



INL SCHATKAMER VAN
DE NEDERLANDSE TAAL

Het lemma in het GiGaNT-lexicon

Marijke Mooijaart

INL Working Papers - Taalbank Nederlands 4

Leiden, 2012

Inhoudsopgave

1. Ter inleiding.....	3
1.1 Het lemma in GiGaNT	3
1.2 Het lemma en de andere entry-elementen	3
1.3 Het lemma voor de gebruiker: lexicon als tool en thesaurus	4
1.4 Deze nota.....	5
2. De vorm van het lemma	6
2.1 Canonieke vorm	6
2.2 Spellingvorm	6
2.3 Klank- en woordvorm (zie verder hoofdstuk 3)	6
2.4 De lemmagrenzen (zie verder hoofdstuk 4)	7
2.5 De weergave van affixen in het lemma (zie verder hs. 5)	7
2.6 De weergave van vormen in de paradigma's van voornaamwoorden en werkwoorden (zie verder hs. 6).....	8
2.7 Vreemde woorden (leenwoorden met uitheemse kenmerken) (zie verder hs. 7).....	8
2.8. Eigennamen (zie verder hs. 8)	8
3. Het lemma en de daarmee aangeduide varianten	8
3.1 Gewone tegenover gemarkeerde varianten	9
3.2 Variatie in woordfamilies	9
3.3 Desambiguering	10
3.4 Behoud van de oorspronkelijke woordstructuur.....	10
3.5 Splitsing en samenvatting	11
3.6 Extra lemma's.....	12
4. De lemmagrenzen: splitsen en verbinden van vormen; onvolledige vormen	12
4.1 Samenkoppelingen: het moderne lemma is één woord.....	12
4.2 Vaste verbindingen: meerdere moderne lemma's.....	13
4.3 Clisivormen: meerdere moderne lemma's	14
4.4 Meerwoordlemma's (multi-word units)	14
4.5 Afkortingen.....	14
4.6. Foutieve vormen.....	15
5. De weergave van affixen	15
5.1 Handhaving van oorspronkelijke affixen.....	15
5.2 Standaardvormen.....	16
6. De weergave van vormen in het paradigma van voornaamwoorden en werkwoorden.....	16
6.1. Persoonlijk voornaamwoord.....	17
6.2. Bezittelijk voornaamwoord	18
6.3. Aanwijzend voornaamwoord	18
6.4. Wederkerend voornaamwoord.....	18
6.5. Wederkerig voornaamwoord	18
6.6. Onbepaalde voornaamwoorden	18
6.7 Overgangsvormen	19
7. Vreemde woorden (leenwoorden met uitheemse kenmerken)	19
7.1 Aansluiting bij de oorspronkelijke vorm	19
7.2 Meervoudige ontleningen	19
7.3 Verouderde vreemde spellingen.....	20
8. Eigennamen	20
8.1 Relatie tot soortnamen.....	20
8.2 Extra informatie.....	20
8.3 De werkwijze per categorie	21
8.3.1 Inheemse (Germaanse) voornamen.....	21
8.3.2. Uitheemse voornamen en andere persoonsnamen	21

8.3.3. Inheemse toenames die herkenbaar bij de appellatieven aansluiten.....	22
8.3.4. Moderne namen (familienamen, namen van zaken)	22
8.3.5. Inheemse toponiemen die herkenbaar bij de appellatieven aansluiten (o.a. perceelnamen, plaatsnamen)	23
8.3.6. Uitheemse toponiemen en toponiemen met uitheemse elementen	23
9. Detailproblemen en voorbeelden bij 3 t/m 5.....	24
Literatuur	35

1. Ter inleiding

1.1 Het lemma in GiGaNT

In GiGaNT (Groot Geïntegreerd Lexicon van de Nederlandse Taal), het computationele lexicon van historisch en modern Nederlands dat door de Taalbank van het INL wordt ontwikkeld, worden alle woorden van het Nederlands, vanaf de zesde eeuw tot nu, opgeslagen.¹

Elk item (*entry*) in GiGaNT bestaat uit een *type-lemma-PoS*-reeks: een geattesteerde woordvorm (*type*), een modern lemma en een aanduiding van de woordsoort (Part of Speech, PoS)². Aan het lemma is een ID gehecht (in het voorbeeld fictief): honech – HONING (ID 12345) – noun. Deze reeksen vormen de kern van GiGaNT; overige informatie zal modulair toegevoegd worden.

Elk woord heeft zo een eigen identiteitsaanduiding, wat een ideaal uitgangspunt biedt voor het opslaan en gebruiken van informatie. Onder *woord* verstaan wij hier ‘lexeem’, de abstracte woordvorm die afziet van grammaticale eigenschappen. Het GiGaNT-lemma representeert dat woord; het is het herkenbare etiket ervoor.

De vaststelling van de lemmavorm volgt vooraf vastgestelde, controleerbare principes. Dit bevordert de voorspelbaarheid en doorzichtigheid van het lemmasysteem voor bewerkers en gebruikers. Zo kan bovendien ook in de toekomst nieuw materiaal op dezelfde manier worden gelemmatiseerd en in GiGaNT worden opgeslagen.

Het uitgangspunt is een vaste identiteit voor een woord gedurende de hele taalgeschiedenis, en dus één hoofdlemma (voortaan “lemma”) per entry. Door voor GiGaNT maar één lemmavorm per woord vast te stellen, willen we het de gebruiker gemakkelijk maken om alle typen varianten van dat woord te verzamelen. Zo ontstaat er ook geen wildgroei aan lemmavarianten. Men kan het lemma rechtstreeks bereiken (bijv. het werkwoord in de infinitiefvorm *bakken*), maar ook vanuit geattesteerde woordvormen (bijv. een vervoegde vorm zoals *bakte*) en vanuit variërende vormen van diverse aard (bijv. de spellingvariant *backen*, de historische preteritumvorm *bieck*).

Slechts in enkele gevallen kan men niet met één lemma volstaan, bijvoorbeeld als een woord zich splitst of als verschillende woorden samenvallen in de loop der tijd. In deze gevallen worden meerdere lemmavormen toegekend. Zie 3.6, Extra lemma’s.

De hoofdvragen in deze nota zijn:

- a. welke vorm krijgt een lemma;
- b. volgens welke criteria worden bepaalde vormen en varianten onder één bepaald lemma ondergebracht.

1.2 Het lemma en de andere entry-elementen

Lemma + woordvorm in context

Geen enkel lemma bestaat geïsoleerd in GiGaNT. Het is altijd gebaseerd op en verbonden met een of meerdere woordvormen in bestaand taalmateriaal, uit de INL-tekstcorpora en woordenboekcitaten, uit te breiden met ander materiaal. Deze woordvormen zijn vanuit het lemma op te roepen, wat mogelijk is doordat ze in de context verrijkt zijn met lemma-informatie.

Voorbeeld: lemma HONING

¹ Voor een beschrijving van GiGaNT zie Ruitenbergh, De Does en Depuydt (2010).

² Op het ID-nummer wordt hier verder niet ingegaan. Voor de PoS-toekenning zie Ruitenbergh en Van pellicom i.s.m. De Does en Depuydt (2012).

- Jeghen die coude maghe nat nemt **onech** [ID 12345 HONING noun acc.sg.] *ende* tempert dat met waremen watre, Maerlant, *Der Naturen Bloeme* (1267)
- Eene kruycke **honichs** [ID 12345 HONING noun gen.sg.], *Statenbijbel*, 1 Kon. 14, 3 [ed. 1688] (WNT).
- Smaak, geur, kleur en consistentie (hardheid of loperigheid) van **honing** [ID 12345 HONING noun sg.] variëren naar gelang van de soort bloemen waaruit de nectar is gewonnen. (2012); *Wikipedia*.

Om het lexicon te vullen moeten wij dus aan de verschillende spelling- en vormvarianten van een woord, ontstaan door dialectische en chronologische ontwikkelingen en door verschillende spellingconventies één moderne lemmavorm toekennen.

Lemma + taalkundige informatie

Aan het lemma wordt, behalve de woordsoort, ook andere taalkundige informatie gehecht, zoals paradigmavormen, morfologische analyse en semantische informatie. Hiervoor verwijzen we naar Ruitenbergh en Van pellicom (2012) en Karsdorp (2011).

Lemma + overige informatie

Aan het lemma in het lexicon kunnen diverse categorieën extra informatie worden gekoppeld, zowel talige als niet-talige. Dit kunnen zijn:

- bij homoniemen van dezelfde woordsoort: een glosse (korte betekenisomschrijving)
VORST 'koning', VORST 'vrieskou'
- bij plaatsnamen en persoonsnamen: de referent aangeduid door de moderne naam
MONSTERHOEK – Monnikenpolder
- bij korte vormen van voornamen: de volledige vorm
BETJE – Elisabeth

De lemma's worden verder gelinkt aan woordenboeken en naslagwerken, zoals de vier historische woordenboeken (ONW, VMNW, MNW en WNT) en het woordenboek van het moderne Nederlands (ANW) van het INL. Daarnaast is koppeling aan andere INL-bronnen en externe bronnen mogelijk, zoals de etymologiebank, taalatlanten en afzonderlijke woordstudies. Ook koppeling aan encyclopedische informatie is mogelijk.

GiGANT zal zich in eerste instantie beperken tot de noodzakelijke informatie genoemd onder 1.1.

1.3 Het lemma voor de gebruiker: lexicon als tool en thesaurus

Het GiGANT-lexicon functioneert als tool (of reeks tools) én als thesaurus (zie ook Ruitenbergh en Van pellicom (2012), 6-7. De *tool*functie houdt wat het lemmadeel betreft o.a. de volgende mogelijkheden in:

- zoeken in historisch tekstmateriaal met een modern lemma als zoekterm;
- verrijking van historisch en modern tekstmateriaal met standaardlemmavormen en daaraan gekoppelde informatie;
- ondersteuning van OCR bij productie en correctie;
- opsporen van tot nu toe onbeschreven woorden, zowel oude als recente (datamining en datasuppletie);
- koppeling van de online woordenboeken van het INL aan elkaar en aan andere woordenboeken;
- koppeling van tekstmateriaal aan de online woordenboeken van het INL.

Als *thesaurus* is GiGANT een databank van velerlei gegevens over de Nederlandse woordenschat en de verbanden daartussen. De gebruiker kan die vanuit elke informatiecategorie die hij bevat, benaderen en desgewenst inzicht krijgen in alle relaties die erin opgesloten liggen. Zo kan het

invoeren van een woordvorm uit een tekst, of van een bekende oudere vorm leiden naar het lemma, en dat weer naar de informatie in de historische woordenboeken.

Zowel voor diachroon als synchroon onderzoek biedt GiGaNT het benodigde materiaal.

Hoe dit alles technisch wordt gerealiseerd, is niet het onderwerp van deze nota; zie hiervoor Ruitenberg, De Does en Depuydt (2010).

1.4 Deze nota

In dit stuk gaan wij in op de op de vorm van de lemma's (2) en op de selectie van de daaronder vallende varianten. Deze varianten worden gevonden onder diverse soorten klank- en vormvarianten (3) en in variatie in structuur, samengevat in de term lemmagrenzen (4).

Daarna volgen besprekingen van enkele bijzondere lemmatypen en probleemcategorieën, namelijk affixen (5), voornaamwoorden (6), vreemde woorden (7) en eigennamen (8). Hoofdstuk 9 bestaat uit verdere illustraties en uitwerkingen van de voorgaande hoofdstukken.

Verklaring van tekens en typografie

Voorbeelden

Voorbeelden (linkerkolom) komen meestal uit de vier historische woordenboeken van het INL, maar ook uit ander historisch en modern tekstmateriaal. Deze woordvormen kunnen ter wille van de duidelijkheid soms gestandaardiseerd zijn. Hierbij staat het GiGaNT-lemma altijd in kapitaal (rechterkolom); een vette eerste letter geldt als hoofdletter.

Aanduidingen van meerledige lemma's in deze nota:

slash (/) betekent: afzonderlijke lemma's voor afzonderlijke woorden;

backslash (\) duidt op ambiguïteit: meerdere verschillende lemma's zijn mogelijk voor één woordvorm;

dubbele slash (//): meerdere lemma's toegekend bij samenvol of splitsing;

dubbele backslash (\\): er is sprake van gelijkwaardige lemmavarianten;

plus (+) verbindt opeenvolgende lemma's voor een clitische vorm;

underscore (_) verbindt elementen van een meerwoordlemma;

koppelteken (-) in een lemma heeft de gewone waarde als verbindingsteken binnen een woord.

Zie voor een inhoudelijk uitleg hs. 3.6.

2. De vorm van het lemma

2.1 Canonieke vorm

Lemmavorm is altijd de canonieke vorm, dus de vorm zonder flexie (bij znw. enkelvoud, nominatief, bij ww. infinitief, etc.), behalve wanneer de geflecteerde vorm gelexicaliseerd is (*geboren, sleets, tropen*).

2.2 Spellingvorm

De spelling volgt de meest recente officiële spellingregels, op dit moment die van 2005, zoals vastgelegd in de Technische handleiding spelling van de Nederlandse Taalunie. Zo wordt de tussen-*n* ingevoerd, en de oude *-sc-* en *-sch-*spellingens gemoderniseerd, alsmede het gebruik van hoofdletters, koppeltekens e.d.

De spelling van de GiGaNT-lemma's worden formeel getoetst aan genoemde spellingregels om het Keurmerk Spelling te verkrijgen.

Bij verouderde woorden die niet volgens de algemene regels herspeld kunnen worden en die niet in het Groene Boekje of de Grote Van Dale staan maar wel in het WNT, geldt de hoofdvorm van het WNT, aangepast aan de recentste spellingregels. De gedachte hierachter is dat het WNT de standaardvorm heeft vastgesteld van regionaal gevarieerde woordvormen.

Spellingens van vreemde woorden komen bij de vormvariatie aan de orde.

2.3 Klank- en woordvorm (zie verder hoofdstuk 3)

Naast spellingvariatie (verschillende weergaven van eenzelfde klank, zoals in *ik - ick*) is er vormvariatie: vormverschillen binnen eenzelfde woord. Ze kunnen betrekking hebben op de klanken van de morfemen, op verbindingsklanken in gelede woorden en op uitgangen, zoals in resp. *vogel - veugel*, *beminnlijk - beminnelijk* en *nergen - nergens*. In principe kiezen we voor het lemma de meest gebruikelijke moderne vorm, als het woord nog bestaat. Hierbij baseren we ons op ons eigen oordeel en op al of niet expliciete oordelen van de Woordenlijst of Groene Boekje (GB), de Grote Van Dale (VD) en het WNT.

Bij nog bestaande woorden

Lemma is de meest gebruikelijke, moderne vorm van het woord in het Standaardnederlands:

mnl. huus	HUIS
vnnl. moêr	MOEDER
mnl. sterre	STER
mnl. bernen	BRANDEN

Soms wijkt de lemmavorm in het simplex volgens bovengenoemd principe af van de lemmavorm van datzelfde element in een samengesteld woord. Verbindingsklanken en uitgangen volgen de moderne vorm.

mnl. arg	ERG
mnl., nnl arglist	ARGLIST
nnl. hemelblauw	HEMELSBLAUW
mnl. nergnen	NERGENS

Bij verdwenen woorden

Lemma is de etymologische vorm met verwerking van regelmatige klankwijzigingen, gespeld volgens de nu geldende regels.

onl. mūrhuodil ‘muurwacht’	MUURHOEDEL
mnl. saen ‘spoedig’	ZAAN
mnl. stiden ‘stevig worden’	STIJDEN
mnl. vergierer ‘wijnroeier’	VERGIERDER

Bij woorden die zich splitsen of samenvallen in de loop van de geschiedenis

Variatie kan leiden tot het ontstaan van verschillende nieuwe woorden uit eenzelfde woord (splitsing), omgekeerd kunnen verschillende woorden samenvallen tot één woord (samenval). Er worden in het eerste geval aan de vormen van vóór de splitsing twee lemma's toegekend.

Splitsing:

mnl. cruut, nnl. kruit, kruid	KRUID ‘plant, specerij’ //
	KRUIT ‘ontplobbare stof’

Bij samenvallende woorden twee homonieme lemma's gegeven bij verschil in betekenis:

mnl. huus, nnl. huis	HUIS ‘bouwwerk’
mnl. huwesc, huus, nnl. huis	HUIS ‘gezin, geslacht’

Bij volledige samenvallende woorden hangt het van het tijdstip daarvan af.

2.4 De lemmagrenzen (zie verder hoofdstuk 4)

Ook woorden die een vaste verbinding vormen, krijgen elk een eigen lemma; lemma's worden in principe gevormd door één woord. In een beperkt aantal gevallen bestaat het lemma uit meerdere woorden, namelijk als een verbinding gevormd wordt door woorden die zelf geen deel uitmaken van het Nederlands en dus niet zelfstandig in de lexicon worden opgenomen. Dit geldt met name ontleende verbindingen.

mnl. kyrie eleison	KYRIE_ELEISON
nnl. toute à toi	TOUTE_À_TOI

Woorden in inheemse verbindingen die nu niet meer maar eertijds wel een zelfstandige status hadden, vormen geen multi-word lemma maar worden afzonderlijk gelemmatiseerd:

nnl. heg noch steg	HEG / NOCH / STEG
nnl. van kindsbeen af	VAN / KINDSBEEN / AF

Bij clitische verbindingen en samenstellingen met Nederlandse (niet hetzelfde als inheemse) elementen worden de moderne woordgrenzen in acht genomen:

mnl. sconinxs	DE + KONING
mnl. tnacht	TE + NACHT
vnnl. lanckx henen	LANGSHEEN
vnnl. wel bedocht	WELBEDACHT

2.5 De weergave van affixen in het lemma (zie verder hs. 5)

Affixen worden altijd in hun eigen vorm in het lemma weergegeven, dus niet weggelaten of vervangen:

onl. gihōrsamighēt	GEHOORZAMIGHEID (niet: gehoorzaamheid)
mnl. hackelen	HAKKELEN (niet: hakken)
mnl. broekacht	BROEKACHT (niet: broekig)

2.6 De weergave van vormen in de paradigma's van voornaamwoorden en werkwoorden (zie verder hs. 6)

Bij voornaamwoorden geldt de nominatief (subjectvorm) als lemmavorm.

mnl. du, di, dijns DU

Bij werkwoorden geldt de infinitief als lemmavorm.

mnl. bem, es, ghewesen ZIJN

Bij deze vormen is de grammaticale informatie (persoon, getal, geslacht, tijd, wijze etc.) zeker zo informatief als de lemmavorm. In de paradigma-informatie die aan deze woorden wordt gehecht, kan men de vormen met afwijkende wortel terugvinden (bij ZIJN bijvoorbeeld *ben, is, geweest*) (Ruitenberg en Van pellicom 2012: 12-13).

2.7 Vreemde woorden (leenwoorden met uitheemse kenmerken) (zie verder hs. 7)

Deze krijgen evenals erfwoorden de modern-Nederlandse vorm:

nnl. vanielje VANILLE

mnl. pore PORIE

vnnl. voyaedse VOYAGE

Voor verdwenen vreemde woorden gelden in principe dezelfde regels als voor inheemse woorden, maar omdat de klankontwikkelingen van vreemde woorden binnen het Nederlands veel onregelmatiger zijn, is hier wel eens een ad hoc-regel toegepast. Zo krijgen *batement* en *esbatement* elk een eigen lemma als frequent voorkomende varianten, hoewel *es-* geen prefix is.

2.8. Eigenamen (zie verder hs. 8)

Eigenamen worden altijd als naam gelemmatiseerd; als PoS krijgen zij NE= named entity. Bij historische namen (tot 1500) kan dat lemma soms samenvallen met het appellatief waarop de naam gebaseerd is. In andere gevallen worden namen in beperkte mate gestandaardiseerd in het lemma, afhankelijk van het type naam.

Alleen indien beschikbaar kunnen bovendien toegevoegd worden als extra informatie bij het lemma:

- de hoofdvorm (hv) bij verkorte vormen of vlevormen van voornamen;
- de moderne naam van de referent (id), als identificatie met de aangeduide persoon of plaats bekend is

vmnl. boengaerd land (perceelsnaam)

BOOMGAARDLAND

vmnl. bettekin

BETJE, (hv) ELISABETH

vmnl. banbrughe

BANBRUG, (id) BAMBRUGGE

onl. osgeresgest (geest van Osgeer)

OSGEERSGEEST, (id) OEGSTGEEST

nnl. Hooft (familienaam)

HOOFT

3. Het lemma en de daarmee aangeduide varianten

Het lemma geldt als koepel voor allerlei diachrone en synchrone varianten. Hieronder worden de verschillende typen varianten besproken. Daarbij komen tegelijk de criteria aan de orde, die bepalen welke van verschillende varianten het standaardlemma wordt.

3.1 Gewone tegenover gemarkeerde varianten

De vorm die in het moderne Nederlands het meest gebruikelijk is, geldt als lemma. Hieronder kunnen vormen vallen die eveneens tot de standaardtaal behoren, maar gemarkeerd zijn in hun gebruik.

rouw, ruw	RUW
einde, eind	EINDE
honig, honing	HONING

Ook als er in het moderne Nederlands een beperkte betekenisdifferentiatie is opgetreden, maar het is moeilijk vast te stellen wanneer dat precies gaat gelden, hanteren we één lemma: *vrouwe* (aanduiding een aanspreekvorm van een adellijke vrouw) en *vrouw* (neutraal) vallen beide onder VROUW

Nog enkele specifieke gevallen:

Voorbeeld 1: *peuluw, poluw, peluw*: hiervan is *peluw* is het meest gebruikelijk: GiGaNT-lemma wordt PELUW, ook in samengestelde woorden.

Voorbeeld 2: bij de *ie-uu*-variatie is per woord voor de moderne vorm als lemma gekozen:

mnl., vnnl. bedieden	BEDUIDEN
mod.nl. beduiden	BEDUIDEN

Zo ook *sturen* en *bestuur* en de vele andere samengestelde verwante woorden. Bij *lui* en *luiden* geldt als standaardvorm en dus lemma: LIEDEN

Reeksen klank- en spellingvarianten van één woord kunnen tot enkele groepen varianten worden teruggebracht.

Voorbeeld 3, *tegen*: spellingen met *t-*, *s-* en *tj-*: TEGEN, spellingen met *j-*: JEGEN.

mnl. tejeghens	TEGENS
mnl. jegenworde	JEGENWOORD

Voorbeeld 4: Veel vormvariatie komt er voor bij dialectwoorden. De varianten *darie*, *derrie*, *daring*, *derring*, *darg*, *derg* 'veenachtige grond' (en daar weer spellingvarianten van) worden onder twee lemma's ondergebracht: DERRIE voor de eerste twee, DERRING voor de overige. Klankwisselingen in de eerste syllabe worden dus genegeerd (*-er-ar*-variatie voor *-r* komt veel voor); bij *-ing* kan men aan een suffix denken (*-g* is daar weer een variant van).

Ook bij tussenklanken en buigingsuitgangen hanteren we de moderne vorm, en als die niet bekend is, aan de vorm van vergelijkbare samenstellingen.

Voorbeelden tussenklanken:

hoerhuis	HOERENHUIS
hemelblauw	HEMELSBLAUW

Is er geen duidelijke analogie, dan blijft de oude tussenklank gehandhaafd. De tussenklank *e* wordt hierbij volgens de moderne spellingregels aangepast. Flexie-uitgangen zoals de adverbiale *-s* kunnen in een ouder woord of lemma voorkomen zonder in het moderne lemma meegenomen te worden en andersom.

Voorbeelden uitgangen:

mnl. sint	SINDS
mnl. elder	ELDERS

3.2 Variatie in woordfamilies

De vorm van een modern lemma dat toegekend is aan een simplex, hoeft niet noodzakelijk te gelden voor de of alle daarmee samengestelde of afgeleide woorden. Dit geldt zowel voor de historische als voor de moderne woordenschat. Zo zijn alle historische varianten van *branden*, o.a. *brennen*, *bernen*, *barnen*, *bornen* in GiGaNT als *-BRANDEN* en *BRAND-* gelemmatiseerd, behalve *BARNSTEEN* en enkele eigennamen. Onder het lemma *STROOP* vallen de vormen *siroop* en *stroop*, maar naast het lemma *APPELSTROOP* is er *LIMONADESIROOP*. Het is lastig precies aan

te geven wanneer in samengestelde woorden geüniformeerd wordt en wanneer ze hun verschillende vormen kunnen behouden. Zie de voorbeelden.

Voorbeeld 1: *lijtteeken, lycteecken, licteken, lidteeken* met als lemma modern-Nederlands LITTEKEN. Toelichting: -TEKEN conform de lemmatisering van het simplex, maar niet LIJK- conform het simplex LIJK, omdat de moderne geassimileerde vorm van de samenstelling LITTEKEN is.

Voorbeeld 2:

mnl. bodeling ‘ingewanden’	BEULING
nnl. beuling ‘ingewanden’	BEULING
Hiernaast mnl. bodel ‘darm’	BODEL (in VMNW geen <i>bodelling</i>)

NB: Uiteraard worden zuivere spellingaanpassingen wel altijd doorgevoerd in samengestelde woorden:
visc, vischbestand, vischfabriek VIS, VISBESTAND, VISFABRIEK

3.3 Desambiguering

Daar waar homoniemen in tekstmateriaal elk dezelfde variante vormen kennen, wordt dit in het GiGaNT-lexicon zoveel mogelijk opgelost door verschillende lemma's toe te kennen. Voorbeeld 1. *gard, garde* ‘roede’, ‘stok’, ‘puntig voorwerp’, naast *gaard, gaarde* ‘tuin’. Beide woorden komen o.a. voor als *gard, gaard, gaerd, garde, gaerde*, etc. Daarnaast is er nog *gard, garde* ‘wachter, wacht’.

Bij lemmatisering is nu strikt onderscheiden in GARD ‘roede’, GAARD ‘tuin’ en GARDE ‘wacht’, ook in de samenstellingen waarvoor geen specifieke modern-Nederlands vorm geldt. Is die er wel, dan wordt dat de lemmavorm, zoals bij *diergaard*:

mnl. garde ‘stok; gaarde’ (maat)	GARD
nnl. gard, garde ‘roede’	GARD
mnl. gaert, garde ‘tuin’	GAARD
nnl. gaarde ‘tuin’	GAARD
nnl. garde ‘lijfwatch’	GARDE

Voorbeeld 2. -SEX- is het moderne lemma voor de leenwoordelementen met de betekenis ‘zes’; tegenover -SEKS- voor alle woorden/woorddelen met de betekenis ‘geslacht’. Dit is conform nieuwe spellingvoorschriften, die alleen voor Engelse leenwoorden (*sexy, sexappeal*) een -x-spelling kennen.

Blijven ze als homoniem lemma gehandhaafd, zoals VORST ‘koning’ naast VORST ‘vriesskou’, dan worden ze in GiGaNT door een verschillend ID en idealiter een glosse onderscheiden en afzonderlijk gelinkt aan citatenmateriaal. (Homoniemnummers zoals gebruikelijk in de historische woordenboeken vervallen.)

3.4 Behoud van de oorspronkelijke woordstructuur

Wanneer een woord in de loop van de tijd in het taalsysteem vervangen wordt door een ander dat daarop lijkt, wordt dat vervangende woord niet aangeduid met het lemma van de voorganger.

Voorbeeld 1: lidwoord onl. *that*, is etymologisch een ander woord dan het moderne lidwoord *het* (oorspronkelijk alleen pers.vnw.). Onl. lidwoord *that* krijgt dus lemma DAT en niet HET. Onl. *thie* en mnl. *die* (als lidwoorden) vallen beide wel onder DE, aangezien *de* zich heeft ontwikkeld uit de volle vorm.

Verschil in afleiding, d.w.z. de samenstellende delen zijn verschillend, kan eveneens voor verschillende woorden zorgen.

Voorbeeld 2: *edellike* wordt EDELLIJK, niet ADELLIJK, basis is *edel* (bnw.) en niet *adel* (znw.).

Ook kan verschil in morfologische structuur een rol spelen:

Voorbeeld 3: *weger* en *wager*, beide ‘persoon die weegt’, zijn eveneens verschillende woorden en geen vormvarianten. Het eerste is afgeleid van het ww. *wegen*, het tweede van het znw. *waag*, dat een ablautsvorm is naast *wegen*. Het zijn dus twee lemma’s, WEGER en WAGER.

3.5 Splitsing en samenvall

Variatie kan leiden tot het ontstaan van nieuwe woorden uit eenzelfde woord (splitsing), omgekeerd kunnen verschillende woorden samenvallen tot één woord (samenvall).

Splitsing

Als van een woord een oude vorm bewaard is naast een nieuwe die hieruit ontstaan is, met in het moderne Nederlands ieder een eigen betekenis, is er in de jongste taalfase sprake van twee verschillende woorden met elk een eigen GiGaNT-lemma. Die twee lemma’s moeten ook op het oorspronkelijke woord toegepast worden.

Voorbeeld 1: Een bekend geval van splitsing is die van onl. *krūt* in nnl. *kruid* ‘plant, specerij’ en *kruit* ‘ontplofbare poeder’. Het eerste woord bevat (onderliggend) de oorspronkelijke eindklank, een stemhebbende dentaal. De vorm met auslautverscherping in het enkelvoud, weergegeven in de officiële spelling, ging voor de stofnaam gelden. In dit geval zijn KRUID en KRUIT samen lemma tot het Nieuwnederlands; als er duidelijke betekenisdifferentiatie is worden afzonderlijke lemma’s toegekend.

Voorbeeld 2: Ook woorden die zonder vormvariatie een ingrijpende betekenisdifferentiatie hebben ondergaan, kunnen in de jongste taalfase als twee verschillende woorden (homoniemen) worden beschouwd, en krijgen bij hun historische vormen een dubbel lemma.

bloem (historisch, bij niet-gedifferentieerde betekenis)

bloem ‘plantdeel’
bloem ‘meel’

BLOEM ‘plantdeel’ // BLOEM ‘meel’
BLOEM ‘plantdeel’
BLOEM ‘meel’

Voorbeeld 3. Ouder Nederlands *pipen* met zowel de betekenis ‘scherp geluid maken’ als ‘fluiten’ werd na de diftongering in het moderne Nederlands *pijpen* ‘op de fluit spelen’, en behield daarnaast de *ie*-vorm, nu gespeld als *piepen*, in de betekenis ‘scherp geluid maken’. Er wordt voor de oude woordenvormen (die vanuit de moderne situatie ambigu zijn) dan met meerdere lemma’s gewerkt.

mnl. *pipen* (2 bet.)
nnl. *pijpen* ‘fluiten’
nnl. *piepen* ‘schel geluid maken’
nnl. *piepen* ‘fluiten’ (bijvorm)

PIJPEN // PIEPEN
PIJPEN
PIEPEN
PIJPEN

Samenvall

In een enkel geval zijn in een modern woord twee verschillende historische woorden samengevallen. In veel gevallen, zoals bij *bijw./bnw.* op *-lijk* en *bijw./vz* zoals *in*, waarvan het bijwoord in het Middelnederlands een *-e*-uitgang had, geeft de woordsoort het verschil aan. In het lemma wordt geen onderscheid gemaakt.

Ook in andere gevallen heeft de samenvall meestal in een vroege fase plaatsgevonden. Het hangt van het woord af, of, en voor welke fase dan twee lemma’s worden toegekend.

Voorbeeld 1

mnl. *plein*, *pleine*

PLEIN

Hier zijn verschillende Franse woorden al vroeg samengevallen in *plein*.

Voorbeeld 2.

mnl. *verblinden* (intrans.) en *verblenden* (trans.) VERBLINDEN

Deze werkwoorden zijn in het Nieuwnederlands samengevallen in *verblinden* maar al in het Middelnederlands liepen de vormen door elkaar.

3.6 Extra lemma's

In principe hanteren we in GiGaNT voor elke woordvorm maar één lemma. Meerdere lemma's kunnen worden toegekend in de volgende gevallen, deels hierboven besproken. Omdat de status van een extra lemma per geval kan verschillen, zijn er ook verschillende notaties. (NB: deze notatie geldt voor deze nota en niet per se voor de uitwerking in GiGaNT).

1) Bij splitsing en samenvatting (zie 3.6).

2) Wanneer een woordvorm, meestal door een bepaalde typografie, twee woorden representeert:

plant(je)	PLANT / PLANTJE
(inter)nationaal	NATIONAAL / INTERNATIONAAL

3) Bij ambiguïteit: het valt, bijvoorbeeld door gebrek aan context, niet te bepalen of lemma 1 of lemma 2 van toepassing is:

mnl. ekel (leenwoord zonder h-, óf germ. woord met h-procope)	EKEL \ HEKEL
--	--------------

4) Bij clitische verbindingen en andere gevallen van aaneengeschreven woorden die gesplitst moeten worden. Er worden aan één woordvorm twee lemma's toegekend, en wel in de volgorde van de elementen in de clisisvorm. Zie ook 4.3.

mnl. slaepstu (< slaeps du)	SLAPEN + DU
mnl. entie (< ende die)	EN + DE

5) Moderne afkortingen van verbindingen die grafisch een eenheid zijn, worden als verbinding gelemmatiseerd (vgl. clitics); elk woord krijgt een eigen lemma. Bij het lemmatiseren wordt de afkorting opgelost.

mod.nl. o.i.v.	ONDER + INVLOED + VAN
mod.nl. z.s.m.	ZO + SPOEDIG + MOGELIJK

Zo ook meer incidentele afkortingen als p.p., p.w., i.h. (midden), en za./zo.

6) In het moderne Nederlands kunnen soms twee geheel gelijkwaardige vormvarianten van één woord voorkomen. Dan kan aan die woordvorm een meervoudig lemma worden toegekend. Het komt veel voor bij *d*-syncope, waarbij minstens de *d*-vorm lemma is.

weer, weder	1. atmosferische toestand,	
	2. nogmaals	WEDER \\ WEER
broeder, broer		BROEDER \\ BROER

Zo ook bij *leder-leer*, en *ledig-leeg*, *voederen-voeren* 'eten geven'. Bij samengestelde woorden geldt niet altijd een meervoudige lemmatisering.

4. De lemmagrenzen: splitsen en verbinden van vormen; onvolledige vormen

Lemma's bestaan in principe uit één woord, maar er zijn uitzonderingen: bepaalde verbindingen. De grammaticale context en de woordbetekenis bepalen of we met een verbinding of samenkoppeling te maken hebben. We passen daar noodgedwongen onze moderne criteria op toe. Het meer dan incidenteel voorkomen van een niet meer bestaande woordcombinatie als één woord wijst op lexicalisering in de desbetreffende taalfase; deze vorm moet dus in principe gehonoreerd worden. De samenstellende elementen worden bepaald bij de morfologische analyse, waardoor de gebruiker de relatie kan vinden tussen verbinding en samenkoppeling.

4.1 Samenkoppelingen: het moderne lemma is één woord

Een samenkoppeling die eertijds één woord was, kan nu nog één woord zijn:

mnl., nnl.	niettemin	NIETTEMIN
mnl.	mijnvrouwe	MEVROUW

maar ook niet meer als koppeling bekend zijn. We lemmatiseren dan toch de koppeling, omdat het geen incidentele clisisvorm is, maar als woord functioneerde.

vmnw tsaermeer (te jare meer 'voortaan') TSAARMEER

Ook als dit woord nog als vaste verbinding bekend is, wordt eventuele assimilatie niet opgelost:

mnl. wattan (bijw. en voegw.) WATTAN

mnl. harentare HARENTARE

mnl. alselpe (<alzo helpe, in aanroep) ALZELPE

vmnl. selp (vz. 'bij') ZELP

De historische koppelingen vertonen vaak assimilatie en andere klankwijzigingen. Dit is zeker het geval bij bastaardvloeken; ook deze worden als de gestandaardiseerde oude vorm gelemmatiseerd en niet geanalyseerd.

mnl. gangswonden GANSWONDEN (niet: gans/god + wonden)

mnl. wetekeye WETEKEI (niet: weten + christus),

mnl. tjacob TJACOB

(de status van *-t-* is hier onduidelijk).

De vorm van het GiGaNT-lemma bij een samenkoppeling de moderne vorm, als die bestaat:

asjebliet < als het je beliet ALSJEBLIEFT

mod.nl. a.u.b. ALSTUBLIEFT

Voegwoorden en bijwoorden die uit twee elementen bestaan, soms los, soms aaneengeschreven, worden als gewone samenkoppelingen opgevat (anders dus dan uitgebreidere verbindingen; zie 4.4):

mnl., vnnl. als dat, alsdat ALSDAT

mnl., vnnl. so dat, sodat ZODAT

mnl., vnnl. te waren, te waar, tewaren TEWAREN

mnl., vnnl. te hant, tehant TEHAND

Samenkoppelingen die in lopende tekst los geschreven zijn, krijgen de moderne lemmavorm:

vnnl. aller liefst ALLERLIEFST

vnnl. boven al BOVENAL

nnl. wel bespraakt WELBESPRAAKT

Hetzelfde geldt voor samenstellingen die ook niet-continu kunnen voorkomen, zoals scheidbaar samengestelde woorden: werkwoorden met een bijwoord als eerste element:

nnl. hij belde me op OPBELLEN

en voornaamwoordelijke bijwoorden:

nnl. daar heb ik geen zin in DAARIN

Sommige lemma's bevatten een koppelteken, als verbindingsteken binnen één woord, conform de moderne spellingregels: *ex-verslaafde*, *wc-bril*, *garage-ingang*.

4.2 Vaste verbindingen: meerdere moderne lemma's

Volgens algemeen lexicografisch gebruik worden vaste verbindingen, w.o. idiomatische uitdrukkingen, in woordenboeken bij een van de hoofdwoorden beschreven; ook in GiGaNT krijgen de afzonderlijke woorden elk een eigen lemma. Verbindingen komen in principe niet als lemma voor. Een verbinding die frequent als zinsdeel voorkomt en (meestal) aan elkaar geschreven wordt, is een samenkoppeling, waaraan uiteraard één lemma wordt gehecht (zie hierboven, 4.1).

Afzonderlijke lemma's gelden dus voor:

mnl. tote dier stont dat TOT / DE / STOND / DAT

mnl. hare no dare HER / NO / DER

Vgl. de equivalente samenkoppeling *harentare* onder 4.1

4.3 Clisisvormen: meerdere moderne lemma's

Incidentele clisis wordt altijd opgelost:

mnl. nien - dat ic <u>niene</u> quam	NIET + NE
mnl. hine - dat <u>hine</u> (= hij hem) sloech HIJ + HIJ	

Bij clisis en assimilatie in verbindingen wordt het resultaat als samenkoppeling opgevat als de vorm regelmatig voorkomt, en krijgt dan een overeenkomstig lemma. Zie hiervoor verder onder 4.1.

mnl. mettesen	METTEZEN
mnl. tenengader	TENENGADER

4.4 Meerwoordlemma's (multi-word units)

In bepaalde gevallen bestaat het lemma uit meer dan één woord. We onderscheiden twee categorieën.

Ten eerste, vreemde woorden die een verbinding vormen die als zinsdeel in het Nederlands functioneert. Deze verbindingen vormen één lemma; de elementen hebben namelijk geen zelfstandige status in de Nederlandse woordenschat. Zie, volgens de moderne spellingregels gespeld, *agnus dei, vis à vis, ex aequo*.

nnl. gloria patri	GLORIA_PATRI
-------------------	--------------

Dit is de naam van een gebed, die bestaat uit de beginwoorden. Andere voorbeelden:

pro Deo	PRO_DEO
nnl. p.m., pro memorie	PRO_MEMORIE
nnl. t. à t., tout à toi	TOUT_À_TOI

Soms worden dergelijke verbindingen als samenkoppeling opgevat, zoals in het volgende geval. Deze spelling wordt ook in VD gevolgd; het is naam van de eerste zondag na Pasen.

mnl., nnl. quasi modo, quasimodo	QUASIMODO
----------------------------------	-----------

De tweede categorie wordt gevormd door eigennamen (zie ook hs. 6) die uit meerdere woorden bestaan. De elementen van de naam worden verbonden, maar ze krijgen elk de standaard-lemmavorm.

In den vergulden Turk (naam van gebouw)	IN_DE_VERGULD_TURK
Vaders wensch (naam van boerderij)	VADER_WENS

Voor andere eigennamen met meer woorden: zie hs. 6.

4.5 Afkortingen

Hieronder zijn inbegrepen symbolen, letterwoorden, initiaalwoorden, verkortingen.

Afkortingen in historische lemma's worden opgelost.

mnl. d (III)	DENARIUS \ DENIER
--------------	-------------------

Toelichting: DENARIUS, wegens de spellingen *d, den, denarj.* voor lat. *denarius*; DENIER (= Franse vorm), want zo waarschijnlijk uitgesproken (of uitgesproken als 'pennink', maar dat heeft niets met de afkorting als zodanig te maken).

Afgekorte woorden krijgen het lemma van het volledige woord.

vnnl. aangen.	AANGENAAM
mod.nl. a.u.b.	ALSTUBLIEFT
mod.nl. wed.	WEDUWE
mod.ndl. j.l.	JONGSTLEDEN
mod.nl. w.o.	WAARONDER

Voor afkortingen die op verbindingen teruggaan, geldt dat ze als clitics worden beschouwd. Zie ook onder 4.3.

vnnl. ul (uwe liefde)	UW + LIEFDE
mod.nl. m.a.w.	MET + ANDER + WOORD
mod.nl. H.M.	HAAR + MAJESTEIT

Afkortingen, verkortingen en letterwoorden kunnen als zodanig gestandaardiseerd zijn en krijgen dan een dubbel lemma waarbij de volle vorm koepellemma wordt:

mod.nl. w.c., wc	WATERCLOSET \\ WC
mod.nl. w.w., ww (werkloosheidswet)	WERKLOOSHEIDSWET \\ WW

Voor symbolen (zoals *m* voor *meter* en *Hz* voor *herz*) en voor afkortingen voor organisaties, instellingen, verordeningen etc. zoals *havo*, *VPRO* kiezen we vooralsnog dezelfde benadering.

Verkortingen zijn dikwijls ingeburgerd en krijgen dan die vorm als GiGaNT-lemma:

mod.nl. plee (mog. < fr. <i>plâit-il</i>)	PLEE
mod.nl. auto (<automobiel)	AUTO
mod.nl. bieb (<bibliotheek)	BIEB

Echter, als in historische teksten een volle vorm voorkomt, die later in verkorte vorm bekend wordt, krijgen beide een eigen lemma, bijv. *schoutheet* naast (het zich daaruit ontwikkelde) *schout*.

4.6. Foutieve vormen

Vormen die per vergissing zijn ontstaan, zoals kopieerfouten, drukfouten, typfouten en spelfouten, komen zowel in handschriften als in drukwerken voor.

Hiervoor geldt het onderscheid

- evident fout, gevallen waarvan de correcte lezing duidelijk is, en die als lemma de correcte vorm krijgen:

mod.nl. botcr (lees: boter)	BOTER
-----------------------------	-------
- niet te interpreteren; mogelijk corrupt/fout. Hier is het lemma conform de aangetroffen vorm.

mnl. <i>savret</i>	SAVRET
--------------------	--------
- vermoedelijk corrupt/fout, waarbij een alternatief kan worden gepresenteerd. Deze vorm kan een dubbel lemma krijgen; nl. lemma 1 = lemma voor de corrupte vorm; lemma 2 = omspelling van het alternatief (het lemma is ambigu).

mnl. bedeesd	BEDEESD \ BETEEST
--------------	-------------------

(“waarschijnlijk bedorven lezing voor *beteest*.”)

5. De weergave van affixen

Affixen worden altijd in hun eigen vorm in het GiGaNT-lemma weergegeven (bijv. in een verkleinwoord, een afgeleide vrouwelijke persoonsnaam, een frequentatief van een ww.). Verschillen in woordstructuur zien we dus altijd terug in de lemmavorm. (Zie ook de voorbeelden in 2.5).

5.1 Handhaving van oorspronkelijke affixen

Verdwenen affixen worden ook opgenomen en niet vervangen door moderne:

mnl. weghelijjn	WEGELIJN (niet: wegje)
(zie voor diminutiefsuffixen verder onder 9.5.2.5).	
mnl. erfacht	ERFACHT (niet: erfachtig)
mnl. broekacht	BROEKACHT (niet: broekig)

Over dit *-acht* is het WNT niet zo duidelijk, maar zie EWN *-achtig* en W. Pijnenburg (1993). *-el* wordt als affix beschouwd; het komt o.a. voor in gereedschapsnamen in:

mnl. seisel, sessel	ZEISEL
Daarentegen:	
mnl. seisen, seisene:	ZEIS

want de status van *-en* is hier onduidelijk.

5.2 Standaardvormen

Voor affixen zijn standaardvormen vastgesteld. Zie voor specifieke gevallen 9.5.2.2. t/m 9.5.2.5.

onl. diupitha	DIEPTE (niet: diepede)
mnl. drapenierigghe	DRAPENIEREGGE
mnl., westvl. sonderlanghe	ZONDERLING

Verschillende (klank-/spelling)vormen van structureel eenzelfde afleiding worden tot één vorm teruggebracht, namelijk de vorm die we in moderne afleidingen zien:

lengede, langde	LENGTE
ternink, teerling	TEERLING

Als eenzelfde affix verschillende moderne standaardvormen kan hebben, kiezen we in GiGaNT per woord één vorm; ze worden dus niet per se tot één suffix teruggebracht.

Voorbeeld 1: *-erie* en *-erij*.

rabauwerie	RABAUWERIJ
clownerie	CLOWNERIE

Voorbeeld 2: *-erd* en *-aard*.

mnl. mostaert, mostert	MOSTERD
mnl. bastaert, bastert	BASTAARD

Voorbeeld 3: *-aar* en *-er* worden niet als verschillende affixen beschouwd, en een oudere vorm op *-er* kan dus een suffix *-aar* krijgen:

mnl. <i>wanderer</i>	WANDERAAR
----------------------	-----------

Dit woord volgt de regels van het moderne Nederlands *-aar* na *ww.*-stammen die eindigen op *sjwa* + *l, n, r*. Het volgende woord volgt niet de moderne regel, maar de moderne standaardvorm (die zelf ook een uitzondering op de regel is, Morfologisch Handboek p. 170-171).

mnl. <i>lerere</i>	LERAAR
--------------------	--------

Zo ook modern *winnaar* i.p.v. *winner*. Het simplex *winner* (tot aan de 20ste eeuw was dit de gewone vorm) krijgt als lemma daarom WINNAAR. In de samenstellingen met *winner* blijft de *-er*-vorm staan (vgl. *kostwinner*).

6. De weergave van vormen in het paradigma van voornaamwoorden en werkwoorden

Bij de GiGaNT-lemmavorm voor voornaamwoorden laten we grammaticale eigenschappen prevaleren boven de fonetische en etymologische. Dit houdt in dat we één hoofdvorm voor een paradigma aanhouden. Voor de gebruiker zijn alle andere, ook qua wortel afwijkende, vormen van het paradigma vindbaar via de vormen die bij de morfosemantische kenmerken zijn opgenomen. Hieronder volgen beslissingen en bijzonderheden over de voornaamste categorieën.

6.1. Persoonlijk voornaamwoord

Alle vormen in eenzelfde paradigma krijgen als GiGaNT-lemma de subjectsvorm

mnl., vnnl. du, dijns, di DU

mnl., nnl. wi, wij, we, onzes, ons WIJ

Gereduceerde en clitische vormen krijgen de volle lemmavorm:

'r - hij heeft 'r (< haar) gezien ZIJ

-et - (hi) naemt (< hi nam-et) NEMEN + HET

Problematisch zijn vooral de voornaamwoorden van de 2^{de} persoon, omdat er verschuivingen binnen de paradigma's hebben plaatsgevonden.

a. u - gij

U, oorspronkelijk alleen objectvorm van de 2^{de} persoon en dus een vorm van het paradigma van *gij*, verkreeg hiernaast subjectfunctie, met overgangsvormen als *ue*, *uwee*. Het is praktisch om voor objectvormen in het Nieuwnederlands t/m de 19^{de} eeuw een dubbel (ambigu) lemma toe te kennen.

Lemmavoorbeelden:

mnl. u, ju, jou (altijd subject) GIJ

vnnl. u (meestal object) GIJ \ U

nnl. u, ou, oe (indien subject) GIJ \ U

nnl. u, uw (indien object) U

Een bijzonder geval is:

nnl. U.E., UE., UE, u E., ue. UE.

Deze afkorting kan niet opgelost worden, omdat niet duidelijk is of zij voor *Uwe Edelheid* of *U Edele* staat. Ze kan in deze (korte) vorm als persoonlijk voornaamwoord opgevat worden.

Omdat niet uitgesloten kan worden dat ze in bepaalde gevallen voor de klankvorm *uwee* staat, geldt deze laatste vorm als tweede lemmavorm.

Dan zijn er nog de subjectvormen:

nnl. uwee, uwé UWEE

nnl. uwees (nominativering bez.vnw. uwees) UWEE

b. gij - jij

Jij, oorspronkelijk nevenvorm van *gij*, kreeg in het Nieuwnederlands gaandeweg een eigen status als voornaamwoord voor informeel gebruik of voor aanspreking van minderen. Ook de objectvormen werden *j*-vormen. Zo verwijzen de lemma's *ju* en *jou* in het MNW naar de hoofdvorm *gij*, in het WNT wordt bij dezelfde vormen naar *jij* verwezen. Vanaf het Vroegnieuwederlands (16^{de} -17^{de} eeuw) t/m de 19^{de} eeuw zou men alle *j*-vormen bij GIJ \ JIJ kunnen onderbrengen, daarvóór bij GIJ. Bij ondubbelzinnig gebruik wordt één lemma toegekend.

Lemmavoorbeelden:

mnl. ge, gy GIJ

vnnl. jij, je GIJ \ JIJ

mod.nl. je, jij, jou JIJ

c. gijlieden - jullie

De meervoudige pers. vnw. kennen uitgebreide vormen met *-lieden*, waarvan die van de tweede persoon, *gijlieden*, *jijlieden* bijvormen had als *jullie* en *jelui*. *Jullie* werd een nieuwe standaardvorm voor de 2^{de} persoon meervoud. Het WNT behandelt dit *jullie* onder het lemma *jij*, vanuit 19^{de}-eeuws perspectief, maar dit past niet bij het moderne voornaamwoordelijke systeem. Voor de lemmatisering is het voorstel alle vormen waarin het element *-luiden*, *-lui*, *-lieden* herkenbaar zijn als *-lieden*-vorm te lemmatiseren, met, al naar gelang de beginconsonant, *gij*- of *jij*- als eerste element. Overige vormen vallen onder JULLIE. Sterk gereduceerde vormen krijgen ook het standaardlemma.

Lemmavoorbeelden:

gijlieden GIJLIEDEN

jelijeden, jeluiider JIJLIEDEN

jijlui, joului, jelui, jullie JULLIE

julder JIJLIEDEN

6.2. Bezittelijk voornaamwoord

Zij krijgen een afzonderlijk lemma, hoewel ze voortkomen uit de persoonlijke voornaamwoorden. Gereduceerde vormen als *m'n*, *je*, worden gestandaardiseerd. Er zijn zelfstandige gebruikte bez.vnw.: het *mijne*, de *zijnen*.

6.3. Aanwijzend voornaamwoord

In historisch materiaal vallen deze voor een deel qua vorm samen met het lidwoord. De context is bepalend.

6.4. Wederkerend voornaamwoord

In ouder Nederlands (en nu soms nog in regionaal en informeel taalgebruik) vervult de objectsvorm van de voornaamwoorden deze functie. Daarom hebben woorden als *hem* en *haar* in deze functie wel een eigen lemmavorm. Ze krijgen dus niet de subjectvorm als lemma, en zeker niet ZICH.

haer - si rechte haer op	HAAR
'em - dat hij 'em vergist had	HEM

De relatie met *zich* wordt vastgelegd in de PoS-toekenning.

6.5. Wederkerig voornaamwoord

Het moderne *elkaar* heeft als voorgangers *elk-ander* en *manlijk-ander*, die ook niet-continu kunnen voorkomen en vele verkorte en geassimileerde vormen hebben. Alle *elk*-vormen woorden als *elkaar* en alle *manlijk*-vormen als *mekaar* gelemmatiseerd.

6.6. Onbepaalde voornaamwoorden

Vooraf oudere vormen van onbepaalde voornaamwoorden kennen door verbuiging en combinatie met diverse woorden allerlei variantie. Doorgaans worden verbogen vormen niet afzonderlijk gelemmatiseerd, maar wordt de woordstructuur wel gehandhaafd. Als een woord nog bekend is, kan de huidige vorm wel invloed op het moderne GiGaNT-lemma hebben.

Enkele voorbeelden:

als 'alles', alles	ALLES
soghedaen	ZOGEDAAN (niet <i>zodanig</i>)
diesgelijk(s), desgelijk(s)	DESGELIJKS
el 'ander'	EL
elkendeen	ELKENDEEN
letter, lettelt, luttelt	LUTTEL
negeen, egeen, engeen	NEGEEN
niemen, niemant, niemanden	NIEMAND
selck, zulk	ZULK
zulks	ZULKS

6.7 Overgangsvormen

In het verleden werden in diverse functies *d*-vormen gebruikt terwijl nu *w*-vormen gewoon zijn: (de plaats) *daer hi sat* 'waar hij zat', (hij vertrok) *dat zij betreurde*. Hier worden de oorspronkelijke vormen gelemmatiseerd, aangezien het om woordvervanging gaat en niet om klankwisseling (Van der Horst 2008).

7. Vreemde woorden (leenwoorden met uitheemse kenmerken)

Leenwoorden met uitheemse klanken en spellingen krijgen als ze nog bekend zijn, de moderne vorm als GiGaNT-lemma. Soms gelden naarnaast specifieke regels.

7.1 Aansluiting bij de oorspronkelijke vorm

Woorden afkomstig uit een vreemde taal in hun oorspronkelijke vorm, voorzover tenminste als Nederlands woord gebruikt, krijgen de standaardvorm uit die taal.

Voorbeeld 1

nnl. intrade 'intoct'(sp., it.)	INTRADE
nnl. negligé (fr.)	NEGLIGÉ

Als een woord in het Nederlands merendeels aansluiting laat zien bij de oorspronkelijke vreemdtalige vorm, dan wordt deze zo overgenomen.

Voorbeeld 2, bij fr. *aise*:

mnl. eise	AISE
mnl. nnl. asegment	AISEMENT
nnl. aisement	AISEMENT
mnl. aaisieren	AISIEREN

Belangrijke varianten krijgen een eigen lemma. De mnl. en nnl. vormen *abatement*, *batement*, *badiment*, *esbatement* worden gelemmatiseerd als ABATEMENT, BATEMENT (ook voor *badiment*), ESBATEMENT.

7.2 Meervoudige ontleningen

Leenwoorden die in een verschillende periode en/of via verschillende talen ontleend worden, en naast vormverschil ook verschillende betekenissen hebben, worden als verschillende woorden opgevat en krijgen verschillende GiGaNT-lemma's.

Voorbeeld 1:

nnl. occasioen 'gelegenheid' < fr. occasion	OCCASIOEN
mod.nl. occasion 'tweedehandsartikel' < eng.	OCCASION

Voorbeeld 2:

mnl., nnl. trop < lat. troppa	TROP \\ TROEP
nnl. troep < fr. troupe	TROEP \\ TROP

De simplicia krijgen in dit geval een extra lemma omdat men wegens het geringe vormverschil aan vormvariatie zou kunnen denken). In de samengestelde woorden is dat niet gebeurd.

7.3 Verouderde vreemde spellingen

Deze worden niet gehonoreerd. Leenwoorden als mnl., vnnl. *bataelge*, *bottelgier*, *selgeren*: de *-g-* hierin is een weergave van de gemouileerde *-l-* in het Frans; die spelling verdwijnt in de loop van de geschiedenis (Salverda de Grave (1906, 279)). De lemmavormen hebben daarom altijd *-l-* of *-ll-*.

mnl. selgeren (asselgeren, asselieren)	SELEREN
mnl. faelgieren	FAILLEREN
mnl. bataelge	BATAILLE

Dit geldt ook voor affixen: het Franse suffix *-age* dat in het oudere Nederlands ook als *-aedse*, *-aige*, *-adze* etc. voorkomt, wordt in het GiGaNT-lemma steeds als *-age* weergegeven.

8. Eigennamen

Eigennamen worden wegens de bijzondere status altijd als namen getagd in GiGaNT door de code NE (Named Entity) in plaats van een gewone PoS-aanduiding. Voor de lemmavorm gelden regels die kunnen afwijken van die voor de soortnamen.

Namen van personen (voornamen en toenames/familienamen) en plaatsen (dorpen, steden, rivieren, percelen, landen etc., zowel inheemse als uitheemse) maken van elke tekst deel uit. Ze komen als lemma vooral voor in het ONW en VMNW. Naast deze hoofdcategorieën van eigennamen zijn er nog namen van andere entiteiten zoals organisaties, gebouwen, schepen, enz.

8.1 Relatie tot soortnamen

De historische namenvoorraad is bij het toekennen van standaardvormen te onderscheiden in diverse categorieën voor wat hun aansluiting bij het niet-onomastische Nederlands betreft:

1. inheemse namen die sterk tegen de appellatieven aanleunen: van de persoonsnamen zijn dat tot aan de moderne tijd enkele typen toenames (o.a. beroeps(bij)namen als *de bakker*, eigenschapsnamen als *witkop*), van de toponiemen de inheemse plaatsnamen, perceelnamen etc. Voorts kunnen dat namen van organisaties, gebouwen etc. zijn. Overigens kunnen deze namen zelf weer zuiver onomastische elementen bevatten: *Jans stuk* (perceelnaam met daarin voornaam), *van Antwerpen* (toenaam met daarin toponiem).
2. inheemse namen die meer een naamkarakter dragen, doordat hun oorsprong niet meer (goed) herkenbaar is: voornamen (*Geertrui*); familienamen in de moderne tijd (*De Vooyoys*);
3. vreemde namen: uitheemse antroponiemen en toponiemen zonder relatie met Nederlandse soortnamen (*Adam*, *Indië*). Ook gelatiniseerde inheemse namen vallen hieronder.

8.2 Extra informatie

Naast het toekennen van lemma's kan nog extra informatie toegevoegd worden bij namen. Dit valt in feite buiten de eigenlijke lemmatisering en het gebeurt alleen als die informatie gemakkelijk beschikbaar is; het kan dus slechts voor een deel van de namen toegepast worden. De waarde van deze informatie is dat de teksten zo een grotere bijdrage aan cultuurhistorisch onderzoek kunnen leveren.

Het betreft:

- a. bij vlevormen of afkortingen van voornamen: de hoofdvorm;

indien gemakkelijk beschikbaar: een vorm die overeenkomt met de naam waarmee desbetreffende entiteit tegenwoordig wordt aangeduid. Bij plaatsnamen ligt de toevoeging van de moderne naam van de referent als tweede lemma voor de hand.

8.3 De werkwijze per categorie

(De vormen hieronder zijn in hoofdzaak 13^{de}-eeuws woordmateriaal).

8.3.1 Inheemse (Germaanse) voornamen

Deze bevatten vaak twee elementen die in het eerste lemma min of meer gestandaardiseerd woorden. Voorbeelden zijn *-bert/brecht*, *her-*, *ijs-*, *-rik*, *-ger/geer*, *-mar*, *-burg*, *-zoete*. Er zijn ook eenstammige namen en verkorte vormen.

Opmerkingen:

- Vormen die slechts weinig variatie vertonen worden samengenomen.
- Gelatiniseerde/geromaniseerde vormen van Germaanse namen krijgen een eigen lemma. Daarom worden vormen op *-e*, *-a* en *-o* onderscheiden.
- Vleivormen / verkleinvormen krijgen hun eigen lemma; de vleisuffixen en diminutiefsuffixen hebben een standaardvorm: *-zoete*, *-man*, en *-tje/-je* (uit *-kin* etc.), *-ijn*, *-sijn*, *-lijn*.

historische vorm	modern lemma 1 eigenaam cf. (germ.) stamvormen	[extra informatie: eigenaam, volle vorm (indien lemma 1 geen volle vorm is)]
gherbrecht	GERBRECHT	--
aleburgh	AALBURG	ADELBURG
heinin	HEINIJN	HENDRIK
heinekin	HEINTJE	HENDRIK
henc	HENK	HENDRIK
henrick	HENRIK	HENDRIK
heinricus	HEINRICUS [gelat. vorm]	HENDRIK
aleit	ALEID	ADELHEID
adelisa	ADELISA	ADELHEID
aleidis	ALEIDIS [gelat. vorm]	ADELHEID
adelheid	ADELHEID	--
addin	ADDIJN [1ste lid mog. adel, 2de lid dim.-suffix]	--
boetsoete	BOETZOETE	--

8.3.2. Uitheemse voornamen en andere persoonsnamen

Ingeburgerde uitheemse namen hebben veelal een aan het Nederlands aangepaste vorm, die, eventueel aangepast, lemma wordt. Hier kunnen verkortingen en vleivormen bij zijn.

Bijbelse namen: hiervoor is de richtlijn de vorm in de meest recente lijst van het Nederlands Bijbelgenootschap: <http://www.nbv.nl/bijbelse-namen/>. Als echter de naamvorm uit deze lijst sterk

afwijkt van de historische vorm, wordt als eerste lemma een vorm gekozen die bij de laatste aansluit, en wordt de officiële vorm als extra informatie toegevoegd.

Voor klassieke namen en heiligennamen kunnen we Wikipedia of andere encyclopedieën raadplegen.

Zo mogelijk berust de volle vorm op de naamvorm van de geïdentificeerde persoon.

historische vorm	modern lemma 1 eigennaam, in structuur overeenkomende vorm	[extra informatie: eigennaam, volle vorm (indien lemma 1 geen volle vorm is)]
griet	GRIET	MARGARETHA
piterken	PIETERTJE	PETRUS
lisebette	LIESBET	ELISABET
janne	JAN	JOHANNES
germain	GERMAIN	GERMANUS
femine	FEMINE	EUFEMIA
iram	IRAM	CHIRAM (bijb.)
facee	FACEE	PEKACH (bijb.)
fierin	VIERIJN	OLIVIER
kerst	KERST	CHRISTUS
apollo	APOLLO	--
apollin	APOLLIJN	APOLLO

Zie ook de opmerkingen over samennemen, gelatiniseerde namen en vleivormen onder 8.3.1.

8.3.3. Inheemse toenames die herkenbaar bij de appellatieven aansluiten

Zeker in het Oud- en Middelnederlands is bij toenames het onderscheid tussen overerfbare namen en individuele namen niet te maken. In beide gevallen is nog geen sprake van een vaste vorm en lijkt het verantwoord er de huidige vorm- en spellingregels op toe te passen.

historische vorm	modern lemma 1 = vorm eigennaam, gestandaardiseerd als soortnaam	modern lemma 2: GEEN
hoenredief	HOENDERDIEF	
felghebuur	FELGEBUUR	
pankuoko	PANNENKOEK	
tytebedde	TIJTEBED	
de bakker	DE / BAKKER	

8.3.4. Moderne namen (familienamen, namen van zaken)

Hoewel de appellatieven hierin vaak herkenbaar zijn, worden de namen vanaf de moderne tijd niet meer als appellatieven gelemmatiseerd, en de spelling ongewijzigd uit de bron overgenomen. In geval dat uit de bron duidelijk is dat ze naar eenzelfde referent verwijzen, worden ze wel geüniformeerd. Persoonsnamen die uit meer woorden bestaan worden als één vorm gelemmatiseerd, andere namen als verbinding.

hist.vorm/	modern lemma 1	modern lemma 2: GEEN
-------------------	-----------------------	-----------------------------

lemma	eigenaam (= huidige naam/namen indien bestaand)	
de Ruijter	DE RUIJTER	
Van der Heijden	VAN DER HEIJDEN	
de Meijboom (naam schip)	DE_MEIJBOOM	
in den vergulden Turk (naam gebouw)	IN_DE_VERGULD_TURK	

8.3.5. Inheemse toponiemen die herkenbaar bij de appellatieven aansluiten (o.a. perceelnamen, plaatsnamen)

Historische toponiemen hebben vaak overgeleverde vormen die als appellatief gezien kunnen worden (en die herkenbaar zijn in de standaardlemma's van de historische woordenboeken). Daarnaast hebben de meeste een gefixeerde, moderne naam die de plaats identificeert. Soms heeft de moderne naam geen of maar gedeeltelijk een taalkundige relatie met de oude naamvorm. Daarom is toevoeging van twee lemma's, één aansluitend bij de vorm en één identificerende, van belang.

Voorbeeld 1: de plaatsnaam Oegstgeest:

osgeresgest (opgetekende vorm 10^{de} eeuw) 'geest(grond) van Osgeer'; *oesghest*, *oesgeist*, *oesgheest* (opgetekende vormen 13^{de} eeuw): *oegstgheest*: latere vormen met herinterpretatie van *oest* > *oegst*, *oogst*.

Oegstgeest: de huidige naam.

Voorbeeld 2: de poldernaam Monsterhoek (polder bij Krabbendijke op Zuid-Beveland, Zeeland): *monsterhoec* (opgetekende vorm 13^{de} eeuw) 'kloosterhoek (stuk land van/bij klooster)'.

Monnikenpolder: de tegenwoordige naam.

Toponymische elementen worden in de lemma's gestandaardiseerd: *-dijk*, *-lo*, *-voorde*. Bij bepaalde elementen wordt ook de gereduceerde vorm of vormvariant als lemmadeel opgenomen (*-em*, *-heem*, *-heim*).

historische vorm	modern lemma 1 = vorm soortnaam, indien afwijkend van eigenaam	modern lemma 2 = eigenaam (= huidige naam/namen indien bestaand)
dorremael	DOENMAAL	DORMAAL
achtienhoven	ACHTTIENHOVEN	ACHTTIENHOVEN
scaepettinc	SCHAAPETTING	--
rēngērasdīk	RENGEERSDIJK	REIGERSDIJK
gheetscure	GEITENSCHUUR	GEETSCHURE (plaats bij Evergem)
boekhout	BEUKHOUT	BOEKHOUT, BOEKHOUTE (twee verschillende plaatsen)
osgeresgest	OSGEERSGEEST	OEGSTGEEST

8.3.6. Uitheemse toponiemen en toponiemen met uitheemse elementen

Namen die uitheemse elementen bevatten, zoals heiligennamen, worden niet aangepast aan de oorspronkelijke vorm van die elementen.

Voor bijbelse plaatsnamen, zie hierboven bij de uitheemse persoonsnamen (8.3.2.). Voor andere aardrijkskundige namen waarvoor een moderne standaardspelling geldt, wordt die overgenomen (Lijst Taalunie). Met oudere, inheemse naamvormen van buitenlandse plaatsen wordt echter rekening gehouden. Net als bij de inheemse toponiemen (8.3.5) kan de moderne aanduiding sterk afwijken van de historische vorm. Ook hier is het toekennen van een extra lemma's zinvol.

hist.vorm/ lemma vmnw	modern lemma 1 eigenaam, in structuur overeenkomende vorm	modern lemma 2 = eigenaam (= huidige naam/namen indien bestaand)
sente-pieters	SINT-PIETERS	SINT-PIETERS
belleem	BELLEEM	BETLEHEM (bijb. plaats) \ BETHLEHEM (naam klooster)
ebela	EBELA	ABEL-BET-MAÄKA (stad in Galilea)
hackengnis	HACKENGNIS	HACQUEGNIES (plaats in Henegouwen)
endi	ENDI	INDIA
pleymuyden	PLEIMUIDEN	PLYMOUTH

9. Detailproblemen en voorbeelden bij 3 t/m 5

Bij 3. Het lemma en de daarmee aangeduide varianten

9.3.1 Het al of niet volgen van regelmatige klankveranderingen

Bij verdwenen woorden geldt de regel voor klankverandering die algemeen heeft gewerkt (zie hierboven), bijvoorbeeld *e*-apocope:

are 'arend' AAR (homoniem)

of diftongering en *e*-apocope:

mnl. mute 'ruien van vogels, ruitijd' MUIT

Is een woord nu niet meer bekend, maar wel opgenomen in het WNT in de spelling De Vries en Te Winkel (op basis van de "beschaafde uitspraak" van de 19^{de} eeuw, d.w.z. de klankvorm van de toen beoogde standaardtaal), dan geldt die spelling/vorm.

Voorbeeld 1, wisseling o - eu:

mnl. becromen BEKREUMEN

nnl. bekreumen BEKREUMEN

Bepaalde klankwijzigingen hebben maar in een deel van de woordenschat gewerkt. In etymologisch nauw verwante woorden kan ook het ene de klankwijziging vertonen en het andere niet: *bruisen* (met *oe*>*ui* en diftongering) *tgov. broes* 'tuit van een gieter' (*oe*-relict); *vouwen* (met *d*-syncope) *tgov. eenvoudig* (zonder *d*-syncope). Voor deze woorden hanteren we de moderne variant, als die nog bestaat. Zo niet, dan construeren we de meest regelmatige vorm:

Voorbeeld 2:

vuldelic, voldelijk VOUDELIIK

namelijk met regelmatige overgang *old*>*oud*, zoals nog in *eenvoudig*, en niet met *d*-syncope, zoals in *vouwen*.

Bij dialectisch *mieg* wordt niet als *mijg* gelemmatiseerd: een diftongische variant is niet aangetroffen.

In bepaalde gevallen bewaart het moderne lemma de oude vorm van een verdwenen woord.

Voorbeeld 3 (waarin *s* niet wordt omgespeld naar *z*):

mnl. besuect BEZIEKT

mnl. besukt BESUKT

moderne Nederlands voorkomt. Bij de lemmatisering moet verder rekening gehouden worden met homoniemen: *leer* 'ladder' en *leer* 'doctrine' hebben geen *d*-vormen zoals *leer/leder* 'dierenhuid', evenmin als *weer* 'man' en *weer* 'inrichting voor visvangst', tgov. *weer/weder* 'atmosferische toestand'; 'terug, nogmaals'. Ook dit pleit ervoor *d*-vormen toe te kennen waar die etymologisch thuishoren (desambiguering).

Voorbeeld 1: blo-blode

mnl. blode	BLODE \ BLO
nnl. bloede, bloo, blo	BLODE \ BLO

Voorbeeld 2:

nnl. leder, leer 'dierenhuid'	LEDER \ LEER
-------------------------------	--------------

Toelichting bij het laatste dubbellemma: beide staan als gelijkwaardig in het GB, hoewel *leer* als simplex nu de meest gebruikte vorm is. In de samenstellingen kan er verschil zijn in behoud van de *-d-* in het moderne lemma. Zo zijn er nieuwe woorden die (waarschijnlijk) nooit een "oude" vorm gehad zouden kunnen hebben, zoals *weestation* (nooit *wederstation*), want historisch gezien bestond die (= het verschijnsel) weervrouw niet. Andersom zijn sommige oudere vormen met *-d-* in die vorm gefixeerd:

Zie:

nnl. weder bijw./vz.	WEDER \ WEER
en nnl. wedergeboorte	WEDERGEBOORTE

Voorbeeld 3: dezelfde regel geldt ook voor hypercorrecte *d*-vormen die nu algemeen Nederlands zijn geworden: *kade* naast ouder *ka* en *kaai* < fr. *caye* (zie Schönfeld par. 37). Alles krijgt trefwoord KADE en daarnaast KAAI of KA, althans in het simplex. Bij de samengestelde woorden komt in het WNT veel *kaai* voor, de oudere fasen hebben ook *ka* (*kadijk*).

Voorbeeld 4:

nnl. vader, vaar	VADER
nnl. broeder, broer	BROEDER \ BROER

Er zijn hier modern-Nederlandse vorm- en betekenisverschillen: *broeder* 'zoon van zelfde ouders' (formeel), 'verpleger', 'kloosterling' (neutraal), naast *broer* 'zoon van zelfde ouders' (neutraal).

(Hetzelfde speelt bij *zus/zuster*.) Het is niet mogelijk die betekenisverschillen altijd te honoreren.

Bij *broer/broeder* is de volgende lijn gevolgd:

in principe is voor BROEDER \ BROER (met extra lemma) gekozen in het simplex en in samengestelde woorden die ook nu de *-broer*-variant kennen, zoals *halfbroer*, *stiefbroer*, *schoonbroer*, geregistreerd door GB of VD. Enkele idiomatische samenstellingen krijgen alleen -BROER- : *drinkebroer*, *jabroer*. Verder steeds -BROEDER-. Bij *vader* behouden we de *d*-loze vorm in *bestevaar*.

Voorbeeld 5: *vlade/vla/vlaai*. Betekenissen en vormen lopen hier door elkaar. De oorspronkelijke woordvorm is *vlade*. In het moderne Nederlands bestaat de vorm-betekenisverdeling: *vlaai* 'speciale (Limburgse) taart' en *vla* 'koude gladde pap'. Hieronder de historische en moderne woordvormen met de belangrijkste betekenissen:

mnl. vlade, vlaey, vla 'dunne koek'	VLADE \ VLAAI
nnl. vla 'koude, gladde pap'	VLADE \ VLA
nnl. vlade (var. vlaai(e), vla) 'dunne koek'	VLADE \ VLA \ VLAAI
nnl. vlaai	VLADE \ VLAAI

In de samengestelde woorden komen we alle drie de varianten tegen.

9.3.4 Onregelmatige en niet goed traceerbare klank- en vormwijzigingen

Deze leiden niet altijd tot een afzonderlijk lemma.

Voorbeeld 1. De dialectvormen *wessel-wissel-wiezel(tje)-weizel(ke)*, door het WNT alle beschouwd als vormen van *wezel* 'klein roofdier'. Zo nu ook voor GiGaNT gelemmatiseerd, inclusief het eerste lid in

mnl. visselcijs	WEZELCIJNS
-----------------	------------

Voorbeeld 2:

nnl. gheroffel, grofffel	GROFFEL (cf. WNT)
--------------------------	-------------------

nl. genoffel	GROFFEL (id.)
maar: nnl. karnoffel	KARNOFFEL (mogelijk hetzelfde woord maar te afwijkend van vorm)

Voorbeeld 3: gevallen van (incidenteel optredende) dissimilatie: de moderne vorm is standaard.

mnl. toverare	TOVENAAR
mnl. camerier	KAMENIER

Diverse andere voorbeelden:

nl. vermelen, vermemelen 'vermolmen'	VERMELUWEN
mnl. ruggelde 'gerugd'	RUGD

want met anorganische *-l-* uit *ruggede*.

mnl. edeken 'herkauwen'	EDIKKEN
-------------------------	---------

Toelichting: het wordt door het VMNW als bijvorm van *edericken* beschouwd.

Voorbeeld 4: vocaalvarianten.

Soms zijn *i-* en *e-*vormen varianten van elkaar maar niet altijd:

mnl. sencken 'doen zinken'	ZENKEN (en niet <i>zinken</i>)
----------------------------	---------------------------------

Toelichting: dit ww. heeft ablaut met causatieve functie; vgl. *drenken*.

rinnen	RINNEN (en niet <i>rennen</i>)
--------	---------------------------------

Toelichting: hoewel dit ww. verouderd is, en de betekenis deels door *rennen* overgenomen, krijgt het een eigen lemma wegens eigen paradigma (sterk tgov. zwak) en eigen betekenisontwikkeling.

Voorbeeld 5:

Als in het moderne Nederlands alleen de gewijzigde vorm nog te vinden is, wordt niet de etymologische vorm gelemmatiseerd.

mnl., nnl. achterskind	ACHTERSKIND
(en niet <i>achterstkind</i> , alhoewel etymologisch * <i>achterstkind</i>)	
gesternte	GESTERNTE
gesternd	GESTERND

De *n*-vormen zijn hier bijvormen die algemeen zijn geworden. De basis is *ster*.

9.3.5 Gewone tegenover gemarkeerde varianten in de moderne standaardvorm

De vorm die in het moderne Nederlands het gebruikelijkst is, wordt lemma.

Voorbeeld 1: *calander*, *kalander*. In de woordenboeken *kalander* (gewone vorm) en *klander* (gemarkeerde vorm). Het moderne lemma wordt KALANDER.

Voorbeeld 2: Bij minder bekende woorden is een frequente historische spelling soms belangrijker dan de vorm in moderne woordenboeken/-lijsten. Het woord *kateil* heeft historisch vooral *ei*-spelling. Het woord staat niet in GB; VD heeft *kateil*: zie *kateel*. Dus anders dan bij het WNT is hier *kateel* hoofdlemma. Er zijn weinig aanwijzingen waarom dat zo is. Gezien daarbij de *ei*-overheersing in de attestaties is het lemma KATEIL geworden.

9.3.6 Volksetymologische wijzigingen en woordspelingen

Volksetymologische wijzigingen worden in het lemma gehonoreerd, we gaan dus niet "retrolemmatiseren".

Voorbeeld 1:

Bij mnl. *schaerwetter*, met *-wetter* < *aweit* e.d., Frans leenwoord, is het lemma SCHAARWETTER (*wetter* 'natmaker').

Voorbeeld 2:

mnl., nnl. heeft *aarzelmaand* 'oktober' kan afgeleid worden van het ww. *aarzelen*. Ook al noemen VMNW en MNW de mogelijkheid van vorming met *herse-* 'gierst', het lemma is die conform de aangetroffen vormen: AARZELMAAND.

Voorbeeld 3:

Bij leenwoorden zijn zowel de volksetymologische vervorming als het oorspronkelijke woord relevant; de laatste natuurlijk alleen als die nog in het Nederlands wordt aangetroffen.

proncke-fykaal	PRONKFIKAAL \\ PONTIFICAAL
----------------	----------------------------

Hier komt dus een dubbel lemma: a. volksetymologisch vorm; b. vorm overeenkomstig het oorspronkelijke woord.

Voorbeeld 4:

mnl. warant 'borg, garantie'	WARANT
mnl. waerhant 'borg'	WAARHAND \\ WARANT

Voorbeeld 5:

Hierbij horen ook opzettelijke "creatieve" analogische varianten (woordspelingen): diverse *war*-varianten van 'wereld', die refereren aan het begrip 'war', 'warrig' krijgen het GiGaNt-lemma WARRELD.

NB: Dit type is te onderscheiden van verbasteringen die niet gemotiveerd zijn door betekenis of vormanalogie (al is het onderscheid soms moeilijk te maken):

Voorbeeld 6:

nnl. variabel	VARIABEL
---------------	----------

9.3.7 Morfonologie: huidige regels worden gevolgd

Voorbeeld 1:

Volgens het Morfologisch Handboek (De Haas en Trommelen 1990: 245) wordt voor het suffix -(e)nis de -v- waar de ww.-stam op eindigt, stemloos (*droefenis, erfenis*). Bij -(e)lijk is dat soms ook het geval: *erfelijk*. Daar geldt het ook voor de z: *vreselijk* (p. 296).

mnl. lovenesse	LOFENIS
mnl. erloofnesse	ERLOFENIS
mnl. gelofnesse (hier ook vormen met korte vocaal)	GELOFENIS

Voorbeeld 2:

d-epenthese na -r- bij de comparatieven van adjectieven en bij znw. op -er: *raarder, spaarder*. Deze -*d*- treedt in historische vormen ook na andere consonanten op, bijv. na -n en -l-, maar wordt dan niet in het moderne lemma overgenomen (gezien het eerste voorbeeld ook bij flexie-uitgangen):

vnnl. veelderhande	VELERHANDE
WNT twijnder	TWIJNER

Toelichting: moderne citaten hebben hier de *d*-loze vorm.

Uitzonderingen zijn gelexicaliseerde woordvormen als *diender* en *vilder* (zo in VD), en vaktermen als *steunder*; zij behouden de -*d*-.

9.3.8 Specifieke gevallen naar woordsoort 1: deelwoord als bnw. - vorm

Een als bijvoeglijk naamwoord gebruikt voltooid deelwoord krijgt de historische vorm, ook al verschilt de vervoeging van die van nu (sterk/zwak).

Dus:

nnl.gewezen 'voormalig'	GEWEZEN
-------------------------	---------

en niet *geweest*.

Vergelijk

mnl., nnl. gewijsde 'vonnis' (znw.)	GEWIJSDE (homoniem)
-------------------------------------	---------------------

hoewel *gewezen* het nu gebruikelijke deelwoord is bij het oorspronkelijk zwakke *wijzen*.

9.3.9 Specifieke gevallen naar woordsoort 2: deelwoord als bnw. - woordsoort

De regel is: een voltooid of tegenwoordig deelwoord met als woordsoort bnw. wordt ook gelemmatiseerd als bnw. Ook al omdat het soms een pseudodeelwoord is: *gebloemd* = 'van bloemen voorzien', *geleverd*: 'met een lever', dus niet afgeleid van de ww. *bloemen* of *leveren*.

Een speciaal geval is *geboren* in de bet. 'op de wereld gebracht'. Oorspronkelijk waren de vormen en lemma's:

onl. giberan ww.	GEBAREN
mnl. geboren bnw.	GEBOREN

mnl. geboren ww.
WNT geboren ww.

GEBOREN
GEBOREN

In MNW en WNT wordt het woord zowel als deelwoord (van *beren* resp. *geberen*), als als adjectief beschreven. Overeenkomstig ONW en EWN wordt als moderne infinitief GEBAREN (naast BAREN) gekozen.

9.3.10 Samenval en splitsing

Voorbeeld 1

horn heeft als kernbetekenis ‘dierenhoorn’ en heeft allerlei betekenissen rond ‘spits’. Het komt in het historische materiaal voor als *horn*, met rekking van de vocaal als *hoorn*, met sjwa-insertie (svarabahktivocaal) als *horen*. Er heeft zich samen met de vormwisseling betekenispecialisatie voorgedaan, waardoor het huidige Nederlands deels verschillende woorden heeft. De lemmavormen zijn nu:

‘hoek’, ‘spits’, ‘punt’ etc. HORN

cf. VD. ‘hoek lands’, niet cf. VD, waar ‘hoek, spits’ (algemeen) ook bij *hoorn* staat; niet in GB. Argument voor HORN is ook frequent HORNIK in oude bronnen.

‘hoorn van dieren’, ‘blaasinstrument’ HOORN \ HOREN

Hoorn is cf. VD, GB; HOREN cf. GB, waar er twee gelijkwaardige lemma’s zijn.

In de samenstellingen en afleidingen idem.

Voorbeeld 2

foelie en *folie* (< lat. *folia*, abl. van *folium* ‘blad’) betekenden beide heel lang zowel ‘dun blad metaal of ander materiaal’ als ‘specerij van het blad van de muskaatnoot’. Voor *foelie* gelden nog steeds 2 betekenissen (Van Dale 2005), in het ANW ook, maar daar is *folie* de hoofdvorm voor ‘dun blad materiaal’. Verder zijn er nog oude lemmavormen met *g*: *fulgie*, *foelge* (niet gebaseerd op de oorspronkelijke vormen; niet goed verklaarbaar, maar waarschijnlijk gewoon als spellingvariant te beschouwen). Op basis van de moderne betekenisonderscheiding zijn de volgende GiGANT-lemma’s toegekend. De Latijnse vorm *folio* heeft nu nog de specifieke betekenis ‘groot formaat papier’, dit blijft/wordt daarvoor het lemma.

mnl. fulgie ‘specerij’	FOELIE
mnl. folie ‘specerij’	FOELIE
mnl. foelge 2 bet.	FOELIE // FOLIE
nnl. foelie 2 bet.	FOELIE // FOLIE
nnl. folie ‘dun vel’	FOLIE
nnl. folie ‘blad uit koopmansboek’	FOLIO

De samenstellingen hebben de oorspronkelijke lemmavorm behouden. Vandaar bijv. FOELIESLAGER.

Folie ‘dwaasheid’ heeft een andere klankvorm en is dus een homoniem naast *folie* ‘dun vel’.

Voorbeeld 3.

Vergelijkbaar zijn de doubletten: soms zijn uit één woord twee gelijkwaardige varianten ontstaan door invloed van de flexievormen (bijv. een op basis van nominatief en een op basis van verbogen vormen). Beide vormen vinden we in historisch materiaal. Als echter in het Standaardnederlands maar één vorm gebruikelijk is, zoals *geel* bij *geel-geluw* en *schaduw* bij *schade-schaduw*, geldt die vorm als lemma voor alle voorkomens.

onl. skado	SCHADUW
mnl. scade, schaduwe	SCHADUW
nnl. schade, schaduw	SCHADUW

Bij samengestelde woorden wordt eveneens de moderne vormen als lemma toegekend. Zijn beide vormen in het moderne Nederlands nog bekend met betekenisdifferentiatie, dan worden die als twee verschillende woorden behandeld:

onl. nahtskado ‘nachtelijke schaduw’	NACHTSCHADUW
mnl. nachtscade ‘plant’ (met vormen <i>nachtschade</i> en <i>nachtschaduw</i>)	NACHTSCHADE
mnl., nnl. nachtschade ‘plant’ (met vormen <i>nachtschade</i> ,	

nascaye, nachtschaduw NACHTSCHADE
 nnl. nachtschaduw 'nachtelijke schaduw' NACHTSCHADUW

Tot dit type doublet horen *geluw, zaluw, scheluw, schaduw, eluw, deluw, valuw, kaluw, meluw, zenuw*. (*Peul-peuluw-peluw* is een ander geval: het is een leenwoord met diverse varianten, waarvan *peluw* nu de gewone vorm is.)

Bij 4. De lemmagrenzen

9.4.1. Enkele specifieke verbindingen 1: type genitief-znw.-znw.

Flexievormen geven vaak uitsluitsel over de structuur:

(die here van) tsgravensteen = des graven steen DE / GRAAF / STEEN
 tsgravensteen (in Leiden) DE / GRAAF / STEEN

(genitiefliedwoord kan niet op *gravensteen* betrekking hebben en hoort dus bij *graaf*)
 tegenover

(die here) tsgravensteens HET / GRAVENSTEEN
 (het) gravensteen (in Leiden) GRAVENSTEEN

9.4.2. Enkele specifieke typen 2: vz./bijw./znw./bnw. + ww.

Of een verbinding met een werkwoord een samenkoppeling is, kan niet door een regel vastgesteld worden. Het criterium is de lexicalisering van de verbinding anno nu. Zo heeft het GB *platliggen, platleggen, schoonwassen, schoonwrijven, zwartkijken, zwartmaken* (hier, willekeurig, alleen voorbeelden van adjectief + ww.) Aangezien het GB niet pretendeert volledig te zijn in het weergeven van woorden met dezelfde structuur kunnen we uit deze voorbeelden niet per se afleiden hoe *bijv. schoon()lebbere*n geschreven moet worden. Als VD en WNT het desbetreffende woord ook niet hebben, gelden de gangbaarheid en de eenheidsbetekenis als criterium.

Analogie kan ook een rol spelen:

de autoruit schoon krabben SCHOONKRABBEN
 hij dweilde de lei schoon SCHOONDWEILEN

Een probleem is dat veel samenkoppelingen in de taalhistorie nog verbindingen waren, wat kan leiden tot nogal onsystematische lemmatisering:

(iemand, iets) lief hebben LIEFHEBBEN (want modern Nederlands)
 (iemand, iets) liever (liefst) hebben LIEVER (LIEFST) / HEBBEN
 (iemand) leed hebben LEED / HEBBEN

NB: de infinitief wordt niet noodzakelijk aaneen geschreven als er een aaneengeschreven deelwoord is: *kaalgegraasd* houdt niet per se in *kaalgrazen*.

9.4.3. Enkele specifieke verbindingen: type vz./bijw./znw./bnw. + deelwoord

De norm bij de als adjectief gebruikte deelwoorden is ook steeds de lexicalisering van de verbinding anno nu. In de eerste plaats geldt de huidige spelling:

wel bespraakt WELBESPRAAKT
 vast besloten VASTBESLOTEN

Geschreven als één woord want zo in het Groene Boekje, net als *blauwgeverfd, witgekalkt, witgelakt, witgepleisterd*. Verder wordt het deelwoord altijd aaneengeschreven als de infinitief aaneengeschreven wordt: *schoonpoetsen* (GB), dus *schoongepoetst*.

De spellingregels kunnen hier niet eenduidig zijn en het GB bevat uiteraard niet alle mogelijke vormen, maar op basis van analogie zouden we kunnen concluderen dat *groengeverfd* (niet in GB) de schrijfwijze is, analoog aan *blauwgeverfd* (zo in GB). Maar gezien de geringere frequentie en de ongewoonheid van het concept moeten we misschien kiezen voor losschrijven bij samenstellingen met minder gewone kleurnamen, dus:

de groen geverfde deur GROENGEVERFD

de <u>olijfgroen</u> <u>geverfde</u> deur	OLIJFGROEN / VERVEN
Zie ook	
<u>hoog opgegroeid</u> gewas	HOOG / OPGROEIEN
<u>slecht opgevoede</u> kinderen	SLECHT / OPVOEDEN
Ook hier zorgt de diachronie voor inconsequenties. Bijv.:	
dit <u>schreit ten hemel</u>	SCHREIEN / TE / DE / HEMEL
dit is <u>ten hemel schreiend</u>	TENHEMELSCHREIEND
Dit laatste is een samenkoppeling volgens het GB.	

Bij 5.2. Affixen

9.5.2.1. Afwijkingen

In bepaalde woorden kan er afgeweken worden van de gebruikelijke standaardisering en wordt een bijzondere vorm van het affix gebruikt:

Voorbeeld 4:

mnl. *begard*, *beggaert*, *bogaert*, nnl. *begard* 'lekenbroeder'. Dit woord bevat het suffix *-aard*, later ook geëvolueerd tot *-erd*, maar heeft veelal de vorm *-ard* (mogelijk omdat het via het Latijn of Frans herontleend is?). De meest recente vorm is *begard*; vandaar hebben alle genoemde woordvormen het lemma BEGARD.

9.5.2.2. Enkele specifieke suffixen 1: *-e*

-e is in de volgende gevallen een suffix en wordt dus gehandhaafd in het GiGaNT-lemma:

- bij (doorgaans uitheemse) afgeleide vrouwelijke persoonsnamen: *scholiere*, *kameniere*, *assistente*.
- (uit *ī*, nu verouderd) bij abstracta afgeleid van adjectieven: *hoge* 'hoogte', *koude* 'kou'. Dus mnl., nnl. *evenhoge* znw. 'bouwsel van gelijke hoogte' (als verdedigingswerk) heeft het lemma EVENHOGE. (Hier zou het evt. ook gesubstantiveerd adj. kunnen zijn.)
- bij verouderde afleidingen voor mannelijke persoonsnamen

mnl. herde	HERDE
mnl. schutte	SCHUTTE
mnl. wrechte	WRECHTE

In deze woorden is *-e* later vervangen door *-er*. Verder nog o.a. *schinke*, *kempe*, *erve*, *winne*; zie De Vooys 1967: 203. Soms komen er *-e*-loze vormen voor, zoals in het WNT *win* (I), het lemma wordt toch *winne*.

NB 1: eveneens een suffix is het mnl. bijwoordelijke *-e* (onl. *-o*), maar dit wordt niet gelemmatiseerd; dit is dus een uitzondering op de regel van het handhaven van oude affixen! Deze uitgang is afgesleten en niet vervangen door een ander. Dus o.a.

onl. armliko	ARMELIJK
mnl. gevankelike	GEVANKELIJK

en vele andere woorden op mnl. *-like* / onl. *-liko*

NB 2: geen suffix, maar wel gehandhaafd in het lemma is de gelexicaliseerde uitgang van gesubstantiveerde adjectieven zoals *schone* 'mooie vrouw', *blinde* 'niet-zierende persoon' en in (op adjectieven gebaseerde) vrouwelijke inwonersnamen als *Leidse*, *Berlijnse*.

Zie verder de voorbeelden onder *e*-apocope, 9.3.2.

9.5.2.3. Enkele specifieke suffixen 2: *-(e)de/-te*

Voor abstracta op *-(e)de/-te* (uit *-itha*) (zie De Vooys 1967: 215-216) wordt in de moderne lemmatisering één vorm gehanteerd: *-te* (resultaat van een klankontwikkeling). (In de historische woordenboeken zelf is dit verschillend gedaan.)

Voorbeelden:

nnl. hoogde	HOOGTE
-------------	--------

mnl. hogede	HOOGTE
mnl. lengede	LENGTE
onl. diupitha	DIEPTE

Afleidingen van vergelijkbare bnw., zoals *oud, lief, vol, dik, oud, rond, koud, stil, zwaar*, worden systematisch gelemmatiseerd met *-e* of *-te*.

Soms staat een *-e*-abstractum naast een *-te*-abstractum en is er dus sprake van twee verschillende woorden: bij *lang*: LANGE en LENGTE, bij *lief*: LIEVE en LIEFDE.

Soms is er geen *-te*-variant: bij *koud*: alleen KOUDE (staat naast KOU ook in GB), bij *gezond* alleen GEZONDE, bij *waard* alleen WAARDE, of er is géén *-e*-variant: bij *steil* alleen STEILTE.

In het volgende geval is er ook een aan het Duits ontleende vorm en zijn er dus drie woorden met elk een GiGaNT-lemma: bij *gemeen*: GEMENE (vgl. VMNW *gemene* (III), MNW *gemene* (II), WNT *gemeen* (III) als 2^{de} lemma), GEMEENTE en GEMEENDE (vgl. MNW, Duits leenwoord).

OPM: soms valt de *-e* vorm samen met gesubstantiveerd adjectief op *-e* of komt ernaast voor, of is er alleen maar dat gesubstantiveerd adjectief (MNW *groene*).

9.5.2.4. Enkele specifieke suffixen 3: feminiserende suffixen

Afgeleide vrouwelijke persoonsnamen krijgen altijd een eigen GiGaNT-lemma en worden naar suffix ook altijd onderscheiden (dat gebeurde in de woordenboeklemma's niet altijd). Dus:

lererse	LERERSSE (niet: lerares)
---------	--------------------------

De volledige reeks suffixen is in standaardvorm: *-in, -es* (heel soms *-esse*), *-ersse* (< *-eresse*), *-se* (< *-ske, -sche*), *-egge* (< *ige, -icghe, -ege* etc.), *-e, -nede, -ster*. Bij het onderscheid tussen *-se* en *-ersse* moet soms naar de citaatvorm worden gekeken. Er kunnen stapelsuffixen voorkomen (bijv. *-steregge*).

9.5.2.5. Enkele specifieke suffixen 4: diminutiefsuffixen

Suffixen die geëvolueerd zijn naar het moderne *-tje* krijgen steeds die vorm (of een allomorf). Dit geldt voor *-kijn, -kin, -ken, -sken, -gien, -gen, -tjen*, etc. In diminutiefsuffixen die verdwenen zijn in het Standaardnederlands en die ook het element *-īn* bevatten, wordt dit als *-ijn* weergegeven: *-sijn, -(e)lijn*.

9.5.2.6. Leensuffixen, suffixen in leenwoorden

Hier wordt per woord omgezet, omdat woorden met deze suffixen vaak via verschillende wegen ontleend worden; de suffixen hebben geen vaste vorm.

Voorbeeld 1: fra. *-on* > ndl. zowel *-on* als *-oen*.

Bastoen (ofr. *baston*) in VMNW, MNW, *baston* in WNT; aldaar ook *bastonneren* en *bastonnade*.

Het GiGaNT-lemma is dus BASTON.

Voorbeeld 2: fra. *-ir*: Mnl. meestal *-ieren*, later in de regel *-eren*. In de lemma's hangt het af van hoe het nu (of in het WNT) voorkomt.

VMNW <i>purgieren</i>	PURGEREN
VMNW MNW WNT <i>purgeren</i>	PURGEREN
VMNW <i>obligieren</i>	OBLIGEREN
VMNW MNW WNT <i>obligeren</i>	OBLIGEREN
MNW <i>faelgieren, faelgeren</i>	FAILLEREN
WNT <i>tabagieren</i>	TABAGIEREN
MNW WNT <i>drapenieren</i>	DRAPENIEREN
WNT <i>kamenieren</i>	KAMENIEREN

De laatste twee woorden zijn afgeleid van het znw. op *-ier*.

Voorbeeld 3: romaans/ofra. suffix *-oel* (zie VMNW *halsbergoel, stassoel*) komt ook via jonger Frans als *-ole, -ool* voor (MNW *citole, WNT kamizool*). Daarnaast is er ook een woord *karakol* dat via het

Spaans is ontleend (met mv. op *-ols* en *-ollen*). Deze worden niet alle op *-ool* gelemmatiseerd maar per woord bekeken:

VMNW halsbergoel	HALSBERGOEL
VMNW wamsoele	WAMZOOOL (naar MNW)
MNW wamsool	WAMZOOOL
MNW citole	CITOOOL
	(moderne aanpassing: e-apocope)
WNT karakol, krakol, karkol	CARACOLE (want zo in GB)

Voorbeeld 4: fr. *-ien* bij inwonersnamen; hier wordt in het moderne lemma het Franse suffix aangepast, analoog aan latere ontwikkelingen in het Nederlands.

VMNW galilien	GALILIJN
Vgl. MNW cahorsijn	CAHORSIJN, KAUWERSIJN
WNT parizijn	PARIZIJN

Bij 5.3. Vreemde woorden

9.5.3.1. Voorbeelden van aanpassingen

Voorbeeld 1, bij ofr. *vaissel(e)ment* ' vaatwerk':

VMNW *facelment*, MNW *facelment*, verwijslemma *vaselment*; WNT *vassel(le)ment*, verwijslemma *vaselment*: dit wordt één lemma: VASELMENT (soort grote gemene deler).

Voorbeelden 2, aanpassing aan modern Nederlands:

Een oorspronkelijke, maar in het Nederlands weggevalen klinker of lettergreep wordt niet toegevoegd. Er komen dus aparte lemma's: SELEREN naast ASSELEREN.

Als beide vormen van een woord gelexicaliseerd zijn en in de woordenboeken opgenomen, blijven dat afzonderlijke lemma's:

mnl., nnl. dangier (< fr. dangier)	DANGIER
mnl., nnl. angier (< fr. dangier)	ANGIER

(Bij *angier*: metanalyse; *d-* is opgevat als lidwoord).

Per geval moet gekeken worden of het woord (of een deel ervan) moderne varianten kent:

mnl. bascoert > BASKOERT, en niet BASCOURT (< ofr. *basse-court*)

Toelichting: wegens simplex in WNT KOERT (II).

Voorbeeld 3:

Een leenwoord kan de vorm hebben van het oorspronkelijke vreemde woord, maar later daarnaast meer/geheel vernederlandst zijn. Het gaat dan om verschillende woorden met elk een eigen GiGaNT-lemma.

mnl. fenigrecum	FENIGRECUM
(< lat. <i>foenum graecum</i> 'Grieks hooi')	
mnl. fenigreec, venigriec	FENEGRIEK
nnl. fenegriek	FENEGRIEK

9.5.3.2 Verdwenen vreemde woorden

Verdwenen vreemde woorden volgen in principe ook de klankwijzigingen, zoals e-apocope, als er een vorm met zo'n klankwijziging aangetroffen is:

mnl. *daet* (< fr. *date*, lat. *datum*)

DAAT

Zo ook mnl., nnl. *date*

DAAT

Als van een verdwenen leenwoord niet goed uit te maken is waarheen het geëvolueerd zou zijn, dan moeten we behoudend lemmatiseren en de historische vorm overnemen.

mnl. <i>seme</i> 'barbeel'	SEME
mnl. <i>semeil</i> (onzekere betekenis)	SEMEIL
mnl. <i>bise</i> 'hinde' (< fr. <i>biche</i>)	BISE

9.5.3.3. Voorbeelden van meervoudige ontlening

Voorbeeld 1, meervoudige ontlening met minimaal klankverschil, en toch verschillende lemma's:

mnl., nnl. porie	PORIE
mnl. pore, porre	POOR

Pore < lat. *porus*. Wsch. is later het lat. mv. *pori* als ev. opgevat, waarna stapelmv. *poriën*, rückbildung ev. *porie*.

Niet alle details van de etymologie wordt uitgedrukt in het lemma. Zie *mole* – *molen* (met verschillende etymologieën!).

Voorbeeld 2:

onl. mola (< ofr. mole < lat. mola)	MOOL
onl. mulina (< mlat. molina, mulina)	MOLEN
mnl. muele (< lat. molinae, oude ontl.)	MOOL
mnl. muelin (< mlat. molina, jonge ontl.)	MOLEN
mnl. mole	MOOL
mnl. molene	MOLEN
nnl. molen	MOLEN

Samenst. als *watermolen* krijgt steeds dit lemma, ook al komt het met en zonder *-n* voor als historisch lemma. *Molegaren* bevat een waarschijnlijk een ander eerste lid. Het EWN interpreteert *mole* overigens als resultaat van herinterpretatie van de buigingsuitgang.

Voorbeeld 3:

In enkele gevallen zijn de vormen niet goed uit elkaar te halen en wordt één lemma toegekend. Bijv. bij de vormen van *glos*, *glosse*, *gloze* (gebaseerd op 2 lat. vormvarianten die elkaar in de tijd afwisselden) en hun afleidingen en samenstellingen wordt het GiGaNT-lemma GLOS etc. (GB heeft *glos* en *glosse* als varianten in één lemma).

Voorbeeld 4: Ook als het om een vormvariant gaat, zonder betekenisverschil, kennen we verschillende lemma's toe:

mnl. epelence, epilenci (ook varianten met -m-)	EPILENTIE
mnl. hipilentie	EPILENTIE
nnl. epilentie	EPILENTIE
nnl. epilepsie	EPILEPSIE

De vorm *epilepsie* heeft *epilentie* in de loop van de tijd vervangen. (EWN hierover: “De vormen met *-ms-* en *-nc-* zijn waarschijnlijk overgenomen uit Oudfrans *epilence* (1265; Rey), dat via het vulgair Latijn is ontstaan uit hetzelfde Latijnse woord [*epilepsia*], en door de medische wetenschap daar later door is vervangen.”)

Literatuur

Bree, Cor van, *Historische taalkunde* (1996/2003). Leuven / Amersfoort:Acco; 2^{de} druk
http://www.dbnl.org/tekst/bree001_histo1_01/

Haas, W. de en M. Trommelen (1993), *Morfologisch Handboek van het Nederlands*. 's-Gravenhage: SDU Uitgeverij.

Horst, J.M. van der (2008), *Geschiedenis van de Nederlandse syntaxis*. Leuven: Universitaire Pers, 2008. Deel 1.

Karsdorp, Folgert B. (2011). *Achtergronden bij de morfologische module van GiGaNT*. Leiden, Instituut voor Nederlandse Lexicologie. Interne nota.

Loey, A. van (1964⁷), *Schönfelds Historische Grammatica van het Nederlands*. Zutphen: Thieme.

Loey, A. van (1980, 1976), *Middel nederlandse spraakkunst*. I. *Vormleer*. Groningen: Wolters-Noordhoff. 9^{de} druk. II. *Klankleer*. Groningen: H.D. Tjeenk Willink, 1976. 7^{de} herziene druk.

Mooijaart, Marijke (2010). 'The complete history? Dutch words in four historical dictionaries'. In: John Considine (ed.), *Current projects in historical lexicography*. Newcastle: Cambridge Scholars Publishing. p. 83-98.

Mooijaart, Marijke (2011), 'Van acht (507) tot zwangerschap (1851): historische lexicografie van boek tot internet'. In *Nederlandse Taalkunde* 16, 2 (Discussie Historische woordenboeken op internet), 141-152.

Pijnenburg W. (1993). '100 jaar achtig. Over de wording van een samengesteld suffix', in: *ABäG* 37, 33-48.

Ruitenbergh, Tilly†, Jesse de Does & Katrien Depuydt (2010). 'Developing GiGaNT, a lexical infrastructure covering 16 centuries'. In: *Proceedings of the XIV Euralex International Congress*. Leeuwarden: Fryske Akademy. 468-476.
http://www.inl.nl/images/stories/onderzoek_en_onderwijs/publicaties/euralex2010/ruitenbergh_doe_s_depuysdt.pdf

Ruitenbergh, Tilly en Katrien Van pellicom i.s.m. Jesse de Does en Katrien Depuydt (2012). *De morfosyntactische component van het GiGaNT-lexicon*. Leiden, Instituut voor Nederlandse Lexicologie. Interne nota.

Salverda de Grave, J.J. (1906). *De Franse woorden in het Nederlands*. Amsterdam.

Sijs, Noline. van der (2004), *Taal als mensenwerk. Het ontstaan van het ABN*. Den Haag: Sdu Uitgevers.

Svensén, Bo, *A Handbook of Lexicography* (2009). *The Theory and Practice of Dictionary-Making*. Cambridge [etc.]: Cambridge University Press.

Toorn, M.C. van den et al. (red.) (1997), *Geschiedenis van de Nederlandse taal*. Amsterdam:

Amsterdam University Press. <http://www.dbnl.org/titels/titel.php?id=tooroo4gescoi>

Vooys, C.G.C. de (1967). *Nederlandse Spraakkunst*. Herzien door M. Schönfeld. Groningen: J.B. Wolters. Ongewijzigde zevende druk, van vierde druk 1957.

Wal, Marijke van der & Cor van Bree (2008). *Geschiedenis van het Nederlands*. Houten: Het Spectrum.

Willemyns, Roland, & Wim Daniëls (ed.) (2003), *Het verhaal van het Vlaams, de geschiedenis van het Nederlands in de Zuidelijke Nederlanden*. Antwerpen/Utrecht: Standaard/Het Spectrum.

Verder geraadpleegd:

Spelling:

Technische handleiding: http://taalunieversum.org/spelling/download/technische_handleiding.pdf

Woordenlijst Nederlandse Taal: <http://woordenlijst.org/>

Historische woordvormen en etymologie:

Historische woordenboeken van het Nederlands met inleidingen (ONW, VMNW, WNT, MNW):

<http://gtb.inl.nl>.

Etymologische woordenboeken:

EWN en andere woordenboeken op www.etymologiebank.nl.