

```
<?xml version="1.0" encoding="iso-8859-1"?>
```

```
<!-- Alpino Dependency Structures DTD -->
```

```
<!--
```

```
  $Id: alpino_ds.dtd,v 1.1 2005/11/25 14:43:27 geertk Exp $
```

We hebben het toplevel element

```
  alpino_ds
```

Daaronder komen voor:

```
  sentence (PCDATA)
```

```
  comments
```

```
  node
```

Ik kan niet zeggen dat ik ze allemaal 1x wil zien, maar dat de volgorde niet uitmaakt. Zie o.a. <http://www.xml.com/pub/a/98/07/dtd/>.

Het kan wel in SGML (de "ampersand connector") en met RELAX NG.

We moeten hier kiezen voor of een te ruime optie of we moeten de volgorde vastleggen:

```
<!ELEMENT alpino_ds ( (node | sentence | comments)+ )>
```

of

```
<!ELEMENT alpino_ds (node, sentence, comments?)>
```

We kiezen voor de stricte variant.

```
-->
```

```
<!--           Alpino Dependency Structure           -->
<!ELEMENT alpino_ds (node, sentence, comments?)      >
<!ATTLIST alpino_ds
  version          NMTOKEN          #IMPLIED>
```

```
<!--           Node           -->
<!--           een knoop van een dependency tree     -->
<!ELEMENT node (node*) >
<!--
```

```
  rel           de dependency relatie van de knoop
  cat           de categorie van de knoop
  pos           de part of speech tag van de knoop (het woord)
  begin        de beginpositie van de bijbehorende woordgroep in de gehele
zin
  end          de eindpositie van de bijbehorende woordgroep in de gehele z
in
  root         de stam van het woord
  word        het woord zoals het in de zin voorkomt
  index       een co-index om gedeelde structuren mogelijk te maken
  id          een per zin uniek nummer om de knopen te kunnen identificere
n
  case        naamval van nomina (nom,acc,dat,..)
```

| | |
|-------------|---|
| comparative | type complement van comparatives (dan,als,dat,...) |
| def | definitietheid van nomina (def,indef,...) |
| frame | volledige Alpino POS-tag |
| gen | gender van nomina (de,het,...) |
| infl | inflection van adjectieven, determiners |
| neclass | classificatie van namen (org, loc, per, misc) |
| num | getalinformatie van nomina (sg,pl,...) |
| per | persoonsinformatie van nomina (fir,thi,...) |
| refl | reflexiviteit van pronomina (refl) |
| sc | subcategorizatie frame |
| special | vergaarbak voor andere speciale features |
| wh | geeft vraagwoordinformatie voor pronomina (ywh,nwh,rwh) |

misschien dat er meerdere #REQUIRED zijn, of dat in sommige NMTOKEN beter gebruikt kan worden ipv. CDATA.

Het is ook mogelijk een opsomming van de mogelijke waarden te geven. Voor sommige attributen is dat misschien wenselijk.

```
-->
<!ATTLIST node
  rel      (hdf|hd|cmp|sup|su|obj1|pobj1|obj2|se|pc|vc|svp|predc|ld|me|pred
m|         obcomp|mod|body|spec|det|part|app|whd|rhd|cnj| crd|nucl|sat|tag
|         dp|top|mwp|dlink|--)
          #REQUIRED
  cat      (smain|np|ppart|pp|ssub|inf|cp|du|ap|advp|ti|rel|whrel|whsub|con
j|         whq|oti|ahi|detp|svl|svan|mwu|top)
          #IMPLIED
  pos      CDATA          #IMPLIED
  begin    CDATA          #IMPLIED
  end      CDATA          #IMPLIED
  root     CDATA          #IMPLIED
  word     CDATA          #IMPLIED
  index    CDATA          #IMPLIED
  id       CDATA          #IMPLIED

  case     CDATA          #IMPLIED
  comparative CDATA          #IMPLIED
  def      CDATA          #IMPLIED
  frame    CDATA          #IMPLIED
  gen      CDATA          #IMPLIED
  infl     CDATA          #IMPLIED
  neclass  CDATA          #IMPLIED
  num      CDATA          #IMPLIED
  per      CDATA          #IMPLIED
  refl     CDATA          #IMPLIED
  sc       CDATA          #IMPLIED
  special  CDATA          #IMPLIED
  wh       CDATA          #IMPLIED
>
```

```
<!--          Sentence          -->
<!--          de tekst van de gehele zin          -->
<!ELEMENT sentence (#PCDATA)          >

<!--          Comments          -->
<!--          een of meerdere comment elementen          -->
```

```
<!ELEMENT  comments      (comment+)      >

<!--          Comment          -->
<!--          een commentaar entry  -->
<!ELEMENT  comment      (#PCDATA)      >
```