

Allgemeiner Korpusaufbau und Metadaten

Korpusaufbau

Das Bonner Frühneuhochdeutsche Korpus ist nach folgender Struktur aufgebaut:

Man findet auf der Internetseite: „www.Korpora.org/Fnhd“ eine riesige Auswahl, sortiert nach Ort, Zeit und sprachlicher Struktur, verschiedene Korpora.

Diese sind konkret in sinnvoll übersichtliche Zeilen gegliedert, welche wiederum mit Zeilenangabenummern versehen sind. Jedes einzelne Wort in der Überschrift erhält jeweils eine eigene Zeile. Das Korpus wird des Weiteren nach Seiten des Original-/ Ursprungstextes strukturiert:

Wenn in der Originalfassung die Seite eins beendet ist, wird innerhalb des Korpus eine Pause in der „Textlandschaft“ ersichtlich. Außerdem wird in dieser Pause der Hinweis: „Seite 1“ in eckigen Klammern angegeben, sodass man erneut daran erinnert wird die in folgendem Falle Seite eins vorliegen zu haben.

Spezielle Metadaten

Grundsätzlich wird im Rahmen des Bonner frühhochneudeutschen Korpus- Programms zwischen XML- und der HTML- Version unterschieden.

Die Schriftsammlungen haben zunächst den allumfassenden einheitlichen Aufbau:

Die als Überschriften kategorisierenden Elemente sind in größerer und „fetter“ Schrift wiedergegeben. Regelmäßige Eingriffe in den Text, das heißt, reformulierte, später hinzugefügte Passagen sind *serifenlos* (= eine spezifische Schriftart) gestaltet.

Während im Korpus Hervorhebungen mittels eines `>emph<` erfasst sind, wird in der HTML Version die „fette“ Form der auszudrückenden Wörter abgebildet. Zitate sind in kursiver Schrift oder aber in gestellter Schrift realisiert. Namen innerhalb des Korpus, wie Aristoteles beispielsweise, werden unterstrichen. Die Zeichen insgesamt werden in beiden Versionen als Uni- Code Zeichen wiedergegeben, was einen bestimmten Browser erforderlich macht. Die Beschließung dieser „Machart“ geht auf die ersten Realisierungsformen aus dem Jahre 2007 zurück und haben seitdem keine notwendigen Änderungen erfahren. Am Text selbst sind weitere, erläuterungswürdige Hilfen installiert:

Begibt man sich mit dem „Mauspfeil“ auf einzelne Wörter des Textes erhält man einen umfassenden Kriterienkatalog zur grammatischen Einordnung des Wortes. In einem neben dem

Text erscheinenden Kästchen werden explizite Auskünfte die Lemmata, Prä- und Suffixe, Vokale, Worttypen, Kasus, Numerus und Genus gegeben.

Einen ähnlichen Umfang betragen die Angaben des Verbs, wobei hier die Wortklasse, das Tempus und der Modus des Verbs im Vordergrund stehen.

Annotationen im Bonner Frühneuhochdeutschkorp

Im Bonner Frühneuhochdeutschkorp findet man in jedem einzelnen Text zu jedem einzelnen Wort Angaben über dieses, wenn man dies anklickt. So findet man Informationen über Wortart, Flexion, (je nach Wortart) Geschlecht und anderen Merkmalen.

Ein Substantiv kann so zum Beispiel wie hier gezeigt annotiert sein:

→ *pueche*

The screenshot shows a web browser window with the URL <http://www.korpora.org/Fnhd/>. The page title is "FnhdC Das Bonner Frühneuhochdeutschkorp" with a link to "Korpora.org 2007". The main content is a text block starting with "02 ZUM" and "03 RATIONALE DIVINORUM OFFICIORUM". The text is from Aristotle's "Rationale Divinorum Officiorum" and includes a quote from Moses: "04 Aristoteles der schreybet in dem pueche von der auzrichtung der gemaine: *We dem lannde, dez cbünig ein cbünd ist und des fürsten frue ezzen*, wann chindes synne und unweishait lannt und lëutte rechtlich nicht gerichteten mogen und fürsten, di da frue ezzen, dabei wir all unmezichait versten, irem chünige nicht geraten mugen, waz des lanndes notdürft sei, sunder si gedenkchent anders nicht denn ired lustes, wie si den volpringen, und daz mag ane lanndes schaden nicht geschehen. Wann an dem ersten seint si gabegierig, darumb daz recht vercheret wirt, wider daz Moyses gesprochen hat zu dem Volkch von Israhel: *Tu solt nicht gabe nemen, di auch di weisen plendent und vercherent der gerechten wort*. Darumb auch bei alter romischer macht der grosten peen vervallen waz, der umb gerechtichait der gabe gerte. 16 Si seint auch ane in selber waich, darumb si des chuniges ere und des

On the right side, a tooltip for the word "pueche" is displayed with the following information:
pueche
Lemma: buch
Vokal: ue
Typ: substantiv
Kasus: dativ
Numerus: singular
Genus: neutrum

So sieht man, dass

die Wörter umfangreich annotiert sind, wie auch folgendes Verb:

→ *schreybet*

The screenshot shows the same web browser window as above. The text block is identical, but the tooltip for the word "schreybet" is displayed on the right with the following information:
schreybet
Lemma: schreiben
Typ: verb
Klasse: stark_1a
Numerus: singular
Tempus: praesens
Modus: indikativ
Person: 3
Form: finit

Die Annotationen sind also hilfreich, anhand des Bonner Frühneuhochdeutschkorpus zahlreiche Aspekte zu untersuchen.

Kurze Anleitung zur Benutzung

Das Korpus wurde für die Erforschung der Flexionsmorphologie im Frühneuhochdeutschen zusammengestellt. Es reicht aber für bestimmte Phänomene (z.B. für seltene Konstruktionen) nicht aus. Das Korpus kann man auch für andere Arten der Untersuchungen benutzen. So empfehlen z.B. Solms und Wegera, es für Erforschungen lexikalischer Entwicklungen zu verwenden, z.B. für Untersuchungen zur Wortbildung, Entlehnungen, Einzelwortkonkurrenz usw.

Das Korpus ist online verfügbar. Man kann alle 40 Texte durchsuchen oder nur bestimmte Texte auswählen, wenn man Phänomene nur aus einem Zeitraum oder aus einer Gegend suchen will.

Die Suchmaske scheint auf den ersten Blick sehr kompliziert und unübersichtlich zu sein. Es gibt zwar überall kleine Anleitungstexte zu bestimmten Funktionen, aber sie sind teilweise schwer ohne weiteres zu verstehen. Eine andere Besonderheit, die auf den Zweck der Zusammenstellung des Korpus zurückzuführen ist, ist, dass man die Lemmata in die Suchmaske ohne Affixe angeben muss. Die Affixe kommen dann erst in anderen Feldern der Suchmaske dazu, wobei auch die Mehrfachwahl möglich ist.

Ausgefilterte Daten lassen sich nicht exportieren. Sie können nur rauskopiert und manuell z.B. in einer Word-Datei bearbeitet werden. Allerdings gibt es eine Hilfsfunktion, die für lexikalische Untersuchungen besonders passend ist: wenn man in den Ergebnissen auf den Link zu einem bestimmten Text klickt, wird der Text auf der Stelle geladen, wo das gesuchte Wort vorkommt. Das ist besonders dann nützlich, wenn man genauere lexikalische Bedeutung herausfinden muss und dafür mehr Kontext braucht. Die Annotationen zu allen annotierten Wörtern in den Texten werden beim Überfahren mit der Maus in einem separaten Fenster angezeigt.

Kommentierte Literaturhinweise

Ausgehend vom Bonner Frühneuhochdeutschen Korpus empfehlen wir exemplarisch zwei Literaturen, die zur Bearbeitung des Korpus geeignet sind.

Zunächst lässt sich Walter Hoffmanns und Friedrich Wetters *Bibliographie frühneuhochdeutscher Quellen* vorstellen. Hierbei handelt es sich um ein kommentiertes Verzeichnis von Texten des 14.-17. Jahrhunderts, welches sich ganz konkret auf das Bonner Korpus bezieht und weitere Hinweise zu Sekundärliteraturen liefert.

Des Weiteren ist Heinz Peter Prells *Die Ableitung von Verben aus Substantiven in biblischen und nichtbiblischen Texten des Frühneuhochdeutschen* empfehlenswert, da die Veröffentlichung von 261 Seiten die vergleichende Untersuchung von Texten des 14. bis 17. Jahrhunderts mit beispielsweise graphischen Darstellungen behandelt. Fachtermini wie *Denominativ, Ableitung, Desubstantiv, Denominativum, Denominales* und *Verbalssystem* und weitere finden dort sehr gut erläutert ihre Aufmerksamkeit.