

HiTS: ein Tagset für historische Sprachstufen des Deutschen

1 Einleitung¹

Im Kontext der Projekte zur Erstellung historischer Sprachkorpora „Referenzkorpus Altdeutsch“ und „Referenzkorpus Mittelhochdeutsch“ entstand ein Tagset für die Wortartanalyse, *HiTS* („Historisches Tagset“). Im Projekt „Referenzkorpus Frühneuhochdeutsch“ wird eine vereinfachte Version davon angewendet, da v.a. die späteren Zeiträume schon nahe am neuhochdeutschen (nhd.) Stand sind.²

HiTS orientiert sich am „Stuttgart-Tübingen Tagset“ (STTS, Schiller et al., 1999), dem Standardtagset für nhd. Korpora, und übernimmt — neben einer ganzen Reihe von Tags — auch das hierarchische Design der Tagnamen. Ursprünglich sollte das Tagset komplett auf STTS aufbauen und dieses lediglich um einige neue Tags erweitern. Es stellte sich jedoch heraus, dass neben einigen notwendigen feineren Unterscheidungen (z.B. bei den Pronominaladverbien) auch die Tagnamen des STTS nicht immer geeignet schienen. Z.B. sind in HiTS der definite und indefinite Artikel eine Unterklasse der Determinativa — die Sonderstellung des Artikels, die im STTS durch ein eigenes Tag ‘ART’ betont wird, ist eine neuere Entwicklung.

Das Tagset dient zur Annotation diachroner Daten. Einige wenige Tags finden dabei nur in den alt(hoch)deutschen Daten Anwendung. Um diachrone Untersuchungen bis in die jetzige Zeit zu ermöglichen, werden im vorliegenden Artikel den HiTS-Tags die jeweils entsprechenden STTS-Tags gegenübergestellt. Allerdings ist nicht in jedem Fall eine eindeutige Abbildung möglich. So wird z.B. in HiTS zwischen attributivisch und substantivisch verwendeten Zahlen unterschieden, nicht aber im STTS. Umgekehrt unterscheidet HiTS nicht zwischen elliptischen („kopflosten“) Nominalphrasen und substantivierten Adjektiven (s. Abschnitt 5), was das STTS aber tut.

¹Die vorliegende Arbeit wurde finanziell unterstützt durch die Deutsche Forschungsgemeinschaft (DFG), DO 544/5-1/2 (Projektgruppe Altdeutsch), und KL 472/3-3/5, WE 1318/14-1/2, WE 1318/16-1, DI 1558/1-1/2 (Projektgruppe Mittelhochdeutsch). Außerdem möchten wir den anonymen Reviewern für ihre hilfreichen Kommentare danken.

²Das Referenzkorpus Altdeutsch (<http://www.deutschdiachrondigital.de>) deckt den Zeitraum 750–1050 ab, das Referenzkorpus Mittelhochdeutsch (<http://referenzkorpus-mhd.uni-bonn.de>, <http://www.rub.de/wegera/rem>) den Zeitraum 1050–1350 und das Referenzkorpus Frühneuhochdeutsch (<http://www.rub.de/wegera/ref>) den Zeitraum 1350–1650. Das Referenzkorpus Altdeutsch erfasst dabei die gesamte althochdeutsche und altniederdeutsche Textüberlieferung mit 650 000 Belegen. Die Referenzkorpora Mittel- und Frühneuhochdeutsch enthalten jeweils eine strukturierte Auswahl der Textüberlieferung mit etwa 2,1 Mio (mittelhochdeutsch) bzw. 4,4 Mio (frühneuhochdeutsch) annotierten Wortformen. Alle Korpora sind bzw. werden annotiert mit Metainformation sowie Information zum Lemma, zur Morphologie und zur Wortart.

Da die muttersprachliche Intuition für die älteren Sprachstufen grundsätzlich fehlt, können die Annotatoren z.B. keine linguistischen Tests durchführen. An die Stelle von Tests tritt daher in der historischen Sprachwissenschaft die Synopse und statistische Analyse möglichst vieler Belege entsprechender oder alternativer Konstruktionen. Bei den verbleibenden nicht auflösbaren ambigen Strukturen muss festgelegt werden, welche der möglichen Analysen annotiert werden soll. Prinzipiell wird dabei die historisch ältere bevorzugt. Beispielsweise wird bei Konstruktionen, die (noch) als pränominaler Genitiv oder (schon) als Komposituserstglied analysiert werden könnten, die Genitivsart annotiert (s. Abschnitt 2).

Den Wortartannotationen liegen als Basiseinheit Wörter zu Grunde. Der Bestimmung der Wortgrenzen („Tokenisierung“) ist ein eigener Abschnitt gewidmet (Abschnitt 2). Im STTS wird dieser Punkt nicht weiter thematisiert (nur der Sonderfall Mehrwortlexeme wird kurz angesprochen (Schiller et al., 1999, S. 9)). In modernen Korpora werden üblicherweise Wortgrenzen anhand von Leerzeichen bestimmt; zusätzlich zählen Satzzeichen als eigene Wörter. In historischen Daten finden sich jedoch zahlreiche Beispiele, bei denen Leerzeichen als Kriterium nicht genügen, um (nach heutigem Verständnis sinnvolle) Worteinheiten zu bestimmen. Die unterschiedlichen Schreibungen können wichtige Hinweise auf den Entwicklungsstand der Sprache liefern, z.B.: werden Komposita (noch) in zwei Wörtern oder (schon) in einem Wort geschrieben? Daher enthält HiTS auch ein Tagset für das Tagging von historischen Wortgrenzenauszeichnungen.

Eine weitere Besonderheit von HiTS ist die konsequente Unterscheidung von Lemma- vs. Beleg-spezifischer Annotation (Abschnitt 3). Beispielsweise kann ein Adjektiv (Lemma) in der Funktion eines Adverbs verwendet werden (Beleg); beide Wortarten werden in den Referenzkorpora annotiert. Diese Doppeltannotation eignet sich insbesondere für die Untersuchung von Sprachwandelprozessen, die mit einem Wortartwechsel einhergehen. Eine ähnliche Unterscheidung findet sich im STTS unter den „lexikalischen Kategorien“: z.B. wird hier angegeben, dass das Nomen *Alter* lexikalisch gesehen ein Adjektiv ist; in STTS-Darstellung: *der Alte*/NN<ADJ (Schiller et al., 1999, S. 13).

Für die Beispielsätze wurde für die althochdeutschen (ahd.) und altsächsischen (as.) Beispiele eine editionsnahe, für die mittelhochdeutschen (mhd.) Beispiele eine handschriftennahe Transkription gewählt.³ Vereinzelt sind auch konstruierte Beispiele darunter; diese sind normalisiert wiedergegeben und ohne Quellenangabe. Handschriftennah ist z.B. die Form *vf* ‘auf’, die normalisiert als *ūf* wiedergegeben wird. Die Quellenangaben

³Die ahd. Beispiele werden nach den jeweiligen Editionen der Texte zitiert. Das Referenzkorpus Altdeutsch hat hierbei, wenn immer möglich, die handschriftengetreueste Ausgabe verwendet. Aus diesem Grund differieren die Schreibweisen in den ahd. Beispielen zum Teil sehr stark. Eine Anpassung der einzelnen Schreibungen an eine gewissermaßen künstlich konstruierte „normal-althochdeutsche“ Schreibung ist schon aus Gründen der Authentizität und der phonologischen dialektalen Differenzen, deren Abbildung wünschenswert ist, nicht sinnvoll.

Die mhd. Beispiele folgen in der Regel der Handschrift, insbesondere dort, wo sie dem MiGraKo (s.u.) entnommen sind; korpuserterne Beispiele werden dagegen meist nach normalisierenden Editionen zitiert.

Das MiGraKo ist das Korpus der entstehenden neuen Mhd. Grammatik; es entspricht in seiner Textkonstitution und Struktur dem „Bochumer Mittelhochdeutsch-Korpus“ und ist in Bonn komplett tokenisiert, lemmatisiert und morphologisch annotiert worden. Es wird einen separaten Teil des Referenzkorpus Mittelhochdeutsch bilden.

referieren auf Siglen, die im Anhang aufgelöst werden. Wir haben uns entschieden, weitgehend auf die Angabe von Sekundärliteratur zu verzichten, und verweisen stattdessen auf die Standardlexika und -wörterbücher.

Der Artikel ist wie folgt aufgebaut: Abschnitt 2 thematisiert die Wortgrenzenbestimmung, Abschnitt 3 die Unterscheidung Lemma- vs. Belegannotation. Die beiden nachfolgenden Abschnitte 4 und 5 beschreiben einzelne Wortarten im Detail: Determinative und Adjektive; bei diesen beiden Wortarten ergeben sich die meisten Unterschiede zwischen HiTS und STTS. Den Abschnitten (oder auch Unterabschnitten) geht jeweils eine Überblickstabelle voraus, die die Tags der vorgestellten Wortart auflistet. In Abschnitt 6 wird das Tagging lateinischer Passagen thematisiert. Im Anhang finden sich Überblickstabellen aller Tags in HiTS.

2 Token (Worteinheiten)

Beispiel		HiTS
original	modernisiert	
<i>indaz</i>	<i>in daz</i>	MS (<i>Multiverbierung mit Spatium</i>)
<i>Schutz Gott</i>	<i>Schutzgott</i>	US (<i>Univerbierung mit Spatium</i>)
<i>Liebes=Ohnmachten</i>	<i>Liebesohnmachten</i>	UH (<i>Univerbierung mit Hyphen</i>)
<i>LandGraff</i>	<i>Landgraff</i>	UB (<i>Univerbierung mit Binnenmajuskel</i>)

Token, d.h. die Einheiten, die der Wortartannotation zu Grunde liegen, werden in modernen Korpora weitgehend aufgrund von Leerzeichen (Spatien) bestimmt. Zusätzlich werden Satzzeichen als eigene Token aufgefasst, und häufig werden klitisierte Formen ebenfalls als eigenständige Token analysiert, so z.B. im TIGER-Korpus: *gibt 's*.⁴

In älteren Handschriften und Drucken gibt es einige wichtige Unterschiede zu modernen bzw. gegenwärtigen Texten, die bei der Tokenisierung zu beachten sind und die diese beeinflussen. Die Zusammen- und Getrennschreibung sowie die Worttrennung am Zeilenende weichen vielfach von dem ab, was aus heutiger Sicht vornehmlich aus syntaktischen Gründen als Worteinheit und damit als Grundlage für die Tokenisierung zu betrachten ist. Interpunktionen bestehen nur in Ansätzen und unterscheiden sich hinsichtlich ihrer Funktion von moderner Interpunktion.

Die Wortgestalt unterliegt diachron verschiedenen Wandelprozessen. In HiTS gilt daher der Grundsatz, dass der handschriftliche Befund erkennbar bleiben soll. Daher unterscheiden wir zwei Ebenen der Textrepräsentation: eine diplomatische, d.h. handschriftennahe Ebene und eine Ebene mit moderner Tokenisierung. Im Folgenden beziehen wir uns auf die beiden Ebenen mit den Begriffen „Text-Ebene“ und „Token(isierungs)-Ebene“ und stellen die beiden Ebenen durch einen senkrechten Strich getrennt dar,

⁴TIGER-Korpus: <http://www.ims.uni-stuttgart.de/forschung/ressourcen/korpora/tiger.html>

gefolgt von einem „Token(isierungs)-Tag“, das die Art der Differenz zwischen den beiden Ebenen benennt:

- (1) a. Schema: Text-Ebene | Token-Ebene | Token-Tag
 b. Beispiel: *Schutz Gott* | *Schutzgott* | US

Wichtigstes Indiz für Getrenntschreibung ist die relative Größe eines Spatiums: Ist ein Spatium zwischen zwei Buchstabenfolgen in Relation deutlich größer als der „normale“ Abstand zwischen zwei Buchstaben, kann Getrenntschreibung angenommen werden. Hier kann es auch innerhalb derselben Handschrift oder auch desselben Drucks erhebliche Schwankungen geben, so dass sich oft nicht sicher entscheiden lässt, welches Spatium (schon) als Getrenntschreibung zu werten ist. Weitere Indizien können sein: Majuskeln am Wortanfang, doch können sie (bes. *J*, *N*, *R*) auch im Wortinneren vorkommen; ab der späteren mhd. Zeit steht außerdem meist <s> statt <f> (Schaft-s) am Wortende.

Die getrennt geschriebenen Einheiten entsprechen zwar mehrheitlich „Wörtern“ im modernen Verständnis, doch wird häufig davon abgewichen: Einerseits können auch Kompositionsglieder oder gebundene Morphe wie Präfixe getrennt geschrieben werden, z.B. *golt vaz*, *ge lebet*, *un verzaget* (‘Goldgefäß, gelebt, unverzagt’). Andererseits werden oft Wörter zusammengeschrieben, vor allem bei Pro- oder Enklise, z.B. *enwoltern* = *en wolte er in* (‘er wollte ihn nicht’); häufig ist Zusammenschreibung von Präpositionen mit dem folgenden Wort, z.B. in Nib *inir* = *in ir* ‘in ihr’, *indaz* = *in daz* ‘in das’, *inscharpfen* = *in scharpfen* ‘in scharfen’, *zeiungest* = *ze iungest* ‘jüngst/zuletzt’, *zeminne* = *ze minne* ‘zu(r) Minne’, *zeritter* = *ze ritter* ‘zu(m) Ritter’.

2.1 Univerbierung getrennt geschriebener Worteinheiten

Zusammen- und Getrenntschreibung von Worteinheiten, die heute univerbiert sind, schwanken stark bis in die Neuzeit hinein. Gelegentlich wird hier die Schreiberintention sichtbar. Dies betrifft insbesondere Komposita, bei denen die Entwicklung von einer Kontaktstellung (etwa in einer Genitivkonstruktion) über eine lose Verbindung (graphisch als Bindestrich oder Doppelbindestrich umgesetzt) oder durch Zusammenrückung schließlich zur Zusammenschreibung führt. Dabei können sich Binnenmajuskeln ergeben.

Auf der Token-Ebene werden die zusammengehörigen Wortteile gemäß heutiger Konventionen zusammengeschrieben. Die Tokenisierungstags (s. Tabelle S. 87) kennzeichnen die Zusammenfügungen mit einem ‘U’ (für „Univerbierung“) sowie mit weiteren Tags ‘S’ (durch „Spatium“ getrennte Wortbestandteile), ‘H’ (durch „Hyphen“ getrennt) und ‘B’ (durch „Binnenmajuskel“ markiert).

- (2) a. Kontaktstellung (getrennte Schreibung): *Schutz Gott* | *Schutzgott* | US
 b. (Doppel-)Bindestrich: *Liebes=Ohnmachten* | *Liebesohnmachten* | UH
 c. Binnenmajuskel durch Zusammenrückung: *LandGraff* | *Landgraff* | UB

Zusammengeschriebene Einheiten, die bereits als Komposita angesehen werden, werden als solche übernommen. Bei Getrennschreibung ist jedoch häufig kaum zu entscheiden, ob es sich um eine Genitivfügung oder ein Kompositum handelt. Lediglich bei eindeutiger Disambiguierung durch Substantivbegleiter kann zwischen Komposita und Genitivfügungen unterschieden werden, s. Bsp. (3a). Lässt der Substantivbegleiter nicht eindeutig erkennen, ob es sich um eine Genitivfügung oder um ein Kompositum handelt, so wird Getrennschreibung beibehalten, s. (3b):⁵

- (3) a. *die*[Fem] *herzen*[Neut] *königinne*[Fem] | *herzenköniginne* | US
 ‘die Herzenskönigin’ (Kompositum)
- b. *der*[Fem/Masc] *sunnen*[Fem] *schîn*[Masc] | |
 ‘der Schein der Sonne/der Sonnenschein’ (ambig; in HiTS analysiert als Genitivfügung)

2.2 Multiverbierung komplexer Worteinheiten

Vielfach finden sich Einheiten aufgrund eines individuellen Schreibstils („fehlerhaft“) oder bedingt durch Klitisierung bis hin zur Krasis (Verschmelzung, s.u.), die eine Trennung in zwei oder mehrere Einheiten nahelegen. Solche Trennungen werden durch das Token-Tag ‘MS’ (für „Multiverbierung mit Spatium“) markiert. (4b) zeigt ein Beispiel, das sowohl Uni- als auch Multiverbierung involviert. Hierfür werden zwei Token-Tags in der Reihenfolge ihres Auftretens verwendet.

- (4) a. *indaz* | *in daz* | MS ‘in das’
enwolde | *en wolde* | MS ‘nicht wollte’
lebeter | *lebet er* | MS ‘lebt er’
- b. *be durfeter* | *bedurfet er* | US-MS ‘bedürft ihr’ (mhd., GRud *βb* 12)

Werden direkt aufeinanderfolgende Wörter nicht nur ohne Spatium geschrieben, sondern miteinander verschmolzen (Krasis), so gelten für die Trennung gesonderte Regeln. Beispielhaft seien hier zwei Fälle angeführt: zum einen die Verschmelzung der Verbendung der 2.Pers.Sg mit nachfolgendem Personalpronomen, bei der aufgrund der unterschiedlichen Verbstämme einmal nach und einmal vor dem Dental getrennt wird (5); zum andern die Verschmelzung von Präposition und nachfolgendem Artikel (6), bei der ausschlaggebend ist, ob der Vokal im Bereich der Krasis eindeutig (auch) zum zweiten Wort gehört, dem er dann zugeschlagen wird (6a).

- (5) a. *wildu* | *wil du* | MS ‘willst du’
 b. *mahtu* | *maht u* | MS ‘(ver)magst du’
- (6) a. *zen* | *z en* (< *ze den*) | MS ‘zu den’
 b. *zun* | *zu n* (< *zuo den*) | MS ‘zu den’

⁵In diesen Fällen bleibt die Token-Ebene und das entsprechende Tokentag-Feld leer, da die modernen Wortgrenzen mit dem Original übereinstimmen.

2.3 Worttrennung am Zeilenende

Beispiel		HiTS
original	modernisiert	
<i>in-↔ dem lande</i>	<i>in dem lande</i>	MLH (<i>Multiverbierung am Zeilenende mit Hyphen</i>)
<i>ou=↔ gen</i>	<i>ougen</i>	ULH (<i>Univerbierung am Zeilenende mit Hyphen</i>)
<i>ou↔ gen</i>	<i>ougen</i>	ULS (<i>Univerbierung am Zeilenende mit Spatium</i>)

Die Worttrennung am Zeilenende kann in älteren Handschriften an jeder Stelle eines Wortes erfolgen (*ou-gen*, *Johan-nes*, *na-ch*), wenn sie auch überwiegend schon den modernen Trennregeln entspricht. Sie wird in den Handschriften und frühen Drucken entweder durch verschiedene Trennzeichen wie „-“ oder „=“ markiert oder bleibt unmarkiert; auch Doppelmarkierung am Ende einer und am Anfang der nächsten Zeile kommt vor.

In allen Fällen wird auf der Token-Ebene ein einziges Token angesetzt. Als Token-Tags dienen ‘ULH’ (für „Univerbierung am Zeilenende, mit Hyphen“) bzw. ‘ULS’ (für „Univerbierung am Zeilenende, mit Spatium (d.h. ohne Hyphen)“; der Zeilenumbruch wird im Folgenden durch „↔“ repräsentiert.)

- (7) a. *ou=↔ gen* | *ougen* | ULH
 b. *ou↔ gen* | *ougen* | ULS

Bei „falsch“ gesetzten Trennzeichen in der Handschrift, mit denen über die Zeilengrenze hinweg Wörter als zusammengehörig gekennzeichnet werden, die (modern) nicht als ein Wort aufgefasst werden, werden auf der Token-Ebene mehrere Token angesetzt. Zur Markierung dient ‘MLH’ (für „Multiverbierung am Zeilenende mit Hyphen“; s. auch nächsten Abschnitt).

- (8) *in-↔ dem lande* | *in dem lande* | MLH

2.4 Partikelverben (trennbare Präfixverben)

Im STTS werden separat stehende Partikeln von Partikelverben mit ‘PTKVZ’ („Partikel, Verbzusatz“) annotiert. Mit welchem Verb im Satz die Partikel zusammen eine Einheit bildet, wird nicht explizit markiert. Im Gegensatz dazu werden in HiTS beide Wortbestandteile als zusammengehörig ausgezeichnet.

Für historische Texte wie etwa mhd. werden die Grenzen für Partikelverben in HiTS enger gezogen als im STTS für das Nhd.: Nur Präpositionen und Adverbien können Ausgangspunkt für Partikelverben sein, nicht jedoch Adjektive oder Substantive. Dabei sind als Partikeln alle Adverbien möglich, zu denen es eine gleichlautende

Präposition gibt: insbesondere *abe*, *ane*, *bī*, *nāch*, *ūf*, *umbe*, *ūz*, *vor(e)*, *vür(e)*, *zuo* ('ab, an, bei, nach, auf, um, aus, vor, für, zu'), außerdem *in* 'ein' (zur Präp. *in*), *nider* 'nieder' und *wider* 'zurück'. Wie im Nhd. können (Pronominal-)Adverbien dem Verb in Kontaktstellung unmittelbar vorangehen oder in Fernstellung dem Verb nachfolgen (s. auch Abschnitt 3.2).

Um die Partikelverben von festen Präfixbildungen zu unterscheiden und um der möglichen Fernstellung von Basisverb und Partikel Rechnung zu tragen, werden beide Bestandteile — auch bei Kontaktstellung oder Zusammenschreibung — als zwei Token analysiert (im Gegensatz zu Präfixverben). Für die Markierung der Zusammengehörigkeit von Partikel und Verb werden Indizes verwendet (hier als Subskripte auf der Token-Ebene dargestellt).

- (9) a. *als er den brief **ane sach** | ... **ane**₁ **sach**₁*
 'als er den Brief ansah' (mhd., GrR βb 29)
 b. *er **sach** den brief **ane** | **sach**₁ ... **ane**₁ |*
 'er sah den Brief an'

Das folgende Beispiel zeigt das Verb *zurückziehen* einmal als Partikel- und einmal als Präfixverb.

- (10) a. *sī **zōch**₁ in **wider**₁*
 'sie zog ihn zurück' (mhd., Iw 1484), Partikelverb
 b. *hieten si dic reht erchant, si hieten **biderzogen** doh an dier ier hant*
 'hätten sie dich recht erkannt, sie hätten ihre Hand auch von dir zurückgezogen' (mhd., AlexiusA 991f), Präfixverb

Da nicht wie im Nhd. in der Regel die Zusammen- bzw. Getrenntschreibung als Unterscheidungsmerkmal dienen kann, ergeben sich Abgrenzungsprobleme zwischen Verbpartikeln einerseits und Adverbien (etwa bei *nider*) oder Pronominaladverbien andererseits. Der Ansatz eines Pronominaladverbs und nicht eines Partikelverbs lässt sich oft darauf stützen, dass die Lesart mit Pronominaladverb kontextuell sinnvoll, die Lesart mit Partikelverb dagegen sinnwidrig ist:

- (11) a. *ein teppet ... **da** sie **vf** solde sizzē*
 'ein Teppich, darauf sie sitzen sollte' (mhd., GRud αb 16), mit dem Pronominaladverb *dār ūf*
 und nicht: '*ein Teppich, wo sie aufsitzen sollte', mit dem Partikelverb *ūfsitzen* (auch mhd. meist vom Besteigen eines Pferdes)
 b. *wilt vnde zam man da **vure**₁ **truch**₁*
 '(Fleisch von) wilden und domestizierten/zahmen (Tieren) trug man da auf' (mhd., GRud A 19), mit dem Partikelverb *vüre tragen*
 und nicht: '*dafür trug man Fleisch von wilden und zahmen Tieren'

2.5 Interpunktionszeichen

Die Interpunktion historischer Handschriften und Drucke folgt in der Regel (noch) nicht modernen satzlogischen Gesichtspunkten, sondern ist weitgehend — wenn auch keinesfalls ausschließlich — noch rhetorisch (nach dem sog. rhythmisch-intonatorischen Prinzip) organisiert.

Im Altdeutschen herrscht der auf antike Vorbilder zurückgehende Punkt vor; andere Zeichen sind dagegen selten. Die Punkte können relativ zur Zeile verschieden positioniert sein: *comma* (*subdistinctio* = tiefgestellter Punkt), *colon* (*media distinctio* = Punkt auf der mittleren Höhe der Zeile) und *periodus* (*distinctio* = hochgestellter Punkt). Dieses System wird im Mittelalter durch Striche und Linien, verbunden mit Punkten, erweitert und dient bis in die Frühneuzeit hinein zur Markierung von (Vor)Lesepausen bzw. Sprechpausen bei Sinneinheiten.

Im Spätmittelalter und in der Frühneuzeit setzt sich ein detaillierteres System von Interpunktionszeichen durch: neben dem Punkt und der Virgel („/“) findet sich zunehmend die Kleinform der Virgel, das Komma, das sich gegenüber der Langform zur Kennzeichnung von Teilsätzen zum Nhd. hin durchsetzt, zunehmend auch in Konkurrenz zum Semikolon (Strichpunkt).

Der Doppelpunkt mit ursprünglich (so bereits im 9. Jh.) allgemeiner (teil-)satztrennender Funktion erscheint häufiger erst im 16. Jh. Das Fragezeichen ist bereits im Ahd. bekannt, findet sich aber ebenfalls erst seit dem 16. Jh. häufiger. Das Ausrufezeichen ist in seiner heutigen Form seit dem 16. Jh. in Grammatiken erwähnt und im 17. Jh. zur Kennzeichnung von Ausrufen, Wünschen, Verwunderung und des Nachdrucks als „Ausrufungszeichen“ verwendet. Ebenfalls zunehmend finden sich die Klammer „()“ zur Markierung von Parenthesen (Einschlusszeichen) und — seit Schottelius (1641) — der aus Poetiken entlehnte Apostroph (Hinterstrich) (vgl. dazu bes. Höchli, 1981).

Für Satzzeichen (in der Originalhandschrift bzw. -druck) wie . : , / ? ! verwenden wir das Tag ‘\$ _’, für sonstige satzinterne Zeichen wie ‘(’ das Tag ‘\$(‘.

3 Lemma- vs. belegspezifische Annotation

In HiTS wird die Wortart einer jeden Wortform zweifach annotiert, und zwar zum einen mit Blick auf das Lemma und zum anderen mit Blick auf den konkreten Beleg, also der Verwendung einer Wortform in einem spezifischen Kontext. Natürlich sind Lemma- und Belegwortart in sehr vielen Fällen identisch. Sie können aber auch divergieren, z.B. wenn ein appellatives Nomen (NA)⁶ als Eigenname (NE) verwendet wird: so etwa das bair. Appellativum *Hollerbirl* (‘Holunderbeere’) als Name eines Bioladens. Die Annotation von *Hollerbirl* nur als NA nach seiner Lemmawortart wäre in diesem Verwendungskontext definitiv irreführend. Und auch bei der Annotation nur nach der Belegwortart geht wertvolle Information verloren, nämlich die, dass es sich hier um keine Eigenschaft des Lemmas handelt.

⁶Im STTS: NN, „normales Nomen“.

Abgesehen davon ist die auf den ersten Blick sehr aufwändige doppelte Wortartannotation speziell für historische Korpora in besonderer Weise attraktiv, weil sie einen unmittelbaren Zugriff auf die Gebrauchsdynamik von Wörtern bietet und auf Sprachwandelprozesse, die mit einer Veränderung der Wortart einhergehen. Dabei werden Verwendungsbereiche und Zeiträume recherchierbar, in denen sich diese Veränderung vollzieht. Dazu im Folgenden zwei einschlägige Beispiele.

Die beiden POS-Ebenen werden dabei wie in (12b,c) gezeigt dargestellt. Steht nur eine Wortartkategorie hinter der Wortform, handelt es sich immer um die Belegwortart, vgl. (12a).

- (12) a. Schema: Wortform / Beleg-POS
 b. Schema: Wortform / Lemma-POS > Beleg-POS⁷
 c. Beispiel: *Hollerbierl* / NA > NE

3.1 Entwicklung der Negationspartikel *nicht*

Die Negationspartikel nhd. *nicht* entwickelt sich aus der Akkusativform des negativen Indefinitpronomens bzw. Pronominals substantivs ahd. *niowiht*, das ursprünglich folgende vier Kasusformen aufweist: *niowiht* (Nom./Acc.), *niowihtes* (Gen.), *niowihti/e* (Dat.), *niowihtu* (Instrumental). Anders als das nhd. Indefinitpronomen *nichts* ist ahd. *niowiht* zunächst nicht selbst Träger der Satznegation, sondern die Kennzeichnung der Negation erfolgt in der Regel durch die Negationspartikel ahd. *ni*, die in frühen althochdeutschen Texten allem Anschein nach im negativen Satz obligatorisch ist (13). Entsprechend annotieren wir *ni* als Negationspartikel mit PTK > PTKNEG und *niowiht* als negatives Indefinitpronomen mit PI > PNEG.⁸

- (13) *inti sie ni/PTK > PTKNEG quedent imo niowuht/PI > PNEG:Sg.Acc*
 ‘und sie nicht sagen ihm Nichts’ = ‘sie sagen ihm nichts’ (ahd., T 104, 7)

Im späten Althochdeutschen wird die Partikel *ni* lautlich zu *ne* abgeschwächt. Dabei treten in Verbindung mit *ni* bzw. *ne* zunehmend adverbiale Verstärker auf, darunter der adverbial gebrauchte Akkusativ des negativen Indefinitpronomens *niowiht* > *niwihht* > *nihht*. Letzteres findet sich in spätmittelhochdeutschen und frühmittelhochdeutschen Texten dann auch ohne *ne* in der Funktion als alleiniger Träger der Negation:

- (14) *swer nihht/PI > PTKNEG gloubet, der ist iu verteilet*
 ‘wer nicht glaubt, der ist schon verurteilt’ (ahd., BGB3)

⁷Das Zeichen > verwenden wir in zweifacher Bedeutung: In Verbindung mit Wortformen wie im Beispiel unten *niowiht* > *niwihht* > *nihht* bedeutet es, dass sich die zweite Wortform aus der ersten diachron entwickelt hat. In Verbindung mit POS-Tags bezeichnet das erste POS-Tag die Kategorie des Lemma, das zweite die des Belegs. Wie im Text argumentiert, stellt die Beziehung zwischen der Lemma- und der Belegkategorie auch (oft) eine diachrone Entwicklung dar.

⁸PTK: Partikel, PTKNEG: Negationspartikel; PI: Pronomen, indefinit; PNEG: Pronomen, negativ.

Die Annotation trägt dem Rechnung, indem sie *nicht* in dieser Verwendung auf der Belegebene mit dem POS-Tag PTKNEG versieht. Die Wortartzuweisung auf Lemmaebene (PI) bleibt unverändert mit dem Effekt, dass Belege wie diese diachron aus verschiedenen Perspektiven recherchierbar sind: aus der Perspektive von PI mit Blick auf die Entwicklung der negativen Indefinitpronomina des Deutschen ebenso wie aus der Perspektive von PTKNEG, d.h. unter dem Aspekt, wie der Wechsel von ahd. *ni* zu nhd. *nicht* von stattem geht. Die Interpretation der Daten bleibt so auf ein Minimum beschränkt.

3.2 Partikelverben im Althochdeutschen

Das zweite Beispiel, das wir in diesem Zusammenhang diskutieren wollen, ist im Grundsatz vergleichbar, aber im Detail etwas anders gelagert. Auch hier handelt es sich um einen Sprachwandelprozess, der eine Veränderung des POS-Tagging (AVD⁹ > PTKVZ) involviert. D.h. es geht um die Herausbildung von Partikelverben im Deutschen, bei deren Entstehung selbständige Morpheme — nämlich Adverbien — als Verbzusätze in das verbale Lexem integriert werden. Anders als im Falle der Entstehung der Negationspartikel *nicht* geht es hier allerdings nicht um einen Wandelprozess, der am Ende der althochdeutschen Zeit einsetzt, sondern um Veränderungen, die schon in althochdeutscher Zeit in vollem Gange sind. Jedenfalls verzeichnen die Referenzwörterbücher des Althochdeutschen (insbesondere Splett, 1993) bereits eine Vielzahl von Partikelverben, gleichzeitig aber immer auch noch die formgleichen Adverbien, die weiterhin zum Lexembestand des Althochdeutschen gehören. Dazu kommt eine weitere Erschwernis, die Annotationsentscheidungen in diesem Bereich problematisch macht, nämlich eine relativ hohe Zahl von Belegen, die nicht eindeutig interpretierbar sind, bzw. deren Interpretation umfangreiche formale und inhaltliche Analysen notwendig macht.

Wir bleiben hier nichtsdestotrotz bei der gleichen Grundstrategie und weisen alle betroffenen Wortformen auf der Lemmaebene durchgehend als AVD aus. Bei der Bestimmung der Wortart auf Belegebene orientieren wir uns hier im Sinne einer für Nutzer jederzeit nachvollziehbaren und transparenten Strategie an dem Referenzwörterbuch von Splett. D.h. wenn dieser — wie im Falle des folgenden Beispiels für das Verb ahd. *duon* einer Partikelverbvariante *uz-duon* (Splett, 1993, S. 1032) verzeichnet, wird *uz* in Belegen wie dem folgenden auf der Belegebene als PTKVZ ausgezeichnet:

⁹In HiTS lautet der Tag für Adverbien AVD (nicht ADV). Daneben gibt es interrogative (AVW), generalisierende (AVG, s. Abschnitt 4) und negative Adverbien (AVNEG).

(15)	AHD	<i>Uz</i>	<i>fon</i>	<i>iro</i>	<i>samanungu</i>
		aus	von	sie:gen.pl.masc	Versammlung
	Lemma-POS	AVD	AP	PPER	NA
	Beleg-POS	PTKVZ	APPR	DPOSGEN	NA
	AHD	<i>duont</i>	<i>sie</i>	<i>iuuwih</i>	
		tun	sie	euch	
	Lemma-POS	VV	PPER	PPER	
	Beleg-POS	VVFIN	PPER	PPER	

‘Sie schließen euch aus ihrer Gemeinde aus’ (ahd., T 171,3)

Das führt dazu, dass die hier auch wieder aus beiden Perspektiven (AVD) und (PTKVZ) möglichen Recherchen zu Ergebnissen mit einem relativ hohen Überlappungsbereich führen und nicht ohne weitere Analysen interpretierbar sind. Dies ist im Sinne einer Minimierung nicht reversibler und intransparenter interpretatorischer Vorentscheidungen durch die Annotatoren aber nicht zu vermeiden. In jedem Fall erhält der Recherchierende so eine zuverlässige und kalkulierbare Ausgangsbasis für seine Untersuchungen. Die interpretatorische Endentscheidung bleibt ihm hier aus guten Gründen im Detail selbst überlassen.

3.3 Synchrone Schwankungen

Neben der Darstellung von Sprachwandelprozessen in einem historischen Korpus ermöglicht die Trennung der POS-Annotation synchronen Erscheinungen gerecht zu werden, bei denen ein Lemma unter bestimmten Umständen die Funktion einer anderen Wortart als der ursprünglichen einnimmt. Die uns bekannten STTS-basierten Korpora bieten für derartige Fälle keine adäquate Recherchemöglichkeit.¹⁰ Entweder würde bei einer lemmabezogenen Annotation ausschließlich die grundlegende Wortart eines Lemmas angegeben, und davon abweichende Verwendungsweisen müssten so ignoriert werden. Oder die POS-Kennzeichnung erfolgte grundsätzlich kontextbezogen, was bedeuten würde, dass u.U. keine klaren kategorialen Zuweisungen mehr möglich sind und zudem quantitative Recherchen verfälscht werden würden. Hingegen kann die Unterscheidung zwischen einer lemma- und einer belegspezifischen Annotation kontextabhängigen Schwankungen bei der Wortart gerecht werden.

Im Nhd. treten besonders in der mündlichen Sprache Konstruktionen auf, in denen ein Indefinitum durch ein Fragepronomen wiedergegeben wird, vgl. (16).

(16) *Ich habe wen gesehen.*

Hier entspricht das w-Pronomen semantisch einem Indefinitum. Diese Möglichkeit, indefinite Elemente durch w-Pronomina darzustellen, kennt bereits das Altdeutsche. In

¹⁰Wie in Abschnitt 1 erwähnt, sieht das STTS mit den „lexikalischen Kategorien“ eine ähnliche Unterscheidung vor, die unseres Wissens aber aktuell von keinem Korpusprojekt genutzt wird.

HiTS-annotierten Korpora erhalten indefinit verwendete w-Pronomina das lemmabezogene Tag PW, als belegbezogenes Tag wird PI vergeben.¹¹

- (17) *ef man huuemu/PW* > **PI:Masc.Sg.Dat** *saca sokea* ‘falls man mit jemandem Streit suche’ (Hel 1521)

4 Determinativa, Pronomen und Pronominaladverbien

4.1 Artikel

Beispiel	HiTS	STTS
<i>der liehte tac</i> ‘der helle Tag’	DD > DDA (ad.), DDART (mhd.) <i>Determinativ, definit, (artikelartig)</i>	ART <i>bestimmter Artikel</i>
<i>ein tier</i> ‘ein Tier’	DI > DIA (ad.), DIART (mhd.) <i>Determinativ, indefinit, (artikelartig)</i>	ART <i>unbestimmter Artikel</i>

Der definite Artikel hat sich seit germanischer Zeit aus dem Demonstrativpronomen, der indefinite Artikel erst einzelsprachlich aus dem Kardinalzahlwort *ein* entwickelt. Beide Prozesse sind im Altdeutschen noch im Gange, wobei der erste schon deutlich weiter gediehen ist als der zunächst noch in ersten Anfängen steckende zweite. Daher werden im altdeutschen Referenzkorpus *der* und *ein* noch als ‘Determinativ, definit’ (DDA) bzw. ‘Determinativ, indefinit’ (DIA) getaggt.

Im Mhd. wird dagegen regelmäßig DDART als Beleg-POS für den Definitartikel und DIART für den Indefinitartikel vergeben. Die Wortartabgrenzungen innerhalb der „Wortfamilie“ *ein* (mit Verwendungen als Artikel, Indefinitpronomen, unbestimmtes Zahladjektiv oder Kardinalzahl, vgl. Dudenredaktion (2009, § 446)) bereiten hier allerdings noch größere Probleme als im Nhd.

- (18) a. *that/DD* > **DDA** *helaga barn* ‘das heilige Kind’ (as., Hel 804)
b. *der/DD* > **DDART** *liehte tac* ‘der helle Tag’ (mhd., Iw 644)
- (19) a. *tho gisaz er múader bi einemo/DI* > **DIA** *brunnen* ‘da saß er müde bei einem Brunnen’ (ahd., O II,14,8-9)
b. *Ave ist ein/DI* > **DIART** *tier, daz heizit helphant* ‘weiter gibt es ein Tier, das heißt Elefant’ (mhd., Phys 138v,9 [13,11])

¹¹PW: Pronomen, interrogativ; PI: Pronomen, indefinit.

4.2 Reflexiv verwendete Formen des anaphorischen Pronomens

Beispiel	HiTS	STTS
<i>im(o)</i> (masc) ‘ihm’	PPER > PRF:3.Sg.Masc.Dat	<i>sich</i> PRF:3.Sg.Dat
<i>im(o)</i> (neut) ‘ihm’	PPER > PRF:3.Sg.Neut.Dat	<i>sich</i> PRF:3.Sg.Dat
<i>ir(u)</i> (fem) ‘ihr’	PPER > PRF:3.Sg.Fem.Dat	<i>sich</i> PRF:3.Sg.Dat
<i>in</i> (pl) ‘ihnen’	PPER > PRF:3.Pl.*.Dat	<i>sich</i> PRF:3.Pl.Dat
	<i>Pronemem der 3. Person in reflexivem Gebrauch</i>	<i>Reflexivpronomen</i>

Nur das Gotische und Altnordische verfügen noch über alle drei Formen des germ. Reflexivpronomens: got.: Gen. *seina*, Dat. *sis*, Acc. *sik*; an.: Gen. *sín*, Dat. *sér*, Acc. *sik*. Das Hochdeutsche hat dagegen lediglich die Akkusativform ahd. *sih*, mhd. *sich* in reflexiver Funktion bewahrt. Der Dativ **sir* < germ. **siz* ist untergegangen. Die Genitivform *sín* des Reflexivs hat sich zwar formal erhalten, ist aber schon ahd. auch für den Gen.Sg.Masc *is* des anaphorischen Pronomens eingetreten und ersetzt im Mhd. zunehmend auch *es* Gen.Sg.Neut: *der eselære wart sin/PPER:3.Sg.Masc.Gen gewar* ‘der Eseltreiber wurde seiner gewahr’ (Kchr 1716); *wi wol ich dir sin/PPER:3.Sg.Neut.Gen gan!* ‘wie sehr gönne ich es dir!’ (Kchr 3186).

Das bedeutet, dass im Ahd. und Mhd. im Gen., vor allem aber im Dativ ausdrucksseitig zwischen reflexiver und nicht reflexiver Verwendung nicht unterschieden werden kann: *ir(e)* ‘ihr’ Dat.Sg.Fem, *im(e)* ‘ihm’ Dat.Sg.Masc/Neut und *in* ‘ihnen’ Dat.Pl können daher, wenn sie numeruskongruent und im Sg. auch genuskongruent mit dem Subjekt sind, neben der nicht-reflexiven auch eine reflexive Bedeutung haben; reflexiv ist z.B. *im* ‘ihm/sich’ in *er enwolde in vor im lāzen niht komen in den strit* ‘er wollte ihn nicht vor **sich** in den Kampf kommen lassen’ (Nib 2274,3), doch wäre in anderem Kontext auch die nicht-reflexive Lesart möglich: ‘er wollte ihn nicht vor **ihm** [d.h. einem andern, dritten] in den Kampf lassen’. Nur sehr vereinzelt tritt in mhd. Zeit schon *sich* als Reflexivpronomen auch im Dativ ein und ersetzt die reflexiven *ir(e)*, *im(e)* und *in*; vergleichsweise häufig geschieht dies allerdings mit bereits ahd. Anfängen nach der Präposition *ze* ‘zu’.

Im HiTS erhalten ahd. *imo*, *iru*, *in* und mhd. *im*, *ir*, *in* in reflexiver Verwendung als allgemeines POS-Tag PPER, als POS-Tag des Belegs dagegen PRF im Unterschied zu PPER bei nicht-reflexivem Gebrauch.

- (20) a. *than beginnid he imu/PPER > PRF:3.Sg.Masc.Dat wuiti andreden*
 ‘dann beginnt er, **sich** vor Strafe zu fürchten’ (as., Hel 3495)
- b. *si erwarp ir/PPER > PRF:3.Sg.Fem.Dat lop vil grozen*
 ‘sie erwarb **sich** sehr großen Ruhm’ (mhd., Nib 1417,4)

4.3 Das reziproke Pronomen *einander*

Das Pronomen *einander* hat stets nur reziproke Bedeutung ('sich gegenseitig, wechselseitig'), während *sich* neben der reflexiven auch eine reziproke Lesart haben kann, z.B. *sie zogen sich die Stiefel aus*: 'sich selbst' bzw. 'sich gegenseitig'. Es erscheint dennoch angemessen, wenn im STTS *sich* und *einander* gleichermaßen als PRF:3.Pl. . . getaggt werden (Schiller et al., 1999, S. 35).

Im Alt- und Mittelhochdeutschen ist die Situation komplizierter, da das Pronomen *einander* wohl erst seit der späteren ahd. Zeit durch Zusammenrückung von *ein* und flektiertem *ander* entstanden ist. Vorläufer sind Fügungen wie *zimit úú . . .thaz ein ándremo fúazi wasge* 'es geziert sich für euch, dass einer dem anderen die Füße wasche' (O IV 11,50; s. AWB 1,476). Im Unterschied dazu ist mhd. (wie nhd.) *einander* zwar subjektbezüglich, kann jedoch nie zugleich auch das Subjekt bilden. Im Ahd. verbindet sich in aller Regel das endungslose *ein* mit dem flektierten *ander*. Wo *ein* nicht Subjekt ist wie im Otfried-Beispiel (O), ist schwer entscheidbar, seit wann mit Univerbierung zu *einander-* zu rechnen ist; jedoch ist dies sicher noch nicht in den seltenen spätahd.-mhd. Belegen der Fall, in denen auch *ein* flektiert ist, z.B. *dar beualen Constantinis man Einen/DIS:Dat.Pl anderen/DIS:Dat.Pl die kint vnde wiph* 'da vertrauten Konstantins Männer einander ihre Kinder und Frauen an' (Roth 2655).

Auch wenn in den Handschriften meist getrennt geschrieben, werden *ein ander*, *ein andere*, *ein anderen* (mit unflektiertem *ein*) ansonsten als unverbirtetes *einander* lemmatisiert und sie erhalten einheitlich PRF als POS-Tag. Anders als im Nhd. muss jedoch zwischen unflektiertem *einander*/PRF:3.Pl.* und flektiertem *einanderen*/PRF:3.Pl.Dat unterschieden werden.

- (21) a. *Súu*/PPER:3.Pl.Neut.Nom *tílegônt ouh einánderúu*¹²/PPER > PRF:
3.Pl.Neut.Acc 'sie vernichten auch einander' (ahd., N Cat.77,8)
 b. *dc erzentûm . . .dc die gûten ainanderen*¹³/PPER > PRF:**3.Pl.Dat**
gebent 'die Arznei . . . , die die Guten einander geben' (mhd., TrHL 109r,17)

4.4 Possessiv verwendete Formen des anaphorischen Pronomens

Im Altdeutschen (wie schon im Germanischen) gab es für die 3. Person kein Possessivpronomen im Sg. des Femininum und im Pl. aller Genera. Stattdessen wurden die Genitivformen des anaphorischen Pronomens *ira* Gen.Sg.Fem bzw. *iro* Gen.Pl in possessivischer Funktion verwendet. Daran änderte sich zunächst auch nichts, als beide Formen in spätahd. Zeit in der Form des Gen.Pl *iro* zusammenfielen (wie auch sonst in der pronominalen Deklination) und *iro* später zu frühmhd. *ire* > *ir* reduziert wurde.

Im Laufe des 12. Jh. muss dieses *ire* (> *ir*) von der Genitivform des Personalpronomens zu einem zunächst unflektierbaren Possessivpronomen umgedeutet worden sein; diese Reanalyse war die Voraussetzung für die Neubildung flektierter Formen wie *iren* (Dat.Pl),

¹² *ein ánderúu* | *einánderúu* | US

¹³ *ain anderen* | *ainanderen* | US

die vom 3. Viertel des 12. Jh. im Mittelfränkischen belegt sind: *iren swein sunnen* ‘ihren zwei Söhnen’, *irs lîves* ‘ihres Lebens’ (Kölner Schreinskarten der Laurenzpfarre, 1159–1172); noch älter sind einige flektierte Belege im altniederländischen Leidener Willeram (um 1100), z.B. *uz heran lando* ‘aus ihrem Land’.

Es sind also drei Entwicklungsphasen zu unterscheiden:

Phase 1: possessiv verwendeter Genitiv des anaphorischen Pronomens

Phase 2: >> indeklinables Possessivpronomen

Phase 3: >> a) flektierbares >> b) flektiertes Possessivpronomen

Wie sind diese unterschiedlichen Verhältnisse jeweils angemessen zu taggen? Für die altdeutsche 1. Phase empfiehlt sich eine Regelung, die einerseits die unveränderte Zugehörigkeit zu PPER, andererseits aber auch die possessive Funktion berücksichtigt, also PPER als allgemeinen POS-Tag, DPOSGEN als POS-Tag des Belegs. Die Grenzziehung zwischen Phase 1 und 2 lässt sich nicht anhand klarer formaler Kriterien vornehmen: Vor allem späthd.-frühmhd. Texte wie WNot oder Will können ebenso gut der 1. wie der 2. Phase zugewiesen werden. So ist es eine letztlich willkürliche Entscheidung, gemäß der üblichen Epochengrenze alle bis zur Mitte des 11. Jh. datierbaren (altdeutschen) Texte der 1. Phase und alle danach datierbaren (mhd.) Texte der 2. bzw. 3. Phase zuzuweisen. Phase 2 und 3 unterscheiden sich nur noch in den morphologischen Merkmalen, z.B.:

Phase 1: *in iro*/PPER > DPOSGEN:*.Gen.Pl.0 *lande*

Phase 2: *in ir(e)*/PPER > DPOSA:Neut.Dat.Sg.0 *lande*
(oder: *ir(e)*/PPER > DPOSA:*.*.0 *lande*)

Phase 3: *in irem*/DPOS > DPOSA:Neut.Dat.Sg.st *lande*

Der Unterschied zwischen den Phrasen 3a und 3b besteht darin, dass *ir* in 3b regelmäßig wie die anderen Possessivpronomen vom Typ *mîn* ‘mein’ flektiert wird, während in 3a-Texten flektierte und unflektierte Formen in weithin regellosem Wechsel vorkommen, so z.B. in den Predigtfragmenten PrF (Ende 12. Jh.) einerseits *ire gebetes* ‘ihres Gebetes’ (3,6), andererseits *ires gebetes* (4,19). Wo — wie in diesem Text — unflektiertes *ire* noch nicht zu *ir* verkürzt ist, lässt sich im Acc.Sg.Fem und Nom./Acc.Pl.Neut allenfalls einer Statistik des gesamten Formengebrauchs entnehmen, ob *ire* eher als flektiert oder als unflektiert anzusehen ist: *ire*/DPOSA:*.*.0 *afterkumelinge* ‘ihre Nachkommen’ (PrF 4,14) oder *ire*/DPOSA:Masc.Nom.Pl.st *afterkumelinge*.

4.5 Die generalisierenden Pronomen und Adverbien

Beispiel	HiTS	STTS
(s)waz ‘was immer’	PG > PG <i>Relativpronomen, generalisierend</i>	PWS <i>subst. Int.pron.</i>
(s)wanne ‘wann (immer)’	AVG > AVG <i>Relativadverb, generalisierend</i>	PWAV <i>adv. Int.- o. Rel.pron.</i>

Das STTS sieht vor (Schiller et al., 1999, S. 49, 51), dass *wer* in den beiden folgenden Beispielen gleichermaßen als substituierendes Interrogativpronomen (PWS) getaggt wird:

- (22) a. *wer/PWS* *das sagt, weiß nicht, was los ist* (nhd.)
 b. *er will wissen, wer/PWS* *mit welchem Zug kommt* (nhd.)

Aus sprachhistorischer Sicht und für die älteren Sprachstufen des Deutschen müssen beide Fälle jedoch kategorial unterschieden werden: Während beim indirekten Fragesatz in (22b) auch im Mhd./Ahd. das Interrogativpronomen *wer* vorausgeht, beruht *wer* im freien Relativsatz in (22a) auf dem „verallgemeinernden Relativpronomen“ ahd. *sō hwer sō* bzw. mhd. *swer*:

Im Ahd. (wie im übrigen Westgermanischen) konnten verallgemeinernde Ausdrücke durch die Verbindung von Indefinitpronomina (*hwer* ‘jemand’, *hwelih* ‘(irgend)jener’) oder Adverbien (*hwanne* ‘irgendwann’, *hwār* ‘(irgend)wo’, *hwio* ‘wie’ etc.) mit zweifachem *sō* gebildet werden, z.B. *sō hwer sō* ‘wenn jemand; wer auch immer’, *sō hwār sō* ‘wo auch immer’. Bereits in ahd. Zeit beginnt das zweite *sō* zu entfallen und im Mhd. ist es mit seltenen Ausnahmen verschwunden. Außerdem verschmilzt das zu *s* reduzierte erste *sō* im Großteil des Mhd. mit dem Folgewort, so dass sich die Formen *swer*, *swaz*, *swelich*, *swanne*, *swie* usw. ergeben. Nur im Mittelfränkischen bleibt *sō* erhalten, sofern es nicht ganz schwindet, z.B. *du bit mir, so wat du wolt* ‘tu mit mir, was immer du willst’ (RhMl 419). Im Laufe der mhd. Zeit beginnt das anlautende *s* der *sw*-Formen zu schwinden und dieser Schwund hat sich im Spätmhd. weitgehend durchgesetzt. Damit ist beim verallgemeinernden Pronomem *wer* < *swer* im Grunde schon der nhd. Stand erreicht, denn die freie Relativsätze einleitenden nhd. *wer*, *was* stehen nicht nur formal in der direkten Nachfolge von mhd. *swer*, *swaz*, sie haben auch die generalisierende Bedeutung bewahrt, die durch optional beigefügtes (*auch*) *immer* nur hervorgehoben, aber nicht erst hergestellt wird; und dasselbe gilt teils auch für Relativadverbien, vgl. z.B. nhd. (23d) und den entsprechenden mhd. Satz in (23b) unten.

Zur Kennzeichnung der generalisierenden Bedeutung dieser Ausdrücke dient im HiTS das Tag-Element ‘G’: PG für die generalisierenden Pronomina *sō hwer sō*, *sō (h)welih sō* (ahd.), *swer*, *swelich* (mhd.), AVG für die generalisierenden Adverbien *sō (h)wanne sō*, *sō (h)wā sō* (ahd.) (23a), *swanne*, *swār*, *swar(e)*, *swie* (mhd.) etc. (23b). Und dasselbe gilt konsequenterweise auch für die spätmhd. Formen mit *s*-Schwund *wer*, *welich*, *wanne*

usw. (23c). Ob und wann im Frühnhd. ggf. stattdessen dem STTS entsprechend ‘PW...’ als POS-Tag vergeben wird, bleibt noch zu klären.

- (23) a. *sie tuont, sōwazsō/PG*¹⁴ *sie wellen, inti sōwārsō/AVG, sōwannesō/AVG inti sōwiesō/AVG sie wellen* (ahd.)
 b. *sie tuont, swaz/PG sie wellen, unde swā/AVG, swanne/AVG unde swie/AVG sie wellen* (mhd.)
 c. *sie tuont, waz/PG sie wellen, unde wā/AVG, wanne/AVG unde wie/AVG sie wellen* (spätmhd.)
 d. *sie tun, was/PWS (immer) sie wollen, und wo/PWAV (immer), wann/PWAV (immer) und wie/PWAV (immer) sie es wollen* (nhd.)

Ein Problem stellen die Formen dar, die zwischen ahd. *sō (h)wer sō* und mhd. *swer* liegen: zum einen das im hochdeutschen Raum nahezu ausschließlich mittelfränkische *sowe*, z.B. *sowe*¹⁵ *on beruret, he is selich* ‘wer immer ihn berührt, er ist glücklich’ (Rhein. Marienlob 224); obgleich meist getrennt geschrieben, wird *sowe* in HiTS lediglich als Variante von *swer* aufgefasst und entsprechend lemmatisiert.

Zum anderen kann oberdeutsch wie mitteldeutsch bis zur ersten Hälfte des 13. Jh., danach nur noch vereinzelt das zweite *so* erhalten bleiben: *swer so*. In diesen Fällen taggen wir *so* als Partikel (PTK), z.B. *Swer/PG so/PTK wil sin der erste, der wirt der alleriungiste*¹⁶ ‘wer (immer) der Erste sein will, der wird der Allerletzte sein’ (WMEv 32,2). Wo *sō* dagegen die adverbiale Bedeutung ‘so, auf solche Weise’ hat, wird es als Adverb (AVD) getaggt, z.B.: *wer/PG so/AVD horit, dastit, sihit, ruchit ader smackeit, ...* ‘wer auf diese Weise hört, tastet, sieht, riecht oder schmeckt’ (Erlös 6926).

¹⁴ *sō waz sō | sōwazsō | US; sō wārā sō | sōwārā | US etc.*

¹⁵ *so we | sowe | US*

¹⁶ *aller iungiste | alleriungiste | US*

4.6 Pronominaladverbien

Beispiel	HiTS (zweiteilig)		STTS
	pronominal	präpositional	
<i>dā pāgant siu umpī</i> ‘darum streiten sie’ (ahd.); <i>dā was vil volkes inne</i> ‘darin waren viele Leute’ (mhd.)	AVD > PAVD <i>Pron.adverb</i>	+ AP > PAVAP	PAV <i>Pron.adverb</i>
<i>thār wir ana lāgun</i> ‘woran wir erlagen’ (ahd.); <i>der palas, dā ... inne was</i> ‘der Palast, in dem’ (mhd.)	AVD > PAVREL <i>Pron.adverb, relativisch</i>	+ AP > PAVAP	—
<i>wā ... umbe</i> ‘warum’ (mhd.)	AVW > PAVW <i>Pron.adverb, interrogativ</i>	+ AP > PAVAP	PWAV <i>adv. Int.- o. Rel.pron.</i>
<i>swā ... umbe</i> ‘warum (auch immer)’ (mhd.)	AVG > PAVG <i>Pron.adverb, generalisierend</i>	+ AP > PAVAP	—

Im STTS werden Pronominaladverbien den Verhältnissen in der nhd. Standardsprache gemäß als stets unverbirt aufgefasst und als solche mit dem POS-Tag PAV versehen. In den älteren Sprachstufen findet sich dagegen nicht selten Distanzstellung des zweiten (präpositionalen) Teils, und zwar als Folge von Topikalisierung nur des ersten (pronominalen) Teils im Hauptsatz, z.B. *dā was vil volkes inne = vil volkes was darinne*, und durch Nachstellung des zweiten Teils bei relativischer Verwendung des Pronominaladverbs, z.B. *der palas, dā vil volkes inne was*. Beim häufigsten Typ mit pronominalem *dā(r)* belaufen sich die Belege mit Distanzstellung im Korpus der mhd. Grammatik (MiGraKo)¹⁷ auf immerhin 10,6% (= 783 Belege).

Pronominaladverbien werden daher in HiTS (mhd.) grundsätzlich als „zweiwortig“ aufgefasst, auch wo sie in der Handschrift zusammengeschrieben sind: Ein handschriftliches *darinne* wird in *dar* und *inne* aufgespalten, die getrennt getaggt werden (s. Abschnitt 2). Entsprechend wird auch mit den relativischen, interrogativen und generalisierenden Pronominaladverbien verfahren:

- (24) a. *dā*/AVD > PAVD *umbe*/AP > PAVAP
 b. *dā*/AVD > PAVREL *umbe*/AP > PAVAP
 c. *wā*/AVW > PAVW *umbe*/AP > PAVAP
 d. *swā*/AVG > PAVG *umbe*/AP > PAVAP

¹⁷Vgl. Fußnote 3.

4.7 Interrogativadverbien

Beispiel	HiTS	STTS
<i>(h)wār</i> ‘wo’	AVW > AVW <i>Interrogativadverb</i>	PWAV <i>Inter.pron., adverbial</i>
<i>sō (h)wār sō</i> ‘wo auch immer’	AVG > AVG <i>Relativadverb, generalisierend</i>	PWAV <i>Rel.pron., adverbial</i>

Im STTS werden *wo*, *wann*, *wie* etc. nicht als Interrogativadverbien, sondern als adverbiale Interrogativ- bzw. Relativpronomen (PWAV) aufgefasst (Schiller et al., 1999, S. 53):

- (25) a. **Wo/PWAV** *wohnt er?* (nhd.)
- b. *er fragt, wo/PWAV er wohnt* (nhd.)
- c. *der Ort, wo/PWAV er wohnt* (nhd.)

Schon hinsichtlich des Nhd. kann man sich aber fragen, warum *wo*, *wann* und *warum* Pronomen, die korrespondierenden *da*, *dann* und *darum* aber Adverb sein sollen (Schiller et al., 1999, S. 56). Für das Mhd. jedenfalls wird man diese Frage wohl verneinen müssen, denn hier gibt es ein — auch formal — sehr ausgeprägtes System von zusammengehörigen phorisch-deiktischen, interrogativen und generalisierenden Wörtern, die üblicherweise (z.B. Splett (1993); Lexer (1878)) als Adverbien klassifiziert werden. Dem schließt sich HiTS an:

Adverb (AVD)			Interrogativadverb (AVW)		
ahd	mhd		ahd	mhd	
<i>dār</i>	<i>dā(r)</i>	‘da, dort’	<i>(h)wār</i>	<i>wā(r)</i>	‘wo’
<i>dara</i>	<i>dar(e)</i>	‘dorthin’	<i>(h)wara</i>	<i>war(e)</i>	‘wohin’
<i>danān</i>	<i>dannen</i>	‘von da weg’	<i>(h)wanān</i>	<i>wannen</i>	‘woher’
<i>dō</i>	<i>(dō)</i>	‘da(mals)’	<i>(h)wanne</i>	<i>wanne</i>	‘wann’
<i>sō</i>	<i>(sō)</i>	‘so’	<i>(h)wio</i>	<i>wie</i>	‘wie’

generalisierendes Relativadverb (AVG)		
ahd	mhd	
<i>sō (h)wār sō</i>	<i>swā(r)</i>	‘wo auch immer’
<i>sō (h)wara sō</i>	<i>swar(e)</i>	‘wohin auch immer’
<i>sō (h)wanān sō</i>	<i>swannen</i>	‘woher auch immer’
<i>sō (h)wanne sō</i>	<i>swanne</i>	‘wann auch immer’
<i>sō (h)wie sō</i>	<i>swie</i>	‘wie auch immer’

Noch weniger akzeptabel wäre es für die historischen Sprachstufen, wie im STTS zu den PWAV unterschiedslos auch die interrogativen Pronominaladverbien des Typs mhd.

wār umbe (nhd. *warum*) zu zählen. Da auch sie im Ahd. und Mhd. trennbar sind, müssen ihre beiden Teile entsprechend den deiktisch-determinativen Pronominaladverbien im Beleg das POS-Tag PAVW bzw. PAVAP erhalten (s. Abschnitt 4.6). Dasselbe gilt für die relativischen generalisierenden Pronominaladverbien des Typs *swār umbe*; als allgemeine POS-Tags fungieren also PAV bzw. AP, als POS-Tags des Belegs PAVG bzw. PAVAP.

5 Adjektive und Adjektivadverbien

5.1 Prädikative Adjektive und Adjektivadverbien

Beispiel	HiTS	STTS
<i>daz almuofen ist reht</i> 'das Almosen ist gerecht'	ADJ > ADJD <i>Adjektiv, prädikativ</i>	ADJA <i>präd. oder adv. Adjektiv</i>
<i>lebet rehto</i> 'lebt gerecht!'	ADJ > AVD <i>Adjektivadverb</i>	ADJD; ADV; PIS (s.u.)

Wie die Beispiele in der Tabelle¹⁸ zeigen, werden prädikativ verwendete Adjektive im HiTS so wie im STTS bestimmt,¹⁹ adverbial verwendete dagegen anders, nämlich in der Regel wie die prädikativen und nicht als Adjektivadverbien. Das ist für das Nhd. inzwischen üblich: Statt einer anderen Wortart wird eine weitere Verwendungsweise angenommen und neben die attributive und prädikative gestellt. Denn es ist kaum zu rechtfertigen, zwar zwischen der attributiven und prädikativen Form keinen Wortartwechsel anzunehmen, wohl aber zwischen der prädikativen und adverbialen, die fast immer gleich lauten (eine Ausnahme wird unten genannt). Früher jedoch wurde für die adverbiale Verwendung ein anderes Wort genommen als für die prädikative. Der Wortunterschied kam meist durch Derivation, ausnahmsweise durch Suppletion zustande. Das zweite Beispiel in der Tabelle zeigt die Derivation mit dem Adjektivadverbsuffix ahd. *-o* > mhd. *-e*. Ein Beispiel für Suppletion ist ahd. *wola* > mhd. *wol(e)* zum Adjektiv ahd., mhd. *quot: fuie ser gîlêbet hâbe wôla*/ADJ > AVD *alder ûbelo*/ADJ > AVD 'wie er gelebt habe, gut oder schlecht' (frühmhd., BaGB 139,21).²⁰ Ein stärkerer Formunterschied zwischen prädikativ und adverbial verwendeten Adjektiven bestand auch insofern, als erstere im Ahd. und Mhd. gelegentlich noch mit Flexionsendung auftraten, vgl. *uuanda er fuozer unde rehtîr ist* 'dass er liebenswert und gerecht ist'

¹⁸WNot 6va und 22va. Die Belege in den Tabellen stammen mit einer Ausnahme alle aus dem frühmhd. WNot und enthalten alle den Stamm *reht* als Beispiel, damit die Beispiele möglichst vergleichbar sind und alt- wie mittelhochdeutsch annähernd gleich berücksichtigen.

¹⁹Allerdings wäre *er hält geheim* nach dem HiTS nicht als PTKVZ („adjektivische abgetrennte Verbzusätze“, Schiller et al., 1999, S. 23) zu bestimmen, sondern als ADJD, was sich wohl aus einer anderen Tokenisierung ergibt.

²⁰Beide Mittel, Derivation und Suppletion, waren aus dem Germanischen ererbt. Denn sie finden sich nicht nur im Westgermanischen (zu dem das Deutsche gehört), sondern auch im Nord- und Ostgermanischen, und zwar zum großen Teil mit ursprungsgleichen Zeichen ausgedrückt.

(frühmhd., WNot 40va), hier vielleicht von lat. *dulcis et rectus dominus* beeinflusst, was kurz vorher zitiert wird.

Es läge nun nahe zu sagen, die Adjektivadverbien seien nach dem HiTS deswegen anders als nach dem STTS zu bestimmen, weil sie im Ahd. und Mhd. noch durch eine eigene Wortbildungsweise von Adjektiven unterschieden wurden, im Nhd. nicht mehr. Aber wo man solch scharfe Grenzen zwischen Sprachstufen und Wortarten sucht, findet man in der Regel fließende Übergänge. Auch die Unterscheidungen zwischen Adjektiv und zugehörigem Adjektivadverb wurden allmählich, nicht plötzlich abgebaut: Noch heute kann das alte Adjektivadverb *lange* nur in adverbialer Verwendung, *lang* dagegen in adverbialer und prädikativer Verwendung gebraucht werden. Vgl. *das dauert ihr zu lang(e)*, kaum aber *das Kleid ist ihr zu lange*. Umgekehrt setzt schon in mhd. Zeit der Schwund des Adverbsuffixes *-e* ein, und zwar so allmählich, dass Formen mit und ohne *-e* in ein- und demselben Text, in selber Bedeutung und sogar in selber lautlicher Umgebung vorkommen:

- (26) a. *d' quā rechte*/ADJ > AVD *als ein helt gerant vf nampotenisen*
 'der rannte **recht** als ein Held gegen Nampotenis an' (mhd., HTri 6236)
- b. *h' triftan lac ab' recht*/ADJ > AVD *als ein ron*
 'Tristan lag wieder **ganz** wie ein gefällter Stamm (da)' (mhd., HTri 3714)

Zudem gab es auch abgesehen vom Schwund des *-e* schon zu mhd. und sogar ahd. Zeit den Fall, dass Adverb- und Adjektivform ausdrucksseitig nicht unterschieden waren, vgl. ahd. *filu* > mhd. *vil(e)* 'sehr, viel, gänzlich', ein alter Acc.Sg.Neutr.

Im HiTS werden solche adverbial verwendeten Fälle, die wie im Nhd. mit prädikativen gleich lauteten, ebenfalls als Adverb angesetzt. Es wird also letztlich nicht nach der Wortbildungs-, sondern nach der Verwendungsweise unterschieden: Trägt die fragliche Einheit nichts zum Inhalt einer Nominalphrase, sondern einer Verbal-, Adjektiv- oder Adverbphrase bei, wird sie als Adjektivadverb bestimmt.

Auch im STTS gibt es Fälle von Formen, deren Stamm wie ein Adjektivstamm lautet, die aber als ADV bestimmt werden: wenn keine Kopulakonstruktion üblich ist oder ein erheblicher Bedeutungsunterschied gegenüber der Verwendung in der Kopulakonstruktion besteht (vgl. Schiller et al., 1999, S. 57).²¹

²¹Die Bindung an Beleglage, Verwendungsweise und an einen mehr oder minder großen Bedeutungsunterschied lässt absehen, dass die Unterscheidung nicht immer jedem nachvollziehbar ist, die Anzeichnung erschwert und ihr Ergebnis uneinheitlicher wird als nötig. Sogar schon anhand der ausgewählten und von Schiller et al. selbst analysierten Beispiele in den STTS-Richtlinien wird das deutlich. Zwei Beispiele:

1. In *Er ist heute früher gekommen* wird *früher* als ADJD bestimmt, *reichlich* in *er hat reichlich gelacht* dagegen als ADV (vgl. Schiller et al., 1999, S. 45, 57, 58), obwohl die Bedeutung kaum abweicht und auch eine Kopulakonstruktion nicht unüblich ist, vgl. *Das Essen war gut und reichlich*. In *Wir haben reichlich gegessen* wird es als PIS (substituierendes Pronomen) bestimmt (vgl. Schiller et al., 1999, S. 58), also vermutlich als Akkusativobjekt gedeutet. Man kann hier aber sicher eine Deutung als Adverbiale vorziehen, vgl. *Wir haben gut und reichlich gegessen*, wo *reichlich* wohl von gleicher Satzgliedart ist wie *gut*. Nach dem HiTS würden *früher* und *reichlich* in allen genannten Beispielen außer in der Kopulakonstruktion als AVD zu bestimmen sein.

5.2 Attributive Adjektive, vorangestellt und nachgestellt

Beispiel	HiTS	STTS
<i>der reht rihtari</i> 'der gerechte Richter'	ADJ > ADJA <i>Adjektiv, attributiv, vorangestellt</i>	ADJA <i>attr. Adjektiv</i>
<i>die brugge reht</i> 'die richtige Brücke'	ADJ > ADJN <i>Adjektiv, attributiv, nachgestellt</i>	ADJD (s.u.)

Die Beispiele in der Tabelle²² zeigen zwei im Neuhochdeutschen nurmehr ausnahmsweise vorkommende Fälle: Endungslosigkeit bei vorangestelltem attributivem Adjektiv und Nachstellung des attributiven Adjektivs. Ersteres wird nicht über die Wortartbestimmung erfasst, die sich zwischen HiTS und STTS auch nicht unterscheidet. Zu letzterem Schiller et al. (1999, S. 18): „Mit ADJD werden prädikativ und adverbial (auch wenn andere Adjektive modifiziert werden) gebrauchte, sowie nachgestellte, nicht flektierte Adjektive bezeichnet.“ Das hieße, die nachgestellten, endungslosen attributiven Adjektive nicht von prädikativen (und nach dem STTS auch nicht von adverbialen) zu unterscheiden, wohl aber von nachgestellten attributiven Adjektiven mit Endung (die allerdings nur sehr selten vorkamen).

5.3 Substantivierte Adjektive und Substantive

Beispiel	HiTS	STTS
<i>die rehten kinent</i> 'die gerechten/Gerechten strahlen'	ADJ > ADJS <i>Adjektiv, substituierend</i>	ADJA; NN <i>attr. Adj.; normales Nomen</i>
<i>daz opfer des rehtes</i> 'das Opfer des Rechtes'	NA <i>Nomen, appellativ</i>	NN <i>normales Nomen</i>

Wie die Beispiele in der Tabelle²³ mit denen des vorigen Abschnittes zeigen, werden im HiTS substantivierte Adjektive sowohl von attributiven Adjektiven als auch von Substantiven unterschieden. Von ersteren dadurch, dass sie der Kopf ihrer Nominalphrase sind. Das wird auch dort angenommen, wo ein Substantiv als Bezugswort mehr oder minder leicht ergänzt werden könnte — „mehr oder minder leicht“ deutet den Grund an:

2. Zu *reht* und anderen heißt es „diese Wortformen sind niemals ADJD, weil sie keine Kopulakonstruktion bilden können“ (Schiller et al., 1999, S. 57). Diese Begründung gilt für *reht* aber nicht, vgl. *Das ist mir sehr reht*; hier wäre vielleicht das Kriterium eines Bedeutungsunterschiedes heranzuziehen, das aber bei den meisten anderen Beispielen für dieselbe Unterscheidung zwischen ADV und ADJA nicht wirkt (vgl. *ich habe ihn kürzlich/ADV gesehen* vs. *der Anlaß meines kürzlichen/ADJA Besuches*) und auch erst bei einer anderen Unterscheidung, der zwischen ADV und ADJD, angeführt wird.

²²WNot 10rb, Rapp 17582

²³WNot 21va und 6va

Die Unterscheidung wäre oft Spekulation, daher zeitaufwändig und für die Auswertung unergiebig. Ein kaum entscheidbarer Fall ist etwa der, dass in einer Aufzählung bald ein substantivisches Bezugswort beim Adjektiv steht, bald nur das Adjektiv:

- (27) *Er ín hat noch nít vernōmen waz blínden lāmen ftōmen Maladen dode lvde vnde waz vns daz bedvde*
'Ihr habt noch nicht vernommen, was Blinde, Lahme, Stumme, Aussätzige, tote Menschen (sind), und was uns das bedeute (was ihre übertragene Bedeutung ist)'
(PrRei 18b,7)

Hier wäre man wohl eher geneigt, sich für leichte Ergänzzbarkeit des Bezugswortes und damit gegen Substantiviertheit der Adjektive zu entscheiden. Das zeigt aber, wie unsicher solche Einschätzungen sind: Denn im westmitteldt. Dialekt dieses Textes sind die (gegen die Formregel gesetzten) schwachen Endungen der Adjektive *blínden*, *lāmen*, *ftōmen*, *Maladen* ein starkes Zeichen dafür, dass es sich um Substantivierungen handelt, und entsprechend hat nur das offensichtlich nicht substantivierte Adjektiv in *dode lvde* die starke Endung.

Was die Unterscheidung zwischen Substantiven und substantivierten Adjektiven angeht, gibt es einige nicht ganz selten vorkommende Stämme wie *reht*, *guot*, *vil*, die als Adjektiv und als Substantiv vorkommen können. Sie sind in vielen Kasus durch die Endung unterschieden, wie das zweite Beispiel der Tabelle für den Gen.Sg.Neutr zeigt, in anderen Kasus aber nicht, vgl. *daz íft recht* 'das ist recht / (das) Recht' (Rupr 96,13). Dann bleiben unsicherere Anhaltspunkte: die Syntax (kongruierendes Attribut oder Bezugswort mit Genitivattribut) und die Semantik ('gut, Gutes' oder 'Gutes; Besitz, Güte, gute Absicht').

Im STTS sind substantivierte Adjektive entweder als ADJA oder, und zwar bei Großschreibung, als NN zu bestimmen (vgl. Schiller et al., 1999, S. 18f). Die Bindung an die Schreibung macht Spekulation unnötig, ist aber zumal für ältere Sprachstufen nicht sinnvoll, vielleicht auch für jüngere zu sehr von Außersprachlichem (wie der gerade gültigen Rechtschreibung) abhängig. — Ein Sonderfall: Wie üblich wird nach dem STTS in einem Fall wie *die Schweizer Schokolade* ein ADJ angesetzt (vgl. Schiller et al., 1999, S. 19); auf älteren Sprachstufen hat man hier ein Substantiv im Genitiv Plural anzusetzen, worauf Endung und Großschreibung noch im Nhd. hindeuten.

6 Lateinische Passagen

Die frühe deutsche Schriftlichkeit entwickelte sich in direkter Abhängigkeit von den lateinischen christlichen und theologisch-philosophischen Texten, d.h. ein Großteil der altdeutschen Literatur ist mehr oder weniger direkt vom Lateinischen beeinflusst. Diese Beeinflussung wirkt sich über das Inhaltliche hinaus z.B. auf syntaktische Konstruktionen oder auf graphische Repräsentationsformen aus. Ein Teil der lateinabhängigen ahd. Schriften ist in textueller Verquickung mit der lateinischen Vorlage überliefert, es handelt sich dabei u.a. um Interlinearglossierungen (z.B. B), Interlinearübersetzungen (z.B. T) oder um lat.-deutsche Mischtexte (z.B. DH).

(28) zeigt ein Beispiel für eine zeilengemäße lat.-ahd. Interlinearübersetzung. Das Beispiel illustriert, dass die ahd. Interlinearübersetzung immer wieder vom lat. Original abweicht, so z.B. in der zweiten Zeile, wo im Lateinischen *circumcideretur* ‘beschnitten wurde’ vor seinem Subjekt *puer* ‘Kind’ steht, während in der ahd. Übersetzung die umgekehrte Reihenfolge gewählt wurde. Die Zeilenumbrüche im Beispiel entsprechen dem Original der Handschrift, die Zuordnung der entsprechenden lat. und ahd. Konstituenten erfolgt nur der besseren Lesbarkeit wegen.

(28)	Lat. Z1	<i>et</i>	<i>postquam</i>		<i>consummati</i>	<i>sunt</i>	
	Lat.-POS	KON	KOUS		VVPP	VAFIN	
	Ahd. Z1		<i>after</i>	<i>thiu</i>	<i>tho</i>	<i>argangana</i>	<i>uuarun</i>
	Ahd.-POS		KOUS	DDA	AVD	VVPP	VAFIN
	Lat. Z2	<i>dies</i>	<i>octo</i>	<i>ut</i>	<i>circumcideretur</i>	<i>puer</i>	
	Lat.-POS	NA	CARDN	KOUS	VVFIN	NA	
	Ahd. Z2	<i>ahtu</i>	<i>taga</i>	<i>thaz</i>		<i>thaz</i>	<i>kind</i>
	Ahd.-POS	CARDA	NA	KOUS		DDA	NA
	Lat. Z3		<i>vocatum</i>	<i>est</i>	<i>nomen</i>	<i>eius</i>	<i>Ihesus</i>
	Lat.-POS		VVPP	VAFIN	NA	PPER	NE
	Ahd. Z3	<i>bisnitan</i>	<i>vvurdi</i>	<i>uuard</i>	<i>imo</i>	<i>genemnit</i>	<i>namo</i>
	Ahd.-POS	VVPP	VAFIN	VAFIN	PPER	VVPP	NA

‘Nachdem acht Tage vergangen waren, dass das Kind beschnitten wurde, wurde er Jesus/Heiland genannt’ (ahd., T 7,1)

Die ahd. Paralleltexte werden gemeinsam mit ihrer lateinischen Vorlage in das Referenzkorpus Altdeutsch aufgenommen. Die lateinischen Texte werden dabei ebenfalls voll annotiert, so dass ohne weiteres lexikalische und grammatikalische Übersetzungsstrategien abrufbar sind. Die Übernahme des POS-Tagsets erfolgt dabei weitestgehend problemlos, lediglich die Hinzunahme eines Tags für das Gerundium (‘VVINFG’) bzw. Gerundivum (‘VVPG’) war notwendig.

(29)	a.	Lat.	<i>Et</i>	<i>cum</i>	<i>stabit</i>	<i>ad</i>	<i>orandum</i>
			und	wenn	stehen	zu	beten
		Lat.-POS	KON	KOUS	VVFIN	APPR	VVINFG
			‘Und wenn ihr steht und betet’ (ahd., T 121,4)				
	b.	Lat.	<i>filius</i>	<i>hominis</i>	<i>tradendus</i>	<i>est</i>	<i>in</i>
			Sohn	Mensch	ausliefern	sein	in
		Lat.-POS	NA	NA	VVPG	VAFIN	APPR
		Lat.	<i>manus</i>	<i>hominum</i>			
			Hand	Mensch			
		Lat.-POS	NA	NA			
			‘der Menschensohn wird in die Hände der Menschen ausgeliefert’ (ahd., T 93,1)				

7 Schluss

In diesem Artikel wurde HiTS vorgestellt, ein Tagset für die Annotation historischer Sprachstufen des Deutschen.

Der Artikel legte den Schwerpunkt auf wichtige Unterschiede zum STTS. Dies beginnt bei der Tokenisierung der Daten, für die HiTS eigene Richtlinien und Tags bereitstellt, da die Verwendung von Spatien in historischen Sprachdaten ungleich inkonsistenter ist als in modernen Sprachdaten. Oft interagiert die Bestimmung der Wortgrenzen mit der entsprechenden grammatischen Analyse, so z.B. bei der Analyse komplexer Nomensequenzen als Genitivkonstruktion vs. Komposita.

Ein weiterer auffallender Unterschied ist die konsequente Unterscheidung zwischen der Kategorie des Lemmas und des konkreten Belegs. Diese Art der Doppelauszeichnung eignet sich insbesondere für die Untersuchung von Sprachwandelprozessen, die mit einem Wortartwechsel einhergehen, wie z.B. im Fall von Verbpartikeln, die sich aus Adverbien herausbilden.

Verglichen mit modernen Daten ergeben sich überdies in den historischen Daten weitaus häufiger Ambiguitäten, die auch nicht im Kontext auflösbar sind. Der Hauptgrund dafür ist, dass wir heutigen Leser und Annotatoren keine muttersprachlichen Intuitionen haben und deshalb z.B. sprachliche Umformungstests nicht zur Disambiguierung genutzt werden können. In solchen Fällen wird die historisch ältere Version angenommen (so sinnvoll bestimmbar). Beispielsweise wird *der sunnen schön* ‘der Sonne Schein’ als Genitivkonstruktion analysiert und nicht als Kompositum. Damit wird erreicht, dass Sprachwandelprozesse nicht zu früh angesetzt werden können, sondern erst ab dem Zeitpunkt, zu dem eindeutige Evidenz vorliegt.

Man könnte argumentieren, dass es zu bevorzugen wäre, ambige Fälle als solche zu markieren. Dann könnte z.B. im Falle der Komposita anhand der Korpora untersucht werden, ob ab einem gewissen Zeitpunkt ambige Fälle gehäuft auftreten, bis dann die ersten eindeutigen Kompositakonstruktionen in Erscheinung treten.

Zwei Punkte sprechen gegen ein solches Vorgehen: Zum einen würde dadurch die Annotationsaufgabe deutlich komplexer und die resultierenden Annotationen (häufig) unzuverlässiger, so dass die Korpusnutzer sich letztlich nicht darauf verlassen könnten, dass tatsächlich *alle* ambigen Fälle als solche annotiert sind. Zum anderen erfordert eine ambige Annotation eine deutlich aufwändigere Art der Repräsentation, da sich Ambiguitäten über mehrere Annotationsebenen erstrecken können. Im schon genannten Beispiel würde in der Kompositalesart die Sequenz *sunnen schön* beispielsweise als ein Token analysiert, so dass folglich nur ein POS-Tag und nur eine morphologische Annotation vergeben würde (30b).²⁴ In der Genitivanalyse hingegen handelt es sich um zwei Token mit ihren entsprechenden Annotationen (30a). Außerdem wird der Artikel *der* einmal als feminin und einmal als maskulin analysiert. Es ist nicht offensichtlich, wie man diese Informationen innerhalb einer Annotation quer über die Ebenen als zusammengehörig auszeichnen könnte.

²⁴Die Analyse in (30b) entspricht nicht den HiTS-Richtlinien. (30a) zeigt die HiTS-konforme Annotation.

(30)	a.	MHD	<i>der</i>	<i>sunnen</i>	<i>schîn</i>	(nach HiTS)
		POS	DD > DDART	NA > NA	NA > NA	
		MORPH	Fem.Gen.Sg	Fem.Gen.Sg	Masc.Nom.Sg	
	b.	MHD	<i>der</i>	<i>sunnen</i>	<i>schîn</i>	
		TOK	<i>der</i>	<i>sunmenschîn</i>		
		TOKTAG		US		
		POS	DD > DDART	NA > NA		
		MORPH	Masc.Nom.Sg	Masc.Nom.Sg		

Will man die Entwicklung von Komposita im Deutschen untersuchen, so bietet ein nach HiTS annotiertes Korpus dennoch nützliche Information, vorausgesetzt, es enthält auch morphologische Information (die Referenzkorpora des Altdeutschen und Mittelhochdeutschen sind morphologisch annotiert):

Eindeutige Kompositakonstruktionen wie in (31), ‘die Herzenskönigin’, sind z.B. auffindbar durch eine Suche nach Wortformen mit der POS-Annotation „NA > NA“ und dem Token-Tag „US“; (31) wäre hierfür ein Treffer. Potenziell ambige Konstruktionen können durch eine Suche nach einem Artikel und (Adjektiv und) Nomen im Genitiv, direkt gefolgt von einem weiteren Nomen abgefragt werden; hier wäre (30a) ein Treffer. Die Treffer müssen dann manuell daraufhin gesichtet werden, ob der Artikel sich alternativ auch auf das zweite Nomen beziehen könnte. D.h. die Annotation nach HiTS kombiniert mit morphologischer Annotation liefert immerhin genug Information, um sehr gezielt nach ambigen Kompositakandidaten zu suchen.

(31)	MHD	<i>die</i>	<i>herzen</i>	<i>küniginne</i>	(nach HiTS)
	TOK	<i>die</i>	<i>herzenküniginne</i>		
	TOKTAG		US		
	POS	DD > DDART	NA > NA		
	MORPH	Fem.Nom.Sg	Fem.Nom.Sg		

Wünschenswert wäre natürlich, dass die Korpora von solchen Untersuchungen profitieren könnten, also dass beispielsweise derjenige, der die Kompositakandidaten einzeln überprüft, das Resultat seiner Überprüfung („ja, ambig“ oder „nein, eindeutig Genitivkonstruktion“) zum Korpus hinzufügen könnte. Die aktuellen Korpusarchitekturen unterstützen solche Wiki-ähnlichen Beitragsmöglichkeiten allerdings noch nicht.

Quellen

Altdeutsch

- B *Benediktinerregel*. — Edition: E. Steinmeyer (Hg.). Die kleineren althochdeutschen Denkmäler. Dublin/Zürich 1971 (Nachdruck von 1916). 190–289.
- BGB3 *Benediktbeurer Glauben und Beichte 3*. — Edition: E. Steinmeyer (Hg.). Die kleineren althochdeutschen Denkmäler. Dublin/Zürich 1971 (Nachdruck von 1916). 357–362.
- DH *De Heinrico*. — Edition: E. Steinmeyer (Hg.). Die kleineren althochdeutschen Denkmäler. Dublin/Zürich 1971 (Nachdruck von 1916). 110–114.
- FP2 *Freisinger Paternoster 2*. — Edition: E. Steinmeyer (Hg.). Die kleineren althochdeutschen Denkmäler. Dublin/Zürich 1971 (Nachdruck von 1916). 43–48.
- Hel *Heliand*. — Edition: B. Taeger (Hg.). Heliand. Tübingen 1984.
- I *Isidor*. — Edition: H. Eggers (Hg.). Isidor. Tübingen 1964.
- M *Muspilli*. — Edition: E. Steinmeyer (Hg.). Die kleineren althochdeutschen Denkmäler. Dublin/Zürich 1971 (Nachdruck von 1916). 66–81.
- O *Otfrid: ‘Evangelienbuch’*. — Edition: O. Erdmann (Hg.). Otfrids Evangelienbuch. Tübingen ⁶1973.
- T *Tatian*. — Edition: E. Sievers (Hg.). Tatian. Paderborn 1966 (Nachdruck der zweiten Ausgabe von 1892).
- WK *Weissenburger Katechismus*. — Edition: E. Steinmeyer (Hg.). Die kleineren althochdeutschen Denkmäler. Dublin/Zürich 1971 (Nachdruck von 1916). 29–38.

Mittelhochdeutsch

- AlexiusA *Alexius A*. Handschrift: Graz, Universitätsbibl., Ms. 1501, Bl. 70–134. — Edition: H. F. Maßmann (Hg.). Sanct Alexius Leben in acht gereimten mittelhochdeutschen Behandlungen. Quedlinburg/Leipzig 1843. 45–67.
- ArnoltSieb *Priester Arnold: ‘Von der Siebenzahl’*. Handschrift: Voralpe, Stiftsbibl., Cod. 276, Bl. 129vb–133vb. — Edition: F. Maurer (Hg.). Die religiösen Dichtungen des 11. und 12. Jahrhunderts. Bd. III. Tübingen 1970. 53–85.
- BaGB *Bamberger Glaube und Beichte*. Handschrift: München, BSB, Cgm 4460. — Edition: E. von Steinmeyer (Hg.). Die kleineren althochdeutschen Sprachdenkmäler. Berlin 1916. Nr. 28B.
- Erlös *Die Erlösung*. Handschrift: Krakau, Bibl. Jagiellońska, Berol. mgq 1412; Laubach, Graf zu Solms–Laubach’sche Bibl., Fragm. T. — Edition: F. Maurer (Hg.). Die Erlösung. Eine geistliche Dichtung des 14. Jahrhunderts. Leipzig 1934. 302–308.
- GRud *Graf Rudolf*. Handschrift: Braunschweig, Stadtbibl., Fragm. 36; Göttingen, Staats- und Universitätsbibl., 4° Cod. Ms. philol. 184.VII. — Edition: C. von Kraus (Hg.). Mittelhochdeutsches Übungsbuch, Heidelberg 1912. 54–71.

- HTri *Heinrich von Freiberg: 'Tristan'*. Handschrift: Florenz, Nationalbibl., Cod. B.R. 226, Bl. 103ra–139vb. — Edition: D. Buschinger (Hg.). Heinrich von Freiberg, Tristan. Göppingen 1982.
- Iw *Hartmann von Aue: 'Iwein'*. Handschrift: Gießen, Universitätsbibl., Hs. 97. — Edition: G. F. Benecke u. K. Lachmann (Hg.). Iwein. Eine Erzählung von Hartmann von Aue. Neu bearb. von L. Wolff. 7. Ausg. Bd. 1: Text. Berlin 1968.
- Kchr *Kaiserchronik A (V)*. Handschrift: Voral, Stiftsbibl., Cod. 276, Bl. 1ra–73vb. — Edition: E. Schröder (Hg.). Kaiserchronik eines Regensburger Geistlichen (MGH Deutsche Chroniken I,1). Berlin 1895.
- Nib *Nibelungenlied*. Handschrift: Karlsruhe, LB, Cod. Donaueschingen 63, fol. 1r–89r (= Nibelungenlied C). — Edition: U. Hennig (Hg.). Das Nibelungenlied nach der Handschrift C (Altdeutsche Textbibliothek 83). Tübingen 1977.
- Phys *Wiener (Jüngerer) Physiologus*. Handschrift: Wien, Österr. Nationalbibl., Cod. 2721, Bl. 130r–158v. — Edition: F. Wilhelm (Hg.). Denkmäler deutscher Prosa des 11. und 12. Jahrhunderts, Abteilung A: Text. München 1914/16. 5–28.
- PrF *Frankfurter Predigtfragmente*. Handschrift: Frankfurt, SUB, Fragm. germ. I 1. — Edition: L. Diefenbach. Mitteldeutsche Predigtbruchstücke. In: Germania 19 (1874). 305–314.
- PrRei *Hessische Reimpredigten*. Handschrift: Hamburg, SUB, Cod. 99 in scrin. 12–312. — Edition: B. Lenz-Kemper (Hg.). Die Hessischen Reimpredigten. Bd. 2: Text. Berlin 2009.
- Rapp *Rappoltsteiner Parzifal*. Handschrift: Karlsruhe, LB, Codex Donaueschingen 97 (Parzival Gd, Hs. D der frz. Perceval-Forschung). — Edition: K. Schorbach (Hg.). Parzifal von Claus Wisse und Philipp Colin (1331–1336). Eine Ergänzung der Dichtung Wolframs von Eschenbach. Straßburg/London 1888.
- RhMI *Rheinisches Marienlob*. Handschrift: Hannover, Landesbibl., Ms. I 81, Bl. 1r–93v. — Edition: A. Bach (Hg.). Das Rheinische Marienlob. Eine deutsche Dichtung des 13. Jahrhunderts. Leipzig 1934.
- Roth *König Rother (H)*. Handschrift: Heidelberg, Universitätsbibl., Cpg 390. — Edition: Th. Frings u. J. Kuhnt (Hg.). König Rother. Bonn/Leipzig 1922.
- Rupr *Ruprecht von Freising: 'Rechtsbuch'*. Handschrift: München, StadtA, Zimelie 1. — Edition: H.-K. Claußen (Hg.). Freisinger Rechtsbuch. Weimar 1941.
- TrHL *St. Trudperter Hohes Lied (A)*. Handschrift: Wien, Österr. Nationalbibl., Cod. 2719. — Edition: H. Menhardt (Hg.). Das St. Trudperter Hohe Lied. Halle a.d. Saale 1934. 1–42.
- VAlex *Lambrechts Alexander (Vorauer Alexander)*. Voral, Stiftsbibl., Cod. 276, Bl. 109ra–115va. — Edition: K. Kinzel (Hg.). Lamprechts Alexander. Nach den drei Texten mit dem Fragment des Alberic von Besançon und den lateinischen Quellen. Halle 1884.
- Will *Williram von Ebersberg: 'Hoheliedkommentar'*. Breslau / Wrocław, Stadtbibl., Cod. R 347 [Kriegsverlust]. — Edition: E. H. Bartelmez (Hg.). The „Expositio in Cantica Canticorum“ of Williram, Abbot of Ebersberg 1048–1085. A Critical Edition. Philadelphia 1967.

- WMEv *Wien-Münchner Evangelienübersetzung*. Handschrift: München, Staatsbibl., Cgm 5250/1; Oxford, Bodleian Libr., MS Germ. b. 3, f. 15; Wien, Österr. Nationalbibl., Cod. Ser. nova 249. — Edition: H. Kriedte (Hg.). Deutsche Bibelfragmente in Prosa des XII. Jahrhunderts. Halle a.d. Saale 1930. 11–14, 64–122.
- WNot *Wiener Notker*. Handschrift: Wien, ÖNB, Codex 2681 („Wiener Notker“ = Notker Y). — Edition: R. Heinzel u. W. Scherer (Hg.). Notkers Psalmen nach der Wiener Handschrift. Straßburg/London 1876.

Literatur

- Dudenredaktion (Hg.) (2009). *Die Grammatik*. Duden 4. Dudenverlag, Mannheim/Wien/Zürich. 8., überarb. Aufl.
- Höchli, S. (1981). *Zur Geschichte der Interpunktion im Deutschen*. de Gruyter, Berlin/New York.
- Lexer, M. (1872–1878). *Mittelhochdeutsches Handwörterbuch*. Hirzel, Leipzig.
- Schiller, A., Teufel, S., Stöckert, C. und Thielen, C. (1999). Guidelines für das Tagging deutscher Textcorpora mit STTS (Kleines und großes Tagset). Technischer Bericht, Universitäten Stuttgart und Tübingen, <http://www.ims.uni-stuttgart.de/forschung/ressourcen/lexika/TagSets/stts-1999.pdf>.
- Schottelius, J. G. (1641). *Teutsche Sprach Kunst*. Braunschweig. (2. Aufl. 1651).
- Splett, J. (1993). *Althochdeutsches Wörterbuch*. de Gruyter, Berlin/New York. 2 Bände.

Appendizes

Appendix I: Token-Tags

Appendix II: Wortart-Tags

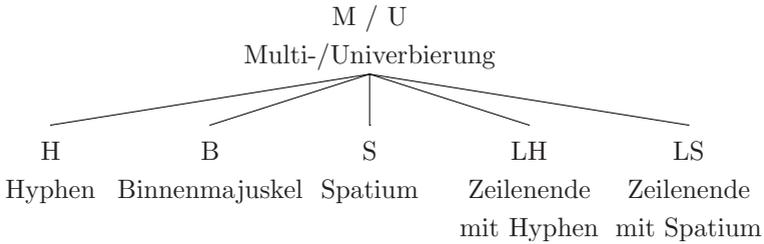
1	ADJ	Adjektive
2	AP	Adpositionen
3	AV	Adverbien
4	CARD	Kardinalzahlen
5	D, P	Determinativa und Pronomen
6	KO	Konjunktionen
7	N	Nomen
8	PAV	Pronominaladverbien
9	PTK	Partikeln
10	V	Verben
11	ITJ, FM	Verschiedenes
12	\$	Interpunktion

Appendix III: Alphabetische Auflistung aller Beleg-Tags

Hinweise zu den Tabellen I und II in den Appendizes

- Die folgenden Tabellen geben eine komplette Auflistung aller in HiTS vorgesehenen Tags. Jeder Tabelle ist eine graphische Darstellung der verwendeten Abkürzungen in den Tagnamen vorangestellt.
- Die Tabellen zu den Wortart-Tags beginnen jeweils mit den „Normalfällen“, d.h. typischerweise solchen Fällen, in denen die Hauptwortart des Lemmas und des Belegs übereinstimmen (ein Beispiel aus der Adjektiv-Tabelle: ADJ > ADJA). Als nächstes werden (exemplarisch) Fälle gelistet, in denen die Lemmawortart abweicht (VVPP > ADJA), gefolgt von Fällen, in denen die Belegwortart abweicht (ADJ > AVD).
- Gelegentlich stehen unter den Tabellen noch Anmerkungen, die auf Unterschiede zum STTS hinweisen (z.B. wird darauf hingewiesen, dass und warum es in HiTS kein Tag APPRART gibt).

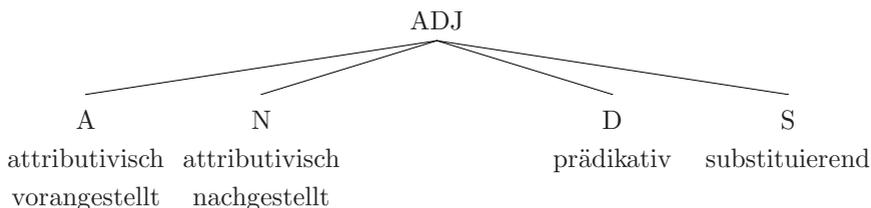
Appendix I: Überblickstabelle Token-Tags



original	Beispiel modernisiert	HiTS
<i>indaz</i>	<i>in daz</i>	MS (<i>Multiverbierung mit Spatium</i>)
<i>LandGraff</i>	<i>Landgraff</i>	UB (<i>Univerbierung mit Binnenmajuskel</i>)
<i>Liebes=Ohnmachten</i>	<i>Liebesohnmachten</i>	UH (<i>Univerbierung mit Hyphen</i>)
<i>Schutz Gott</i>	<i>Schutzgott</i>	US (<i>Univerbierung mit Spatium</i>)
<i>in- ↔ dem lande</i>	<i>in dem lande</i>	MLH (<i>Multiverbierung am Zeilenende mit Hyphen</i>)
<i>ou= ↔ gen</i>	<i>ougen</i>	ULH (<i>Univerbierung am Zeilenende mit Hyphen</i>)
<i>ou ↔ gen</i>	<i>ougen</i>	ULS (<i>Univerbierung am Zeilenende mit Spatium</i>)

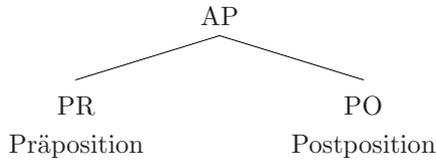
Appendix II: Überblickstabellen Wortart-Tags

1 Adjektive



Beispiel	HiTS	STTS
<i>der</i> [reht] <i>rihtari</i> ‘der gerechte Richter’	ADJ > ADJA <i>Adjektiv, attributiv, vorangestellt</i>	ADJA <i>attr. Adjektiv</i>
<i>die brugge</i> [reht] ‘die richtige Brücke’	ADJ > ADJN <i>Adjektiv, attributiv, nachgestellt</i>	(ADJD; selten)
<i>daz almuofen ist</i> [reht] ‘das Almosen ist gerecht’	ADJ > ADJD <i>Adjektiv, prädikativ</i>	ADJD <i>präd. oder adv. Adj.</i>
<i>die</i> [rehten] <i>skinent</i> ‘die gerechten/Gerechten strahlen’	ADJ > ADJS <i>Adjektiv, substituierend</i>	ADJA; NN <i>attr. Adj.; normales Nomen</i>
<i>uf ainem</i> [getoweten] <i>chle</i> ‘auf einem betauten Klee’	VVPP > ADJA <i>Partizip Präteritum, adjektivisch</i>	ADJA <i>attr. Adj.</i>
<i>daz</i> [brinnent] <i>ole</i> ‘das brennende Öl’	VVPS > ADJA <i>Partizip Präsens, adjektivisch</i>	ADJA <i>attr. Adj.</i>
Außerdem:		
<i>lebet</i> [rehto] ‘lebt gerecht!’	ADJ > AVD <i>Adjektiv, adverbial</i>	ADJD; ADV <i>präd. oder adv. Adj.; Adverb</i>

2 Adpositionen



Beispiel	HiTS	STTS
[<i>mit</i>] <i>der hant</i> ‘mit der Hand’; <i>da</i> [<i>z</i>] <i>en</i> (= <i>zen</i>) <i>Hunin</i> ‘dort bei den Hunnen/im Hunnenland’	AP > APPR <i>Präposition</i>	APPR; APPRART <i>Präposition (mit Artikel)</i>
<i>inan</i> [<i>úbari</i>] ‘über ihn’	AP > APPO <i>Postposition</i>	APPO <i>Postposition</i>
Außerdem:		
<i>dâ págant siu</i> [<i>umpi</i>] ‘darum streiten sie’	AP > PAVAP <i>Pron.adverb, präpositionaler Teil</i>	—

Anmerkungen:

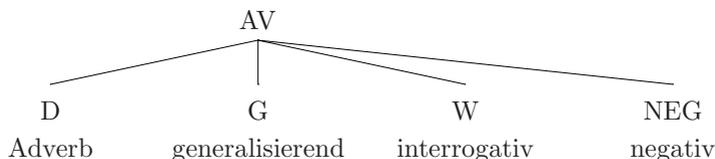
- Kontrahierte Präpositionen (im STTS: APPRART) existieren auch in den früheren Sprachstufen, werden jedoch in HiTS immer als zwei Token (Artikel + Präposition) analysiert:

zen | *z/APPR en/DDART* (< *ze den*) | MS ‘zu den’

- Der nachgestellte zweite Teil moderner Zirkumpositionen (im STTS: APZR) entspricht in früheren Sprachstufen einem Nomen:

durch der liebe willen/NA ‘um der Liebe willen’

3 Adverbien



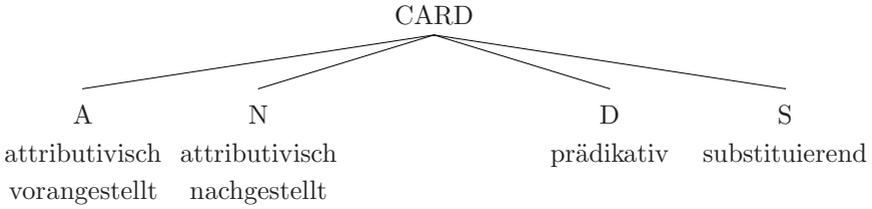
Beispiel	HiTS	STTS
<i>ein gotishus [uīl] mechtic</i> ‘eine sehr mächtige Kirche’	AVD > AVD <i>Adverb</i>	ADV <i>Adverb</i>
[do] <i>begagenda imo min trohtin</i> ‘da begegnete ihm mein Herr’	AVD-KO > AVD <i>Konjunkionaladverb, adverbial</i>	ADV <i>Adverb</i>
[swie] <i>sie wellen</i> ‘wie immer sie wollen’	AVG > AVG <i>Relativadverb, generalisierend</i>	PWAV <i>adv. Int.- o. Rel.pron.</i>
[hwār] ‘wo’; [wie] <i>fro</i> ‘wie froh’	AVW > AVW <i>Adverb, interrogativ</i>	PWAV <i>adv. Int.- o. Rel.pron.</i>
<i>der daz tages lieht [nie] negesah</i> ‘der das Tageslicht nie (nicht) sah’	AVD > AVNEG <i>Adverb, negativ</i>	AVD <i>Adverb</i>
<i>lebet [rehto]</i> ‘lebt gerecht!’; [groze] <i>willechomen</i> ‘sehr willkommen’	ADJ > AVD <i>Adjektiv, adverbial</i>	ADJD <i>präd. oder adv. Adj.</i>
Außerdem:		
[dâ] <i>pâgant siu umpi</i> ‘darum streiten sie’	AVD > PAVD <i>Pron.adverb, pronominaler Teil</i>	—
[uf] <i>huob er die hende</i> ‘hoch hob er die Hände’	AVD > PTKVZ <i>Verbzusatz</i>	PTKVZ <i>abgetr. Verbzusatz</i>

Anmerkung:

- Relativische Verwendungen werden ebenfalls mit AVD getagged:

nv was div iwchfrowwe genomen her vz, da/AVD si gefangen lac ‘nun war das Mädchen (dort) herausgeholt worden, wo sie gefangen gelegen hatte’

4 Kardinalzahlen



Beispiel	HiTS	STTS
[<i>hunderet</i>] unde [<i>uifzech</i>] tage '150 Tage'	CARD > CARDA <i>Kardinalzahl, attributiv, vorangestellt</i>	CARD <i>Kardinalzahl</i>
<i>daz waren ceichen</i> [<i>sibeniu</i>] 'das waren sieben Zeichen'	CARD > CARDN <i>Kardinalzahl, attributiv, nachgestellt</i>	—
(ohne eindeutigen Korpusbeleg)	CARD > CARDD <i>Kardinalzahl, prädikativ</i>	CARD <i>Kardinalzahl</i>
<i>in [driu] getailet 'in drei geteilt'; siben</i> [<i>hundert</i>] <i>siner manane</i> <i>diu [eine] 'die eine'</i>	CARD > CARDS <i>Kardinalzahl, substituierend</i>	CARD <i>Kardinalzahl</i>

5 Determinativa und Pronomen

Anmerkungen zu den Determinativa und Pronomen:

- Tags, deren Name mit „D“ beginnt (\approx Determinativa), haben zwei Unterkategorien, für den Typ und die Position, z.B.:

ein/DIS (Determinativ, indefinit, substituierend)

- Tags, deren Name mit „P“ beginnt (\approx Pronomen), gelten stets als substituierend und haben deswegen nur eine Unterkategorie für den Typ, z.B.:

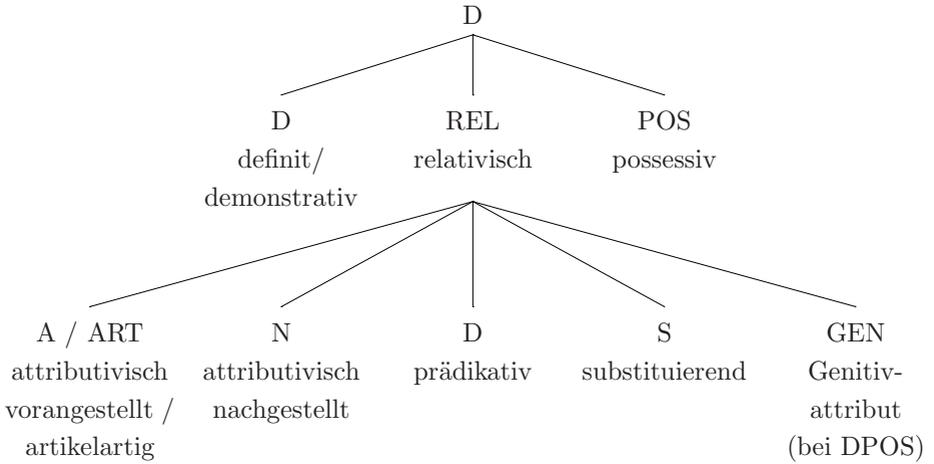
man/PI (Pronomen, indefinit)

- Attributive Verwendungen im Genitiv, die im STTS mit „AT“ markiert werden (PRELAT, PWAT), werden in HiTS als substituierend annotiert (DRELS, PW), z.B.:

Sanctus Johannes, des/DRELS tac wir hivte begen ‘Sankt Johannes, dessen Tag wir heute begehen’

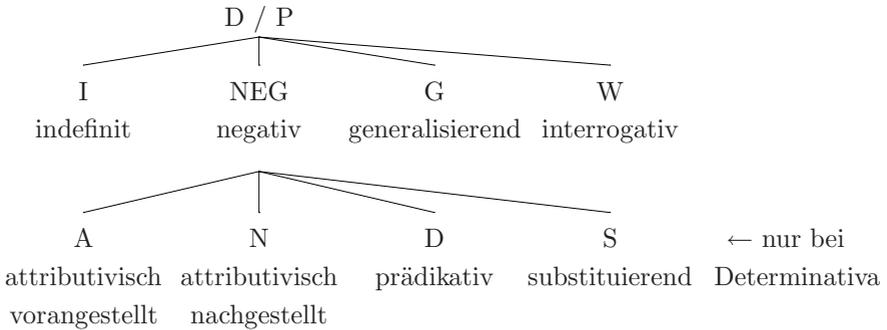
wes/PW sun ‘wessen Sohn’

- Es stehen nicht zu allen theoretisch denkbaren Subklassen Beispiele in den Tabellen. Beispielsweise fehlen bei vielen prädikativen Verwendungen die Einträge (z.B. DDD, DID etc.). Sie sind aber prinzipiell nicht ausgeschlossen.



Beispiel	HiTS	STTS
Definita/Demonstrativa: DD		
[<i>der</i>] <i>liehte tac</i> ‘der helle Tag’	DD > DDA (ad.), DDART (mhd.) <i>Determinativ, definit, vorangestellt bzw. artikelartig</i>	ART <i>bestimmter Artikel</i>
[<i>dise</i>] <i>rede</i> ‘diese Rede’; <i>ober alle [die] lant</i> ‘über alle diese Länder’; <i>vf [ienem] sal witen</i> ‘in jenem weiten Saal’	DD > DDA <i>Determinativ, definit/demonstrativ, vorangestellt</i>	PDAT <i>attr. Dem.pron</i>
<i>der got [selbe]</i> ‘Gott selbst’	DD > DDN <i>Determinativ, definit/demonstrativ, nachgestellt</i>	—
[<i>dizze</i>] <i>ist ein anphanlich zit</i> ‘dieses ist eine angenehme Zeit’; <i>in [des] gewalt</i> ‘in dessen Gewalt’	DD > DDS <i>Determinativ, definit/demonstrativ, substituierend</i>	PDS <i>subst. Dem.pron</i>

Beispiel	HiTS	STTS
Relativpronomen: DREL		
<i>di haiden</i> , [<i>dɪ</i>] <i>dort wa- ren</i> ‘die Heiden, die dort waren’; <i>Sanctus Johan- nes</i> , [<i>des</i>] <i>tac wir hɪvte begen</i> ‘Sankt Johannes, dessen Tag wir heute be- gehen’	DD > DRELS <i>Determinativ, relativisch, substituie- rend</i>	PRELS; PRELAT <i>subst. oder attr. Rel.pron</i>
Possessiva: DPOS		
<i>in</i> [<i>mɪnem/irem</i>] <i>hús</i> ‘in meinem/ihrer Haus’	DPOS > DPOSA <i>Determinativ, possessiv, vorangestellt</i>	PPOSAT <i>attr. Poss.pron.</i>
<i>der name</i> [<i>sin</i>] ‘sein Na- me’	DPOS > DPOSN <i>Determinativ, possessiv, nachgestellt</i>	—
<i>thaz dar thin</i> [<i>thin</i>] <i>ist</i> ‘das da dein ist’	DPOS > DPOSD <i>Determinativ, possessiv, prädikativ</i>	PPOSS <i>subst. Poss.pron.</i>
<i>alle die</i> [<i>sine</i>] ‘all die Seinen’	DPOS > DPOSS <i>Determinativ, possessiv, substituie- rend</i>	NN <i>normales Nomen</i>
<i>mid</i> [<i>iro</i>] <i>handon</i> ‘mit ihren Händen (= mit den Händen ihrer/von ihnen)’	PPER > DPOSGEN <i>Determinativ, personal-possessiv, Ge- nitiv (aus Pronomen)</i>	—



Beispiel	HiTS	STTS
Indefinita: DI/PI		
[<i>ein</i>] <i>tier</i> ‘ein Tier’	DI > DIA (ad.), DIART (mhd.) <i>Determinativ, indefinit, vorangestellt bzw. artikelartig</i>	ART <i>unbestimmter Artikel</i>
[<i>alliv</i>] <i>iar</i> ‘alle Jahre’; <i>thúruh</i> [<i>thehéin</i>] <i>wóroltruam</i> ‘für irgend-einen irdischen Ruhm’	DI > DIA <i>Determinativ, indefinit, vorangestellt</i>	PIAT <i>attr. Indef.pron.</i>
under disen chunigen [<i>allen</i>] ‘unter all diesen Königen’	DI > DIN <i>Determinativ, indefinit, nachgestellt</i>	—
<i>der redet</i> [<i>vil</i>] ‘der redet viel’; [<i>theheín</i>] <i>thero fórasagono</i> ‘(irgend-)einer der Propheten’	DI > DIS <i>Determinativ, indefinit, substituierend</i>	PIS <i>subst. Indef.pron.</i>
[<i>íaman</i>] <i>hiar in lán</i> te ‘jemand hier im Land’	PI > PI <i>Pronomen, indefinit</i>	PIS <i>subst. Indef.pron.</i>
[<i>huer</i>] ‘(irgend)wer’	PW > PI <i>Pronomen, indefinit</i>	PIS <i>subst. Indef.pron.</i>
[<i>man</i>] ‘man’	NA > PI <i>Pronomen, indefinit (aus Substantiv)</i>	PIS <i>subst. Indef.pron.</i>
Außerdem:		
<i>di en salt du</i> [<i>nít</i>] <i>schouwen</i> ‘die sollst du nicht anschauen’	PI > PTKNEG <i>Negationspartikel (aus Pronominal-substantiv)</i>	PTKNEG <i>Negationspartikel</i>

Beispiel	HiTS	STTS
Negative: DNEG/PNEG		
<i>ther heilant ni gab imo [nohhein] antuurti</i> ‘der Heiland (nicht) gab ihm keine Antwort’	DI > DNEGA <i>Determinativ, negativ, vorangestellt</i>	PIAT <i>attr. Indef.pron.</i>
<i>ni si mán [nihein] so véigi</i> ‘kein Mann soll so gottlos sein’	DI > DNEGN <i>Determinativ, negativ, nachgestellt</i>	PIAT <i>attr. Indef.pron.</i>
<i>[nihéinan] ni gifångun</i> ‘sie nahmen keinen/nie-manden gefangen’	DI > DNEGS <i>Determinativ, negativ, substituierend</i>	PIS <i>subst. Indef.pron.</i>
<i>[niowih] ‘nichts’; [nioman] giuwisso in taugle uuaz tuot</i> ‘gewiß tut niemand etwas im Verborgenen’	PI > PNEG <i>Pronomen, indefinit, negativ</i>	PIS <i>subst. Indef.pron.</i>

Anmerkungen:

- Ad. *nohhein* und mhd. *nehein* treten nur in negativen Kontexten (d.h. mit Negation) auf und werden mit DI > DNEGA etc. annotiert:

ther heilant ni gab imo nohhein/DI > DNEGA antuurti ‘Der Heiland (nicht) gab ihm keine Antwort’ (ad.)

daz ne saget uns nehein/DI > DNEGA puch daz deheiner so riche ware ‘davon (nicht) berichtet uns kein Buch, dass jemand so mächtig gewesen wäre’ (mhd.)

- Ad. *thehein* kommt nur in nicht-affirmativen Kontexten (z.B. in Fragesätzen, hypothetischen Sätzen, mit Negation; wie Englisch *any(one)*) vor und wird mit DI > DIA/DIN etc. annotiert.

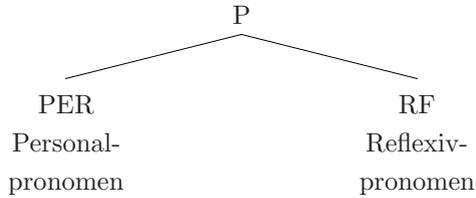
nist ist thehein/DI > DIA tuala ‘es gibt keinen Zweifel (= nicht ist irgendein Zweifel’ (ad.)

war imo súlih man thihéin/DI > DIN so quami wisheiti héim ‘woher käme ihm ein solcher Mann (= solch Mann (irgend)ein) der Weisheit heim (ad.)

- Mhd. *dehein* (der Vorläufer von nhd. *kein*) wird teilweise noch wie ad. *thehein* in nicht-affirmativen Kontexten (wie Englisch *any(one)*) verwendet, teilweise drückt es aber auch schon allein die Negation aus (wie Englisch *no(one)*). Da die Unterscheidung teilweise nur schwer zu treffen ist, werden beide Fälle mit DI > DNEGA/DNEGS etc. annotiert.

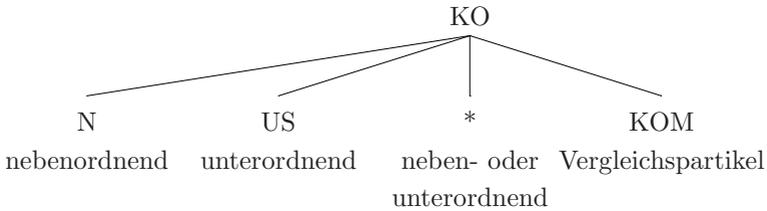
daz ne saget uns nehein puch daz deheiner/DI > DNEGS so riche ware ‘davon (nicht) berichtet uns kein Buch, dass jemand so mächtig wäre’ (mhd.)

Beispiel	HiTS	STTS
Generalisierende: DG/PG		
<i>uernemet ...</i> , [<i>swelh</i>] <i>rat iv baz geualle</i> ‘vernehmt ... , welcher Rat euch besser gefällt’	DG > DGA <i>Determinativ, generalisierend, vorangestellt</i>	PWAT <i>attr. Interr.pron.</i>
[<i>swelh</i>] <i>baz entwíchen mag, der sol ouch entwíchen</i> ‘welcher [sc. Wagen] besser ausweichen kann, der soll auch ausweichen’	DG > DGS <i>Determinativ, generalisierend, substituierend</i>	PWS <i>subst. Interr.pron.</i>
[<i>swer</i>] <i>mir niene tvot, der sol óvch mich zefrívnde han</i> ‘wer mir nichts tut, der soll auch mich zum Freund haben’; <i>ín [swez] wiltbanne</i> ‘in wessen Jagdrevier’	PG > PG <i>Pronomen, generalisierend</i>	PWS; PWAT <i>subst. oder attr. Interr.pron.</i>
Interrogativa: DW/PW		
<i>owi, [welh] mort da gefrumet wart!</i> ‘oh weh, welch ein Mord wurde da verübt!’	DW > DWA <i>Determinativ, interrogativ, vorangestellt</i>	PWAT <i>attr. Interr.pron.</i>
[<i>wélih</i>] ? ‘welcher?’	DW > DWS <i>Determinativ, interrogativ, substituierend</i>	PWS <i>subst. Interr.pron.</i>
[<i>wer</i>] ? ‘wer?’; [<i>wes</i>] <i>bedunchet whc von criste, [wes] sun er si?</i> ‘Was meint ihr von Christus, wessen Sohn er sei?’	PW > PW <i>Pronomen, interrogativ</i>	PWS; PWAT <i>subst. oder attr. Interr.pron.</i>



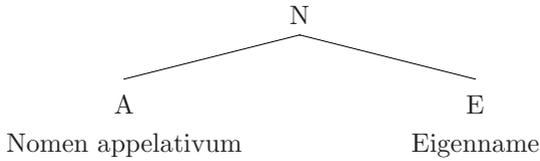
Beispiel	HiTS	STTS
Sonstige Pronomen		
[er] 'er'	PPER > PPER <i>Pronomen, personal, irreflexiv</i>	PPER <i>irrefl. Pers.pron.</i>
<i>im(o)</i> 'sich'; [<i>ein-ander</i>] 'einander'	PPER > PRF <i>Pronomen, personal, reflexiv</i>	PRF <i>refl. Pers.pron.</i>
[<i>sich</i>] 'sich'	PRF > PRF <i>Pronomen, personal, reflexiv</i>	PRF <i>refl. Pers.pron.</i>
Außerdem:		
<i>mid</i> [<i>iro</i>] <i>handon</i> 'mit ihren Händen (= mit den Händen ihrer/von ihnen)'	PPER > DPOSGEN <i>Pronomen, personal-possessiv, Genitiv</i>	—

6 Konjunktionen



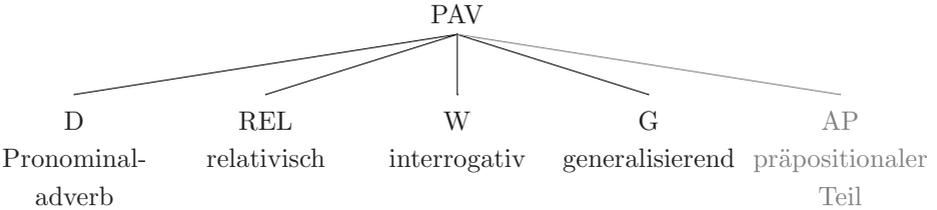
Beispiel	HiTS	STTS
<i>ih</i> [unde] <i>iohan</i> ‘ich und Johann’	KO > KON <i>Konjunktion, nebenordnend</i>	KON <i>nebenordn. Konj.</i>
[sundir] <i>du solt in mich verwandelot werden</i> ‘sondern du sollst/wirst in mich verwandelt werden’	AVD-KO > KON <i>Konjunkionaladverb, nebenordnend</i>	KON <i>nebenordn. Konj.</i>
[daz] <i>si giengen heim</i> ‘dass sie heimgingen’	KO > KOUS <i>Konjunktion, unterordnend</i>	KOUS <i>unterordn. Konj.</i>
<i>do sint si inne</i> , [want] <i>her betit</i> ‘darin sind sie, denn/da er betet’	KO > KO* <i>Konjunktion, neben- oder unterordnend</i>	KON; KOUS <i>neben- o. unterordn. Konj.</i>
<i>mer</i> [danni] <i>ein iar</i> ‘mehr als ein Jahr’	KO > KOKOM <i>Vergleichspartikel</i>	KOKOM <i>Vergleichskonj.</i>
Außerdem:		
[do] <i>begagenda imo min trohtin</i> ‘da begegnete ihm mein Herr’	AVD-KO > AVD <i>Konjunkionaladverb, adverbial</i>	ADV <i>Adverb</i>

7 Nomen



Beispiel	HiTS	STTS
<i>der</i> [<i>kuninc</i>] 'der König'	NA > NA <i>Nomen appellativum</i>	NN <i>normales Nomen</i>
<i>ze</i> [<i>rome</i>] 'nach Rom'	NE > NE <i>Eigenname</i>	NE <i>Eigennamen</i>
<i>mit</i> [<i>suften</i>] und <i>mit</i> [<i>weinen</i>] 'mit Seufzen und mit Weinen'; <i>ez ist wol ze</i> [<i>lobenne</i>] 'es ist sehr zu loben'	VVINF > NA <i>Infinitiv, substantiviert</i>	NN <i>normales Nomen</i>

8 Pronominaladverbien

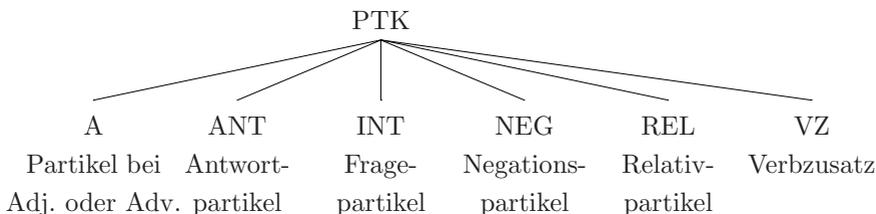


Beispiel	HiTS (zweiteilig) pronominal präpositional	STTS
Pronominaladverbien		
[<i>dā</i>] <i>pāgant siu</i> [<i>umpī</i>] 'darum streiten sie'; [<i>dā</i>] <i>was vil volkes</i> [<i>inne</i>] 'darin waren viele Leute'	AVD > PAVD + AP > PAVAP <i>Pron.adverb</i>	PAV <i>Pron.adverb</i>
<i>thaz fūndament ...</i> , [<i>thar</i>] <i>thiu érda ligit</i> [<i>úfe</i>] 'das Fundament, worauf die Erde ruht'; <i>der palas</i> , [<i>dā</i>] ... [<i>inne</i>] <i>was</i> 'der Palast, in dem'	AVD > PAVREL + AP > PAVAP <i>Pron.adverb, relativisch</i>	PWAV <i>adv. Int.- o. Rel.pron.</i>
[<i>wā ... umbe</i>] 'warum'	AVW > PAVW + AP > PAVAP <i>Pron.adverb, interrogativ</i>	PWAV <i>adv. Int.- o. Rel.pron.</i>
[<i>swā ... umbe</i>] 'warum (auch immer)'	AVG > PAVG + AP > PAVAP <i>Pron.adverb, generalisierend</i>	PWAV <i>adv. Int.- o. Rel.pron.</i>

Anmerkung:

- Anstelle der Präpositionen können auch bestimmte Adverbien in Pronominaladverbien vorkommen. Diese erhalten als Lemmawortart ebenfalls AP zugewiesen:
ez chnite fuor im dar/PAV > PAVD nider/AP > PAVAP 'es kniete vor ihm nieder'

9 Partikeln



Beispiel	HiTS	STTS
[ze] <i>lanc</i> ‘zu lang’; [ze] <i>uul</i> ‘zu viel’	PTK > PTKA <i>Partikel bei Adjektiv oder Adverb</i>	PTKA <i>Part. bei Adj. o. Adv.</i>
[<i>nêin</i>], <i>sprach der herre Gawêin</i> ‘‘nein’’, sagte Herr Gawein’	PTK > PTKANT <i>Antwortpartikel</i>	PTKANT <i>Antwortpartikel</i>
[<i>eno</i>] <i>ni lâsut ir in giscribum</i> ‘Last ihr etwa nicht in der Heiligen Schrift?’	PTK > PTKINT (ad.) <i>Fragepartikel</i>	—
<i>di</i> [<i>en</i>] <i>salt du nit schouwen</i> ‘die sollst du nicht anschauen’	PTK > PTKNEG <i>Negationspartikel</i>	PTKNEG <i>Negationspartikel</i>
<i>than is im sô them salte</i> , [<i>the</i>] <i>man bi sêes stade uuidô teuuirpit</i> ‘dann geht es ihm wie dem Salz, das man am Meeresufer weithin zerstreut’	PTK > PTKREL <i>Relativpartikel</i>	—
[<i>uf</i>] <i>huob er die hende</i> ‘hoch hob er die Hände’	AVD > PTKVZ <i>Verbzusatz</i>	PTKVZ <i>abgetr. Verbzusatz</i>
<i>di en salt du</i> [<i>nit</i>] <i>schouwen</i> ‘die sollst du nicht anschauen’	PI > PTKNEG <i>Negationspartikel (aus Pronominalsubstantiv)</i>	PTKNEG <i>Negationspartikel</i>

Anmerkungen:

- Anders als im STTS werden in HiTS Verbpartikeln und Basisverben immer getrennt annotiert, selbst bei Kontaktstellung:

(als er den brief) anesach | ane/PTKVZ sach/VVFIN ‘als er den Brief ansah’

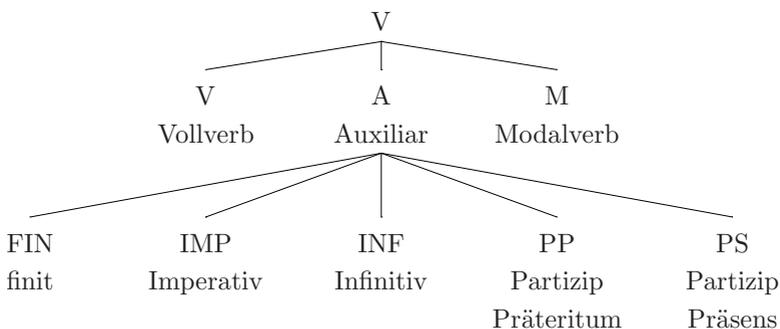
- Zum STTS-Tag PTKZU für „zu vor Infinitiv“ gibt es in HiTS keine Entsprechung. In den älteren Sprachstufen gibt es eine ähnlich aussehende Vorgängerkonstruktion,

in der der Infinitiv allerdings flektiert ist (Dativ) und *zu* daher als Präposition zu analysieren ist:

ez ist wol ze/AP > APPR *lobenne/VVINF* > NA 'es ist sehr zu loben'

daz si ime ze/AP > APPR *helfen/VVINF* > NA *chomen* 'dass sie ihm helfen kommen/zu Hilfe kommen'

10 Verben



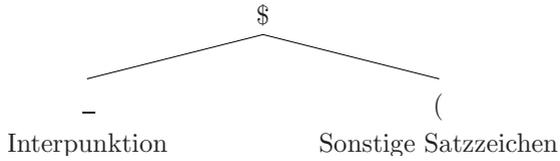
Beispiel	HiTS	STTS
Vollverben: VV		
<i>er [sprach] 'er sprach'</i>	VV > VVFIN <i>Vollverb, finit</i>	VVFIN <i>finites Verb, voll</i>
<i>[gloubet] mir 'glaubt mir'</i>	VV > VVIMP <i>Vollverb, Imperativ</i>	VVIMP <i>Imperativ, voll</i>
<i>si begunden in [fragen] 'sie begannen ihn zu fragen'</i>	VV > VVINF <i>Vollverb, Infinitiv</i>	VVINF <i>Infinitiv, voll</i>
<i>der winter was [vergan] 'der Winter war vergangen'</i>	VV > VVPP <i>Vollverb, Partizip Präteritum, im Verbalkomplex</i>	VVPP <i>Partizip Perfekt, voll</i>
<i>dannen wirt der lip [sennende] 'deshalb beginnt der Leib, sich zu sehnen'; ...unt wir da sin [mendent] 'und wir uns dort freuen mögen (= uns freudig sein)'</i>	VV > VVPS <i>Vollverb, Partizip Präsens, im Verbal-komplex</i>	ADJD <i>adv. oder präd. Adj.</i>
Außerdem:		
<i>mit [suften] und mit [weinen] 'mit Seufzen und mit Weinen'</i>	VVINF > NA <i>Infinitiv, substantiviert</i>	NN <i>normales Nomen</i>
<i>uf ainem [getouweten] chle 'auf einem betauten Klee'</i>	VVPP > ADJA etc. <i>Partizip Präteritum, adjektivisch</i>	ADJA etc.
<i>daz [brinnent] ole 'das brennende Öl'</i>	VVPS > ADJA etc. <i>Partizip Präsens, adjektivisch</i>	ADJA etc.

Beispiel	HiTS	STTS
Auxiliare und Modalverben: VA/VM		
<i>der winter [was] vergan</i> 'der Winter war vergan- gen'	VA > VAFIN <i>Auxiliar, finit</i>	VAFIN <i>finites Verb, aux</i>
...	VA > VAFIN etc.	VAFIN etc.
<i>er [scol] unser helfare</i> <i>wesen</i> 'er soll unser Hel- fer sein'	VM > VMFIN <i>Modalverb, finit</i>	VMFIN <i>finites Verb, modal</i>
...	VM > VMINF etc.	VMINF etc.

11 Verschiedenes

Beispiel	HiTS	STTS
[o we] 'o weh'	ITJ > ITJ <i>Interjektion</i>	ITJ <i>Interjektion</i>
<i>Dú súdoze dínoro</i> [gratię] <i>ist bézzera</i> <i>dánne dú scárfe déro</i> [legis] 'die Süße deiner Gnade (gratia) ist besser als die Schärfe des Gesetzes (legis)'	FM > FM <i>Fremdsprachliches Material</i>	FM <i>Fremdspr. Material</i>

12 Interpunktion



Beispiel	HiTS	STTS
[. : , / ? !]	$\$__ > \$__$ <i>originale Interpunktion</i>	\$. \$. <i>Komma; satzbeendende Interpunktion</i>
[() ' " "]	$\$(> \$($ <i>sonstige Satzzeichen</i>	\$(<i>sonstige Satzzeichen</i>

Appendix III: Alphabetische Auflistung aller Beleg-Tags

Tag	Beschreibung
ADJA	<i>Adjektiv, attributiv, vorangestellt</i>
ADJD	<i>Adjektiv, prädikativ</i>
ADJN	<i>Adjektiv, attributiv, nachgestellt</i>
ADJS	<i>Adjektiv, substituierend</i>
APPO	<i>Postposition</i>
APPR	<i>Präposition</i>
AVD	<i>Adverb</i>
AVG	<i>Relativadverb, generalisierend</i>
AVNEG	<i>Adverb, negativ</i>
AVW	<i>Adverb, interrogativ</i>
CARDA	<i>Kardinalzahl, attributiv, vorangestellt</i>
CARDD	<i>Kardinalzahl, prädikativ</i>
CARDN	<i>Kardinalzahl, attributiv, nachgestellt</i>
CARDS	<i>Kardinalzahl, substituierend</i>
DDA	<i>Determinativ, definit, attributiv, vorangestellt (ad.)</i>
DDART	<i>Determinativ, definit, artikelartig (mhd.)</i>
DDD	<i>Determinativ, definit/demonstrativ, prädikativ</i>
DDN	<i>Determinativ, definit/demonstrativ, attributiv, nachgestellt</i>
DDS	<i>Determinativ, definit/demonstrativ, substituierend</i>
DGA	<i>Determinativ, generalisierend, attributiv, vorangestellt</i>
DGD	<i>Determinativ, generalisierend, prädikativ</i>
DGN	<i>Determinativ, generalisierend, attributiv, nachgestellt</i>
DGS	<i>Determinativ, generalisierend, substituierend</i>
DIA	<i>Determinativ, indefinit, attributiv, vorangestellt (ad.)</i>
DIART	<i>Determinativ, indefinit, artikelartig (mhd.)</i>
DID	<i>Determinativ, indefinit, prädikativ</i>
DIN	<i>Determinativ, indefinit, attributiv, nachgestellt</i>
DIS	<i>Determinativ, indefinit, substituierend</i>
DNEGA	<i>Determinativ, negativ, attributiv, vorangestellt</i>
DNEG	<i>Determinativ, negativ, prädikativ</i>
DNEGN	<i>Determinativ, negativ, attributiv, nachgestellt</i>
DNEGS	<i>Determinativ, negativ, substituierend</i>
DPOSA	<i>Determinativ, possessiv, attributiv, vorangestellt</i>
DPOSD	<i>Determinativ, possessiv, prädikativ</i>
DPOSGEN	<i>Determinativ, personal-possessiv, Genitiv</i>
DPOSN	<i>Determinativ, possessiv, attributiv, nachgestellt</i>
DPOSS	<i>Determinativ, possessiv, substituierend</i>
DRELS	<i>Determinativ, relativisch, substituierend</i>
DWA	<i>Determinativ, interrogativ, attributiv, vorangestellt</i>
DWD	<i>Determinativ, interrogativ, prädikativ</i>
DWN	<i>Determinativ, interrogativ, attributiv, nachgestellt</i>
DWS	<i>Determinativ, interrogativ, substituierend</i>

FM	<i>Fremdsprachliches Material</i>
ITJ	<i>Interjektion</i>
KO*	<i>Konjunktion, neben- oder unterordnend</i>
KOKOM	<i>Vergleichspartikel</i>
KON	<i>Konjunktion, nebenordnend</i>
KOUS	<i>Konjunktion, unterordnend</i>
NA	<i>Nomen appellativum</i>
NE	<i>Eigennamen</i>
PAVAP	<i>Pronominaladverb, präpositionaler Teil</i>
PAVD	<i>Pronominaladverb, pronominaler Teil</i>
PAVG	<i>Pronominaladverb, pronominaler Teil, generalisierend</i>
PAVREL	<i>Pronominaladverb, pronominaler Teil, relativisch</i>
PAVW	<i>Pronominaladverb, pronominaler Teil, interrogativ</i>
PG	<i>Pronomen, generalisierend</i>
PI	<i>Pronomen, indefinit</i>
PNEG	<i>Pronomen, indefinit, negativ</i>
PPER	<i>Pronomen, personal, irreflexiv</i>
PRF	<i>Pronomen, personal, reflexiv</i>
PTKA	<i>Partikel bei Adjektiv oder Adverb</i>
PTKANT	<i>Antwortpartikel</i>
PTKINT	<i>Fragepartikel (ad.)</i>
PTKNEG	<i>Negationspartikel</i>
PTKREL	<i>Relativpartikel</i>
PTKVZ	<i>Verbzusatz</i>
PW	<i>Pronomen, interrogativ</i>
VAFIN	<i>Auxiliar, finit</i>
VAIMP	<i>Auxiliar, Imperativ</i>
VAINF	<i>Auxiliar, Infinitiv</i>
VAPP	<i>Auxiliar, Partizip Präteritum, im Verbalkomplex</i>
VAPS	<i>Auxiliar, Partizip Präsens, im Verbalkomplex</i>
VMFIN	<i>Modalverb, finit</i>
VMIMP	<i>Modalverb, Imperativ</i>
VMINF	<i>Modalverb, Infinitiv</i>
VMPP	<i>Modalverb, Partizip Präteritum, im Verbalkomplex</i>
VMPS	<i>Modalverb, Partizip Präsens, im Verbalkomplex</i>
VVFIN	<i>Vollverb, finit</i>
VVIMP	<i>Vollverb, Imperativ</i>
VVINF	<i>Vollverb, Infinitiv</i>
VVPP	<i>Vollverb, Partizip Präteritum, im Verbalkomplex</i>
VVPS	<i>Vollverb, Partizip Präsens, im Verbalkomplex</i>
\$_	<i>originale Interpunktion</i>
\$(<i>sonstige Satzzeichen</i>