

## **Manuelle Abgleichung bei automatisierter Vorannotation:** Das Tagging grammatischer Kategorien im *Referenzkorpus Altdeutsch*

---

### **Abstract**

Das Referenzkorpus Altdeutsch überführt die grammatischen Informationen der etablierten Textwörterbücher zu den ältesten deutschen Texten in ein Annotationstool, wobei die so automatisiert gewonnenen Daten für jede einzelne Wortform manuell überprüft und angeglichen werden müssen. In diesem Artikel werden verschiedene Fälle vorgestellt, in denen die grammatische Annotation des Referenzkorpus Altdeutsch von den aus den Grammatiken und Glossaren implementierten Angaben abweicht. Hierbei handelt es sich vor allem a) um Alternativdeutungen zu den traditionellen Interpretationen bestimmter Formen, b) um Lücken in der Beschreibung bestimmter Formen, die aus einem ausschließlich historisch ausgerichteten Grammatikverständnis entstehen und c) um Fehler, die aufgrund einer nicht eindeutigen Beschreibung entstehen.

### **1 Einleitung**

Das Referenzkorpus Altdeutsch<sup>1</sup> digitalisiert und annotiert sämtliche Texte der althochdeutschen und altsächsischen Überlieferung (ca. 750 – 1050 n. Chr.). Die Annotation umfasst linguistische und strukturelle Informationen wie auch die Metadaten eines Textes.

Zur Bearbeitung werden die möglichst handschriftentgetreuesten Textausgaben in das Annotationstool Elan<sup>2</sup> überführt und die jeweiligen Textwörterbücher und Glossare<sup>3</sup> implementiert. Die so automatisiert gewonnenen lemmabezogenen grammatischen Informationen müssen für jede einzelne Wortform manuell überprüft und dem Kontext entsprechend angeglichen werden. Die präzisen, umfangreichen und mit ungemein großer Sachkenntnis zusammengetragenen Angaben der zum Großteil noch aus dem 19. Jhd. stammenden Glossare, die für alle größeren altdeutschen Texte verfügbar sind, bieten dem Korpus somit eine wertvolle Grundlage.

Für die Auszeichnung der PoS-Informationen und der morphologischen Werte unterscheidet das Referenzkorpus Altdeutsch zwischen einer lemma- und einer belegbezogenen Kennzeichnung. Während auf der PoS-Lemmaebene (L) grundsätzlich nur die Werte angegeben werden, die im Lemma selbst festgelegt sind – also die Grundmarkierungen der Wortarten bzw. die unveränderlichen, im Lexikon verzeichneten morphologischen Informationen – werden auf der PoS-Belegebene (B) diese Werte spezifiziert. Prinzipiell werden dabei für die Lemmaebene (L) die in den einzelnen Wörterbüchern aufgeführten Informationen zur Wortart übernommen. Lediglich bei wenigen Ausnahmen werden diese angeglichen, z.B. bei abweichenden grammatischen Annahmen des Referenzkorpus Altdeutsch zu den Informationen der Glossare und Wörterbücher (vgl. hierzu Abschnitt 3).

## 2 Manuelle Angleichung bei automatisierter Vorannotation

Zur Differenzierung zwischen Lemma- und Belegebene werden die PoS-Kategorien hinsichtlich ihrer Hauptwortart und ihrer konkreten syntaktischen Verwendungsweise unterschieden.<sup>4</sup>

(1)	Wortform	<i>lag</i>	<i>liggian</i>	<i>liggiandi</i>	<i>liggeandan</i>
	Lemma	<i>liggian</i>	<i>liggian</i>	<i>liggian</i>	<i>liggian</i>
	PoS L	VV	VV	VV	VV
	PoS B	VVFIN	VVINF	VVPS	VVPSA

Bei der morphologischen Beschreibung ist die Kennzeichnung der Flexionsklasse des Belegs dort präziser, wo die Flexionsklasse des Lemmas nicht eindeutig ist, der Beleg aber eine Festlegung erlaubt. So flektiert z.B. das ahd. Lemma *man* ‚Mann‘ ursprünglich konsonantisch, in den Singularformen erscheint kein Flexionsmorphem (NGDA *man-Ø*). Diesem Lemma wird entsprechend für die konsonantische Deklination der Wert C zugewiesen. Daneben erscheinen aber bereits zu althochdeutscher Zeit Formen der maskulinen a-Stämme (G *mannes*, D *manne*) und in diesen Fällen erhält *man* den Wert a für die Deklinationsklasse. Da *man* im Ahd. also sowohl konsonantisch als auch analog zu den a-Stämmen flektiert, erhält das ahd. Lemma die Werte beider Flexionsklassen (Fl(exions)klasse L(emma)). Erst bei eindeutigen Wortformen wird die Angabe der Flexionsklasse präzisiert (Fl(exions)klasse B(eleg)).

(2)	<i>Yrthugis</i>	<i>thar</i>	<i>thoh</i>	<i>éines</i>	<i>man</i>	(O 2,42)
	gedenkst	da	doch	eines	Mann-SG.GEN	
	„(Du) gedenkst da doch eines Mannes“					

Lemma	<i>man</i>
PoS L, B	NA
Fl.klasse L	C,a_Masc
Fl.klasse B	C_Masc

(3)	<i>inti</i>	<i>ther</i>	<i>mannes</i>	<i>sun</i>	<i>uuirdit</i>	<i>giselit</i>	(T 153, 2)
	und	der	Mann-SG.GEN	Sohn	wird	verraten	
	„Und der Menschensohn wird verraten werden“						

Lemma	<i>man</i>
PoS	NA
Fl.klasse L	C,a_Masc
Fl.klasse B	a_Masc

Durch die Differenzierung zwischen Lemma und Beleg bestehen Möglichkeiten von statistischen Abgleichen bestimmter Verteilungen wie z.B. von Vorkommen von Deklinationsklassen eines Subkorpus. Das Lemma *man* ‚Mann, Mensch‘ flektiert ursprünglich konsonantisch und tritt in althochdeutscher Zeit zunehmend stark flektierend nach dem Vorbild der

## Manuelle Abgleichung bei automatisierter Vorannotation

maskulinen a-Stämme auf, weswegen beide Flexionsformen im Wörterbuch verzeichnet sind und entsprechend für die Flexionsklasse des Lemmas (Flexion L) aufgeführt sind.

Im Tatian tritt das Lemma *man* ‚Mann, Mensch‘ 266mal auf. Davon werden 157 Belege sowohl als konsonantisch als auch als a-Stamm beschrieben, da es sich hier um formgleiche Kasus handelt. Die übrigen Belege können aufgrund ihres eindeutigen Flexionsmorphems einer Deklinationsklasse zugewiesen werden, die a-Stämme (70mal) treten dabei fast doppelt so häufig wie die konsonantischen Formen (39mal) auf. Die folgende Tabelle zeigt die Belegzahlen für einzelne Kasusformen von *man* im Tatian.

Lemma / Form	C, a	C	a	gesamt
<i>man</i>	157	39	70	266
Sg. Nom	75	-	-	75
Gen	-	-	57	57
Dat	-	6	13	19
Akk	29	-	-	29
Pl. Nom	-	28	-	28
Gen	19	-	-	19
Akk	-	5	-	5
Dat	34	-	-	34

**Tabelle 1:** Vorkommen von *man* im Tatian nach Flexionsklasse und Form

Neben der Spezifizierung der von den Wörterbüchern dem Lemma zugewiesenen Werte dient die Unterscheidung zwischen Lemma- und Belegebene auch der Beschreibung von Sprachwandlungsprozessen.

Im Ahd. und As. gibt es für den Plural und das Femininum kein Possessivpronomen, in diesen Fällen wird der jeweilige Genitiv des Personalpronoms (Fem. Sg. *ira*, Pl. *iro*) verwendet. Durch die Trennung der PoS-Kennzeichnung in Lemma und Beleg kann so z.B. die kontextbezogene Verwendung von ahd./as. *iro* ‚ihr‘ als nicht flektierendes Possessivpronomen mit dem Tag DPOS auf der Ebene PoS B(eleg) dargestellt werden. Die lemmabezogene formale Bedeutung als Genitiv Plural des Personalpronoms wird mit dem Tag PPER auf der Ebene PoS L(emma) und durch die Angabe der veränderlichen flexivischen Werte auf einer morphologischen Ebene als Genitivform beschrieben.

- (4) *After thiu her uuosc iro fuozzi.* (T 156,1)  
 danach er wusch er-MASC.GEN.PL. FüÙe-PL.ACC.  
 „Danach wusch er ihre FüÙe“ (eigentlich: die FüÙe ihrer = die FüÙe von ihnen)

Lemma	ër
PoS L	PPER
PoS B	DPOS
Flexion	<i>Masc.Gen.Pl.3</i>

Bei der Annotation der meisten Wortformen folgt das Referenzkorpus Altdeutsch weitestgehend den Angaben der Wörterbücher, d.h. die im Annotationstool für die jeweilige Ebene angegebenen Werte entsprechen den Angaben in den Glossaren und auf der Belegebene werden von den Annotatoren ggf. Entscheidungen zugunsten des einen oder anderen Wertes getroffen.

Überdies erscheinen aber Fälle, bei denen das Referenzkorpus Altdeutsch von den Beschreibungen der Textwörterbücher und Grammatiken<sup>5</sup> abweicht.

Zum einen kann es sein, dass das reichhaltige Wissen der Wörterbücher und der Grammatiken nicht mehr dem aktuellen Vorgehen bei der Beschreibung von grammatischen Phänomenen entspricht, und deswegen eine alternative Kennzeichnung entwickelt werden muss. Teilweise treten systematisch Lücken in der Beschreibung auf, zumeist, weil entweder das Wissen von den Verfassern vorausgesetzt wurde oder weil bestimmte Unterschiede nicht erfasst wurden. In einigen (wenn auch wenigen) Fällen führt die durchgehende Systematisierung im Sinne einer Vereinfachung von verschiedenen Beschreibungsmöglichkeiten bei durchweg komplexer Beleglage zu fehlerhaften Vorannotationen.

Mit anderen Worten: In welchen Fällen stößt die Nutzung der altdeutschen Wörterbücher und Grammatiken trotz deren Kenntnisreichtums, deren Detailgenauigkeit und Fülle an Belegen für ein computerbasiertes, modernen Ansprüchen genügendes Korpus an ihre Grenzen?

### 3 Konflikte zwischen traditioneller Grammatikbeschreibung und moderner Theorie: Alternativentscheidungen

Bereits in den ältesten Texten des Deutschen treten mehrgliedrige Konjunktionen und Adverbien auf.

Die Instrumentalis-Form *thiu* des Demonstrativpronomens ahd. *ther*, as. *the* ‚der‘ erscheint oftmals in Verbindung mit einer Adposition oder einem Adverb. Diese Zusammensetzungen mit *thiu* haben häufig zwei Lesarten und treten dann sowohl als Adverb bzw. Präpositionaladverb als auch als Subjunktion auf.

- (5) *Than mag man dragan* ADV[*afar thiu*] *līhtlicora* *lid*  
dann kann man-PI auftragen danach leicht-COMP Wein  
(H 2054)

„dann soll man danach den leichteren Wein auftragen“

## Manuelle Abgleichung bei automatisierter Vorannotation

- (6) SUBJ[*Aftar thiu*] *sie inan erhiengun, intfiengun sin giuuati (...)*  
 nachdem sie ihn erhängen ergriffen sein Gewand  
 (T 203,1)  
 „Nachdem sie ihn erhängt hatten, ergriffen sie sein Gewand (...)“  
 vgl. Lat. Tatian:  
*Postquam autem crucifixerunt eum, acceperunt vestimenta eius ...*

Die Bestandteile derartiger Konstruktionen werden in den älteren Wörterbüchern als Einzelformen aufgefasst, so dass nur die Einträge für die einzelnen Bestandteile durch die automatisierte Vorannotation erstellbar sind.

Für *afar thiu* in (1) werden nach der automatisierten Vorannotation, in diesem Fall mit SEHRT (1966), folgende Werte angeführt:

Wortform	<i>afar</i>	<i>thiu</i>				
Lemma	afar	the	the	the	thiu	thiwa
Übersetzung	zeitlich	der	der	der	Magd, Dienerin	Dienerin, Magd
PoS Lemma	AP	DD	DD	DD	NA	NA
PoS Beleg	APPR?	DDA	DDA	DDA	NA	NA
Flexion Lemma					jo_Fem	n_Fem
Flexion Beleg 1					jo_Fem	n_Fem
Flexion Beleg 2		Fem Sg Nom	Masc Neut Sg Ins	Neut Pl Nom Acc	Sg Nom,Acc	Sg Nom

**Tabelle 2:** Vorschläge für as. 'after thiu' nach SEHRT (1966)

Anders erscheint die Vorannotation von *afar thiu* im ahd. Tatian wie in (2) nach dem Glossar von SIEVERS (1892):

Wortform	<i>afar</i>		<i>thiu</i>
Lemma	afar	afar	dër
Übersetzung	hinten, hinter ... her, von hinten, nach hinten; (da)nach, später, entsprechend. m. Gen.: nach. m. Dat.: entlang, über ... hin, hinter ... her, auf, durch, nach, zu; hinsichtlich, entsprechend, zuzolge,	hinten, hinter ... her, von hinten, nach hinten; (da)nach, später, entsprechend. m. Gen.: nach. m. Dat.: entlang, über ... hin, hinter ... her, auf, durch, nach, zu; hinsichtlich, entsprechend, zuzolge,	dieser, der(selbe); diese, die; dieses, das; wer, welcher; welche; was, welches

	gemäß. m. Akk.: hinter, nach, gemäß. subst.: das Hinten	gemäß. m. Akk.: hinter, nach, gemäß. subst.: das Hinten	
PoS Lemma	ADV	ADV   AP	DD
PoS Beleg	ADV	ADV   APPR?	DDA
Flexion Lemma			
Flexion Beleg 1			
Flexion Beleg 2			_Sg_Gen

**Tabelle 3:** Vorschläge für ahd. 'after thiu' nach SIEVERS (1892)

Beide Beschreibungen entsprechen allerdings nicht unserem Verständnis derartiger Konstruktionen, wonach *afar thiu* – und ähnliche Verbindungen – als eine Worteinheit aufgefasst und entsprechend annotiert werden.

Einen Eintrag aber über die möglichen Verbindungen von *thiu* in der Funktion eines Adverbs oder einer Subjunktion sucht man in den meisten Glossaren zu den altdeutschen Texten häufig vergeblich. Werden Verbindungen von einer Adposition mit *thiu* angeführt, erscheinen diese unvollständig und unsystematisch. Zudem wird meist lediglich eine Bedeutungsvariante genannt, jedoch wird dieser keine eigenständige grammatische Funktion zugewiesen.

Dabei tritt gerade *thiu* sehr häufig in verschiedenen Verbindungen auf, in denen eine demonstrative Funktion, wie die älteren Glossare *thiu* auch in diesen Fügungen wenigstens implizit geben, kaum noch zugewiesen werden kann und es sich oft auch formal nicht um einen Instrumentalis handeln kann.<sup>6</sup> In diesen Fällen werden Verbindungen aus *thiu* und Adposition vom Referenzkorpus Altdeutsch als lexikalisiert angesehen.

Für das vorliegende Korpus werden sämtliche Einzelwörter einer mehrteiligen Subjunktion oder eines Adverbs zusammengeführt und als eine Wortform annotiert.

#### 4 Eine Frage der Perspektive: Lücken in der Beschreibung

Ein herausragendes Anliegen des Referenzkorpus Altdeutsch ist es, ein möglichst genaues Bild vom tatsächlichen Sprachstand der ältesten Überlieferung des Deutschen zu geben. Dieses Vorhaben erscheint trivial, erweist sich aber bei der zeitlichen und räumlichen Streuung der Texte und vor dem Hintergrund der grammatischen Beschreibungen der relevanten Glossare, auf die sich das Referenzkorpus Altdeutsch stützt, als relativ vielschichtige Aufgabe.

Die altdeutschen Sprachen zeigen eine relativ große Anzahl an substantivischen Deklinationsklassen auf, die jedoch zahlreichen Wandelerscheinungen unterliegen: Neben Kasusschwund (Instrumentalis) und formalem Kasuszusammenfall (z.B. Genitiv und Dativ Singular bei den o-Feminina) fallen verschiedene Deklinationsklassen zusammen oder gehen ineinander über, und es besteht bei vielen Substantiven eine Tendenz zum Genuswechsel.

Die Glossare greifen diese Heterogenität in Teilen auf, indem sie zwar von den ursprünglichen germanischen Deklinationsklassen ausgehen, jedoch abweichende Genera und Stammbildungen verzeichnen. Allerdings beschränken sich die Wörterbücher auf die Angaben von starker und schwacher Deklination, ohne die entsprechenden Stämme anzuführen. Hier ergänzen für die automatisierte Vorannotation die Grammatiken (BRAUNE 2004, GALLÉE 1993) die einzelnen Stammklassen. Diese zeigen grundsätzlich eine diachrone Anlage in der Einteilung der Deklinationsklassen ausgehend von der germanischen Stammbildung. So wird z.B. das Lemma *sunta* ‚Sünde‘ entsprechend dem ursprünglichen Thema als *jo*-Stamm angeführt.<sup>7</sup>

Die Belege zeigen aber, dass neben den *j*-haltigen Formen auch Flexionsformen auftreten, die den *o*-Stämmen entsprechen, dass also das Paradigma von *sunta* sich den *o*-Stämmen formal angleicht und damit die Voraussetzungen für einen Klassenwechsel (oder Klassenzusammenfall) gegeben sind.<sup>8</sup>

- (7) *dir uuirdu ih pigihitik allero minero suntiono*  
dir werde ich „beichtig“ all-GEN.PL. mein-GEN.PL. Sünde-GEN.PL.  
(AB,1)

- (8) *dir uuirdu ih pigihitik allero mindero suntiono*  
(BG1,1)  
„Ich beichte dir alle meine Sünden“

Im Vergleich dazu die Genitiv-Formen ursprünglicher *o*-Stämme:

- (9) [Fortsetzung von (7)]  
*meinsuartio enti lugino, kiridono enti unrehteru*  
Meineide-GEN.PL und Lügen-GEN.PL. Begierde-GEN.PL. und unrechte-GEN.SG.  
i-Stamm a-Stamm o-Stamm  
  
*fizusheiti, huorono*  
Hinterlist-GEN.SG. Ehebruch-GEN.PL.  
i-Stamm o-Stamm  
(AB,5)

[ich beichte:]

„Meineide und Lügen, Begierden und unrechte Hinterlist, Ehebrüche“

Wenn wir also den Wandel bei den Deklinationen der Substantive abbilden wollen, ist es notwendig, die synchron im Alrdeutschen tatsächlich auftretenden Formen zu kennzeichnen. Demnach werden Formen, die noch Reflexe des ursprünglichen Themas zeigen, als *jo*-Feminina gekennzeichnet, während abweichend von den Grammatiken alle anderen starken Formen auf der Belegebene als *o*-Feminina getaggt werden.

- |      |   |                 |         |
|------|---|-----------------|---------|
| (10) | <i>dir uuirdu ih pigihtik allero minero</i> | <i>suntiono</i> | (AB,1)  |
|      | Lemma                                       | sunta           |         |
|      | Flexion L                                   | jo_Fem          |         |
|      | Flexion B                                   | jo_Fem          |         |
|      |   |                 |         |
| (11) | <i>dir uuirdu ih pigihtik allero minero</i> | <i>suntono</i>  | (BGa,1) |
|      | Lemma                                       | sunta           |         |
|      | Flexion L                                   | jo_Fem          |         |
|      | Flexion B                                   | o_Fem           |         |

## 5 Grammatisches Gemenge: Inadäquate Beschreibungen aufgrund der automatisierten Vorannotation

Trotz sorgfältiger Prüfung bleibt es nicht aus, dass bei der automatisierten Vorannotation Fehler entstehen. Diese zeigen sich vor allem bei Formen mit mehreren grammatischen Funktionen oder bei unterschiedlichen Lemmata, die formgleich sind.

Das ad. Lemma *filu* ‚viel, sehr‘ geht auf ein u-Neutrum *filu* ‚Vieles, Vielheit‘ zurück, welches in den altdutschen Texten zum einen in Argumentposition als indeklinables Substantiv, zumeist mit abhängigem Genitiv (12), und zum anderen als graduelles Adverb (13) auftritt.

- (12) *so sculun gi undar iuuua fiund faren, undar filu theodo*  
 so sollt ihr unter eure Feinde gehen unter ‚Viel‘-ACC.SG. Volk-GEN.PL.  
 (H 1875)  
 „So sollt ihr unter eure Feinde gehen, unter viele Völker (=unter eine Vielheit der Völker)“
- (13) *thaz lioht ist filu war thing* (O II 2, 13)  
 das Licht ist viel wahr Ding  
 „Das Licht ist eine sehr wirkliche Sache“

In den ahd. Textwörterbüchern wird *filu* unterschiedlich beschrieben, in den meisten Fällen werden beide Varianten mit den entsprechenden Belegstellen genannt. Ausgehend von SPLETT (1993) für das Gesamthochdeutsche wurde es jedoch als Adverb vorannotiert, auch wenn dies nicht immer der tatsächlichen Lesart entspricht. In den Fällen, in denen es als Substantiv auftritt, muss deswegen die Vorannotation korrigiert werden.

Eine Analyse als Adjektiv wie in GALLÉE (1993:228) für das Altsächsische ist zumindest fraglich, jedoch in die Vorannotation aufgenommen. Für den Heliand annotiert das Referenzkorpus Altdeutsch in Übereinstimmung mit SEHRT (1966:180) *filu* als Adverb oder Substantiv, so dass auch hier von der Vorannotation abgewichen wird.

Eine andere Problematik birgt die Vorannotation der Homonyme ahd. *noh* a) ADV ‚noch, auch, außerdem‘ und b) KONJ ‚und nicht, auch nicht‘. Aufgrund der Formgleichheit werden beide Lemmata in der Vorannotation zusammengefasst, in einigen Texten nur in den Fällen, bei denen die Lesart als Konjunktion möglich ist, in einigen Texten aber in allen Fällen. Dies ist von den Angaben im jeweiligen Textwörterbuch abhängig, das der Vorannotation zugrunde lag. Im Zweifelsfall muss der Annotator dann aufgrund des Kontextes entscheiden.

### 6 Fazit

Die Erfahrungen der Arbeiten des Referenzkorpus Altdeutsch zeigen, dass die zum Teil schon sehr alten Glossare zu den ältesten Texten des Deutschen und die Grammatiken für ein digital aufbereitetes, tiefenannotiertes Korpus von großem Nutzen sind. Die zunächst zeitaufwändige Methode, sämtliche Textwörterbücher jeweils zu den einzelnen Überlieferungen einzulesen und dann die daraus gewonnenen Informationen manuell im Annotationstool zu bearbeiten, erweist sich als ausgesprochen wirkungsvoll. Da so gut wie jede Wortform der altdeutschen Texte in den entsprechenden Glossaren detailliert verzeichnet ist, ergeben sich bei der lemmabezogenen Vorannotation nur dann Abweichungen, wenn die Bearbeiter/innen des Referenzkorpus Altdeutsch grundsätzlich eine Form anders als vorgeschlagen deuten. Diese Fälle können wie gezeigt systematisch behandelt und begründet werden.

---

<sup>1</sup> Das Forschungsprojekt Referenz Korpus Altdeutsch <http://www.deutschdiachrondigital.de> wird von der Deutschen Forschungsgemeinschaft (DFG) gefördert und ist am Lehrstuhl für Sprachgeschichte an der Humboldt-Universität zu Berlin, am Institut für Empirische Sprachwissenschaft an der Goethe-Universität Frankfurt a.M. und am Lehrstuhl für Indogermanistik an der Schiller-Universität Jena angesiedelt.

<sup>2</sup> <http://www.lat-mpi.eu/tools/elan/>

<sup>3</sup> HEFFNER (1961), HENCH (1890), HENCH (1893), KELLE (1881), SEHRT (1955), SEHRT (1966), SIEVERS (1874), SIEVERS (1892) und WADSTEIN (1899). Bisher standen keine digitalisierten Versionen zur Verfügung.

<sup>4</sup> Zur Kennzeichnung der PoS-Kategorien verwendet das Referenzkorpus Altdeutsch ein an das STTS (Stuttgart Tübingen Tagset, SCHILLER et al., 1999) angelehntes Tagset, welches dieses für die Annotation historischer Daten modifiziert. Ein Ausschnitt des DDDTS (Deutsch Diachron Digital Tagset) mit den in diesem Aufsatz aufgeführten Tags befindet sich im Anhang.

<sup>5</sup> BRAUNE/REIFFENSTEIN (2004), GALLÉE (1993).

<sup>6</sup> Dies erscheint allenfalls noch in Verbindung mit der Präposition as. *mid*, ahd. *mit*, die u.a. den Instrumentalis fordert, möglich; in der Bedeutung eines Pronominaladverbs *damit*. Vgl.: *endi fōdie is hundos mid thiu* ‚und füttere die Hunde damit‘ (Hel 3017), wobei *thiu* sich hier auf das vorerwähnte Neutrum *brōd* ‚Brot‘ bezieht und damit formal einem substituierenden Demonstrativpronomen entspricht.

---

<sup>7</sup> Das germ. Substantiv besteht ursprünglich aus drei Morphemen: Wurzel, Stammbildungselement (Thema) und Flexionsendung. Die Wurzel trägt die inhaltliche Bedeutung und ist zunächst eine grammatisch neutrale Abstraktion, welche normalerweise nicht isoliert auftritt. Aus der Wurzel werden durch Stammbildungselemente Lexeme abgeleitet, welche in der Regel der Wurzel eine grammatische Kategorie zuweisen. Wurzel und Stammbildungselement bilden den Wortstamm, welcher nicht alleine auftreten kann, sondern eine Flexionsendung erhalten muss. Die Einteilung in Stammklassen erfolgt für das Ahd. aufgrund des ursprünglichen Themas, auch wenn dieses häufig nicht mehr sichtbar ist. Allerdings zeigen sich in vielen Formen noch die Reflexe der ursprünglichen Stammbildungselemente, die dann üblicherweise als Teile der Endung interpretiert werden. So zeigt Nom.Sg. *suntea* ‚Sünde‘ mit dem *-e-* noch einen Rest des Themabestandteils *-j-*, während in der Form Nom.Sg. *sunta* dieses geschwunden (oder vollkommen verschmolzen) ist.

<sup>8</sup> Nach dieser Auffassung werden die ursprünglichen Stammbildungssuffixe synchron Althochdeutsch und Altniederdeutsch als Teil des Flexionsmorphems betrachtet; die Umlaute werden spätestens für den Übergang zum Mittelhochdeutschen als nicht mehr phonologisch determiniert angesehen. Die Annahme der Reanalyse des Themas als Flexiv und der Lexikalisierung des Umlauts bildet eine der Voraussetzungen für den möglichen Zusammenfall von substantivischen Deklinationsklassen im Deutschen. Beim Tagging der altdeutschen Daten wird von dieser Annahme ausgegangen, um morphologische Schwankungen präzise darstellen zu können.

## Anhang: DDDTS (Deutsch Diachron Digital Tagset)

Ausschnitt

PoS Beleg	PoS Lemma	Beschreibung
DPOS	DPOS	Determinativ, possessiv
DPOS	PPER	Sonderfall: Genitiv des PPER als DPOS
NA	NA	Nomen, Appellativ
PPER	PPER	Pronomen, personal, irreflexiv
VVFIN	VV	finites Verb, Vollverb
VVINFINF	VV	Infinitiv, Vollverb
VVPS	VV	Partizip Präsens, Vollverb, im Verbalkomplex
VVPSA	VV	Partizip Präsens, Vollverb, attribuerend

## Literatur

### Textausgaben

- AB Altbairische Beichte. In: E. Steinmeyer (Hg.). (1971). Die kleineren althochdeutschen Denkmäler. Dublin, Zürich: Weidmann, 309. (Nachdruck von 1916).
- BGa Bairisches Gebet A. In: E. Steinmeyer (Hg.). (1971). Die kleineren althochdeutschen Denkmäler. Dublin, Zürich: Weidmann, 310-314. (Nachdruck von 1916).
- H Heliand. (1984). Hg. Burkhard Taeger, Tübingen: Niemeyer. 9. Aufl.
- O Otfriids Evangelienbuch. (1973). Hg. Oskar Erdmann, Tübingen: Niemeyer. 6. Aufl.
- T Tatian. (1966). Hg. Eduard Sievers. Paderborn: Schöningh. (Nachdruck der zweiten Ausgabe von 1892).

### Grammatiken und Wörterbücher

- Braune, W./Reiffenstein, I. (2004). Althochdeutsche Grammatik. Tübingen: Niemeyer. 15. Aufl.
- Gallée, J. H. (1993). Altsächsische Grammatik. Tübingen: Niemeyer. 3. Aufl.
- Heffner, R.-M. S. (1961). A Word-Index to the Texts of Steinmeyer. Die kleineren althochdeutschen Sprachdenkmäler. Madison: The University of Wisconsin Press.
- Hench, G. A. (1890). The Monsee Fragments. Strassburg: Trübner.
- Hench, G. A. (1893). Der althochdeutsche Isidor. Strassburg: Trübner.
- Kelle, J. (1881). Glossar der Sprache Otfriids. Regensburg: Manz.
- Sehr, E. (1955). Notker-Wortschatz. Halle: Niemeyer.
- Sehr, E. (1966). Vollständiges Wörterbuch zum Heliand und zur altsächsischen Genesis. Göttingen: Vandenhoeck & Ruprecht.

Sievers, E. (1874). Die Murbacher Hymnen. Halle: Buchhandlung des Waisenhauses.

Sievers, E. (1892). Tatian. Lateinisch und althochdeutsch mit ausführlichem Glossar. 2<sup>nd</sup> edition.  
Paderborn: Schöningh.

Splett, J. (1993). Althochdeutsches Wörterbuch. Berlin: de Gruyter.

Wadstein, E. (1899). Kleinere altsächsische Sprachdenkmäler. Norden/Leipzig: Soltau.