

MGR. SYLVA VALIHRACHOVÁ

DIE MÖGLICHKEITEN DER COMPUTERNUTZUNG BEI DER SPRACHWISSENSCHAFTLICHEN ANALYSE MITTEL- UND FRÜHNEUHOCHDEUTSCHER TEXTE

0. Einleitung

In diesem Beitrag wird eine neue, auf der elektronischen Datenverarbeitung beruhende Methode zur sprachwissenschaftlichen Analyse mittel- und frühneuhochdeutscher Texte beschrieben. Diese neue Vorgehensweise wurde als Grundlage für die umfassende sprachwissenschaftliche Analyse des frühneuhochdeutschen Textes „Das Leben des heiligen Hieronymus“ entwickelt, die den Gegenstand meiner Dissertation darstellt. Die neue Methode ermöglicht eine präzisere und schnellere sprachwissenschaftliche Analyse vor allem mittelhochdeutscher und frühneuhochdeutscher Texte aller Textsorten, bei der bis zu einem gewissen Grade alle Sprachbereiche mit einbezogen werden können.

Mit Hilfe des entwickelten Computerprogramms können die graphisch — (phonetisch) — phonologischen und morphologischen Gegebenheiten eines sprachlichen Materials wie auch einige wichtige sprachliche Erscheinungen im Bereich der Syntax und des Wortschatzes erforscht werden. Der gewonnene Belegkorpus ist vollständig und kann statistisch zuverlässig ausgewertet werden. Die Struktur der Database wurde mit dem Ziel entworfen, das Bearbeiten und Auswerten des sprachlichen Materials bzw. des Belegkorpusses nach möglichst vielen vielfältigen Kriterien und Gesichtspunkten zu ermöglichen. Um die Effektivität des Computerprogramms zu erhöhen, haben wir uns bemüht, seine Bedienung so einfach wie möglich zu halten.

Die Struktur der Database ist in der ersten Linie für die sprachwissenschaftliche Analyse mittel- und frühneuhochdeutscher Texte entwickelt worden. Soll das entwickelte Computerprogramm zur Analyse eines althochdeutschen Textes verwendet werden, ist im Bereich der Morphologie mit einigen Einschränkungen zu rechnen. Bei den schwachen Verben muß auf die Einteilung in die drei bzw. vier Gruppen der schwachen Verben verzichtet werden. Bei Adjektiven ist in der Database kein Feld für die Kategorie des Stammes vorgesehen.

Seiner Struktur nach ist das Computerprogramm auch für die Analyse neu-hochdeutscher Texte geeignet. Es ist allerdings die Frage, ob eine derartige sprachwissenschaftliche Analyse neuhochdeutscher Texte sinnvoll wäre.

1. Vorarbeit

Der handschriftliche oder gedruckte Text wird zuerst in die elektronische Form überführt. Das kann mit Hilfe des Textverarbeitungsprogramms Microsoft Word 7.0. geschehen oder der Text wird direkt in das Computerprogramm eingegeben. Der Text wird dem handschriftlichen oder gedruckten Original treu wiedergegeben. Es werden sowohl die Textseiten als auch die Seitenzeilen beibehalten. Die im Originaltext benutzte Groß- und Kleinschreibung wird bewahrt. Die Buchstaben, die im Text als Initialen geschrieben werden, werden mit speziellem Zeichen versehen. Für bestimmte graphische Sonderzeichen wurden ihre elektronischen Entsprechungen zu den Zwecken des Computerprogramms hergestellt.

Der in die elektronische Form überführte Originaltext kann mit maschineller Hilfe weiter bearbeitet werden. Wird der Originaltext nicht direkt in das Computerprogramm geschrieben, muß er zuerst mit dem Befehl „Lies den Text ein“ in das Computerprogramm übernommen werden. Der eingelesene oder direkt in das Computerprogramm geschriebene Text wird mit dem Befehl „Überführe in die Datenform“ in den Bereich des Computerprogramms überführt, in dem die eigentliche sprachwissenschaftliche Analyse des sprachlichen Materials durchgeführt werden kann. Jedem Wort, das im Text vorkommt, werden nun seine sprachwissenschaftlichen Charakteristika zugeordnet. Um eine möglichst umfassende sprachwissenschaftliche Analyse des sprachlichen Materials durchführen zu können, wurde eine Database entworfen, in die bei jedem im analysierten Text vorkommenden Wort Informationen sowohl graphischer und phonologischer als auch morphologischer und zum Teil auch syntaktischer Art angegeben werden können.

2. Eigentliche Analyse

Bei der eigentlichen Textanalyse erscheint der zu analysierende Text in der rechten Hälfte des Bildschirms. In der linken Hälfte des Bildschirms sind die Felder der Database zu sehen. In diese Felder werden die sprachlichen Informationen über einzelne Wörter des Textes eingegeben. Bei jedem Wort sind zunächst lediglich vier Felder zu sehen. Das Feld „Beleg“, in das die belegte graphische Form des aktuell analysierten Wortes vom Text in der rechten Hälfte des Bildschirms automatisch vom Computerprogramm kopiert wird. Das Feld „Phonem“, in das die Information über die phonologische Gestalt des Wortes eingetragen wird. Das Feld „Graphem“, in das die Information über die graphische Gestalt des Wortes eingegeben wird, und das Feld „Wortart“ zur Angabe der Wortart des analysierten Wortes.

Das zu analysierte Wort muß zuerst mit der Maus in der rechten Hälfte des Bildschirms markiert werden. Mit Hilfe der entsprechenden Richtungsbefehle ist es möglich, sich innerhalb des Textes in der rechten Hälfte des Bildschirms zu bewegen. Man kann zum nächsten links oder rechts liegendem Wort übergehen oder im gesamten Text nach Textseiten nach vorne oder nach hinten blättern. Außerdem ist es möglich, auf der aktuellen Seite des Textes mit der Maus ein beliebiges Wort anzuklicken und sich auf diese Weise im Text zu bewegen. Wurde die Analyse eines bestimmten Wortes durchgeführt, muß sie von dem Sprachwissenschaftler mit Hilfe des Befehls „Ich bestätige“ bestätigt werden. Dies ist vor allem in den Fällen wichtig, in denen der Computer selbst einen Vorschlag zur Analyse des aktuell analysierten Wortes macht.

Da bei der Überführung des Originaltextes in die elektronische Form die Textseiten und Seitenzeilen beibehalten wurden, und bei der Entwicklung der Struktur des Computerprogramms daran gedacht wurde, müssen die Informationen, die die Position der Wörter im Textes betreffen, d.h. die Seiten- und Zeilenzahl, dem Computer nicht extra in ein spezielles Feld der Database angegeben werden. Dies stellt vor allem bei dem häufig großen Umfang des zu analysierenden Sprachmaterials eine wesentliche Arbeitserleichterung dar.

2. 1. Phonem

In das Feld „Phonem“ werden alle Phoneme des betreffenden Wortes nach bestimmten Kriterien angegeben, die das spätere Auswerten der sprachwissenschaftlichen Informationen innerhalb des Computerprogramms ermöglichen. Bei der Festlegung dieser Kriterien haben wir uns zum Teil auf die Untersuchungsmethode gestützt, die Prof. Peter Ernst in seiner Habilitationsarbeit „Die Anfänge der frühneuhochdeutschen Schreibsprache in Wien“ verwendet und die einen der ersten Computereinsätze bei der Analyse mittel- und frühneuhochdeutscher Texte darstellt. Den frühesten verdienstvollen Einsatz auf dem Gebiet der Computernutzung bei der sprachwissenschaftlichen Analyse frühneuhochdeutscher Texte stellt jedoch die Arbeit von I. T. Piirainen „Graphematische Untersuchungen zum Frühneuhochdeutschen“ dar, die sich — ähnlich wie die Arbeit von P. Ernst — auf den Bereich der graphischen und phonologischen Erscheinungen konzentriert.

In das Feld „Phonem“ werden diejenige Phoneme eingetragen, die in dem betreffenden Wort auftreten. Die einzelnen Phoneme werden zwischen Bindestriche gesetzt. Die Vokale der Haupttonsilben werden mit dem Zeichen # versehen, um bei dem Suchen und weiterem Bearbeiten von den Vokalen der Nebentonsilben unterschieden werden zu können. Komposita und Nebensilben werden an der Fuge mit dem Zeichen + getrennt. Die Benutzung dieser Ordnungskriterien ist erforderlich, um innerhalb des Computerprogramms gezielt nach einzelnen Phonemen oder Phonemgruppen suchen und diese weiter bearbeiten zu können. Das Bezugssystem der Phoneme kann der Sprachwissenschaftler je nach dem Gegenstand und der Zielsetzung seiner Forschungsarbeit frei wählen. Das gewählte Bezugssystem der Phoneme muß konsequent ange-

halten werden und nur diejenige Zeichen und Sonderzeichen beinhalten, die dem Computerprogramm zur Verfügung stehen.

Aufgrund der Ordnungskriterien, kann man gezielt nach dem Vorkommen bestimmter Phoneme oder Phonemkombinationen suchen, wobei die Position des Phonems oder der Phonemgruppe innerhalb des Wortes berücksichtigt werden kann. Phoneme, die im Anlaut auftreten, sind — den Ordnungskriterien nach — lediglich mit einem Bindestrich hinter dem Phonemzeichen versehen. Mit dem entsprechenden Suchbefehl, in den das Phonemzeichen und der Bindestrich angegeben werden, können alle Worte mit dem betreffenden Phonem im Anlaut gesucht werden. Bei Phonemen, die im inneren Anlaut vorkommen, steht das Zeichen + vor und der Bindestrich hinter dem Phonemzeichen. Phoneme im Inlaut stehen zwischen zwei Bindestrichen. Phoneme im Auslaut haben ein Bindestrich vor dem Phonemzeichen, bei Phonemen im inneren Auslaut, steht hinter dem Phonemzeichen das Zeichen +. Dies garantiert, daß mit entsprechenden Suchbefehlen nur Phoneme in den gewünschten Wortpositionen gesucht werden können.

2. 2. Graphem

Das Angeben der sprachlichen Informationen in das Feld „Graphem“ erfolgt analog zu Feld „Phonem“. Um die Arbeit zu erleichtern, erscheinen im Feld „Graphem“ die bei dem betreffenden Wort belegten graphischen Zeichen automatisch und werden bereits mit den Bindestrichen dazwischen versehen. Der Sprachwissenschaftler kann nach Bedarf die Bindestriche löschen, sie durch + Zeichen ersetzen oder neu einfügen, er kann ebenfalls das Zeichen # zur Bezeichnung der Stammsilbe einfügen. Da die Ordnungskriterien im Feld „Graphem“ identisch mit denjenigen im Feld „Phonem“ sind, geschieht auch das Suchen nach bestimmten Schreibungen (Graphien) innerhalb des Feldes „Graphem“ analog zu Feld „Phonem“. Man muß nur die gewünschte Schreibung mit oder ohne die Information über ihre Wortposition mit dem Suchbefehl im Feld „Graphem“ suchen.

2. 3. Morphem

Dasselbe Suchprinzip kann auch beim Suchen der Morpheme benutzt werden, weil ein und dasselbe Morphem bei der Phonemerfassung der Worte immer auf dieselbe einheitliche Weise eingegeben wird. Um ein bestimmtes Morphem zu suchen, muß man nur das betreffende Morphem mit oder ohne die Informationen über seine Wortposition in den Suchbefehl im Feld „Phonem“ eingeben.

2. 4. Wort

Wie bereits unter den Punkten 2.1. und 2.2. beschrieben wurde, kann in den Feldern „Phonem“ und „Graphem“ gezielt nach einzelnen Phonemen und Phonemgruppen bzw. Graphien gesucht werden. Dabei kann die Wortposition der Phoneme und Grapheme berücksichtigt werden. Da ein bestimmtes Wort bzw. eine bestimmte Wortform in das Feld Phonem immer auf dieselbe einheitliche

Weise eingetragen wurden, ist es im Feld „Phonem“ ebenfalls möglich, nach ganzen Wörtern bzw. Wortformen zu suchen. Dies ist vor allem bei der Wortschatzanalyse eines Textes von Bedeutung. Analog dazu besteht im Feld „Graphem“ die Möglichkeit, nach verschiedenen Schreibungen eines bestimmten Wortes zu suchen.

3. Wortart

Beim Feld „Wortart“ ist die Arbeitsweise anders. Die entsprechende Wortart wird nicht in das Feld mit Hilfe der Tastatur eingeschrieben, sondern dem Sprachwissenschaftler steht bereits eine Liste der Wortarten zur Verfügung, die in einem mittel- oder frühneuhochdeutschen Text vorkommen können. Es genügt, das Feld mit der entsprechenden Wortart mit der Maus anzuklicken, um dem Computerprogramm die gewünschte Information zu vermitteln. Je nach der Wortart erscheinen dann in der linken Hälfte des Bildschirms unter dem Feld „Wortart“ weitere Felder der Database, in die die entsprechenden morphologischen Kategorien der jeweiligen Wortart angegeben werden. Bei diesen Feldern werden ebenfalls Auswahllisten angeboten, oder es besteht die Möglichkeit einer „ja — nein — Antwort“. Auf allen Ebenen der Database ist es möglich, das Feld „?“ zu wählen, denn nicht bei jedem Wort können alle sprachwissenschaftlichen Charakteristika mit Sicherheit bestimmt werden.

3. 1. Die Struktur der Database

Für den Bereich der morphologischen Merkmale haben wir also bei der Strukturierung der Database ein anderes Ordnungsprinzip als für den Bereich der phonologischen und graphischen Erscheinungen verwendet. Jeder morphologischen Kategorie wurde ein Feld der Database zugeordnet. Je nach der Wortart des analysierten Wortes erscheinen die entsprechenden Felder der Database, in die die morphologischen Merkmale des aktuell analysierten Wortes eingetragen werden können. Die Struktur der Database mußte den Suchmöglichkeiten der Elektronischen Datenverarbeitung angepaßt werden. Die Felder der Database entsprechen in manchen Fällen nicht den morphologischen Kategorien, wie sie in den Standardgrammatiken des Mittel- und Frühneuhochdeutschen dargestellt werden. Dies betrifft vor allem die Auswahllisten der einzelnen morphologischen Kategorien, bei denen in manchen Fällen Änderungen vorgenommen werden mußten, um ein effektives programminternes Suchen und Bearbeiten des sprachlichen Materials zu ermöglichen.

Nach der Angabe der Wortart erscheinen bei einigen Wortarten unter dem Feld „Wortart“ außer den Feldern der Database, in die die Informationen über die morphologischen Merkmale der analysierten Wörter eingegeben werden, auch Felder, in die einige Angaben über die syntaktischen Funktionen der analysierten Wörter gemacht werden können. Auch diese Felder sind erst dann auf dem Bildschirm zu sehen, wenn bei dem analysierten Wort bereits seine Wortart

bestimmt wurde. Die Angaben über einige syntaktische Funktionen der analysierten Wörter ermöglichen es, diese bei dem Bearbeiten des Belegkorpusses in Betracht zu ziehen und bei der Auswertung der sprachlichen Erscheinungen zu berücksichtigen. Auf diese Weise können die auf die syntaktische Rolle der Wörter gebundenen Besonderheiten in der Flexion oder Graphik der Wörter getrennt behandelt und ergriffen werden.

3. 1. 1. *Substantive*

Bei Substantiven sind die morphologischen Kategorien des Kasus, Numerus, Genus wie auch des Stammes zu bestimmen. Dementsprechend erscheinen bei den Wörtern, die als Substantive identifiziert wurden, unter dem Feld „Wortart“ die Felder „Kasus“, „Numerus“, „Genus“ und „Stamm“. Analog zum Feld „Wortart“ werden auch bei diesen Feldern die möglichen Antworten vorgegeben. Somit besteht beim Feld „Kasus“ die Möglichkeit, mit der Maus den Nominativ, Genitiv, Dativ, Akkusativ oder „?“ auszuwählen. Im Feld „Numerus“ wird zwischen Singular, Plural und „?“ unterschieden. Beim Feld „Genus“ besteht die Wahl zwischen Maskulinum, Femininum, Neutrum und „?“. Im Feld „Stamm“ werden alle mhd. Stämme, d.h. die *-a-*, *-ja-*, *-wa-*, *-ô-*, *-jô-*, *-wô-*, *-i-*, *-u-*, *-os-/es-*, *-n-*, *-nt-/nd-*, *-r-* Stämme, aufgelistet. Außerdem ist es möglich, in diesem Feld die Kategorie Wurzelnamen (WN), Eigenname (EN), Fremdwort (FW), substantiviertes Adjektiv (SA), substantiviertes Partizipialadjektiv (SPA), substantiviertes Infinitiv (SI) und „?“ zu wählen.

3. 1. 2. *Adjektive*

Wird ein Wort als Adjektiv identifiziert, erscheinen unter dem Feld „Wortart“ analog zu den Substantiven die Felder „Kasus“, „Numerus“, „Genus“, in die der Kasus, Numerus und Genus des aktuell analysierten Adjektivs eingegeben werden können. In das Feld „Stufe“ kann die Angabe über die Steigerungsstufe des analysierten Adjektivs eingetragen werden. Die Auswahlliste dieses Feldes enthält die drei Steigerungsstufen, nämlich den Positiv, Komparativ und Superlativ, wie auch die Alternative „nicht steigerungsfähig“.

Ist ein Partizip I oder II in einem Satz des analysierten Textes als ein Adjektiv verwendet, muß bei der Bestimmung der Wortart im Feld „Wortart“ die Kategorie des Adjektivs gewählt werden. In das Feld „Ursprung“ kann dann die Information eingegeben werden, ob es sich um ein Partizip I oder Partizip II handelt. Dies ermöglicht den Vergleich der Partizipialformen, die zur Bildung der analytischen Verbformen verwendet werden, mit denen, die in einem Satz als Adjektive oder substantivierte Adjektive gebraucht werden.

In das Feld „Stellung“ wird die Information eingegeben, ob ein Adjektiv im Satz in attributiver oder prädikativer Stellung vorkommt. Auf diese Weise können die Adjektive in prädikativer und attributiver Stellung getrennt behandelt werden. Dies ist vor allem bei der Erforschung der adjektivischen Flexion von großer Bedeutung. Die Struktur der Database ermöglicht es, ebenfalls die sub-

stantivierten Adjektive mit einzubeziehen, so daß eine vollständige Untersuchung dieses Phänomens durchgeführt werden kann.

3. 1. 3. *Pronomina*

Bei den Pronomina variieren die Felder der Database nach den Untergruppen. Aus diesem Grund muß zuerst entschieden werden, ob es sich bei dem analysierten Pronomen um ein Personal-, Relativ-, Possessiv-, Demonstrativ-, Interrogativ-, Indefinit- oder Reflexivpronomen handelt. Ist das zu analysierende Pronomen ein Personalpronomen, muß außerdem die Information vermittelt werden, ob es ein geschlechtiges und ungeschlechtiges Personalpronomen ist. Bei den ungeschlechtigen Personalpronomina werden die Angaben zu Person, Numerus und Kasus in die entsprechenden Felder „Person“, „Numerus“ und „Kasus“ gemacht. Bei den geschlechtigen Personalpronomina werden in die Felder „Genus“ und „Kasus“ die Informationen über den Genus und Kasus des betreffenden Personalpronomens eingetragen. Im Feld „Person“ erscheint automatisch die 3. Person, im Feld „Numerus“ die Kategorie des Singulars.

Bei den Relativ-, Demonstrativ- und Indefinitpronomina ist es möglich, entweder die Angaben zu Kasus, Numerus, und Genus in die entsprechenden Felder der Database einzutragen oder die Alternative „Erstarrte Form“ zu wählen. Bei den Possessivpronomina erscheinen die Felder „Person“, „NumerusP“, „GenusP“ sowie „Kasus“, „Numerus“ und „Genus“, in die die jeweiligen morphologischen Merkmale des zu analysierenden Possessivpronomens eingetragen werden. Bei den Interrogativpronomina werden die Informationen über Kasus, Numerus und Genus des betreffenden Interrogativpronomens eingegeben. Bei Reflexivpronomina sind die morphologischen Kategorien der Person, des Numerus und des Kasus relevant. Dementsprechend erscheinen auch die dazugehörigen Felder „Person“, „Numerus“, „Kasus“.

Bei allen Untergruppen der Pronomina sind also die Felder „Numerus“ und „Kasus“ zur Angabe der entsprechenden morphologischen Information vorhanden. Ebenfalls das Feld „Genus“ erscheint mit der Ausnahme der ungeschlechtigen Personalpronomina und des Reflexivpronomens bei allen Untergruppen. Das Feld „Person“ kommt bei den Personalpronomina und bei den Possessivpronomina vor, wobei bei den geschlechtigen Personalpronomina in dieses Feld automatisch vom Computerprogramm die 3. Person eingetragen wird. Das Feld „Erstarrte Form“ ist bei den Relativ-, Demonstrativ- und Indefinitpronomina vorgesehen. Bei allen Feldern ist es möglich, in den Auswahllisten das Feld „?“ anzugeben.

Die Angabe über die Rolle des zu analysierten Pronomens im konkreten Satz ist bei den Possessiv-, Demonstrativ-, Interrogativ- und Indefinitpronomina in das Feld „Gebrauch“ einzutragen. Es wird zwischen den Pronomina unterschieden, die in attributiver Stellung im Satz vorkommen, und denen, die im Satz stellvertretend für Substantive stehen. Diese Unterscheidung ermöglicht es, die eventuellen, an die syntaktische Rolle des Pronomens gebundenen Unterschiede in der Form der jeweiligen Pronomina zu erforschen.

3. 1. 4. *Artikelwörter*

Ist das analysierte Wort ein Artikelwort, muß zuerst im Feld „Artikel“ entschieden werden, ob es sich um einen bestimmten oder unbestimmten Artikel handelt. Danach werden in die Felder „Kasus“, „Numerus“ und „Genus“ die übrigen morphologischen Charakteristika des aktuell analysierten Artikelwortes angegeben.

3. 1. 5. *Verba*

Bei Verben ist die Struktur der Database am kompliziertesten. Nachdem das aktuell analysierte Wort im Feld „Wortart“ als Verb identifiziert wurde, sind in der linken Hälfte des Bildschirms zunächst vier Felder zu sehen: das Feld „Form“, „Funktion“, „Gruppe“ und „Reflexiv“.

Im Feld „Form“ wird abgefragt, ob es sich bei der analysierten Verbform um ein Verbum Finitum oder eine infinite Verbform handelt. Wird dem Computerprogramm die Information vermittelt, daß die analysierte Verbform eine finite Verbform ist, bietet das Computerprogramm dem Sprachwissenschaftler die Möglichkeit, die Angaben über die Person, Numerus, Tempus und Modus der aktuell analysierten Verbform in die entsprechenden Felder der Database einzutragen. Im Feld „Person“ wird die Auswahlliste mit den drei grammatischen Personen, d.h. der 1. 2. und 3. Person, angeboten. Im Feld „Numerus“ stehen die zwei Alternativen Singular und Plural zur Wahl. Beim Feld „Tempus“ besteht die Möglichkeit lediglich zwischen dem Präsens oder dem Präteritum zu wählen, weil diese die einzigen synthetischen Tempusformen im Mittelhochdeutschen wie auch im Frühneuhochdeutschen darstellen. Im Feld „Modus“ kann der Indikativ, Konjunktiv und Imperativ gewählt werden. Im Feld „GenusVe“ wird automatisch vom Computerprogramm die Angabe „Aktivum“ eingetragen, denn das Passivum nur analytisch gebildet werden kann.

Ist die analysierte Verbform eine infinite Verbalformen, sind folgende mögliche Antworten vorgegeben: Infinitiv, Gerundium 2, Gerundium 3, Partizip I, Partizip II., wobei die Alternativen „Gerundium 2“ und „Gerundium 3“ für die flektierten Formen des Infinitivs, d.h. für die ursprünglichen Genitivformen bzw. Dativformen eines Gerundiums, stehen.

Wurden die Angaben über die morphologischen Merkmale der Verbalform gemacht, ist die Verbalform im Feld „Funktion“ unter dem syntaktischen Aspekt zu analysieren. Im Feld „Funktion“ wird zuerst zwischen den Hilfsverben und den Vollverben unterschieden. Kommt eine Verbalform im Satz als Vollverb vor, muß lediglich das entsprechende Feld der Auswahlliste angeklickt werden. Die im Feld „Verbum finitum“ bereits durchgeführte Analyse wird dadurch vollständig und es erscheinen keine weiteren Felder der Database mehr. Ist die Verbalform ihrer Funktion nach ein Hilfsverb, wird im Feld „Hverb“ abgefragt, ob sie zur Bildung eines zusammengesetzten Tempus, zum Ausdruck der Modalität oder des Passivs verwendet wird. In der angebotenen Auswahlliste können in diesem Falle mehrere Varianten gleichzeitig eingegeben werden, denn das analysierte Hilfsverb zur Bildung mehrerer analytischer Formen gleichzeitig

benutzt werden kann. Wird das Verb zur Bildung eines zusammengesetzten Tempus verwendet, fragt das Computerprogramm ab, um welches zusammengesetzte Tempus es sich handelt.

Im Feld „Gruppen“ wird die Einteilung in die Untergruppen „Schwaches Verb“ (SchwV), „Starkes Verb“ (StV), „Präteritopräsentium“ (PP), „Wurzelverb“ (WV), das Verb „haben“, „sein“ und „wollen“ durchgeführt. Bei den starken Verben und den Präterito-Präsentien kann zusätzlich in das Feld „Klasse“ die entsprechende Ablautsreihe eingegeben werden. Bei schwachen Verben ist es möglich, im Feld „Vokalwechsel“ zwischen den Verben mit und ohne Wechsel des Stammvokals zu unterscheiden.

Im Feld „Reflexiv“ wird die Frage beantwortet, ob es sich bei der analysierten Verbform um ein reflexives oder nicht reflexives Verb handelt oder nicht.

3. 1. 6. *Adverbien*

Bei Adverbien als einer unflektierten Wortart wird lediglich die Steigerungsstufe bestimmt. Es erscheint das Feld „Stufe“, in dem die Alternativen Positiv, Komparativ und Superlativ sowie „nicht steigerungsfähig“ gewählt werden können.

3. 1. 7. *Konjunktionen*

Im Falle der Konjunktionen kann im Feld „Art“ zwischen den koordinierenden und subordinierenden Konjunktionen unterschieden werden. Auf weitere Differenzierung der Konjunktionen nach den semantischen Kriterien der Satzbeziehung wurde im Hinblick auf die Zielsetzung der Dissertation verzichtet.

3. 1. 8. *Zahlwörter, Präpositionen, Partikeln und Interjektionen*

Bei Zahlwörtern, Präpositionen, Partikeln und Interjektionen können keine weiteren sprachwissenschaftlichen Informationen angegeben werden. Bei diesen Wortarten wird lediglich die Information über die jeweilige Wortart vermittelt.

4. Suchmöglichkeiten in der Database

In einer so strukturierten Database kann dank dem heutigen Entwicklungsstand der Elektronischen Datenverarbeitung und den Eigenschaften des benutzten Programmiermediums innerhalb der jeweiligen Wortarten mit den entsprechenden Suchbefehlen gezielt nach einzelnen morphologischen Kategorien gesucht werden. Dabei besteht die Möglichkeit, mehrere morphologische Kategorien als Suchkriterien gleichzeitig zu berücksichtigen. Bei den Substantiven kann also z.B. nur nach allen a-Stämmen gesucht werden oder es können auch die Kategorien des Numerus und Kasus als zusätzliche Kriterien mit einbezogen werden, so daß das Computerprogramm dem Sprachwissenschaftler lediglich die Auflistung aller a-Stämme im Nom. Sg. liefert.

Mit dem Suchen innerhalb der jeweiligen Wortarten kann das Suchen in den Feldern „Phonem“ und „Graphem“ kombiniert werden. Innerhalb des bereits gefundenen Belegkorpusses bestimmter sprachlicher Erscheinungen ist also ein weiteres Sortieren möglich. Das Computerprogramm kann innerhalb dieses Belegkorpusses nach bestimmten Wörtern suchen und dabei ihre unterschiedlichen Schreibungen berücksichtigen. Die einzelnen Belege der gesuchten sprachlichen Erscheinungen können alphabetisch eingeordnet werden. Es können Gruppen der identischen Wortformen zusammengestellt werden, und die Anzahl der Belege in der jeweiligen Gruppe kann dem Sprachwissenschaftler vermittelt werden. Bei jedem Suchen kann die Information über die Position des jeweiligen Belegs im Text in Klammern hinter dem Beleg eingegeben werden.

Innerhalb des analysierten Textes kann mit den entsprechenden Suchbefehlen ebenfalls nach linearen Reihenfolgen bestimmter Wörter oder Wortarten gesucht werden. Dabei können wiederum die morphologischen Merkmale der gesuchten Wörter oder Wortarten berücksichtigt werden. Diese Funktion des Computerprogramms ist für die Erforschung der Wortfolge unersetzbar. Sie ermöglicht jedoch auch manche detailliertere Untersuchungen im Bereich der Morphologie. So können z. B. die Reihenfolge bestimmter Artikel, Adjektiv, Substantiv im Nom. Sg. Mask. und die Reihenfolge unbestimmter Artikel, Adjektiv, Substantiv im Nom. Sg. Mask. gesucht und die Unterschiede in der Flexion der Adjektive in den Stellungen nach dem bestimmten und unbestimmten Artikel ergriffen werden.

5. Automatische Analyse

Das Computerprogramm wurde so konstruiert, daß es in bestimmten Fällen dem Sprachwissenschaftler seinen eigenen Vorschlag der Analyse einer Wortform machen kann. Es ist in diejenigen Fällen möglich, in denen die aktuell analysierte Wortform graphisch mit einer Wortform identisch ist, die das Computerprogramm bereits analysiert hat. Der Computer behält in seinem Speicher alle sprachwissenschaftlichen Charakteristika, die dieser Wortform bei der letzten Analyse zugeordnet wurden, und schlägt vor, diese Charakteristika auch der aktuell analysierten Wortform zuzuordnen. Da das Computerprogramm auf dieser Ebene lediglich mit der graphischen Form der Wörter arbeiten kann, trifft seine Analyse nicht immer zu, und muß vom Sprachwissenschaftler oft korrigiert werden. Der Sprachwissenschaftler muß entweder die ganze Analyse der Wortform neu durchführen oder nur einzelne Angaben ändern. Die vom Computer selbst durchgeführte Analyse einer Wortform mit oder ohne einer Korrektur des Sprachwissenschaftlers muß vom Sprachwissenschaftler mit dem Befehl „Bestätigung“ („Ich bestätige“) bestätigt werden. Ebenfalls diese Funktion des Computerprogramms stellt für den Sprachwissenschaftler eine weitere Arbeitserleichterung dar.

6. Zusammenfassung

Bei der sprachwissenschaftlicher Analyse mittel- und frühneuhochdeutscher Texte ist es in der ersten Linie erforderlich, einen vollständigen, statistisch zuverlässigen Belegkorpus zu gewinnen. Die traditionelle, manuell durchgeführte Analyse der Texte ist auf Grund des häufig umfangreichen Sprachmaterials ein sehr mühsames und zeitaufwendiges Verfahren, das nur partiell zu den gewünschten Ergebnissen führt. Mit Hilfe des entwickelten Computerprogramms ist es möglich, einen vollständigen Belegkorpus des analysierten Sprachmaterials zu gewinnen, der statistisch vollkommen zuverlässig ausgewertet werden kann. Das analysierte sprachliche Material kann von verschiedensten Perspektiven und nach beliebigen Kriterien bearbeitet und ausgewertet werden. So kann die traditionelle Methode dank der Möglichkeiten der EDV durch ein schnelleres und zuverlässigeres Verfahren ersetzt werden, daß zu einer erheblichen Arbeitserleichterung führt und ein detailliertes und präzises Bearbeiten des gewonnenen Belegkorpusses unter unterschiedlichen Gesichtspunkten und nach spezifischen Bedürfnissen der jeweiligen sprachwissenschaftlichen Analyse ermöglicht.

LITERATUR

- Ernst, Peter: Die Anfänge der frühneuhochdeutschen Schreibsprache in Wien; Wien 1994.
Piirainen, I. P.: Graphematische Untersuchungen zum Frühneuhochdeutschen; Berlin 1968.

