

1. Inleiding¹

1.1. Het doel van de reeks

In het tegenwoordige onderzoek van de oudere literatuur heeft de traditionele esthetische visie plaatsgemaakt voor een historische benadering. Zag de literatuurwetenschap voorheen de literatuur in de eerste plaats als een vorm van kunst die de taal als uitdrukkingsmiddel gebruikt, nu bekijkt ze de literatuur vooral ook als een cultuurhistorisch verschijnsel. Daardoor gaat de aandacht niet alleen uit naar de tekst zoals de scheppende kunstenaar hem bedoeld heeft, maar ook naar de tekst zoals die concreet overgeleverd is en gefunctioneerd heeft. Afzonderlijke teksten worden bestudeerd in het verband waarin ze in de middeleeuwse handschriften geplaatst zijn. Daarbij komen ook teksten aan bod die uit literair-esthetisch oogpunt minder groots, fraai of kunstig zijn, zoals artesliteratuur, geestelijk proza, gebruiksteksten en de kleine vormen in het algemeen. Ook komen zo spreukenverzamelingen, liedbundels, compilaties en andere vaste verbanden in het vizier, kortom overleveringsgehelen die het niveau van de afzonderlijke teksten overstijgen. Deze literaire verschijnselen konden moeilijk geplaatst worden in de literair-esthetische opvatting van literatuur als oorspronkelijke schepping van een groot kunstenaar.

De auteur zal met zijn oeuvre altijd een belangrijk uitgangspunt en brandpunt van het letterkundig onderzoek blijven. Er zal dan ook behoefte blijven aan uitgaven die de oorspronkelijke tekst van de auteur trachten te reconstrueren. De cultuurhistorische visie op literatuur vraagt daarnaast om uitgaven die de concrete historische vorm weergeven waarin teksten (voor)gelezen en gebruikt werden. Verzamelhandschriften bieden vele invalshoeken voor dergelijk onderzoek. Deze materieel-tekstuele eenheden stellen het historische verband van de overlevering in de plaats van tekstkeuzen die gewoonlijk gebaseerd zijn op moderne, a-historische, vaak literair-esthetische overwegingen. Wanneer de literatuurwetenschap verzamelhandschriften als een van haar uitgangspunten neemt, verplicht ze zich ook aandacht te schenken aan de minder fraaie literatuur en aan de concrete context waarin de teksten indertijd geplaatst werden. Deze nog altijd actuele overwegingen vormden in 1994 de motieven voor de uitgave van middeleeuwse verzamelhandschriften uit de Nederlanden.

Mede door de uitgave van een reeks delen is de discussie over wat wij onder een verzamelhandschrift dienen te verstaan, goed op gang gekomen. Het oorspronkelijke uitgangspunt dat een verzamelhandschrift een codex is die van meet af aan als materiële eenheid bedoeld is en waarin tenminste twee teksten zijn bijeengebracht, bleek, hoewel in de grond niet onjuist, toch bepaalde nuancerings te behoeven. Reflectie op dit onderwerp door verscheidene onderzoekers heeft tot een precieze conceptualisering van het begrip verzamelhandschrift geleid, een begrip waarachter in feite een veelheid aan historische verschijningsvormen van de middeleeuwse codex schuilgaat. [nadere invulling door Hans Kienhorst?]

Ondanks deze formele beschrijving van het begrip ‘verzamelhandschrift’ heeft deze reeks zich tot dusverre altijd in de eerste plaats gericht op handschriften waarvan het verzamelkarakter op de voorgrond treedt. Bekende voorbeelden zijn het in de reeks opgenomen handschrift-Van Hulthem en het Comburgse handschrift, die hoewel ze verzamelhandschriften van volkomen verschillende aard zijn, toch juist het verzamelaspect gemeenschappelijk hebben. Ook voor toekomstige uitgaven zal bij voorkeur worden gekozen voor in dit opzicht vergelijkbare handschriften.

Het doel van de reeks diplomatische uitgaven is op korte termijn verantwoorde grondslagen te bieden voor literair-historisch, taalkundig of ander onderzoek van middeleeuwse teksten in de context waarin ze indertijd overgeleverd zijn. Het gaat er hier vooral om materiaal te leveren voor diepgaande studies en niet zozeer om afgeronde resultaten van dergelijke studies te presenteren. De uitgaven zullen (voor zover dat bij diplomatische uitgaven mogelijk is) vlot leesbaar zijn, echter zonder filologische en cultuurhistorische aantekeningen, die immers gedegen en langdurig vooronderzoek vergen.

De diplomatische editie kan snel op ontwikkelingen in het onderzoek inspelen doordat zij in verhouding weinig onderzoek vereist. En ‘snel’ betekent in dit verband niet kortstondig of vluchtig. De diplomatische editie is immers geen voorlopige of lagere vorm van tekstuitgave. Als basisvorm van de uitgave blijft zij de hechte grondslag voor verdere studie en eventuele verdere, tekstkritische uitgaven. Zij bevat minder elementen die snel kunnen verouderen dan de kritische edities die voorzien zijn van uitgebreide filologische apparaten en letterkundige inleidingen. De combinatie van relatieve snelheid en grote duurzaamheid waren aanvankelijk de motieven om de methode van de diplomatische uitgave te verkiezen voor deze reeks en deze gelden nog steeds.

In de inleiding op de tekstuitgave wordt dezelfde doelmatigheid nagestreefd. Uitsluitend die informatie wordt gegeven die op het uitgegeven handschrift betrekking heeft. Hierbij wordt bijzondere aandacht besteed aan de materiële, codicologische aspecten van het verzamelhandschrift dat uitgegeven wordt.

Nadat in de editiereeks *Middeleeuwse Verzamelhandschriften uit de Nederlanden (MVN)* een aantal publicaties in boekvorm verscheen, is binnen de projectcommissie de behoefte ontstaan om voortaan naast uitgaven in boekvorm ook edities in digitale vorm te publiceren. Met de digitalisering en beschikbaarstelling van het eerste deel uit de MVN-reeks, *Het Geraardsbergse handschrift*, werd in 2003, bij wijze van proef, een nieuwe weg ingeslagen. Om een zo groot mogelijke duurzaamheid te waarborgen werd besloten de digitale edities in te richten in XML-formaat. Voor de codering van de teksten kon worden aangesloten bij de standaarden die werden ontwikkeld door

¹ De tekst van dit hoofdstuk is grotendeels overgenomen uit Mertens 1994.

het Text Encoding Initiative (TEI Guidelines, versie P5). Slechts in gevallen waarin de aanbevelingen van het TEI niet voorzagen moest worden gezocht naar specifieke oplossingen voor coderingsvraagstukken. *[PB: we streven er naar ook edities te kunnen aanmaken binnen het eLaborate platform].*

Deze richtlijnen zijn allereerst bestemd voor editoren die van plan zijn een deel te vervaardigen in de MVN-reeks. Ze zullen tevens als uitgangspunt dienen bij de vervaardiging van een digitale werkomgeving waarin veel van de XML-codering onzichtbaar kan blijven voor de terzake minder ervaren editor.

De basis van een MVN-editie wordt gevormd door een diplomatische uitgave van de teksten uit een handschrift. In een nieuw op te zetten reeks zullen ook kritische edities worden uitgebracht. Vooralsnog is het de bedoeling om de kritische edities in boekvorm te laten verschijnen en de digitale vorm voor te behouden aan de diplomatische uitgaven. Dit sluit overigens niet uit dat diplomatische uitgaven in bepaalde gevallen ook in boekvorm kunnen verschijnen. In elk geval zal aan elke editie, of ze nu alleen digitaal of ook als boek verschijnt, een XML-bestand ten grondslag liggen. *[PB: dit is dus niet altijd het geval].* In deze richtlijnen wordt getoond hoe een MVN-editie gestalte krijgt als XML-bestand.

Voor kritische uitgaven zijn afzonderlijke richtlijnen beschikbaar.

De voorbeelden in dit document gebruiken de XML-syntax. In het algemeen illustreren ze echter conceptuele vragen over de opbouw van de editie. Van toekomstige editoren wordt niet noodzakelijkerwijs diepgaande vertrouwdheid verwacht met de XML-syntax.

Een opmerking nog over de programma's die gebruikt kunnen worden bij de vervaardiging van een digitale MVN-editie. In beginsel wordt er aan editoren een werkomgeving aangeboden binnen de XML-editor Oxygen. Natuurlijk is het ook mogelijk om te werken met andere programma's, voorwaarde is echter wel dat deze weten om te gaan met Unicode (in [hoofdstuk 6](#) worden de Unicode-tekenen opgesomd die voor MVN-edities gebruikt kunnen worden).

1.2. Wat is een diplomatische uitgave?

Onder diplomatische uitgave wordt hier een editie verstaan waaraan slechts één handschrift ten grondslag ligt en waarin de tekst en alle verdere grafische informatie weergegeven worden in overeenstemming met dit handschrift. Daarbij wordt niet gestreefd naar een typografische imitatie van het handschrift, maar naar een functionele weergave van de tekst. De functie van alle lettervormen en overige grafische verschijnselen uit het handschrift die de structuur en de betekenis van de tekst tot stand brengen, wordt zodanig gecodeerd dat er een typografische weergave in een webbrowser mogelijk wordt. Een belangrijke kanttekening hierbij is dat dezelfde grafische verschijnselen in het ene geval tot de tekst kunnen behoren en in het andere geval tot de esthetische en functionele aspecten van de tekstopmaak. Zo kan de kopiist een regel wit laten omdat hij onderaan de kolom niet aan een nieuwe tekst wil beginnen, terwijl hij elders bij het tekstbegin geen regel blanco laat. Zo'n witregel is in dat geval eerder een zaak voor de codicoloog-paleograaf dan voor de filoloog-teksteditor.

2. De onderdelen van een digitale MVN-editie

Een editor van een digitale MVN-editie vervaardigt de volgende onderdelen:

- Een XML-bestand met de transcriptie van de uitgegeven tekstverzameling
- Een XML-bestand met het voorwerk, de inleiding en het bijwerk *[PB: hier gaan we niet meer van uit. Voorwerk, inleiding en bijwerk worden voorlopig aangemaakt in Wordpress. In de toekomst hopen we dat eLaborate hier in voorziet.]*
- Een set digitale foto's van het uitgegeven handschrift. Indien mogelijk wordt een complete set aangeleverd.

De inhoud van de twee XML-bestanden komt tot op zekere hoogte overeen met het onderscheid in de gedrukte delen tussen het deel van het boek met de diplomatische weergave van de teksten uit het handschrift en het overige deel van de editie. Verschillen zijn onder meer de plaats van de Verantwoording van de editie, die nu is opgenomen in de het bestand met de transcriptie en de Registers, die nu zijn geïntegreerd in het transcriptiedeel.

In de volgende hoofdstukken wordt nauwkeurig uitgelegd hoe de beide XML-bestanden dienen te worden vormgegeven ([hoofdstuk 3: De inleiding en het bijwerk](#) en [hoofdstuk 5: De globale structuur van het transcriptiedeel](#)).

De twee bestanden verschillen behalve door hun inhoud ook in de mate van vrijheid die de editor heeft. Voor het bestand met de inleiding is de editor slechts gehouden aan enkele algemene richtlijnen die dienen ter structureren van zijn in principe vrije tekst. Voor het bestand met de transcriptie wordt gebruik gemaakt van een speciaal voor de reeks ontwikkeld MVN-sjabloon (een zgn. 'template'). In het MVN-sjabloon is de hoofdstructuur van het XML-document, zoals dat nodig is voor de vervaardiging van de transcriptie reeds voorhanden.

Elk TEI-document bestaat uit een <teiHeader>, dat is een gedeelte met meta-informatie over de overige tekst, en een <text> element, waarin de eigenlijke inhoud van het document is opgenomen. De wijze waarop de afzonderlijke onderdelen van een TEI-bestand moeten worden vormgegeven wordt besproken in de hoofdstukken 3 en 5. Voor de codering van de transcriptie bestaan nauwkeurige richtlijnen die worden uiteengezet in de hoofdstukken 5 en 6.

3. De inleiding en het bijwerk

PB: bijwerk komt dus in WordPress. Elke MVN-editie bestaat uit een inleiding, een tekstuitgave en verschillend bijwerk (samenvatting, bibliografie, lijst van afbeeldingen, etc.). Deze onderdelen, behalve de tekstuitgave, worden ondergebracht in een afzonderlijk XML-bestand dat wordt gecodeerd in TEI-Lite. Als elk TEI-document valt dit document in twee delen uiteen:

- de <teiHeader> van het bestand met de inleiding en het bijwerk
- een <text> element met daarin de eigenlijke tekst van de inleiding en het bijwerk

3.1. De <teiHeader> van het bestand met de inleiding en het bijwerk

[nog te schrijven]

3.2. Voorwerk, inleiding en bijwerk

Het gedeelte na de <teiHeader> bestaat uit vrije tekst van de editeur. Nadere richtlijnen voor het schrijven van de inleiding en het bijwerk worden gegeven in [hoofdstuk 3: De inleiding en het bijwerk](#).

Hier volgt een korte toelichting op de inhoudelijke aspecten van de diverse onderdelen van voorwerk, inleiding en bijwerk. Vormgevingsaspecten zijn op deze plaats niet aan de orde, maar wel schenken we aandacht aan elementen van de presentatie die inhoudelijke keuzes met zich meebrengen, zoals de benaming van het handschrift, titelconventies, bestanddelen van de inhoudopgave, e.d.

Alle onderdelen van de editie worden in principe in het Nederlands geschreven. Hierbij wordt de voorkeurspelling toegepast. Gebruik van een andere taal dan het Nederlands is mogelijk, indien daar een gegronde reden voor bestaat.

3.2.1. Titel

Als hoofdtitel neemt de editeur de aanduiding voor het handschrift die traditioneel in het onderzoek gebruikt wordt. Als het handschrift nog niet zo'n 'roepnaam' heeft, dan bedenkt de editeur er een. De ondertitels en de vermelding van de editeur hebben een vaste vorm: Hs. [bewaarplaats, signatuur]. Diplomatische webeditie bezorgd door [naam/namen van de editeur (s)]. Als de editeur bepaalde hoofdstukken niet zelf heeft geschreven worden de namen van de auteurs van die hoofdstukken aan de vermelding van de editeur toegevoegd, bijvoorbeeld: met een codicologische beschrijving door [naam].

3.2.1.1. Roepnaam

Meestal bestaat de 'roepnaam' uit het woord handschrift en een nadere omschrijving (bijv. Het Comburgse handschrift, Het handschrift-Van Hulthem). Bij het bedenken van een roepnaam is het niet verstandig de huidige bewaarplaats als (enige) nadere omschrijving te gebruiken omdat dit te weinig kenmerkend is. Dit geldt in het bijzonder voor bewaarplaatsen als Berlijn, Brussel, Den Haag en Wenen.

3.2.2. Inhoudsopgave

Een gedetailleerde inhoudsopgave van de uitgave als geheel hoeft niet door de editeur te worden verzorgd. Dit overzicht wordt voor de webpagina afgeleid door het style-sheet. Een overzicht van de inhoud van het handschrift wordt wel door de editeur gemaakt.

3.2.3. Karakteristiek

Het eerste onderdeel van de inleiding is een samenvattend overzicht van de aard en inhoud van de tekstverzameling. In dit overzicht wordt gestreefd naar een karakterisering van de verzameling als geheel. Het is daarbij niet de bedoeling dat aan (alle) afzonderlijke teksten een bespreking wordt gewijd, maar dat veeleer de grote lijnen en de eigenaardigheden van de verzameling worden belicht.

3.2.4. Geschiedenis van het onderzoek

Op deze plaats geeft de editeur een bondige uiteenzetting van het eerdere onderzoek. Daarbij wordt vooral gelet op het onderzoek dat zich specifiek op het handschrift als tekstueel-materiële eenheid heeft gericht. Publikaties waarin het handschrift slechts gebruikt wordt als getuige van de overlevering, zijn uit dit oogpunt minder van belang.

3.2.5. Taaleigen

In een systematische, maar niet-uitputtende beschrijving wordt het taaleigen behandeld van de verschillende kopiïsten van het handschrift.

3.2.6. Inhoudsopgave van het handschrift

Een gedetailleerde inhoudsopgave van de teksten in het verzamelhandschrift vermeldt de teksten die in het handschrift voorkomen, met hun opschrift en/of beginwoorden en de foliumnummers, bladzijde- en kolomaanduidingen. De teksten worden in dit overzicht genummerd. Als een tekst reeds eerder naar ditzelfde handschrift is uitgegeven, wordt de uitgave hier vermeld. Is een tekst meermalen uitgegeven, dan wordt in elk geval de uitgave

genoemd die het meest verantwoord is. Daarbij kan de editor nog overwegingen laten meespelen, zoals de vraag welke uitgave het meest gangbaar of het recenst is, als het maar om een uitgave gaat die naar hedendaagse wetenschappelijke normen verantwoord wordt geacht. Desgewenst vermeldt de editor meer dan een uitgave. De vermelding van een eerdere uitgave van een tekst gebeurt door middel van de gebruikelijke literatuurverwijzing (zie).

3.2.6.1. Tekstnummering

Heeft het handschrift een oorspronkelijke tekstnummering – zoals in het handschrift-Van Hulthem bijvoorbeeld het geval is – dan wordt deze overgenomen. Bij het coderen wordt het oorspronkelijke tekstnummer in het betreffende `<text>` element als waarde van het `@n` attribuut gegeven. Het `@xml:id` attribuut representeert dan de hiërarchische geleiding van de tekstverzameling, zoals die door de editor is vastgesteld (zie verder). Kent het handschrift geen oorspronkelijke tekstnummering, dan worden de teksten genummerd door de editor. Soms kan het wenselijk zijn dat de editor een in het handschrift voorkomende tekstnummering aanvult of nader uitwerkt, zoals bijvoorbeeld in de editie van het handschrift-Van Hulthem (Brinkman & Schenkel 1999) gebeurt (vgl. aldaar de verzameling spreuken, tekst 108.1-234).

3.2.6.2. Identieke teksten

Bij teksten die meer dan eens in het handschrift voorkomen, worden in de inhoudsopgave links aangebracht. Als teksten tevens in andere webedities binnen de MVN-reeks voorkomen, wordt eveneens een link aangebracht.

3.2.6.3. Tekstgrenzen

Juist in verzamelhandschriften kunnen zich twijfels omtrent de begrenzing van de afzonderlijke teksten voordoen. In dit geval stelt de editor een oplossing voor en vermeldt hij zijn twijfels in de annotatie.

3.2.7. Overzicht van afkortingen

Een lijst van afkortingen met voorbeelden van hun oplossingen wordt, op basis van de codering gegenereerd in de webeditie. Opmerkingen over speciale gevallen maakt de editor op deze plaats.

3.2.8. Literaturopgave

De literaturopgave (element `<listBibl>`) wordt als laatste hoofdstuk opgenomen. Bibliografische referenties in de tekst zelf en in het transcriptiedeel bestaan uit verkorte titels (Berteloot 2002, of Verhandelingen 1883). Door middel van het `<ref>` element wordt in `@target` attribuut naar de bibliografie verwezen. De afzonderlijke titelbeschrijvingen staan in het een `<bibl>` element dat wordt voorzien van een `@xml:id` attribuut dat bij voorkeur de vorm heeft van de verkorte bibliografische referentie, evenwel zonder de daarin voorkomende spaties of het &-teken. In de lopende tekst kan een verijzing er zo uitzien:

De eerste paragraaf van deze inleiding is ontleend aan `<ref target="#Mertens1994">Mertens 1994</ref>`.

Daarbij verwijst `#Mertens1994` naar de bibliografie waar de titel als volgt is beschreven:

```
<bibl xml:id="Mertens1994">
  <author>Mertens, Th.</author>
  <title>Richtlijnen voor de uitgave van Middeleeuwse Verzamelhandschriften uit de Nederlanden</title>.
  <pubPlace>Hilversum</pubPlace>:
  <publisher>Verloren</publisher>,
  <date>1994</date>.
</bibl>
```

3.2.9. Afbeeldingen

Indien mogelijk worden digitale kleurenopnamen aangeleverd van alle bladzijden van een handschrift, alsmede van de schutbladen en de band. De bestanden met de foto's van de beschreven bladzijden worden gekoppeld aan de codering van de bladzijden in het transcriptiedeel. In het bestand met de inleiding kunnen verwijzingen worden opgenomen naar afbeeldingen die ter illustratie dienen.

3.2.10. Samenvatting

De editor voorziet de uitgave van een Engelstalige samenvatting van de inleiding. Als daar goede redenen voor zijn kan de samenvatting ook nog in een andere taal gepubliceerd worden. De uitgave van een verzamenhandschrift uit het Nederlands-Duitse grensgebied verdient bijvoorbeeld ook een samenvattende karakteristiek in het Duits; een handschrift waarin zowel Nederlandse als Franse teksten voorkomen, ook een samenvatting in het Frans.

4. Richtlijnen voor de beschrijving van het handschrift

In de beschrijving van het handschrift moet tenminste aandacht worden besteed aan:

- Materiaal en samenstelling
- Watermerken (indien het handschrift uit papier is vervaardigd)
- Bladsignaturen en custoden
- Opmaak

- Schrift
- Tekstgeleding en rubricering
- Band
- Datering en lokalisering
- Geschiedenis
- Eigendomsnotities
- Gebruikerssporen

In een paragraaf *Besluit* worden de meest relevante bevindingen uit het betreffende hoofdstuk bondig samengevat en worden zo mogelijk conclusies getrokken.

5. Het transcriptiebestand

Het transcriptiebestand valt uiteen in een gedeelte met meta-informatie over de transcriptie en de getranscribeerde bron, en een deel dat de eigenlijke weergave van de teksten uit het handschrift bevat:

- de `<teiHeader>` van het transcriptiebestand, waarin o.m. een beknopte beschrijving van het handschrift en een verantwoording van de editiewijze
- een `<text>` element met de transcriptie, waarin ook de annotatie en de lemmatisering voor de registers worden geïntegreerd.

5.1. De `<teiHeader>` van het transcriptiebestand

De informatie die door het element `<teiHeader>` wordt gemarkeerd, is afhankelijk van het gegeven of de digitale editie wordt vervaardigd naar een reeds bestaande boekpublicatie (zoals het geval was bij de digitalisering van MVN deel 1: het Geraardsbergse handschrift), of dat er rechtstreeks wordt geëditteerd naar een handschriftelijke bron. We bespreken hier dit laatste geval, omdat deze situatie zich naar verwachting het meest zal voordoen. Retro-digitalisering wordt in een aparte paragraaf behandeld [nog te schrijven].

Wanneer een digitale editie wordt vervaardigd op basis van een tekstverzameling in handschrift, is het element `<teiHeader>` in hoofdlijnen volgt gestructureerd.

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>
<!-- Titel -->
      </title>
      <author>
<!-- Auteur -->
      </author>
    </titleStmt>
    <publicationStmt>
      <p>
<!-- Gegevens over uitgave -->
      </p>
    </publicationStmt>
    <sourceDesc>
      <p>
<!-- Beknopte beschrijving van het handschrift -->
      </p>
    </sourceDesc>
  </fileDesc>
  <encodingDesc>
<!--
    <xi:include href="http://www.huygensinstituut.knaw.nl/projecten/mvn/webfiles/charDecl.xml" xmlns:xi="http://www.w3.org/2001/XInclud
      <xi:fallback>
        <xi:include href="charDecl.xml"/>
      </xi:fallback>
    </xi:include>
-->
    <editorialDecl>
      <p>
<!-- Verantwoording van de editie -->
      </p>
    </editorialDecl>
  </encodingDesc>
  <profileDesc>
    <handNotes>
      <handNote xml:id="h1">
        <p>
<!-- Opmerkingen over de tekst- en kopiistenhanden -->
        </p>
      </handNote>
```

```

</handNotes>
</profileDesc>
</teiHeader>

```

De wijze waarop de tekst uitgegeven is, wordt verantwoord in het <editorialDecl> element. De editor verwijst hier naar de algemene richtlijnen voor de uitgave en omschrijft waar hij in deze uitgave van die algemene richtlijnen is afgeweken. Daarbij geeft hij steeds aan waarom hij dat gedaan heeft.

Het aantal handen wordt genoemd als waarde van het attribuut @hands bij het <handDesc> element.

```

<handDesc hands="2">
  <handNote xml:id="SIGLEalfa">
    <p>Het handschrift is geschreven door twee kopiïsten, alfa, die een textualis schreef (textualis)en</p>
  </handNote>
  <handNote xml:id="beta">
    <p>beta, die een cursiva schreef</p>
  </handNote>
</handDesc>

```

PB: ik denk eerder aan een structuur als de volgende:

```

<handDesc hands="2">
  <summary>Het handschrift is geschreven door twee kopiïsten, alfa en beta</summary>
  <handNote xml:id="SIGLEalfa">kopiïst alfa (textualis)</handNote>
  <handNote xml:id="SIGLEbeta">kopiïst beta (cursiva)</handNote>
</handDesc>

```

Een uitgebreide, ‘vertellende’ beschrijving van het handschrift, waaronder ook de opgave van de teksten is begrepen, is onderdeel van de inleiding tot de uitgave en vormt daarmee een onderdeel van het afzonderlijke XML-bestand waar ook de literatuuropgave en de samenvatting in worden opgenomen (zie hoofdstuk 3). Een beknopte, catalogus-achtige beschrijving van het handschrift krijgt een plaats in de teiHeader, en wel binnen het element <msDesc>, dat genest wordt in het element <sourceDesc> (zie hierboven). Een uitgewerkte toelichting op dit onderdeel is opgenomen als hoofdstuk 4 (Richtlijnen voor de beschrijving van het handschrift). [volgt]

5.2. Het deel met de eigenlijke transcriptie

Het geheel van de weergave van de teksten wordt beschouwd als <text> element. De annotatie en de registers worden op basis van lemmatisering buiten het document om gegenereerd en maken dus geen deel uit van de structuur van het XML-bestand.

De kern van de editie, de middeleeuwse tekstverzameling, wordt aangeduid als element <group>. Deze <group> bestaat uit een aantal <text> elementen: dat zijn de afzonderlijke teksten. Wanneer teksten bijeen horen en als het ware een verzameling binnen de verzameling vormen (zoals dat bijvoorbeeld bij spreken het geval kan zijn) worden zij opgevat als geneste <group> binnen de <group> die de tekstverzameling als geheel beschrijft.

Kort gezegd: elke <group> is samengesteld uit een of meer elementen <text>; en elk element <text> binnen de <group> die de tekstverzameling aanduidt, bevat weer het element <body>. Natuurlijk kan een <group> ook andere elementen <group> bevatten, bijvoorbeeld wanneer er sprake is van een deelverzameling binnen de tekstverzameling. Uiteindelijk zal een <group> toch bestaan uit elementen <text>. Die <text> elementen omvatten op hun beurt dan weer verplicht een element <body>.

De interne geleiding van de geëditteerde teksten kan, indien nodig, worden aangegeven door het element <div>. Dat is het geval bij een tekst die uit duidelijk te onderscheiden eenheden is opgebouwd.

De structuur van dit deel van de editie ziet er dus als volgt uit:

```

<teiHeader>...</teiHeader>
<text>
  <group>
    <text>
      <body/>
    </text>
  </group>
</text>

```

De tag van het element <text> die het transcriptiedeel omsluit krijgt een @xml:id attribuut met als waarde een door de editor zelf te bedenken sigle waarmee de editie in het hele document wordt geïdentificeerd. Ter voorkoming van verwarring bij zoek- en vervangopdrachten tijdens het werken aan de editie, verdient het aanbeveling om de sigle uit een (korte) string van hoofdletters te laten bestaan die zo mogelijk niet voorkomt in de uitgegeven tekst.

Aan elk nevenschikt of genest element <group> of <text> wordt ter identificatie een eigen @xml:id attribuut toegekend met een waarde die is samengesteld uit de handschriftsigle en het tekstnummer. Daarnaast wordt aan het element een @n attribuut toegekend met het tekstnummer zoals het in de editie moet verschijnen.

```

<group xml:id="SIGLE45" n="45">...</group>

```

De waarde van het @xml:id attribuut wordt gerelateerd aan de hiërarchische structuur van de tekstverzameling. De waarde van het @n attribuut volgt in principe de oorspronkelijke tekstnummering van het handschrift. Kent het handschrift geen oorspronkelijke tekstnummering dan brengt de editor de waarde van het @n attribuut in overeenstemming met het @xml:id attribuut.

De interne structuur van het <group> element dat de tekstverzameling omvat is dus (versimpeld weergegeven, maar met @xml:id en @n attributen) bijvoorbeeld als volgt (de sigle VVEVM staat voor het ‘Van vrouwen ende van minne-handschrift’):

```

<text xml:id="VVEVM">

```

```

<group>
<!-- Editie van de teksten -->
<text n="1" xml:id="VVEVM1">
<body>
<!-- Tekst 1 -->
</body>
</text>
<group n="2" xml:id="VVEVM2">
<text n="2.1" xml:id="VVEVM2.1">
<body>
<!-- Tekst 2.1 -->
</body>
</text>
<text n="2.2" xml:id="VVEVM2.2">
<body>
<!-- Tekst 2.2 -->
</body>
</text>
</group>
</group>
</text>

```

Het voorbeeld laat zien, dat wanneer er sprake is van zelfstandige teksten die om een inhoudelijke reden bijeen lijken te horen (bijvoorbeeld een collectie rijmspreuken), deze teksten kunnen worden samengenomen als `<group>`. Deze `<group>` bestaat dan weer uit afzonderlijke `<text>` elementen.

5.2.1. De codering van de fysieke structuur van het handschrift

In het editiegedeelte met de transcriptie komen verschillende structuren bij elkaar: de structuur van de tekstverzameling, de fysieke structuur van de codex en een XML structuur. Een MVN-editie wil immers niet alleen een tekstcorpus uitgeven zoals dat is overgeleverd in een bepaalde handschriftelijke bron, maar ook in de uitgave de relatie van de teksten tot hun fysieke bron tot uitdrukking laten komen. In verreweg de meeste gevallen zal de hiërarchische structuur van de tekstverzameling (bestaande uit teksten, die weer uit strofen of paragrafen bestaan) niet samenvallen met de hiërarchische fysieke structuur van de bron (bestaande uit katernen, die weer uit dubbel- of enkele bladen zijn samengesteld, elk met een recto- en versozijde, waarop een of meer kolommen, die elk weer al dan niet zijn gelinieerd).

Nu heeft elk XML-bestand per definitie één hiërarchische structuur die bestaat uit de markering van een groot geheel waarin kleinere eenheden zijn ingebed. Het is echter niet mogelijk om in een XML-document twee ongelijke hiërarchische structuren (hier die van de tekstverzameling en die van de codex) aan te brengen. Om toch beide structuren in het bestand aanwezig te laten zijn, is er gekozen voor een oplossing waarbij de structuur van de tekstverzameling op hiërarchische wijze de structuur van het desbetreffende gedeelte van het XML-document bepaalt (bijvoorbeeld: het tekstgedeelte van regel x tot regel y is een strofe). De structuur van de bron wordt weergegeven door middel van niet-hiërarchische overgangsmarkeringen (bijvoorbeeld: op deze plaats vindt er een overgang plaats van katern 1 naar katern 2).

De verschillende wijzen van coderen zou men kunnen vergelijken met een tolgeweg tussen twee bestemmingen. De teksten zijn trajecten tussen twee toelopen; de overgangsmarkeringen zijn de hectometerpaaltjes langs de route.

5.2.2. De afzonderlijke teksten en hun interne structuur

Elke afzonderlijke tekst wordt gecodeerd door middel van het `<text>` element. Aan dit element wordt een `@xml:id` attribuut toegevoegd dat is opgebouwd uit de sigle van het handschrift, gevolgd (zonder spatie) door het tekstnummer. Tevens wordt een `@n` attribuut toegevoegd met de waarde van het tekstnummer, zoals het in de editie zichtbaar moet worden.

```

<text xml:id="SIGLE63" n="63">...</text>

```

De hoofdtekst van een tekst bestaat bij proza uit een of meer `<p>` (paragraph) elementen en bij dichtwerk uit `<lg>` (line group) elementen (strofen). Voor een titel, opschrift, clauskop of tussenkopje wordt het element `<head>` gebruikt. Het element `<closer>` wordt gebruikt voor onderschriften, bijschriften, verstellingen, groeten, meta-opmerkingen (zoals 'ghecolationneert') of slotopmerkingen (Amen, Explicit) die geen deel uitmaken van de hoofdtekst.

Een prozatekst zou er gecodeerd dus, gesimplificeerd, zo uit kunnen zien:

```

<text xml:id="SIGLE63" n="63">
<body>
<head>Opschrift</head>
<p>
<!-- Tekst van de eerste paragraaf -->
</p>
<p>
<!-- Tekst van de laatste paragraaf -->
</p>
<closer>Explicit</closer>
</body>
</text>

```

Zoals men ziet komt er tussen de openings- en sluittags van het element `<text>` een element `<body>`. Indien gewenst kan een eventueel aanwezige (hiërarchische) substructuur in de tekst worden gecodeerd door middel van `<div>` (division) elementen die dan telkens één of meer `<p>` elementen omsluiten. De `<div>` elementen worden niet genummerd.

Een berijmde tekst van twee strofen zou er zo uit kunnen zien:

```

<text xml:id="SIGLE64" n="64">
<body>
<head>Opschrift</head>
<lg>

```

```

<1>Tekst van de eerste versregel</1>
<1>Tekst van de tweede versregel</1>
<1>Tekst van de derde versregel</1>
<1>Tekst van de vierde versregel</1>
</lg>
<lg>
<1>Tekst van de vijfde versregel</1>
<1>Tekst van de zesde versregel</1>
<1>Tekst van de zevende versregel</1>
<1>Tekst van de achtste versregel</1>
</lg>
<closer>Amen</closer>
</body>
</text>

```

Bij berijmde teksten worden de afzonderlijke versregels dus binnen een `<lg>` (line group) element genest door middel van het `<1>` (line) element. Line groups zelf kunnen ook genest voorkomen, anders gezegd: het is mogelijk dat een line group voorkomt binnen een grotere line group, bijvoorbeeld in het geval van een toneelstuk waarin een refrein voorkomt binnen de claus van een personage. Elk `<1>` element moet deel uitmaken van een `<lg>` element. Om elke versregel te identificeerbaar te maken, wordt aan een `<1>` element een `@xml:id` attribuut toegevoegd dat gelijk is aan het text `@xml:id` attribuut, aangevuld met een punt en een regelnummer. In een `@n` attribuut wordt dit regelnummer nog eens afzonderlijk gespecificeerd. Bijvoorbeeld:

```

<text n="62" xml:id="SIGLE62">
  <body>
    <lg>
      <1 n="1" xml:id="SIGLE62.1">Tekst versregel 1</1>
      <1 n="2" xml:id="SIGLE62.2">Tekst versregel 2</1>
      <1 n="3" xml:id="SIGLE62.3">Tekst versregel 3</1>
      <1 n="4" xml:id="SIGLE62.4">Tekst versregel 4</1>
    </lg>
  </body>
</text>

```

Ook strofen (`<lg>`) kunnen genummerd worden middels het `@n`-attribuut.

5.2.3.

5.2.3.1. Witregels

Zoals in het bovenstaande voorbeeld bij proza is te zien, worden witregels met een leeg `<lb>` (line break) element gemarkeerd, zowel in proza als in berijmde teksten.

5.2.3.2. Inspringing

Een strofe (line group) waarvan de regels inspringen kan men coderen met behulp van het `@rend` attribuut met als waarde 'indent':

```

<lg rend="indent">
  <1 n="1" xml:id="SIGLE62.1">Tekst versregel 1</1>
  <1 n="2" xml:id="SIGLE62.2">Tekst versregel 2</1>
  <1 n="3" xml:id="SIGLE62.3">Tekst versregel 3</1>
  <1 n="4" xml:id="SIGLE62.4">Tekst versregel 4</1>
</lg>

```

Als een dergelijke strofe een opschrift heeft dat zich ter rechter- of linkerzijde van de strofe bevindt, coderen we dit als `<head>` element met attribuut `@place` en waarde `margin-right` of `margin-left`. Het `<head>` element wordt in de transcriptie aangebracht op een plaats die overeenkomt met de plaats waar het in zich het handschrift bevindt.

```

<lg rend="indent">
  <1 n="1" xml:id="SIGLE62.1">Tekst versregel 1</1>
  <head place="margin-left">Bijschrift</head>
  <1 n="2" xml:id="SIGLE62.2">Tekst versregel 2</1>
  <1 n="3" xml:id="SIGLE62.3">Tekst versregel 3</1>
  <1 n="4" xml:id="SIGLE62.4">Tekst versregel 4</1>
</lg>

```

5.2.3.3. Codering van de fysieke structuur van de bron

Tot dusverre hebben de coderingen van de teksten uit het handschrift betrekking op de tekstuele inhoud gehad. Onderdelen van de tekst werden omsloten door open- en sluittags. Daarmee werd aangegeven dat alle tekst tussen deze tags tot een specifiek element behoort.

Zoals hiervoor reeds is uiteengezet wordt de fysieke structuur van de bron op een andere manier gecodeerd, namelijk met milestone-elementen: lege elementen die geen tekst omvatten, maar die slechts een plaats markeren. Vier elementen worden gebruikt: `<pb>`, `<lb>`, `<cb>` en een algemeen element `<milestone>`. *PB: inmiddels bestaat ook het <gb>-element (gathering break)*. De drie eerstgenoemde elementen worden gebruikt voor pagina-, handschriftregel- en kolomovergangen; ze worden voorzien van een attribuut `@xml:id` ter identificatie van de handschriftelijke bron waarnaar verwezen wordt, de bladzijde en het regelnummer. In een `@n` attribuut wordt de waarde weergegeven zoals die in de editie kan verschijnen.

```

<lb n="1" xml:id="SIGLEf146v.1"/>

```

Dit `<lb>` element gaat altijd vooraf aan de tekst op de betreffende regel, dus bijvoorbeeld een `<1>` element.

Voor andere overgangen dan die van bladzijde, kolom en handschriftregel wordt het element `<milestone>` gebruikt. Dit wordt voorzien van een attribuut `@ed` (met verwijzing naar de handschriftelijke bron d.m.v. het handschriftsigele). In het attribuut `@unit`, dat verschillende waarden kan krijgen, wordt aangegeven welke overgang de milestone markeert. Zo wordt een katernscheiding aangegeven door de waarde 'quire'. In een `@n` attribuut wordt het nieuwe nummer of de waarde van de milestone aangegeven. Afhankelijk van de opbouw van handschrift wordt er een bepaald nummer aan het katern toegekend.

```

<milestone n="1.2" ed="SIGLE" unit="quire"/>

```


Daarbij lijkt het zinvol om binnen een convoluut een onderverdeling in groepen katernen te onderscheiden. Een milestone voor een katerngroep wordt gespecificeerd door een *@unit* attribuut met waarde 'msPart':

```
<milestone n="1" ed="SIGLE" unit="msPart"/>
```

Bij conventie worden milestone-tags geplaatst voorafgaand aan elementen waarvan het begin samenvalt met de milestone. Voorbeeld van de codering van een tekst waarbij tussen versregel 14 en 15 een overgang plaatsvindt van het eerste naar het tweede katern van een handschrift:

```
<l xml:id="SIGLE59.14" n="14">versregel 14</l>
<milestone ed="SIGLE" unit="msPart" n="1"/>
<gb xml:id="SIGLEq1.2" n="1.2"/>
<pb xml:id="SIGLEf125r" n="125r"/>
<cb xml:id="SIGLEf125ra" n="a"/>
<lb xml:id="SIGLEf125ra.1" n="1"/>
<l xml:id="SIGLE59.15" n="15">
<!-- versregel 15 -->
</l>
```

Als er naast de digitale editie ook een editie in boekvorm bestaat is het zinvol om in de digitale editie de plaatsen te markeren waar paginaovergangen in het boek plaatsvinden. Ook dat gebeurt met behulp van het lege element *<pb>*. Ter onderscheiding van het element dat gebruikt wordt voor de handschriftelijke bron krijgt hier het 'ed'-attribuut de waarde 'MVNx' (waarbij x staat voor het nummer van de editie in de MVN-reeks).

```
<pb ed="MVNx" n="42"/>
```

Handenscheiding wordt aangegeven door het lege element *<handShift>*. Een daarbij gebruikt attribuut is *@new* (identificeert de nieuwe hand) en *@term* (hier wordt het schrifttype vermeld). De waarde van het *@term*-attribuut is in principe vrij, maar zal in overeenstemming met geldende paleografische conventies worden toegekend (textualis, cursiva, etc.). *PB: Hebben we @term nodig? Bestaat niet in TEI.* Het attribuut *@new* heeft waarden die overeenkomen met de handensiglen zoals die in de *<teiHeader>* worden gedefinieerd.

```
<handShift new="delta" term="textualis"/>
```

Het element *<handNote>* wordt in de TEI header gebruikt (binnen het *<msDesc>* element) om elke unieke hand die door de editor wordt onderkend te definiëren. Voor elke hand moet er een afzonderlijk element aanwezig zijn in de header (zie aldaar).

Het gebruik van milestone-elementen wordt voorbehouden voor de codering van verschijnselen die betrekking hebben op de constructie van de bron: katernen, bladen, aflijning, alles wat gedaan is ter voorbereiding van het eigenlijke beschrijven van de tekstdrager valt hieronder. Voor de fysieke aspecten en eigenschappen van de tekst, zoals paleografische eigenschappen, kleur van de inkt, enz. wordt het gebruik van het milestone element achterwege gelaten.

5.2.3.4. Registers

Anders dan in de gedrukte MVN-delen het geval is, vallen de registers niet onder het nawerk. Registers worden afzonderlijk gegenereerd. Ten behoeve van deze registers (opschriften en incipits) worden het eerste zelfstandig naamwoord, werkwoord of de eerste eigennaam van een opschrift of beginregel in de geëditteerde tekst gelemmatiseerd. Voor woorden die geen eigennamen zijn wordt daarbij gebruik gemaakt van het element *<w>* met attribuut *@lemma*. De waarde van het attribuut is de voluit geschreven, genormaliseerde vorm van het woord, zoals het voorkomt in het Middelnederlandsch Handwoordenboek.

```
<w lemma="stadhuus">stadt huus</w>
```

Persoonsnamen worden gelemmatiseerd met het element *<name>*, het attribuut *@type* met waarde 'person' en een attribuut *@reg* dat als waarde de genormaliseerde vorm van de naam heeft, waaronder zij dient te worden opgenomen in de registers.

```
Sente <name type="person" reg="Bernardus">bernaert</name>
```

Alle opschriften, titels en incipits worden voorzien van een taalcode, zodat afzonderlijke lijsten per taal kunnen worden gegenereerd.

Om bij het genereren van een woordconcordantie te voorkomen dat bij afgebroken woorden die over twee of meer regels zijn geschreven, de woorddelen niet als afzonderlijke woorden door het programma worden opgevat, worden afbrekingsstreepjes gecodeerd door een *<c>* element met attribuut *@type* en waarde 'shy'. Voorbeeld:

```
<l>datmen siet den li<c type="shy">--</c>
<lb/>chame</l>
```

PB: toegevoegd: In het geval dat de kopiist géén afbrekingsstreepje zou hebben geschreven, zou dit voorbeeld er zo uitzien:

```
<l>datmen siet den li<lb/>chame</l>
```

Uit het ontbreken van 'whitespace' (i.e.: spaties, tabs, nieuwe regels) volgt hier dat 'li' en 'chame' hier een enkel woord vormen. In principe kan zo'n woordafbreking over regels, kolommen en pagina's heengaan, zoals hier:

```
<l>datmen siet den li<pb/>
<cb/>
<lb/>chame</l>
```

PB: helaas toont de weergave in html van dit voorbeeld een aantal regelovergangen. In de XML staan die er niet!

5.2.4. De codering van verwijzingen

PB: Bij de verwijzingen gaat het ws. grotendeels om de inleidende teksten, is dus voor transcriptie niet direct relevant. Voor we bezien hoe de diverse onderdelen waaruit deze structuur is opgebouwd, moeten worden vor-

mgegeven en aan welke voorwaarden ze moeten voldoen, nog een enkel woord over een aspect van de editie die in alle onderdelen aan de orde kan zijn: de manier van verwijzen en linken.

Bij een verwijzing kan het gaan om een verwijzing naar een plaats binnen het document of naar een plaats buiten het document. Voor beide gevallen wordt één manier van coderen gebruikt, namelijk het `<ref>` element. Met dit element wordt een tekstdeel gemarkeerd vanwaaruit wordt verwezen naar één of meer andere plaatsen. Dat kunnen plaatsen in hetzelfde document zijn, maar bijvoorbeeld ook plaatsen die zich op het internet bevinden. De plaats waarnaar wordt verwezen wordt aangeduid met een `@target` attribuut. Binnen hetzelfde document wordt dit target aangeduid door middel van een op die plaats aanwezige xml:id. Deze identifier wordt in het `@target` attribuut voorafgegaan door een #-teken. Als binnen hetzelfde document wordt verwezen kan dat er als volgt uitzien:

```
Een ten dele uitgewist eigendomskenmerk vinden we op <ref target="#SIGLEf102v">f.102v.</ref>.
```

De waarde van het target, in dit geval #SIGLEf102v, is dan toegekend door middel van een xml:id aan een specifieke plaats in het document, in dit geval is dat een leeg `<pb>` (page break) element waarmee het begin van f. 102v is gemarkeerd. Bij een externe verwijzing hoeft er geen #-teken te worden gebruikt. In onderstaand voorbeeld wordt verwezen naar een afbeelding op het internet.

```
Een <ref target="#http://www.huysensinstituut.knaw/mvn/SIGLEf102vafb">foto</ref> van het ten dele uitgewiste eigendomskenmerk vinden we op f.10
```

5.2.5. De plaatsing van verwijzingen naar illustraties en niet-textuele grafiek

Als er op een bepaalde plaats in het handschrift een illustratie of een grafisch, niet-textueel verschijnsel voorkomt (zoals bijvoorbeeld een notenbalk), wordt dat ter plaatse aangegeven door een `<figure>` element met een attribuut `@xml:id` dat naar de bedoelde afbeelding verwijst.

```
<figure xml:id="GBHSfig101ruitsnede">
  <graphic url="..." />
  <head>Bijschrift</head>
</figure>
```

Binnen het `<figure>` element zijn ook de verwijzing naar de plaats waar de afbeelding staat (`<graphic>` element met `@url` attribuut) en het element `<head>` voor het fotobijschrift genest.

5.2.6. Muzieknotatie

Muzieknotatie wordt grafisch weergegeven in een afzonderlijk bestand (per geheel van samenhangende notenbalken wordt één afzonderlijk bestand gebruikt), waarnaar wordt verwezen door een leeg `<graphic>` element. Een `<graphic>` element staat dan binnen een `<figure>` element op de lokatie waar de notatie daadwerkelijk voorkomt. Een `<graphic>` element is voorzien van een `@url` attribuut met als waarde de lokatie waar de afbeelding zich bevindt.

6. Mark-up van de transcriptie

De basis van de transcriptie voor zover deze de structuur van de tekstverzameling, de structuur van de afzonderlijke teksten en de fysieke structuur van het handschrift betreft, werd behandeld in [hoofdstuk 2](#) en [hoofdstuk 3](#). Uitgangspunt is dat de teksten uit de bron letterlijk worden weergegeven in het XML-document, volgens dezelfde richtlijnen die gelden voor de diplomatische edities in boekvorm. In een aantal gevallen echter is het noodzakelijk op een andere wijze te werk te gaan. Dat geldt in het bijzonder voor verschijnselen die tot dusverre noodgedwongen (door de beperking van het medium van het gedrukte boek) op bijzondere wijze grafisch werden weergegeven of in annotatie werden beschreven. Deze gevallen kunnen worden ingedeeld in de onderstaande categorieën.

- tekstwijziging en correctie
- tekstbeschadiging en tekstverlies
- onzekere lezingen
- de codering van afkortingen en de oplossing van afgekorte tekst
- niet-linguïstische verschijnselen in de tekst, bijzondere grafische tekstweergave in de bron (onderstreepte tekst, dagletters in kalenders, vorkteksten, marginale tekst (geschreven buiten de tekstkolommen))
- opmaak- en katernopbouw-gerelateerde tekst (representanten, reclamen, custoden)
- bijzondere tekens (initialen, paragraaftekens, afbreektekens)
- afbeeldingen, figuren en tabellen
- paleografisch commentaar

6.1. Coderen of annoteren?

PB: Voor de meeste verschijnselen die in de hier volgende secties besproken worden, geven we een manier om het betreffende verschijnsel in TEI gedetailleerd te coderen. We beseffen echter dat gedetailleerde codering in strijd

kan zijn met het doel van de reeks: snelle publicatie van handschriften. Het is daarom ook mogelijk om deze verschijnselen op de traditionele wijze te behandelen in annotatie. Daarbij wordt gebruik gemaakt van het `<note>`-element. We onderscheiden vooralsnog de volgende typen: *tc* (tekstkritisch) en *pc* (paleografisch-codicologisch).

Om het mogelijk te maken dat de gebruiker switcht tussen een gestructureerde weergave en een weergave d.m.v. annotaties kan ook zowel een annotatie als een gestructureerde weergave van een verschijnsel worden gegeven. In dat geval wordt de noot ingebed in de structuurcodering en krijgt als `@rend`-attribuut *alternative*. Bijvoorbeeld:

```
<del type="strikethrough" rend="diagonal"
  hand="A">boec<note rend="alternative" type="tc">
  <mentioned>boec</mentioned> is doorgehaald</note>
</del>
```

In de noot uit het handschrift aangehaalde tekst wordt ingebed in het `<mentioned>`-element.

6.2. Tekstwijziging en correctie

De tekst zoals hij door een kopiist in eerste instantie is neergeschreven kan door hemzelf of anderen zijn gewijzigd. Daarbij kan het gaan om weglating van tekst zonder dat daar nieuwe tekst voor in de plaats komt (door middel van doorhalen, expungeren of uitraderen), om toevoeging van tekst zonder dat daarbij eerder geschreven tekst wordt vervangen (bijschrijven, al dan niet met gebruikmaking van invoegtekens), om vervanging van eerder geschreven tekst door nieuwe tekst (door middel van doorhalen of uitraderen en bijschrijven, of overschrijven) en om wijzigingen in de volgorde van reeds geschreven tekst. Tenslotte kan het zo zijn dat een correctie in tweede instantie toch weer wordt verworpen.

In al deze gevallen kan het gaan om de vervanging van slechte door betere tekst, of van goede tekst door tekst die niet (vanuit een linguïstisch oogpunt) beter of slechter is, maar eenvoudigweg anders. Over de kwaliteit van de tekstwijziging doet de editor van de diplomatische uitgave geen uitspraak; deze maakt dan ook geen deel uit van de codering.

Voor alle genoemde gevallen is er speciale codering beschikbaar, die hier wordt toegelicht.

6.2.1. Verwijdering van tekst zonder dat er nieuwe tekst voor in de plaats komt

Het element waarmee verwijderde tekst van beperkte omvang wordt gemarkeerd is het `` element. Dit element kan de attributen `@type` voor het soort verwijdering en `@hand` voor de corrector bevatten. (Het attribuut `@status` lijkt in de MVN-reeks niet van toepassing). Toegestane waarden voor het `@type` attribuut zijn: *strikethrough* (doorhaling), *erasure* (rasuur), *expunction* (expungering), *subpunction* (subpungering), *reagent* (reagens), *wash* (afwassen). Daarnaast wordt in de TEI Guidelines ook het `@rend` attribuut gebruikt, waarmee de realisering van de correctie kan worden vermeld. Het `@type` attribuut fungeert dus op een iets abstracter niveau dan het `@rend` attribuut. Het `@type` attribuut gebruiken we voor de wijze van verwijdering, het `@rend` attribuut voor de uitvoering daarvan:

```
Cvcurbita dat sijn <del type="strikethrough" rend="diagonal"
  hand="A">b</del>quoerden ende
```

Mogelijke waarden voor het `@rend` attribuut zijn: *horizontal*, *vertical*, *diagonal*.

Het coderen van weglatingen van langere passages waarin bepaalde tekststructuren (strofegrenzen, paragraafgrenzen) vallen, zou, indien het `` element gebruikt zou worden, conflicteren met het verbod op overlappende codering. PB: S.v.p. geen `<delSpan>` gebruiken zonder voorafgaand overleg. Hiervoor is een oplossing gevonden in de codering met behulp van een element `<delSpan>`, dat het begin van een weggelaten passage markeert en waarin met het attribuut `@to` wordt vooruitgewezen naar een bepaalde plaats waar de weglating eindigt. Die plaats wordt gemarkeerd door een leeg `<anchor>` element of een ander leeg element dat voorzien is van een xml:id, dat tevens wordt opgenomen als waarde van het `@spanTo` attribuut. Ook bij `<delSpan>` kan men met de attributen `@type` en `@rend` het soort markering van de weglating en de realisering ervan preciseren, terwijl met het attribuut `@hand` de corrector wordt geïdentificeerd.

```
<lg>
<1 n="1">Een ionghelinc . vul der <delSpan type="strikethrough"
  rend="horizontal" hand="delta" spanTo="GHHSIII15.2anchor1"/>weereit luste</1>
<1 n="2">Zochte eens<anchor xml:id="GHHSIII15.2anchor1"/> in wandelingen ruste</1>
<1 n="3">So verre hi hem te gane bewant</1>
<1 n="4">Dat hi eenen houden hermite vant</1>
</lg>
```

Als weggelaten tekst geheel onleesbaar is dan kan hij niet getranscribeerd worden. De passage in de transcriptie wordt dan gemarkeerd met het element `<gap>` binnen een `` element. Met het attribuut `@reason` kan middels de waarden 'illegible' of 'deletion' worden aangegeven waarom de weggelaten tekst onleesbaar is. Als gedeelten van de weggelaten tekst toch nog gedeeltelijk leesbaar zijn, dan wordt het `<unclear>` element gebruikt om passages aan te duiden waarvan de lezing niet geheel zeker is.

6.2.2. Toevoeging van tekst zonder dat daarbij eerder geschreven tekst wordt vervangen (bijschrijving al dan niet met gebruikmaking van invoegtekens)

Tekst die is toegevoegd, zonder dat ze voor andere tekst in de plaats komt wordt op vergelijkbare wijze gecodeerd als weggelaten tekst. Hiervoor wordt het `<add>` element gebruikt. Daarbij worden de attributen `@place` en `@hand` gebruikt en, indien van toepassing, het attribuut `@rend`, waarmee het gebruik van een invoegteken kan worden gespecificeerd. Het attribuut `@hand` kan alle waarden aannemen die worden genoemd in het element

<handNotes> in de TEIheader. Het *@rend*-attribuut kan twee waarden aannemen: ofwel er is (op wat voor wijze dan) ook een aanduiding op welke plaats de toevoeging moet worden ingevoegd, ofwel zo'n aanduiding ontbreekt. In het eerste geval neemt het *@rend*-attribuut de waarde 'pointer' aan; in het tweede geval 'nopointer'.

Met het attribuut *@place* wordt aangegeven waar de bijgeschreven tekst te vinden is. Daarvoor kunnen de volgende waarden worden gebruikt, overeenkomstig de situatie in de bron:

- inline (tekst is bijgeschreven op een opengelaten plaats)
- supralinear (bijgeschreven boven de regel)
- infralinear (bijgeschreven onder de regel)
- left (links van een woord)
- right (rechts van een woord)
- margin-left (bijgeschreven in de linkermarge van de kolom)
- margin-right (bijgeschreven in de rechtermarge van de kolom)
- margin-top (bijgeschreven in de bovenmarge)
- margin-bottom (bijgeschreven onderaan de bladzijde)
- opposite (bijgeschreven op de tegenoverliggende bladzijde)
- verso (bijgeschreven aan de ommezijde)
- mixed (bijgeschreven op verschillende plaatsen)

Langere toevoegingen aan een tekst kunnen worden gecodeerd met het <addSpan> element, dat op dezelfde wijze wordt gehanteerd als het <delSpan> element. *PB: <addSpan> alleen na overleg gebruiken.*

[VOORBEELDEN]

6.2.3. Vervanging van eerder geschreven tekst door nieuwe tekst (door middel van doorhalen of uitraderen en bijschrijving, of overheenschrijven)

Een veelvoorkomende manier van corrigeren is het verwijderen van een fout en het schrijven van de juiste tekst. Dit kan op verschillende manieren worden uitgevoerd: het ene woord of de ene letter wordt over een ander heen geschreven; een woord of letter kan worden doorgehaald en het goede woord of de letter kan erboven, eronder, ernaast of in de marge worden geschreven; dit alles kan het werk van een hand zijn of van verschillende handen, actief op verschillende tijdstippen. Ook kan het zijn dat de volgorde van het aanbrengen van correcties, of zelfs de uiteindelijke lezing onduidelijk is.

In zulke gevallen gaan we als volgt te werk: eerst coderen we de tekst die is weggelaten met het element en vervolgens coderen we de tekst die voor de weggelaten lettertekens in de plaats is gekomen met het <add> element.

Hiermee is in de codering nog niet expliciet aangegeven dat de toegevoegde tekst een correctie is van de wegge-laten tekst. Dat kan worden verholpen door gebruikmaking van het <subst> element (substitution). Zo wordt aangegeven dat de verwijdering en de toevoegingen deel uitmaken van hetzelfde correctieproces.

```
<1>Dat ghi hem <subst>
<del>heden</del>
<add>nemmermeer</add>
</subst>latet gaen.</1>
```

Wanneer de correctie plaatsvindt in een woord waarin zich een afkorting bevindt, wordt het <subst> element genest binnen het <choice> element en moet de correctie zowel binnen het <abbr> element als het <expan> element worden gecodeerd.

Als het verwijderde woord of de letters onleesbaar zijn (bijvoorbeeld door rasuur) wordt dit weergegeven door het lege element <gap>.

```
<del type="erasure">
<gap extent="5" unit="characters"/>
</del>
```

6.2.4. Open varianten

Als een corrector letters of woorden heeft bijgeschreven die lijken te fungeren als correctie op reeds geschreven tekst, zonder dat daarbij de eerst geschreven tekst is doorgehaald, hebben we te maken met open varianten. Het is in dat geval de vraag welk van de twee woorden beschouwd moet worden als behorend tot de hoofdtekst. Als in de Richtlijnen voor de diplomatische delen in boekvorm, wordt de op de regel geschreven variant opgenomen in de hoofdtekst, zonder verdere markering. De passage waarvoor twee alternatieven bestaan wordt gecodeerd met behulp van het <choice> element. Het tekstgedeelte in dat hoofdtekst staat wordt vervolgens gecodeerd met het element <seg>, het bijgeschreven alternatief met <add>.

```
<1>dat ghi hem <choice>
<seg>heden</seg>
```

```
<add>nemmermeer</add>
</choice> laten sal</l>
```

PB: *aangepast*

6.2.5. Herstel van correcties (door de kopiïst)

Herstel van een correctie kan worden aangegeven door het element `<restore>`. Het `@hand` attribuut wordt gebruikt om de kopiïst te identificeren die verantwoordelijk is voor het herstel en het `@type`-attribuut kan worden gebruikt om het soort correctie dat wordt hersteld te verduidelijken. Omdat we voor een doorhaling van een toegevoegde tekst (gecodeerd met `<add>`) het `` element kunnen gebruiken, zal het `<restore>` element alleen gebruikt worden voor het herstel van een verwijderde tekst (gecodeerd met ``).

6.3. Tekstbeschadiging en tekstverlies

Als een deel van de getranscribeerde tekst volstrekt onleesbaar is en daardoor niet in de transcriptie kan worden weergegeven, kan dit deel worden gemarkeerd door het lege element `<gap>`. Het attribuut `@reason` kan met de waarden 'illegible' of 'missing' dan de reden van de omissie aangeven. Indien een bekende hand verantwoordelijk is voor de onleesbaarheid kan deze met het attribuut `@hand` worden aangeduid. De attributen `@agent` en `@extent` dienen ervoor om resp. de wijze waarop de tekst onleesbaar is gemaakt, en de omvang van het onleesbare tekstdeel te benoemen. De omvang van het niet getranscribeerde deel wordt aangegeven in een geschat aantal 'letters', of, indien het om een groter gedeelte gaat, in millimeters.

Als de beschadigde tekst voor een deel nog wel leesbaar is, worden de volgende elementen gebruikt: `<damage>` voor het aanduiden van een passage waar de tekst beschadigd is; en `<unclear>` voor woorden, groepen woorden of passages die door de beschadiging niet meer met volledige zekerheid kunnen worden getranscribeerd.

Bij het `<damage>` element kunnen de attributen `@agent`, `@hand`, `@degree` en `@extent` gebruikt worden. Het gebruik van de drie eerstgenoemde spreekt voor zichzelf, bij `@degree` hebben we de keuze tussen de waarden *slight*, *moderate* en *serious*.

Wanneer de omvang van de beschadiging tekststructurende grenzen overschrijdt en de codering daardoor in een situatie van potentiële overlap terecht zou komen, moet ervoor gekozen worden het `<damage>` element een leeg element te maken. Het `@extent`-attribuut geeft dan de omvang aan van het gebied dat beschadigd is.

Tekstverlies doordat een bladrand is afgesneden wordt dan zo gecodeerd:

```
<damage agent="pagecut">
  <gap/>
</damage>
```

Zoals we reeds eerder zagen wordt het `<unclear>` element gebruikt in gevallen dat tekst niet met zekerheid te transcriberen is, omdat het deels onleesbaar is. De attributen `@reason`, `@hand` en `@agent` geven resp. aan waarom het tekstdeel moeilijk leesbaar is, welke hand eventueel verantwoordelijk is voor het onleesbaar maken en, in geval van beschadiging, wat, indien bekend of reconstrueerbaar, de oorzaak van de beschadiging is.

Mogelijke waarden voor het attribuut `@reason` zijn: *damage* en *illegible*. De laatstgenoemde waarde kan bijvoorbeeld gebruikt worden als een in de vouw verdwenen letter of woord ternauwernood zichtbaar is door een strakke binding van het handschrift.

Waar het beschadigde tekstdeel geen tekststructurende coderingsgrenzen kent kan een beschadiging als volgt worden gecodeerd:

```
<damage extent="whole leaf" agent="rubbing"> ... </damage>
```

Zijn er wel tekststructurende grenzen, dan is het mogelijk om de beschadigde stukken in logische partes op te spitsen, waarbij het `<damage>` element een leeg element moet worden gemaakt:

```
<p>
  <pb n="5r"/>
  <damage agent="rubbing"
    extent="whole leaf"/>
</p>
<p>
  <damage agent="rubbing"
    extent="whole leaf"/>
  <pb n="5v"/>
</p>
```

Stel dat een tekst wel beschadigd is, maar desondanks goed leesbaar. Deze tekst wordt dan als volgt getranscribeerd:

```
Daer hij den tempel <damage>o</damage>f hadde verloren
```

Als daarentegen door beschadiging de lezing onzeker wordt, dan kan als volgt worden gecodeerd:

```
Daer hij den tempel <unclear reason="damage">o</unclear>f hadde verloren
```

Wanneer het nodig is iets meer te zeggen over de aard van de beschadiging dan zou de codering er uit komen te zien als een `<unclear>` element dat wordt genest binnen een `<damage>` element:

```
Daer hij den tempel <damage agent="rubbing">
  <unclear>o</unclear>
</damage>f hadde verloren
```

Mocht de letter in het geheel niet meer leesbaar zijn, dan wordt het lege `<gap>` element gebruikt:

```
Daer hij den tempel <gap reason="illegible" agent="rubbing"
  extent="1"/>f hadde verloren
```

Het is mogelijk dat een passage die voor de editeur in het handschrift niet meer leesbaar is, dat nog wel was voor een editeur uit vroeger tijd, bijvoorbeeld doordat deze een reagens gebruikte waardoor de tekst kortstondig leesbaar werd, voordat door de werking van het middel de zwartkleuring intrad. In die gevallen kan een editeur de lezing van zijn voorganger geven, onder diens verantwoordelijkheid. Voor de codering van de oudere lezing tekst kan het element `<supplied>` worden gebruikt. Met het attribuut `@source` wordt dan de verantwoordelijke editeur benoemd.

Het `<supplied>` element kan het best worden genest binnen een `<damage>`- of een `<gap>`-element, waarmee de situatie in het handschrift precies kan worden beschreven. Het `<gap>` element is bijvoorbeeld aan de orde als er in een handschrift werkelijk beschreven materiaal ontbreekt (een hoek van een blad, een compleet blad, e.d.) dat ooit wel aanwezig was.

6.3.1. Voorbeelden van codering van onvolmaakt overgeleverde tekst

In de TEI Guidelines wordt geconstateerd dat het gebruik van de elementen `<gap>`, `<damage>`, `<unclear>`, `<supplied>` en `` dikwijls met elkaar samenhangt. Welk element van met name de eerste vier genoemde men kiest, is voornamelijk afhankelijk van de mate waarin een beschadiging of ingreep in de tekst de leesbaarheid heeft beïnvloed. De volgende aanwijzingen voor de toepassing van deze elementen, waarvan sommigen genest kunnen worden, zijn hierop gebaseerd.

Een tekst is opzettelijk beschadigd: Hiervoor gebruiken we het `` of het `<delSpan>` element.

Een tekst is onopzettelijk beschadigd: Hiervoor gebruiken we het `<damage>` element. Als in een handschrift met reagentia is geprobeerd de tekst leesbaarder te maken beschouwen dit als een geval van onopzettelijke beschadiging.

De staat van de getranscribeerde tekst wordt aangegeven door de elementen `<gap>`, `<unclear>` en `<supplied>`:

- `gap`: de tekst is afwezig;
- `unclear`: tekst is twijfelachtig want moeilijk te lezen;
- `supplied`: tekst is van een eerdere editeur. Wanneer een tekst door reagentia onleesbaar is geworden kan dit element worden ingezet. De eerder editeur heeft, indien hij degene is geweest die de reagentia heeft gebruikt, naar alle waarschijnlijkheid de tekst nog goed kunnen lezen.

`<gap>` en `<supplied>` komen dus alleen voor in combinatie met `` of `<damage>`. Alleen `<unclear>` kan ook zonder `` of `<damage>` voorkomen, namelijk als een tekst gewoon slecht leesbaar is.

Bij al deze elementen moeten toepasselijke attributen worden gebruikt om de oorzaak van en het soort beschadiging te documenteren en de mate van zekerheid van de transcriptie weer te geven.

6.4. Onzekere lezingen

Als er onzekerheid bestaat omtrent een lezing die niet kan worden toegeschreven aan al dan niet opzettelijke beschadiging, maar die een andere oorzaak heeft, zoals slordigheid of haastig schrijven van de kopiist, dichtlopen van letters tijdens het schrijfproces, enz., dan moet bij de codering het `<unclear>` element worden gebruikt. Voor alle andere gevallen zie paragraaf 5.2.

```
D<unclear>e</unclear>s gauen reinouts sone
```

6.5. De codering van afkortingen en de oplossing van afgekorte tekst

Het coderen van afkortingen vormt een probleem waarover veel discussie bestaat en dat moeilijk tot een bevredigende oplossing komt. De middeleeuwse praktijk van afkorten is sterk regionaal bepaald en in de loop van eeuwen gevormd door adaptaties van de tradities in het afkorten van Latijnse woorden. Hierdoor is er een verscheidenheid aan systemen ontstaan die het bijzonder moeilijk maakt deze op een logische grondslag te analyseren. Het coderen van afkortingen kan door de polyvalentie van de meeste afkortingstekens nooit voldoende zijn om 'automatisch' de juiste oplossingen, bijvoorbeeld via een stylesheet, te genereren. Dat betekent dat, wil de codering brongetrouw zijn, zowel de afkorting zelf als de oplossing ervan gecodeerd moeten worden. De werkwijze die in de gedrukte delen van de MVN-reeks werd gehanteerd, namelijk: weergave van de opgeloste vormen, met markering van de onzekere letters (inclusief de letters die in de afkorting zelf voorkomen en dus strikt genomen niet onzeker zijn!) door middel van cursivering, heeft evidente beperkingen. Zo is het, om een voorbeeld te geven, uit de tekst van een gedrukte MVN-editie niet mogelijk op te maken hoe het woord 'persemen' in het handschrift is geschreven: als 'p#semen' of als 'p'semen'. Zo is het evenmin mogelijk om inzicht te krijgen in de frequentie van het gebruik van afkortingen, of om specifiek, wellicht karakteristiek gebruik van afkortingen inzichtelijk te maken, om op eenvoudige wijze tekensystemen van verschillende kopiïsten met elkaar te vergelijken, editorsbeslissingen ten aanzien van oplossingen met elkaar te vergelijken, enzovoort.

De digitale editie biedt de mogelijkheid om alle afkortingen in een handschrift en hun uitgeschreven vormen te coderen, zodat de gebruiker van de editie naar believen zoekacties kan verrichten binnen een cumulerend corpus afkortingen. Het zou hier te ver voeren de ingewikkelde problematiek van het systematiseren en coderen van afkortingen te bespreken. De TEI Guidelines bieden verschillende mogelijkheden om afkortingen te coderen.

PB: Dit wordt een beetje anders: een afkorting mag volgens het hierna gegeven mechanisme met <choice> worden uitgeschreven, maar het is ook toegestaan alleen de opgeloste letters, die we traditioneel cursiveren, te coderen als <ex> (editorial expansion). Dit hangt af van de prioriteiten van de editoren en de omvang van het handschrift. Een woord waarin één of meer afkortingen voorkomt, wordt op twee manieren gecodeerd: in de vorm waarin het door de kopiist is geschreven en in de door de editor opgeloste vorm. Dat kan binnen het <choice> element, dat het mogelijk maakt twee alternatieve lezingen te geven voor een tekstplaats. De verkorte vorm wordt gecodeerd met het <abbr> element, het afkortingsteken met een leeg <g> element (glyph), en de opgeloste vorm komt binnen het <expan> element te vallen. Bij het <g> element is natuurlijk een verwijzing nodig naar het bedoelde teken. Daarvoor gebruiken we het @ref-attribuut. In het geval van een nasaalstreep wordt dat:

```
<g ref="#bar"/>
```

In het bestand charDecl.xml zijn deze afkortingstekens gedefinieerd. Een verwijzing naar dit externe bestand is standaard opgenomen in het MVN-framework.

De codering van de afkorting voor ‘ende’ ziet er zo uit:

```
<choice>
  <abbr>en<g ref="#bar"/>
</abbr>
  <expan>ende</expan>
</choice>
```

en die van ‘lren’ (lettren), met nasaalstreep boven de e, zo:

```
<choice>
  <abbr>lre<g ref="#bar"/>n</abbr>
  <expan>lettren</expan>
</choice>
```

Merk op dat de code voor een teken dat boven of door een letter moet worden weergegeven, zoals bijvoorbeeld de nasaalstreep, verschijnt na dat teken. De afkortingstekens zelf kunnen ook worden weergegeven door middel van tekens met vaste Unicode Code-points, die gemakkelijk kunnen worden aangeklikt op de Character Map die binnen oXygen kan worden gebruikt. (Het ligt in de bedoeling om in de toekomst een speciale selectie van veelgebruikte afkortingstekens als het ware ‘klaar te leggen’ in een gecustomiseerde MVN-character map.)

PB: In de alternatieve codering zien deze voorbeelden er zo uit:

```
en<ex>de</ex>
l<ex>ett</ex>re
<ex>n</ex>
```

Recentelijk is Unicode uitgebreid met tal van tekens die voor de mediëvist nuttig zijn omdat ze het mogelijk maken de feitelijke realisering van de afkortingen in de bron op een betrouwbare wijze weer te geven; het Medieval Unicode Font Initiative (MUFI) heeft voorstellen gedaan die voorzien in de meeste leemten die op dit punt nog in Unicode bestonden. De handleiding van de recentste MUFI-versie, die voor het merendeel van de benodigde tekens de te gebruiken Unicode karakters definieert, is te vinden op <http://www.mufi.info/specs/MUFI-Alphabetic-3-0.pdf>. Als tekens nog niet in Unicode beschikbaar zijn gebruiken we definities in de Unicode Private Use Area, bij voorkeur de tekens die MUFI heeft gedefinieerd (zie ook <http://gandalf.aksis.uib.no/mufi/PUAcoord/PUA-4.0-a-2.html>). *PB: beetje aangepast*

Hieronder volgt een aantal veelvoorkomende afkortingstekens en hun codering:

```
<g ref="#bar"/> : -
<g ref="#apomod"/> : `
<g ref="#usmod"/> : 9
<g ref="#condes"/> : 9
<g ref="#etfin"/> : 3
<g ref="#pflour"/> p (met een krul door de staart)
<g ref="#pbardes"/> : p (met streep door de staart)
```

De lange s met een streep door de schacht is in Unicode: 1E9C.

6.6. Niet-linguïstische verschijnselen in de tekst, bijzondere grafische tekstweergave in de bron en omissie (rubricatie, onderstreping, dagletters in kalenders, vorkteksten, marginale tekst (geschreven buiten de tekstkolommen))

De tekst in het handschrift kan op verschillende manieren door de kopiist zijn opgetekend, zodat bepaalde woorden of passages meer opvallen dan omringende tekst, of door de toevoeging van speciale tekens een bijzondere waarde krijgen toegekend. Daarbij kan gebruik zijn gemaakt van inkt van een andere kleur dan in de hoofdtekst is gebruikt, van onderstreping, samenvoeging met behulp van accolades, enz. Ook kan er tekst op een bijzondere plaats, buiten de kaders van de hoofdtekst of de aflijning zijn geschreven. Een aantal van de belangrijkste van dit soort typografische verschijnselen wordt hieronder, met de gewenste coderingswijze toegelicht.

6.6.1. Rubriek

Tekst geschreven met rode inkt kan worden gecodeerd met behulp van het <hi> element met rend="rubric" als nadere specificatie.

```
<hi rend="rubric">Miserere mei deus</hi>
```

Waar de rubricator paleografisch moet worden onderscheiden van de teksthand, zal tevens gebruik moeten worden gemaakt van het `<handShift>` element (zie par. 3.2.2.2). *PB: eenduidiger zou misschien zijn om op `<hi>` ook het `@resp-attribuut` toe te staan?*

6.6.2. Onderstreping

Onderstreping in de tekst wordt gemarkeerd met behulp van het `<hi>` element, waarbij de waarde `underline` wordt gegeven als `@rend-attribuut`.

6.6.3. Opengelaten ruimte

Het gebeurt wel dat een kopiist in zijn tekst ruimte open laat, bijvoorbeeld voor een initiaal of een miniatuur die in een later stadium zou moeten worden aangebracht, maar nooit werd gerealiseerd. Ook kan een ruimte zijn opengelaten waarvan niet meteen duidelijk is wat daarvan de functie is geweest. Dit soort (mogelijk) betekenisvolle lege plekken worden gecodeerd met een leeg `<space>` element. Te gebruiken attributen hierbij zijn `@dim`, waarmee wordt duidelijk gemaakt of de ruimte horizontaal of verticaal is en `@quantity`, daarmee wordt het aantal geschatte lettertekens (waarvan de eenheid wordt gespecificeerd met 'unit') gegeven. [`@verticaal` lines, horizontaal characters, introduceren] *PB: in plaats van `<space>` kan ook een `<note>` worden gebruikt.*

Als de letters van een woord opzettelijk uiteen zijn geschreven, wordt dat gecodeerd met `<hi>`, attribuut `@rend` en waarde 'spaced'. *PB of een `<note>`*

```
<hi rend="spaced">AMEN</hi>
```

6.6.4. Marginalia

Marginale teksten, geschreven buiten de afschrijving kunnen worden gemarkeerd met het `<note>` element. Als het daarbij gaat om invoegingen of correcties op tekst die in de kolom is geschreven, moet het `<add>` of `<addSpan>` element worden gebruikt. Het `<note>` element wordt nog binnen het `<l>` element geplaatst.

```
<l>
<note place="margin-right" type="ms">te eener huere</note>Den eerste dach van aprilis sijts ghewes
</l>
```

6.6.5. Vorkteksten en accolades

Het komt voor dat een tekst met opeenvolgende tekstregels die met dezelfde woorden beginnen of eindigen (of zelfs in het midden van de regels gelijklopende woorden bevatten) worden weergegeven door middel van een enigszins op accolades lijkend systeem. De woorden die in verschillende regels hetzelfde zijn worden dan slechts eenmaal opgeschreven en vanuit of naar die woorden toe trekt de kopiist dan lijnen naar de regels waar de overige tekst staat. Een voorbeeld daarvan vinden we in tekst 3 van het Geraardsbergse handschrift.

[Afbeelding].

Deze tekst wordt (enigszins vereenvoudigd) op de volgende manier gecodeerd met het (niet in de TEI voorkomende) element `<ditto>`. [UITWERKEN]. *PB: Mike, als je dit tegenkomt, kom even overleggen.*

6.7. Opmaak- en katernopbouw-gerelateerde tekst (representanten, reclamen, custoden)

Representanten zijn tekens die behoren tot de tekst, maar die geschreven zijn met een doel tijdelijk te functioneren tot de rubricator de letter of het teken in een andere vorm heeft gerealiseerd. In de TEI guidelines is voor codering van deze tekens niet voorzien. Wij doen het voorstel hiervoor het `<c>` element te gebruiken met als waarde voor het attribuut `@type`: `guide`.

In een handschrift kan men tekst tegenkomen die buiten de afschrijving is geschreven en die betrekking heeft op de samenstelling van de katernen.

Het voorkomen en de precieze locaties van katernsignatures, reclamen en custoden worden in beginsel in de beschrijving van het handschrift opgenomen, maar in bijzondere gevallen kunnen deze ook worden gecodeerd in de transcriptie. Het kan immers voorkomen dat een reclame van speciale tekstuele betekenis is, bijvoorbeeld als het woord verwijst naar een volgend katern dat is verdwenen of wanneer het een variëteit bevat. Hiervoor kan dan het `<fw>` element (afkorting voor: forme work) worden ingezet. Het attribuut `@place` wordt gebruikt voor de aanduiding van de positie van het gemarkeerde woord (of woorden),

```
<fw type="catch" place="bot-right">den roesen</fw>
```

PB: deze waarde voor place hebben we nog niet

PB: running titles ook met `<fw>`-element?

6.8. Bijzondere tekens (initialen, paragraaftekens, afbreekttekens)

6.8.1. Initialen

Voor initialen en lombarden wordt het `<hi>` element gebruikt. Met het attribuut `@rend` en de attribuutwaarde 'capitalizeN' wordt aangegeven dat het om een (bijzondere) hoofdletter gaat met een hoogte van N regels, waarbij N vanzelfsprekend variabel is.


```
<hi rend="capitalsize3">H<hi/>oert</hi>
```

6.8.2. Paragraafteken

Het paragraafteken, semi-paragraafteken en de representant voor het paragraafteken worden opgevat als noot van de kopiist en als volgt weergegeven (@type=ms om aan te geven dat het gaat om een noot die al in het manuscript aanwezig is):

```
<note place="margin-left" type="ms">
  <g ref="#00B6"/>
</note>
```

Een representant voor een paragraafteken wordt dus zo gecodeerd:

```
<note place="margin-left" type="ms">
  <c type="guide" ref="#dblsol"/>
</note>
```

NB: het teken //, waarnaar dblsol verwijst, bestaat wel binnen het MUFI-voorstel, maar is nog niet geaccepteerd als Unicode. *PB: inmiddels wel*

6.8.3. Afbreektteken

Als een afbreektteken in de tekst van het handschrift geschreven is, dan gebruiken we daarvoor het <c> element met het attribuut 'type' en waarde: shy.

```
<c type="shy">-</c>
<lb/>
```

6.8.4. Getallen, letters met zelfnoemfunctie en acrostichons

Getallen worden gecodeerd met het element <num>. In Romeinse cijfers geschreven getallen worden weergegeven met attribuut @type dat de waarde 'roman' heeft. Het getal wordt niet-genormaliseerd getranscribeerd in minuskel.

```
<num type="roman">ix</num>
```

Letters met zelfnoemfunctie worden gemarkeerd met het element <mentioned> en global attribuut type:

```
<mentioned type="self">v</mentioned>
```

Letters die deel uitmaken van een acrostichon worden gecodeerd met het element <c> en attribuut @type:

```
<c type="acrostic">A</c>
```

6.8.5. Figuren en tabellen

[Dient nader uitgewerkt te worden]

6.8.6. Afbeeldingen

Afbeeldingen van bladzijden worden opgenomen in het <pb> element, met behulp van het @facS-attribuut, dat naar de url van de bedoelde afbeelding wijst.

```
<pb n="f101r" facS="http://www..."/>
```

6.8.7. Paleografisch commentaar en alternatieve lezingen

Opmerkingen over de wijze waarop de kopiist bepaalde letters of lettercombinaties heeft gerealiseerd, voor zover deze niet te rangschikken zijn als (zelf)correcties of onduidelijkheid (zie daarvoor de paragrafen 5.1-5.2) kunnen worden gegeven in een apart gedeelte met paleografisch commentaar. Dit commentaar wordt gescheiden van de tekst in het hoofddocument om primaire tekst gescheiden te houden van secundaire.

We verwijzen daarbij van de noot naar de hoofdtekst. We gebruiken twee manieren van verwijzen, afhankelijk van de vraag of moet worden verwezen naar een tekstfragment of naar een bepaald punt in het document.

Wanneer er commentaar moet worden gegeven op een citeerbaar deel van de tekst gebruiken we het <seg> element, dat wordt geplaatst om de geciteerde tekst. De tags van het <seg> element moeten wel passen binnen de hiërarchische structuur van de XML. In gevallen waar dat niet mogelijk is, waar tekststructurende grenzen bijvoorbeeld binnen het tekstfragment vallen, kan het seg-element worden opgesplitst door middel van het @part attribuut, waaraan dan oplopende numerieke waarden worden toegekend. *PB: Herman, ik vind dit een voorbeeld van waar we dingen te ingewikkeld hebben gemaakt; zullen we niet noten gewoon in de tekst plaatsen?*

De tweede manier van verwijzen gebruiken we als er commentaar gegeven moet worden op een verschijnsel dat zich op een welbepaalde plaats bevindt, die niet zichtbaar is in de hoofdtekst. De editeur plaatst dan daar waar zich het te becommentariëren verschijnsel voordoet een <note> element waarin hij vrije tekst schrijft.

6.8.8. Opschriften, tussenkopjes, clausen en clauskoppen

Opschriften worden gecodeerd door het element <head>, attribuut @type en waarde 'opschrift'.

```
<head type="opschrift">Een vers jeghen gramscip</head>
```

Een tussenkopje krijgt de waarde 'tussenkop':

```
<head type="tussenkop">Hier naer volcht vanden kerstinen . Artuer</head>
```

Clauskoppen in dialoog- of toneelteksten worden gecodeerd door middel van het <head> element met attribuut @type en waarde 'speech'.

PB: dit impliceert waarschijnlijk dat we niet afdwingen dat een <head> aan het begin van een tekstuele eenheid staat (wat TEI wel doet).

6.9. Onzekere lezingen en verantwoordelijkheid

Niet alleen bij de transcriptie van tekst die valt binnen bovengenoemde categorieën, maar bij transcriptie in het algemeen geldt dat er meer of mindere mate van zekerheid kan bestaan omtrent de gekozen interpretatie van een verschijnsel in de bron. De TEI Guidelines bieden een mogelijkheid om een uitspraak te doen over de mate van zekerheid met het `<certainty>` element. (Wat hier niet bedoeld wordt zijn de gevallen waarin de tekst in de bron door wat voor oorzaak dan ook slecht leesbaar is geworden. Daarvoor wordt het element `<unclear>` gebruikt, zie paragraaf 0000. *PB: <certainty> moeten we denk ik niet gaan gebruiken: onzekerheid wordt verantwoord in <note>.*

Indien er onzekerheid over de juiste transcriptie bestaat kan met het `<respons>` element worden aangegeven wie verantwoordelijk is voor een specifieke oplossing. Met name in gevallen waarbij verscheidene editoren aan een transcriptie werken kan dit nuttig zijn. *PB: <respons> moeten we denk ik niet gaan gebruiken: waar van toepassing wel @resp-attribuut.*

Voordat we deze elementen bespreken zullen we ons buigen over tekst die bij de transcriptie speciale aandacht vraagt.

6.10. Interpunctie

Bijzondere interpunctie kan worden gecodeerd met het `<pc>` element. Het `@unit` kan worden gebruikt om aan te geven wat voor teksteenheden hier mee wordt begrensd.

...

PB: Herman nog bekijken

6.11. Kritische ingrepen

Uitgangspunt van de reeks is de diplomatische weergave. Wanneer de editor in overleg met de commissie besluit om toch correcties aan te brengen gebeurt dat met behulp van het `<corr>` element. Samen met de originele weergave (in een `<sic>`-element) worden het `<corr>`-element geplaatst in een `<choice>`. Voorbeeld:

```
<choice>
  <sic>faut</sic>
  <corr>fout</corr>
</choice>
```

PB: Herman nog bekijken

7. Literatuur

- [1]. Mertens, Th. *Richtlijnen voor de uitgave van Middeleeuwse Verzamelhandschriften uit de Nederlanden*. Hilversum: Verloren, 1994.