Redaktionssystem eingebunden ist. Es weist einerseits zusätzliche, teils rein arbeitspraktische Informationen zu den Quellen auf (z.B. Standorte, Bibliothekssignaturen usw.), es umfasst aber auch mehr Einträge, z.B. ausgesonderte Quellen sowie Quellen, die wir zwar bereits erfasst, aber noch nicht in den publizierten Teilen des Wörterbuchs zitiert haben. Die einzelnen Einträge weisen außerdem Felder für Zusatzinformationen auf, darunter Angaben zu Textsorten, Datierungen und Schreibsprachen. Die Daten sind aber nur für das elektronische Grundcorpus vollständig eingetragen; beim Ausbau des Quellenverzeichnisses wurden die entsprechenden Informationen weniger systematisch erhoben. Bei vielen Einträgen fehlen sie.

Im Kontext des BMBF-geförderten Projektes „eHumanities-Zentrum für Historische Lexikographie“ (ZHistLex, <https://zhistlex.de/>) und im Zuge der Entwicklung eines neuen Redaktionssystems für das MWB kam neue Bewegung in die Annotation mit Zusatzinformationen. Nachdem das Quellenverzeichnis in XML-Daten überführt wurde, konnte es einerseits mit Daten des Referenzcorpus Mittelhochdeutsch (ReM, <https://www.linguistics.rub.de/rem/>) abgeglichen und bei übereinstimmen-

den Quellen mit den Symptomwertangaben des ReM angereichert werden. Darüber hinaus konnten wir Hilfskraftmittel gewinnen, um aus dem Verfasserlexikon Angaben zu Textsorten, Schreibsprachen und Datierungen unserer Quellen zu übertragen. Die Mittel waren erschöpft, bevor die Arbeit abgeschlossen und einer Qualitätsprüfung unterzogen werden konnte. Dennoch freuen wir uns, dass ein Format und Vorgehen für die Annotation entwickelt und erprobt und dadurch eine wichtige Grundlage geschaffen wurde. Zukünftige Aufgaben sind einerseits die Vervollständigung der Daten für noch nicht annotierte Quellen, andererseits die Formalisierung der Angaben zu Textsorten und Schreibsprachen; die Datierungen sind bereits streng formalisiert eingetragen. In einem weiteren Schritt müssten dann Schnittstellen bzw. Anwendungsmöglichkeiten geschaffen werden: Denkbar sind nicht nur Filterungen im Quellenverzeichnis selbst, also etwa eine Suche nach mitteldeutschen Predigttexten des 13. Jahrhunderts. Spannend und potenziell richtungsweisend stellen wir uns Verknüpfungen der Symptomwertangaben mit unseren Wörterbuchartikeln oder dem Belegarchiv vor, beispielsweise um unterschiedliche Hervorhebungen oder Sortierungen der Belege nach diachronischen oder diatopischen Kriterien zu ermöglichen.

Leider handelt es sich dabei noch um Zukunftsperspektiven, deren Umsetzung aus unserem Projekt heraus nicht leistbar ist. Wir setzen uns aber weiterhin dafür ein, dass unsere internen Werkzeuge verbessert und der Öffentlichkeit zur Verfügung gestellt werden.

Literatur

- AWB = *Althochdeutsches Wörterbuch*. Auf Grund der von Elias von Steinmeyer hinterlassenen Sammlungen, im Auftrag der Sächsischen Akademie der Wissenschaften zu Leipzig. Bd. 1 ff. Berlin: Akademie (bis 2013), Berlin/Boston: De Gruyter (ab 2014) 1952 ff.
- Bechstein, Reinhold (1860) (Hrsg.): *Eberhard von Erfurt, Heinrich und Kunegunde*, Quedlinburg: Basse.
- BMZ = *Mittelhochdeutsches Wörterbuch*. Mit Benutzung des Nachlasses von Georg Friedrich Benecke ausgearbeitet von Wilhelm Müller und Friedrich Zarncke. 3 Bde., Leipzig: Hirzel 1854–1866.
- Gärtner, Kurt (2000a): Althochdeutsch oder Mittelhochdeutsch? Abgrenzungsprobleme im Bereich der Glossenliteratur und ihre Bedeutung für die Sprachstadienlexikographie. In: Haubrichs, Wolfgang et al. (Hrsg.): *Theodisca. Beiträge zur althochdeutschen und altniederdeutschen Sprache in der Kultur des frühen Mittelalters*. Berlin/New York: De Gruyter, 105–117.
- Gärtner, Kurt (2000b): Quellenauswahl, Corpuskomplexe, Arbeitsverfahren und Kooperation. In: Gärtner, Kurt/Grubmüller, Klaus (Hrsg.): *Ein neues Mittelhochdeutsches Wörterbuch: Prinzipien, Probeartikel, Diskussion*, Göttingen: Vandenhoeck & Ruprecht, 18–33 (362–377).
- Grubmüller, Klaus (2000): Zielsetzung und Darstellungsprinzipien. In: Gärtner, Kurt/Grubmüller, Klaus (Hrsg.): *Ein neues Mittelhochdeutsches Wörterbuch: Prinzipien, Probeartikel, Diskussion*, Göttingen: Vandenhoeck & Ruprecht, 10–17 (354–361).

- Hübner, Alfred (1938) (Hrsg.): *Ulrich von Türlheim, Rennewart. Aus der Berliner und Heidelberger Handschrift*. Berlin: Weidmann.
- Kragl, Florian (2015): Normalmittelhochdeutsch. Theorieentwurf einer gelebten Praxis. In: *Zeitschrift für deutsches Altertum und deutsche Literatur* 144, 1–27.
- Lexer = *Mittelhochdeutsches Handwörterbuch* von Matthias Lexer. Zugleich als Supplement und alphabetischer Index zum Mittelhochdeutschen Wörterbuch von Benecke – Müller – Zarncke. 3 Bde., Leipzig: Hirzel 1872–1878.
- Meier-Brancke, Marlis (1969) (Hrsg.): *Die Rittertreue. Kritische Ausgabe und Untersuchungen*. Hamburg: Buske.
- MWB = *Mittelhochdeutsches Wörterbuch*. Im Auftrag der Akademie der Wissenschaften und der Literatur Mainz und der Akademie der Wissenschaften zu Göttingen hrsg. von Kurt Gärtner, Klaus Grubmüller, Jens Haustein, mitbegründet von Karl Stackmann. Bd. 1 ff. Stuttgart: Hirzel 2006 ff.
- MWB Online. <http://www.mhdwb-online.de/> (letzter Zugriff: 5. 8. 2020).
- Plate, Ralf/Recker-Hamm, Ute (2001): Elektronische Materialgrundlage und computergestützte Ausarbeitung eines historischen Belegwörterbuchs. Erfahrungen und Perspektiven am Beispiel des neuen Mittelhochdeutschen Wörterbuchs. In: Lemberg, Ingrid/Schröder, Bernhard/Storrer, Angelika (Hrsg.): *Chancen und Perspektiven computergestützter Lexikographie*. Tübingen: Niemeyer, 155–177.
- Queens, Frank/Recker-Hamm, Ute (2005): A Net-based Toolkit for Collaborative Editing and Publishing of Dictionaries. In: *Literary and Linguistic Computing* 20, Issue Suppl., 165–175.
- Solms, Hans-Joachim (2014): Die Schimäre einer mittelhochdeutschen Gemeinsprache. Eine grammatikographische Studie auf der Grundlage des Bochumer Mittelhochdeutsch-Korpus. In: *Jahrbuch für Germanistische Sprachgeschichte* 5, 111–134.
- Toischer, Wendelin (1888) (Hrsg.): *Ulrich von Eschenbach, Alexander*. Tübingen: Litterarischer Verein in Stuttgart.
- VL = *Die deutsche Literatur des Mittelalters. Verfasserlexikon*. 2., völlig neu bearb. Aufl., hrsg. von Gundolf Keil, Kurt Ruh (federführend bis Band VIII, 1992), Werner Schröder, Burghart Wachinger (federführend ab Band IX, 1995) und Franz Josef Worstbrock, redigiert von Kurt Illing (bis Band I) und Christine Stöllinger-Löser, XIV Bde. Berlin/New York: De Gruyter, 1978–2008.
- Voorwinden, Norbert (2003): Ist er ze Sahsen oder ze Brabant gewahsen? Beobachtungen zum ‚Flämln‘ des jungen Helmbrecht. In: Pijnenburg, Willy (Hrsg.): *Quod vulgo dicitur. Studien zum Altniederländischen*. Amsterdam: Rodopi, 311–319.
- Wegera, Klaus-Peter (2000): Grundlagenprobleme einer mittelhochdeutschen Grammatik. In: Besch, Werner et al. (Hrsg.): *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. 2. Teilband. 2., vollst. neu bearb. u. erw. Aufl. Berlin/New York: De Gruyter, 1304–1320.

