

260716

Maschinelle Verarbeitung altdeutscher Texte IV

Beiträge zum Vierten Internationalen Symposium
Trier 28. Februar bis 2. März 1988

Herausgegeben von Kurt Gärtner,
Paul Sappeler und Michael Trauth

AM,
T 10
Mas 1

2001/92

Max Niemeyer Verlag
Tübingen 1991

Universität Tübingen
NEUPHIL. FAKULTÄT
BIBLIOTHEK



ze des entstehenden Wörterbuches im Blick zu haben und Entscheidungen zu treffen, die der Einheitlichkeit dieses Ganzen dienen. Natürlich bedeutet die ständige Revidierbarkeit aller Entscheidungen, bedeutet vor allem die Möglichkeit, jederzeit Ergänzungen anzubringen, auch eine Gefahr. Man kann sich dadurch zu einer Arbeitsweise verleiten lassen, die aus dem Wörterbuch ein modernes Pendant zu dem ἀνήνυτον ἔργον der Penelope werden läßt. Diese Gefahr muß man sehen, aber man kann ihr mit Vernunft und gutem Willen begegnen. So meine ich denn, nehme ich alles in allem, ich schulde dem »schnellen Trottel« – so nannte den Computer vor Jahren einmal gesprächsweise ein Mathematiker – großen und ehrlichen Dank.⁸

⁸ Zum Schluß möchte ich den Damen Edith Bock und Sylvia Gross für die Anfertigung des Typoskripts danken.

Ein Wörterbuch zu Gottfrieds 'Tristan'

Wer Datenverarbeitung für lexikographische Zwecke einsetzt und daran denkt, das Erzeugnis zu publizieren und es allgemein benutzbar zu machen, muß entscheiden, wieviel philologische Durcharbeitung er aufwendet.

Er kann einerseits bei einem rein mechanisch hergestellten, rohen Formwörterbuch, d. h. Index oder Konkordanz, bleiben, einer Umformung des Basistextes, die ganz auf sein graphematisches Bild bezogen ist. Das Minimum an Aufwand, dessen Unterschreitung doch auffällt, sollte darin bestehen, das Ergebnis der Maschinentätigkeit, bevor man es aus der Hand gibt, auf programmierte Sinnlosigkeiten zu prüfen.

Zu bedenken ist, daß sich die Veröffentlichung einer ganz unbearbeiteten Konkordanz kaum mehr lohnt, nachdem sich Datenverarbeitungsgeräte rapide ausbreiten; für die Ad-hoc-Herstellung einer einfachen Konkordanz oder – oft noch besser – für gezielte Suchabfragen im Text, die manchmal sogar noch spezifischere Analysegesichtspunkte als das bloße Vorkommen einer Wortform zu berücksichtigen gestatten, genügen schon der maschinenlesbare Text und ein PC. Ansonsten sind die rohen Hilfsmittel insgesamt unverächtlich, reichen sie doch für viele Aufgaben sehr wohl aus.

Nichtsdestoweniger ist andererseits seit langem eine Neigung zum philologischen Bearbeiten rechnererzeugter Formwörterbücher zu beobachten, die sich in Findelisten, morphologischen Verweisen, Lemmatisierung, Homographentrennung und ähnlichem äußert.¹ An dieser Tendenz ist besonders sympathisch, daß sie wieder näher hinführt zur traditionellen Wörterbucharbeit mit ihrem Ziel wortkundlicher Darstellung, innerhalb derer die gliedernde Beschreibung und Bezeugung der Bedeutung, des Gebrauchs und der grammatischen Verhältnisse von Wörtern ihren Platz hat. Sehr weit ist man noch nicht gekommen auf diesem an sich richtigen Weg, das Formwörterbuch dem traditionellen Wörterbuch ähnlich werden zu lassen; die rechte gegenstandsbezogene Gestaltungsfreiheit ist noch nicht erreicht, wie auch umgekehrt die Neugestaltung der technisch-organisatorischen Seite traditioneller Wörterbucharbeit keineswegs abgeschlossen ist.

In manchen Äußerungen klingt an, der Bearbeitungsaufwand für Lexika zu literatur- und geistesgeschichtlich weniger bedeutsamen, anspruchsloseren, nicht so bewußt geformten, weniggelesenen und -interpretierten Texten lohne sich nicht recht.

¹ Die Gegenüberstellung roher und bearbeiteter Maschinenerzeugnisse sieht etwas anders aus und ist nicht ganz so scharf, wenn man Versuche berücksichtigt, manche »höheren« Unterscheidungen, die zunächst philologischer Einsicht vorbehalten scheinen, »hinterherum« mittels Analyseprogramm zu leisten; aufs Ganze gesehen ist aber die Bearbeitung unmittelbar von der Kopfarbeit des Philologen abhängig.

Wie dem auch sei, unser (noch nicht abgeschlossener) Beitrag zu den Versuchen, lexikographische Maschinenarbeit philologisch weiterzuentwickeln und dem Philologen als Wörterbuchbenutzer brauchbare Maschinenerzeugnisse anzubieten, besteht in einem Programmsystem zur Rechnerunterstützung von Wörterbucharbeit und einem Wörterbuch zu Gottfrieds 'Tristan'.

Das Wörterbuch beruht auf RANKES Text (maschinell eingelesen), in dem aber die Vokallängen bezeichnet wurden, und einer nach ziemlich engen Regeln ausgewählten Menge lexikalisch interessanterer Lesarten der Ausgabe MAROLD/SCHRÖDER. In den Artikeln sollen in erster Linie Bedeutungen in ihrer Vielfalt zum Vorschein kommen, dann auch das syntaktische Verhalten der Wörter, ihr Vorkommen in Fügungen und phraseologischen Einheiten dargestellt werden (möglichst mit den Bezügen zwischen diesen Ebenen), und schließlich sollen ausreichende, wohleingegrenzte Belegzitate eine Übersicht über Verwendungszusammenhänge in weitem Sinn geben. Das Wörterbuch ist auf Vollständigkeit angelegt; dies kommt den blasseren, nicht so prägnanten und deutlich herausstellbaren Wortverwendungen zugute, die man sonst doch etwas zu beschneiden in Gefahr ist, und zwingt überhaupt dazu, eher aus dem Material heraus als nach Vorerwartungen zu denken.

Das Programmsystem ist auf TUSTEP aufgesetzt und verwendet bewußt soweit wie möglich dessen Standardmittel und -funktionen, damit man nicht z.B. die Bedienung eines neuen Editors lernen muß, damit man geräteunabhängig und entwicklungsfähig ist und eine gute Datenverarbeitungs-Umgebung hat. Es besteht aus zwei Teilen: Im einen wird die Lemmatisierung erledigt, eingeschlossen ein großer Teil der Homographentrennung (durch Markierungen im Text), im andern, in dem dann stärker bis in die Einzelheiten das philologische Judicium wirken kann, findet die Feinbearbeitung der Artikel statt, die jeweils in kleinen Gruppen dafür bereitstehen. Diese Feinbearbeitung, von der im folgenden hauptsächlich die Rede sein soll, geschieht schrittweise, in allmählicher Annäherung an eine sich herausbildende individuelle Gestalt des Artikels, wie auch der Basistext nicht von Anfang an fertig sein muß, sondern in gewissen Grenzen noch verbessert werden kann. Der Artikel besteht während dieser Zeit, so wird angenommen, aus gegeneinander beweglichen Teilen, nämlich einerseits aus kommentierenden Abschnitten und Gliederungskommentar, die belegstellenunabhängig sind, und andererseits aus Belegstellen mit stellenbezogenem Kommentar, zu welchem auch die Belegzitate gerechnet werden; die Belegstellen sind von Anfang an da und müssen unangetastet bleiben (außer in ihrer Reihenfolge), der kommentierende Text kommt im Lauf der Arbeit hinzu und ist frei formulierbar, soweit man sich nicht terminologische, redaktionelle und stilistische Regeln auferlegt (deren Einhaltung zum Teil wieder mit dem Rechner abprüfbar ist). Der Arbeitsablauf ist auf den Dialog am Bildschirm zugeschnitten; bei keinem der Schritte muß der Rechner verlassen werden.

Wie sieht nun die Arbeit am Tristanwörterbuch mit unserem Programmsystem konkret aus, und zwar dort, wo sich das Bemühen des Philologen und der Einsatz des Computers unmittelbar berühren: eben bei der Abfassung der einzelnen Artikel? Der Arbeitsprozeß und der dabei verwendete Baustein unseres Programm-

```
@-00100 @a= #152.166/0 *S* spil
@-00100 @+ mite/ und dar zuo hovegesindes vil,/ die sô durch er-
nest sô durch >>spil/ in volgeten unz an den kiel,/ und swaz ie-
man dâ geviel,/ dâ in sîn wille
@s 2195 7
@-00100 @+ Tristan der sprach: 'diz sî&v getân!'/ suz sâzen sî
zwên über daz >>spil./ der marschalch sprach: 'Tristan, ich wil/
wider ûf ze herbergen gân;/
@s 2249 7
@-00100 @+ maneger tugent./ swaz vuoge er aber an der stete/ mit
geb#.ârden oder mit >>spil getete,/ daz was in dâ wider also ein
wint:/ sî nam des wunder, daz ein
@s 2280 5
@-00100 @+ in:/ diu sprach er wol und kunde ir vil,/ dâ mite sô
zierter in sîn >>spil./ ouch sang er wol ze prise/ schanzûne und
sp#.âhe wise&s,/ refloit und
@s 2292 7
:
:
@-00200 @u= /* spile
@-00200 @+ swaz ir geredet, ich w#.âne, ich wil/ mit vrumen und
mit êren/ von disem >>spile kâren./ ich h#.âte michel arbeit/
unsinneclîchen an geleit,/ solte ich nu
@s 11 60 3
@-00300 @u= /* spiles
@-00300 @+ verdâht/ an ir spil sô sêre,/ daz sî dô nihtes mêre/
niwan ir >>spiles gedâhten./ nu sîz dô vollebrâhten,/ sô daz
Tristan daz spil gewan,/
@s 2317 3
@-00300 @+was aber unlanc,/ die sî von im h#.âten an der stete,/
wan swaz er in dâ >>spiles getete/ mit handen oder mit munde,/
dazn gie niht von grunde:/ daz herze dazn
@s 7526 6
```

Feldkennungen (vom Programm
eingetragen):

```
@- Ordnungszahl
@a Lemmazeile
@u Unterlemmazeile
@+ Kontext
@s Belegstelle
```

Sonstige Kodierungen:

```
>> Leitwort
&x Homographentrenner
# Seiten-Zeilen-NummerderKonkordanz
*S* Wortart des Lemmas
*/* Unterlemma
```

Abb. 1: Hintergrunddatei vor der Bearbeitung des Artikels
(Ausschnitt)

systems sollen hier möglichst knapp und durchsichtig dargestellt werden, obwohl manche Details nicht unterschlagen werden dürfen.

Die Frage aus der Sicht des Lexikographen ist die: Welche Leistungen muß ein Programm erbringen, das mir für den Hauptteil der Wörterbucharbeit, für die Abfassung der Artikel, von Nutzen sein soll? Es seien fünf Punkte genannt:

1. Das geforderte Programm muß dem Bearbeiter die Möglichkeit geben, die Belege, d.h. Belegzitate und Belegstellen, des Artikels nach seinen Vorstellungen

- anzuordnen und sie während der Arbeit schnell und ohne Aufwand umzustellen.
2. Der Bearbeiter muß die Möglichkeit haben, den Umfang eines Belegzitats in jedem einzelnen Fall schnell und ohne Aufwand festzulegen.
 3. Das Programm muß die Möglichkeit unterstützen, an jeder Stelle eines Artikels Kommentare von beliebiger Länge anzubringen.
 4. Ein so beschaffenes Programm muß andererseits verhindern, daß Belegstellen verändert werden oder verlorengehen können; und es muß
 5. so weit wie möglich verhindern, daß Verschreibungen in den Belegzitaten auftreten können.

Im folgenden wird gezeigt, wie diese Anforderungen von dem Baustein XART unseres Programmsystems unterstützt und eingelöst werden.

Ausgangspunkt für die Bearbeitung eines Artikels ist eine bereits mit einem anderen Programmbaustein erstellte lemmatisierte Konkordanz. Aus dieser werden einer oder auch mehrere zu bearbeitende Artikel (und das bedeutet zu diesem Zeitpunkt: jeweils ein Lemma plus seine nach den Unterlemmata angeordneten Belege) in ein File kopiert, das im folgenden ›Hintergrunddatei‹ heißt. Bei diesem Herauskopieren (vgl. Abb. 1 mit dem Beispiel-Artikel *spil*) werden die Daten in Felder eingeteilt, welche zu ihrer Identifizierung eine aus zwei Zeichen bestehende, hier jeweils fett gedruckte Feldkennung bekommen. So kennzeichnet @a in unserem Fall den Beginn des Lemmafeldes, welches besteht aus der Seiten-Zeilen-Nummer der Konkordanz als ›Heimatadresse‹ des Artikels (#), der Angabe der Wortart (*S*) und dem Lemmawort (*spil*); @u bezeichnet die Unterlemmafelder, @+ den Beginn eines Feldes mit Kontext, @s ein Feld mit der Belegstelle. Außer dieser Einteilung in Felder bekommt jede Einheit, d.h. die Lemmazeile, jede Unterlemmazeile, jeder Beleg, eine Ordnungszahl, die sich bei jedem Unterlemma erhöht. Die Feldkennung für die Ordnungszahl lautet in unserem Fall @-.

Die ›Hintergrunddatei‹ ist, wie ihr Name bereits vermuten läßt, nicht Gegenstand der Bearbeitung, sondern ein stets präsentés Kontrollmittel. Die gesamte Arbeit am Artikel erfolgt in einer Kopie dieses Files, der ›Arbeitsdatei‹. Bei der Erstellung dieser zweiten Datei können die Felder mit dem Kontext beliebig verkürzt werden, was besonders bei umfangreicheren Artikeln den Vorteil hat, daß eine größere Anzahl von Belegen auf einer Bildschirmseite Platz findet.

In Abb. 2 ist unter Punkt 0 ein Ausschnitt aus dem Artikel *spil* in der Arbeitsdatei vor der Bearbeitung wiedergegeben, d.h. in der Gestalt, wie er sich dem Bearbeiter auf dem Bildschirm vor dem ersten Eingriff präsentiert. Wie durch einen Vergleich mit Abb. 1 erkennbar, wurde hier der Kontext um über die Hälfte auf 75 Zeichen verkürzt, gegenüber einer Länge von 160 Zeichen in der Hintergrunddatei. (Die Werte wie auch die Feldkennungen sind variabel.)

Im folgenden seien kurz die vier wichtigsten Operationen in der Arbeitsdatei erläutert.

1. Das Ordnen oder Umstellen der Belege.

Die Abfolge der Belege eines Artikels richtet sich in der unbearbeiteten Datei nach ihrer Zugehörigkeit zu einem Unterlemma bzw. nach der aufsteigenden Reihen-

0) Ausschnitt aus der "Arbeitsdatei" vor der Bearbeitung:
 @-00100 @a= #152.166/0 *S* spil
 @-00100
 @+sô durch ernest sô durch >>spil/ in volgeten unz an den kiel,/ und swaz @s 2195 7
 @-00100
 @+sâzen sî zwên über daz >>spil./ der marschalco sprach: 'Tristan, ich wil/ @s 2249 7
 @-00100
 @+mit geb#.ârden oder mit >>spil getete,/ daz was in dâ wider alse ein @s 2280 5
 @-00100
 @+ir sit&v daz irresameste >>spil./ daz ieman ûf dem brete kan./ er ist @s 9888 5
 @-00200 @u= /* spile
 @-00200
 @+und mit êren/ von disem >>spile kêren./ ich h#.âte michel arbeit/ unsinneclichen @s 11 60 3

1) Ordnen/Umstellen der Belege durch Verändern der Ordnungszahl:
 @-02010
 @+sô durch ernest sô durch >>spil/ in volgeten unz an den kiel,/ und swaz @s 2195 7
 @-03010
 @+mit geb#.ârden oder mit >>spil getete,/ daz was in dâ wider alse ein @s 2280 5

2) Abgrenzung des Belegtextes:
 a) bei ausreichendem Kontext:

@-03020
 @[ir sit&v daz irresameste >>spil./ daz ieman ûf dem brete kan@. er ist @s 9888 5

b) bei ungenügendem Kontext:

@-03020
 @+ swaz ir gerede@[ich wæne, ich wil/ mit vrumen und mit êren/ von disem >>spile kêren@. ich hæte michel arbeit/ unsinneclichen an geleit./ solte ich nu
 @+und mit êren/ von disem >>spile kêren./ ich h#.âte michel arbeit/ unsinneclichen @s 11 60 3

3) Eintrag von Kommentaren:

@-03010 @/ \$ a) }>Schachspiel<:{
 @-03010
 @+sâzen sî zwên über daz >>spil./ der marschalco sprach: 'Tristan, ich wil/ @s 2249 7 @; Brett und Figuren??

4) Verdoppeln einer Belegstelle:

@-03020
 @[ir sit&v daz irresameste >>spil./ daz ieman ûf dem brete kan@. er ist @s 9888 5 @-03051 @r

Feldkennungen (vom Bearbeiter eingetragen, Auswahl):

@[Belegzitat oder stelligegebundener Kommentar
 @. fällt weg (hier: Ende des Belegzitats)
 @/ nichtstelligegebundener Kommentar, Gliederungskommentar
 @; interne Bemerkung
 @r Verdoppeln der vorhergehenden Belegstelle

Abb. 2: Operationen in der Arbeitsdatei

folge der Belegstellen. Für eine Gliederung des Artikels nach semantischen oder syntaktischen Kriterien ist daher in den meisten Fällen eine gründliche Änderung dieser Anordnung erforderlich. Das geschieht nun nicht etwa durch ein mühevolleres Hin- und Herkopieren der zahlreichen Belege, sondern ausschließlich durch die einfache Veränderung der Ordnungszahl im entsprechenden Feld. Das eigentliche Umgruppieren und Umsortieren der Belege in die angestrebte Reihenfolge besorgt erst das Programm XART, das jedoch beliebig oft und zu jedem Zeitpunkt der Bearbeitung gestartet werden kann. Der Bearbeiter hat so die Möglichkeit, schnell und ohne technischen Aufwand die gewünschte Ordnung des Artikels herzustellen, das Ergebnis immer wieder zu überprüfen, zu verfeinern oder zu verwerfen.

2. Die Abgrenzung des Belegzitats.

In aller Regel entspricht der rein mechanisch abgetrennte Kontext zu einer Belegstelle im Umfang nicht dem Zitat, wie es im fertigen Artikel erscheinen soll. Der Bearbeiter ist in der Lage, sein Belegzitat ohne Mühe abzugrenzen, es gewissermaßen aus dem vorliegenden Kontext herauszuschneiden, indem er Anfang und Ende des Zitats durch zwei dafür vorgesehene Feldkennungen markiert (in unserem Fall sind es @ für den Anfang und @. für das Ende des Belegzitats). Das übrige erledigt XART: Es eliminiert die überflüssig gewordenen Kontextreste (*er ist*), holt den alten Kontext aus der Hintergrunddatei und stellt ihn unverändert und in seinem bisherigen Umfang als zusätzliche Kontrolle an seinen alten Platz, nun zwischen ausgewähltem Belegzitat und Belegstelle. Es ist wichtig, daß hierbei vom eigentlichen Kontext kein Wort abgeschrieben oder neu geschrieben werden mußte; eine mögliche Fehlerquelle der traditionell hergestellten Wörterbücher, Verschreibung bei den Textziten, wird damit so gut wie ausgeschlossen. (In Klammer hinzugefügt wurde bei diesem Zitat später nur ein verdeutlichendes *wip* nach *ir*; das Ergebnis der geschilderten Prozedur ist in Abb. 3 bei der zweiten Ordnungszahl 3020 zu sehen: ausgewähltes Belegzitat - alter Kontext - Belegstelle). Auch für den Fall, daß der verkürzte Kontext in der Arbeitsdatei zur Bestimmung des Belegzitats nicht ausreichen sollte, ist gesorgt: Da Arbeits- und Hintergrunddatei dieselbe Zählung aufweisen, kann die Zeile mit dem gleichen, aber umfangreicheren Kontext aus der Hintergrunddatei (wie am Beispiel 2b in Abb. 2 ersichtlich) herüberkopiert werden, ohne daß dabei der Editor verlassen werden müßte. Die Abgrenzung des Belegzitats erfolgt dann in der dargelegten Weise mittels der beiden Feldkennungen @ für und @. (das Ergebnis steht wieder in Abb. 3 Mitte).

3. Das Eintragen von Kommentaren.

Bemerkungen des Bearbeiters zu Morphologie, Semantik, Syntax, Lesarten, Forschungsdiskussion etc. können an jeder Stelle des Artikels eingefügt werden und werden von XART entsprechend verwaltet. Unter Punkt 3 von Abb. 2 sind zwei verschiedenartige Kommentare als Beispiele angeführt: Zuerst ein, von einem Beleg unabhängiger, die Semantik betreffender Gliederungskommentar (»a »Schachspiel«), der nicht direkt mit einem einzelnen Beleg verbunden ist und daher ein

```
@-03000 @/ $ III. ]einzelne Arten der Unterhaltung:{
@-03010 @/ $ 1) ]>Brettspiel, Würfelspiel<:{
@-03010 @/ $ a) ]>Schachspiel<:{
@-03010 @[ swaz vuoge er aber an der stete/ mit geb#.arden oder
mit >>spil getete
@+mit geb#.arden oder mit >>spil getete,/ daz was in da wider
alse ein @s 2280 5 @];
@-03010 @ (zabelwortelîn) sprach er wol und kunde ir vil,/ da
mite so zierter in sin >>spil
@+mite so zierter in sin >>spil./ ouch sang er wol ze prise/
schanzûne @s 2292 7 @];
:
:
@-03011 @/ ]evt. Brett und Figuren assoziierend:{
@-03011 @[ sus sazen si zwên über daz spil
@+sâzen si zwên über daz >>spil./ der marschale sprach: 'Tri-
stan, ich wil/ @s 2249 7 @; Brett und Figuren??
@-03020 @/ $ b) ]allg., übertr.:{
@-03020 @[ ir (wip) sit&v daz irresameste >>spil./ daz ieman uf
dem brete kan
@+ir sit&v daz irresameste >>spil./ daz ieman uf dem brete kan./
er ist @s 9888 5 @];
:
:
@-03020 @[ ich w#.âne, ich wil/ mit vrumen und mit êren/ von di-
sem >>spile kâren
@+und mit êren/ von disem >>spile kâren./ ich h#.âte michel ar-
beit/ unsinneclîchen @s 11 60 3 @];
@-03020 @[ dazt iemer soltest uf gegeben/ alse gar gewonnen
>>spil
@+gegeben/ alse gar gewonnen >>spil./ alsolhes spottes wart da
vil/ getriben @s 11357 4 @];
@-03020 @[ so man enmac, der danne wil./ daz ist ein harte
unw#.âge >>spil;/ so man wol mûge, so welle:/ daz ist guot spil-
gevelle
@+daz ist ein harte unw#.âge >>spil;/ so man wol mûge, so wel-
le:/ daz ist @s 16436 6 @];
@-03021 @[ daz spil git ]bzw.{ hat guot gevelle }La.{
@+so welle:/ daz ist guot >>spilgevelle./ dan lit niht herzelei-
des an./ @d 16438 4
@-03050 @/ $$ ::A{ +irresam{
@-03051
@+ir sit&v daz irresameste >>spil./ daz ieman uf dem brete kan./
er ist @d 9888 5 @];
@-03052 @/ +unw#.âge{
@-01053
@+daz ist ein harte unw#.âge >>spil;/ so man wol mûge, so wel-
le:/ daz ist @d 16436 6
```

Weitere Feldkennungen:

@] stellungsbundener Kommentar nach Belegstelle
@d Doppel einer Belegstelle (entstanden aus @r)

Abb. 3: Arbeitsdatei nach der Bearbeitung

eigenes Ordnungszahlfeld (3010) bekommen hat; die Feldkennung für einen Kommentar mit dieser Funktion lautet @/. Der andere angegebene Kommentar («Brett und Figuren??») gehört untrennbar zum Beleg und hat daher keine eigene Ordnungszahl. Seine Feldkennung @; weist ihn als interne Überlegung des Bearbeiters aus, die zwar die ganze Zeit über mitgeführt, jedoch schließlich nicht gedruckt wird. Es mag bei der Nennung dieser beiden wichtigsten Kommentartypen bleiben.

4. Das Verdoppeln einer Belegstelle.

Soll eine Belegstelle an mehr als einem Ort im Artikel erscheinen – in unserem Wörterbuch ist das besonders der Fall bei mehrfach belegten Fügungen, die zu verschiedenen Bedeutungsgruppen gehören, allgemeiner: wenn sich bestimmte Gliederungskriterien des Artikels überschneiden – so wird diese Stellenangabe, zur Vermeidung von Fehlern, nicht vom Bearbeiter selbst mit Standardmitteln, sondern vom Programm verdoppelt. Dies geschieht auf die denkbar einfachste Art, indem hinter der zu verdoppelnden Belegstelle nur die zukünftige Ordnungszahl des Doppels (3051) und die Feldkennung @r angegeben wird. Das Ergebnis erscheint wieder in Abb. 3: Das Doppel zu unserer Stelle 9888 befindet sich nun mitsamt automatisch beigefügtem Kontrollkontext unter der gewünschten »Adresse« 3051. Die verdoppelte Stelle ist deutlich als solche ausgewiesen: Sie trägt die vom Programm eingetragene Feldkennung @d statt @s.

Ist die Bearbeitung eines Artikels abgeschlossen (diesen Zustand dokumentiert der Ausschnitt in Abb. 3), wird er an seinen Ursprungsort in der lemmatisierten Konkordanz zurückkopiert, genauer gesagt, es wird dort die alte, unbearbeitete Version durch die aktuelle, bearbeitete ersetzt. Im Laufe der Arbeit an den Artikeln nimmt so die Konkordanz nach und nach die Gestalt des späteren Wörterbuches an.

Abb. 4 zeigt das Ergebnis einer möglichen Weiterverarbeitung des Artikels *spil* mit den TUSTEP-Standardprogrammen SATZ und SASIM (Satzsimulation). Es ist zu beachten, daß hier nicht das – teure – Endprodukt einer Belichtung, sondern »nur« der Output eines Laserdruckers abgebildet ist. Man kann leicht ermessen, welche Hilfe für einen Lexikographen diese Möglichkeit darstellt, einen fertigen oder noch in Arbeit befindlichen Artikel sofort in der simulierten Form des zukünftigen Satzes begutachten zu können.

Zum Schluß sei der bisher nur beiläufig erwähnte Ablauf des Programmabbaus XART kurz mit Blick auf das Schema von Abb. 5 zusammengefaßt.

Das Programm kann während der Arbeit an einem Artikel beliebig oft gestartet werden. Seine Aufgabe ist die Herstellung einer aktualisierten Version der Hintergrund- wie der Arbeitsdatei. Zu diesem Zweck vergleicht das Programm diese beiden Dateien bzw. wählt aus ihnen die jeweils gültigen Felder in der Weise aus, wie dies über Parameter vom Bearbeiter verlangt wird. So werden bestimmte Felder, die nicht verändert werden dürfen, also die Lemmazeile, die Unterlemmazeilen und die Belegstellen, zuerst daraufhin geprüft, ob sie in beiden Dateien übereinstimmen, bevor sie für die neue Version übernommen werden. Ist dies nicht der Fall, bricht das Programm ab und meldet den Fehler. Andere Felder wieder-

seitspil	spil
<p>spil) kan 7876; wie si ... / schriben lerne alle wege, / latin unde seitspil 11949; er machte unde vant / an iegelichem seitspil / leiche unde guoter noten vil 19197</p> <p>A edele 7876 S zus. mit buoch 2096 7727 7847 7876; spräche 3724 7703; zunge 2096; latin 11949; schriben 11949; jagen 3724; vuoge an ~ 7703 V kunnen 2096 3657 3667 3684 3724 7876; lëren 7727; lernen 7847 11949; sin emezekeit zuo ~ këren 2096; sîner stunde vil an ~ vertuon 2096; an ~ vlizen 3667</p> <p>spil S. (→ bettespil, hantspil, harpfenspiel, hovespil, minnenspiel, rottenspiel, schächzabelspiel, seitspil, vederspil, wunnenspiel)</p> <p>I. <i>allg.: (spielerisches) Tun, Verhalten:</i> zwô schœne hæten ... / ein spil gemachet under in zwein 17581; (der dâ minnet) hât dem ein gelîchez spil 19449; diz ist doch unser vrouwen spil; / waz nimest dû dich hie mit an? 9916</p> <p>S vrouwen ~ 9916 V ein ~ under in machen 17581; ein gelîchez ~ hân 19449</p> <p>II. <i>Unterhaltung, Vergnügen, Freude:</i> sô kan ich spil. daz tuon ich dir, / des ouch dîn herze lîhte gert: / schœniu cleider unde pfert 3732; sô durch ernest sô durch spil 2195; ist ez ernest, ist ez spil, / er ist ie, swie sô man wil 15743; sô minne an tumben kînden / ir spil gerâtet vinden 12432; arcwân / der begunde ... / an Tristande unde Îsôte / sînen wuocher bern unde spil 16467; (von herze und ouge) swer in daz spil leiden wil, / ... der liebet in daz spil. / sô mans ie harter dannen nimet, / sô si des spiles ie mê gezimet / und sôs ie harter clebent an 17827 17828 17830; Petitcreiu. / sînes herzen spil von Avaliu / und sîner ougen gemach 15798; der vindet dâ des herzen spil 17095 daz beste mîner ougen spil / und mînes herzen wunne 16261</p>	<p>S ernest 2195 15743; ~ und gemach 15798; wuocher und ~ 16467; des herzen ~ 15798 17095; der ougen ~ 16261</p> <p>V kunnen 3732; tuon 3732; lieben (<i>lieb machen</i>) 17828; leiden (<i>verleiden</i>) 17827; nemen 17830; gern 3732; gezemen 17830; an cleben 17830; an em sin ~ bern (<i>streiben</i>) 16467; an em sin ~ vinden 12432</p> <p>III. <i>einzelne Arten der Unterhaltung:</i></p> <p>1) <i>Brettspiel, Würfelspiel:</i></p> <p>a) <i>Schachspiel:</i> swaz vuoge er aber an der stete / mit gebærdem oder mit spil getete 2280; (zabelwortelîn) sprach er wol und kunde ir vil, / dâ mite sô zierter in sîn spil 2292; jene die wâren verdâht / an ir spil sô sêre, / daz si dô nihtes mêre / niwan ir spiles gedâhten 2315 2317; Tristan daz spil gewan 2319; <i>evi. Brett und Figuren assoziierend:</i> sus sâzen si zwên über daz spil 2249</p> <p>b) <i>allg., überr.:</i> ir (wîp) sît daz ir resameste spil, / daz ieman ûf dem brete kan 9888; welle, daz dich welle: / daz spil hât guot gevelle 9924; ir tætet ouch alsô, / ir liezet ouch gewonnen spil 11057; ich wæne, ich wil / mit vrumen und mit êren / von disem spile këren 11060; dazt iemer soltest ûf gegeben / alse gar gewonnen spil 11357; sô man enmac, der danne wil, / daz ist ein harte unwæge spil; / sô man wol mûge, sô welle: / daz ist guot spilgevelle 16436; daz spil gît bzw. hât guot gevelle <i>La.</i> 16438</p> <p>A irresam 9888; unwæge 16436 S gebærdem 2280; (spil)gevelle 9924 16436 16438; ~ ûf dem brete 9888 V gewinnen 2319; kunnen 9888; zieren 2292; über daz ~ sitzen 2249; gedenken 2317; an daz ~ verdâht sîn 2315; guot gevelle hân bzw. geben (<i>gute Chancen haben</i>) 9924 16438; gewonnen ~ ûf geben 11357; gewonnen ~ lâzen 11057; von dem ~ e këren 11060</p> <p>2) <i>Musizieren (instrumental und vokal):</i> swaz er in dâ spiles getete / mit handen oder mit munde, / dazn gie niht von grunde: / daz herze dazn was</p>

Abb. 4: Satzsimulation des Artikels *spil*

1. Vergleichen von Arbeitsdatei (1) und Hintergrunddatei (2) bzw. Auswählen der einzelnen Felder aus (1) und (2).

Die jeweilige Behandlung der Felder ist über Parameter steuerbar; im aktuellen Fall gilt:

- a) Felder, die in (1) und (2) gleich sein müssen (sonst Abbruch und Meldung):
 - @a Lemmzeile
 - @u Unterlemmzeile
 - @s Belegstelle
 - b) Felder, die aus (1) genommen werden:
 - @- Ordnungszahl
 - @[ausgewähltes Belegzitat, stellunggebundener Kommentar
 - @/ sonstiger Kommentar (ebenso: @], @; etc.)
 - c) Felder, die aus (2) genommen werden:
 - @+ Kontext
 - d) Felder, die übergangen werden:
 - @. wegfallender Text
2. Merken bestimmter Felder. Es dient im aktuellen Fall zum Verdoppeln von Belegen: Die Felder @+ (Kontext) und @s (Belegstelle) werden gemerkt und nach der Kennung @r (zu verdoppelnder Beleg) ausgegeben.
 3. Mischen der gültigen Felder. Das Zusammenfassen zu (Sortier-)Einheiten ist steuerbar; im aktuellen Fall wird mit dem Feld @- (Ordnungszahl) eine neue Einheit begonnen.
 4. Prüfen der Reihenfolge und Häufigkeit der Feldkennungen pro Einheit. Fehler werden gemeldet und führen zum Abbruch.
 5. Sortieren der Einheiten nach der Ordnungszahl. Ergebnis ist die neue Hintergrunddatei.
 6. Anlegen einer neuen Arbeitsdatei. Zeileneinteilung und Kontextlänge sind steuerbar.

Abb. 5: Funktionen des Programms XART

um werden gleich entweder aus der einen oder der andern Datei geholt: aus der Arbeitsdatei die Ordnungszahl, das abgegrenzte Belegzitat sowie der Bearbeiter-Kommentar, aus der Hintergrunddatei dagegen das unversehrte Kontextfeld; Felder mit der Kennung @. werden – eine vierte Möglichkeit – bei der Auswahl übergangen. Eine weitere Leistung von XART ist das Merken bestimmter Felder. Dies wird in unserem Fall etwa zum bereits genannten Verdoppeln der Belege eingesetzt: die Felder @+ (Kontext) und @s (Belegstelle) werden gemerkt und dann nach der Kennung @r wieder ausgegeben. Die ausgewählten und die neugebildeten Felder werden in einem dritten Schritt gemischt und anschließend auf Reihenfolge und Häufigkeit der Feldkennungen geprüft (so darf etwa jede Kennung nur einmal pro Einheit erscheinen). Gibt es keine Beanstandungen bei dieser Prüfung, werden die Einheiten nach der Ordnungszahl sortiert; es ist so die neue Version der Hintergrunddatei entstanden. Zum guten Schluß legt XART auch eine neue Arbeitsdatei an, und die Bearbeitung des Artikels kann in der geschilderten Weise fortgeführt werden.