Taalkundig Protocol voor de Lexicologische Koppeling

Richard Piepenbrock
CGN-Subwerkgroep Lexicon
Afdeling Taal & Spraak
KU Nijmegen
Erasmusplein 1
Postbus 9103
6500 HD Nijmegen

Telefoon: +31-(0)24-3611134
Fax: +31-(0)24-3612907
E-mail: R.Piepenbrock@let.kun.nl
Homepage: http://lands.let.kun.nl/index.html

20 februari 2004

Versie 2.5

Inhoudsopgave

1 Definitie

De uitvoering van de lexicologische koppeling betreft het in de part-of-speech annotatielaag aan elkaar verbinden van een aantal los geschreven elementen die in de lexicologische en grammaticale theorie veelal als delen van één meerwoordsuitdrukking worden beschouwd.

In het CGN-project koppelen we als *continue meerwoordsuitdrukkingen* van oorsprong vreemdtalige, maar ingeburgerde meerledige niet-eigennamen, en in- en uitheemse meerledige eigennamen en titels.

Als *discontinue meerwoordsuitdrukkingen* koppelen we de scheidbaar samengestelde werkwoorden.

2 Algemene richtlijnen

1. De reikwijdte van een meerwoordsuitdrukking is steeds één uiting, in de praktijk een door de leestekens '.', '?' of '...' afgebakend fragment getranscribeerde spraak. Het voorzetsel *af* in het als volgt getranscribeerde fragment wordt dan ook niet gekoppeld: "Je belt haar toch nog wel even op hè? Af bedoel ik". Hier is de programmatuur voor de koppeling, `cgntsel.c`, ook op ingesteld.
2. Bij de koppeling voegen we niet handmatig nieuwe lemmata of tags toe, hoewel het koppelingsprogramma ons daartoe in staat stelt. We tekenen slechts dergelijke gewenste aanvullingen of wijzigingen aan op papier of in een aparte file. Dit heeft te maken met het annotatietraject; iedere wijziging moet worden doorgegeven aan de groepen die de orthografische transcriptie, POS-tagging en syntax behandelen. Omdat koppeling geheel aan het eind van het traject staat, is het niet goed in dergelijke eindversies, die immers moeten worden gesynchroniseerd met andere, eerdere versies, veranderingen aan te brengen.
3. We koppelen geen afgebroken of onjuist uitgesproken elementen, dus wanneer er een dergelijk element voorkomt bij de potentiële koppelingsitems, dan vervalt de noodzaak tot koppeling van de gehele uitdrukking.
4. Bij herhalingen of andere aarzelingen koppelen we zo mogelijk de elementen die het laatst door de spreker herhaald of gecorrigeerd zijn. B.v.

   "Ik ga [ga] ervan [uit] dat we samen.." ⇒ **uitgaan**
   "Ik wil uh ga uh d*a [doe] zeker [mee] met die prijsvraag" ⇒ **meedoen**

3 Continue meerwoordsuitdrukkingen

3.1 Vreemdtalige meerledige niet-eigennamen

Voorbeelden: *à la, anorexia nervosa, chili con carne, crème fraîche, eau de cologne, et cetera, nota bene, per se, prima donna, sine qua non, vice versa.*

Deze uitdrukkingen worden gekoppeld indien ze ingeburgerd zijn, d.w.z. indien ze als complete uitdrukking voorkomen in de 13e druk van de Grote Van Dale en minstens één van de afzonderlijke delen niet in Van Dale voorkomt (anders zou het gewoon gaan om een toevallige reeks afzonderlijk geassimileerde leenwoorden). De koppeling van deze elementen wordt geheel gestuurd door de voorhanden zijnde vreemdtalige expressies in het CGN-lexicon.

Het koppelen zal vrijwel altijd eenvoudig zijn, omdat het onwaarschijnlijk is dat homonymie zal optreden met niet-vreemdtalige expressies. Wel moet worden gelet op inbedding van de vreemdtalige expressies in een vreemdtalige context.

In het volgende fragment:

> "D'r werd aangekondigd dat men vanuit de Vlaamse regering met een verkiezingswebsite rechtstreeks naar de burger de resultaten on line zou brengen."

is *on line* een in het Nederlands ingeburgerde expressie in een Nederlandse context. Deze wordt dus gekoppeld.

Maar in:

> "Het is bij ons concern allemaal geautomatiseerd. We zijn nou helemaal de kant op gegaan van on line document storage and retrieval."

maakt *on line* deel uit van een niet geassimileerd Engelstalig begrip. Dit wordt dus **niet** gekoppeld.

3.2 In- en uitheemse meerledige eigennamen en titels

In de koppelingspraktijk worden alle continue sequenties van hoofdletterwoorden aangeboden ter koppeling, of ze nu te boek staan als eigennaam, titel (dus niet-eigennaam), al een lexicale ingang zijn in het CGN-lexicon of niet. De programmatuur is erop ingericht, dat tussenwerpsels en afgebroken woorden de sequentie mogen onderbreken zonder de continuïteit van de uitdrukking aan te tasten.

Bij deze reeksen hoofdletterwoorden geven we steeds de maximale reeks tokens als één geheel aan die in die context zinnig is.

Wanneer meerwoordsstructuren na elkaar voorkomen, zoals in appositie (bijstelling) of in opsomming, scheiden we de relevante delen:

> "Braziliaanse massasteden als [São Paulo] [Rio De Janeiro] en [Porto Alegre]"
> "[Willem Elsschot] [Alfons De Ridder] eigenlijk schreef [Kaas] [Villa Des Roses] en [Lijmen] [Het Been]"

Bij ingesloten meerwoordsstructuren geven we de maximale reeks zinnige tokens aan:

> "vanavond spelen [Acda En De Munnik]"
> "het boek [De Komst Van Joachim Stiller]"
> "de bootleg [An Evening With Frank Zappa And Captain Beefheart]"
> "de wet [Waardering Onroerende Zaken]"

Voorbeelden koppeling:

Input:

```
had                  _                  _
u                    _                  _
nog                  _                  _
uh                   _                  _
een                  _                  _
vraag                _                  _
wa*a                 _                  _
wat                  _                  _
betreft              _                  _
Telcomsel            COMB(eigen)        => Telcomsel_Telcomsel(1)
Telcomsel            COMB(eigen)        => Telcomsel_Telcomsel(1)
```

Output:

```
had              _                  _
u                _                  _
nog              _                  _
uh               _                  _
een              _                  _
vraag            _                  _
wa*a             _                  _
wat              _                  _
betreft          _                  _
Telcomsel        _                  _
Telcomsel        _                  _
```

Input:

```
Raad             COMB(eigen)    => Raad_Van_Commissarissen_Raad_Van_Bestuur
Van              COMB(eigen)    => Raad_Van_Commissarissen_Raad_Van_Bestuur
Commissarissen   COMB(eigen)    => Raad_Van_Commissarissen_Raad_Van_Bestuur
Raad             COMB(eigen)    => Raad_Van_Commissarissen_Raad_Van_Bestuur
Van              COMB(eigen)    => Raad_Van_Commissarissen_Raad_Van_Bestuur
Bestuur          COMB(eigen)    => Raad_Van_Commissarissen_Raad_Van_Bestuur
wij              _              _
willen           _              _
graag            _              _
dat              _              _
u                _              _
dat              _              _
gaat             _              _
doen             _              _
.                _              _
```

Output:

```
Raad             COMB(eigen)        => Raad_Van_Commissarissen(4)
Van              COMB(eigen)        => Raad_Van_Commissarissen(4)
Commissarissen   COMB(eigen)        => Raad_Van_Commissarissen(4)
Raad             COMB(eigen)        => Raad_Van_Bestuur(13)
Van              COMB(eigen)        => Raad_Van_Bestuur(13)
Bestuur          COMB(eigen)        => Raad_Van_Bestuur(13)
wij              _                  _
willen           _                  _
graag            _                  _
dat              _                  _
u                _                  _
dat              _                  _
gaat             _                  _
doen             _                  _
.                _                  _
```

4 Discontinue meerwoordsuitdrukkingen

4.1 Scheidbaar samengestelde werkwoorden

4.1.1 Definitie

Scheidbaar samengestelde werkwoorden zijn combinaties van werkwoorden met een eerste lid van een andere woordsoort. De met het werkwoord compatibele woordsoorten zijn, in termen van het CGN-taggingprotocol [3], finaal voorzetsel (*afhalen*), bijwoord (*weghalen*), adjectief (*loshalen*) en, minder frequent, substantief (*ademhalen*).

In het geval van voorzetsels worden bij de lexicologische koppeling uitsluitend elementen met *VZ(fin)* gekoppeld. Als een element dus fout gemarkeerd is, zoals in “ik tekende in *[VZ(init)]* op *[VZ(init)]* dat tijdschrift”, kan niet worden gekoppeld met het lemma *intekenen*. Hoogstens kan de fout bij toeval worden gesignaleerd en gecorrigeerd, en dan de koppeling worden overgedaan.

In een paar uitzonderlijke gevallen gaan werkwoorden verbindingen aan met andere woordsoorten:

Werkwoord (voltooid deelwoord):

gevangenhouden
gevangenmaken
gevangennemen
gevangenzetten
gevangenzitten
verbeurdverklaren

Voornaamwoord:

daarlaten
daarstellen

Veel combinaties van substantief en werkwoord, zoals *diamantslijpen*, *klootschieten*, *kroeglopen* en *waterfietsen*, komen alleen in de infinitiefvorm voor, zijn dus niet scheidbaar en worden dan ook niet ter koppeling aangeboden.

Als reden voor het aaneenschrijven van combinaties van werkwoorden met adjectiva en substantiva (waardoor ze als scheidbaar samengesteld werkwoord worden opgenomen en kandidaten zijn voor koppeling) wordt vaak een idiomatische lading aangevoerd en/of een hoge gebruiksfrequentie. Zo is *zwartwerken* van een hoger idiomatisch gehalte dan *hard werken* en wordt met ademhalen een overdrachtelijker soort van *halen* bedoeld dan het letterlijke *boodschappen halen*. Omdat geen duidelijke regels te geven zijn voor het al dan niet aaneenschrijven van werkwoordscombinaties, verlaten we ons in de praktijk op de beschikbare lexicale bronnen en spellingwijzers.

4.1.2 Aangeboden lexicale items

De basis van het CGN-lexicon is samengesteld uit het Groene Boekje, diverse edities van Van Dale en additionele bronnen als de Spellingwijzer Onze Taal. Deze bronnen zijn het niet altijd eens over welke scheidbaar samengestelde werkwoorden zijn toegestaan. Van Dale geeft soms een toelichting als “is geen samenstelling”, of “hoewel volg. de Wdl. *(= Woordenlijst Nederlandse Taal)* aaneengeschreven.....”, b.v. bij *meekunnen*, dat wel in het Groene Boekje staat. Zie hiervoor ook de ANS [2], paragrafen 9·2·2 (p. 508-9) en 12·2·2·2/I/2 (p. 609-10). Voor de koppeling is de set bewust vrij ruim genomen, zodat het lexicon niet alleen scheidbare werkwoorden bevat uit de dwarsdoorsnede van alle bronnen. *Meekunnen* staat dan ook in het CGN-lexicon.

In twijfelgevallen, en bij afhankelijkheid van het al dan niet aaneenschrijven (en dus van de status van scheidbaar werkwoord), van een bepaald betekenisonderscheid, beslist altijd de 13e druk van de Grote Van Dale.

4.1.3 Mogelijk ontbrekende lexicale items

Zeer veel zelfstandige werkwoorden die een handeling uitdrukken kunnen gecombineerd worden met voorzetsels en bijwoorden uit de volgende reeks. Zie hiervoor de ANS [2], paragrafen 12·2·2·2/II-III (p. 610-35). Een wezenlijk aantal van deze combinaties is opgenomen, maar vele zullen nog ontbreken, omdat de combinatorische processen productief zijn:

- *aan-* (in de betekenis van "doorlopend op gemakzuchtige of ineffectieve manier doen", zoals *aanknoeien*, *aanleuteren*, ofwel samen met *komen*, in de betekenis van "op een bepaalde manier komen", zoals *hij komt aanfietsen/aanslenteren/aanstiefelen*).
- *af-* (in de betekenis van "heel veel of vaak achtereen doen", zoals *(heel wat winkels) aflopen*, *(nogal wat) afkletsen*, *(me toch een hoop) afzeuren*).
- *bij-* (in de betekenis van "aanvullend doen", zoals *bijschilderen*, *bijschaven*).
- *door-* (in de betekenis van "blijven doen", zoals *doorschrijven*, *doorwandelen*).
- *mee-/mede-* (in de betekenis van "doen samen met", zoals *meeverpakken*, *meelachen*, *meebreien*).
- *na-* (in de betekenis van "na een eerder moment doen", zoals *nazitten*, *nastromen*, ofwel "controlerend doen", zoals *nameten*, *naschilderen*, ofwel "imiteren", zoals *nazingen*, *nadansen*). .
- *neer-* (in de betekenis van "naar beneden (doen) bewegen", zoals *neerplonzen*, *neermeppen*).
- *om-* (in de betekenis van "veranderen", zoals *ombuigen*, *omschakelen*, ofwel "uit zijn evenwicht brengen", zoals *omstoten*, *omzagen*).
- *onder-* (in de betekenis van "(naar) beneden doen", zoals *ondersneeuwen*, *onderploegen*).
- *op-* (in de betekenis van "naar boven doen", zoals *opdreggen*, *oppompen*, ofwel "helemaal weg- of klaarmaken', zoals *opruimen*, *opfokken*).
- *over-* (in de betekenis van "over een afstand (doen) bewegen", zoals *overslepen*, *overhevelen*, ofwel "opnieuw doen", zoals *overschrijven*, *overplamuren*).
- *rond* (in de betekenis van "overal naartoe doen", zoals *rondbrengen*, *rondbabbelen*, ofwel "in de rondte doen", zoals *ronddraaien*, *rondbuigen*).
- *samen-*.
- *terug-* (in de betekenis van "naar het uitgangspunt doen", zoals *terugbellen*, *terugeisen*).
- *toe-* (in de betekenis van "in de richting van doen", zoals *toesnellen*, *toezwaaien*, ofwel "dicht maken", zoals *toepersen*, *toedekken*).
- *uit-* (in de betekenis van "naar buiten doen", zoals *uitgooien*, *uitlopen*, ofwel "tot het einde toe, helemaal doen", zoals *uitwerken*, *uitzoeken*).
- *voor-* (in de betekenis van "van tevoren doen", zoals *voorproeven*, *voorbewerken*).
- *voort-* (zie *door-*).
- *weg-* (in de betekenis van "(doen) verdwijnen", zoals *wegfietsen*, *wegsmijten*).

- Combinaties met complexe bijwoorden als

  aaneen-
  achteraan-
  achterna-
  achterom-
  achterop-
  achterover-
  achteruit-
  bijeen-
  dooreen-
  ineen-
  omhoog-
  omlaag-
  omver-
  onderuit-
  opeen-
  uiteen-
  voorbij-
  voorop-
  voorover-
  vooruit-.

4.1.4 Constructies met adverbiale voornaamwoorden

Als een potentieel gescheiden vorm van een samengesteld werkwoord voorkomt met één van de volgende adverbiale voornaamwoorden (PO *VNW* met PDTYPE *adv-pron*):

er
d'r
waar
hier
daar
ergens
nergens
overal

dan is er heel vaak sprake van een gesplitst voornaamwoordelijk bijwoord, zoals *eraan* of *daarmee* (of een vergelijkbare constructie met *(n)ergens/overal*), waarbij het adverbiaal voornaamwoord verwijst naar een bepaald element dat gerealiseerd kan worden als een prepositioneel complement. In dat geval vindt **geen lexicologische koppeling plaats**.

Test hiervoor of het adverbiaal voornaamwoord vervangen kan worden door een volledig gerealiseerd complement (nl. een naamwoordelijke constituent). Is dat het geval, dan is het **geen scheidbaar samengesteld werkwoord**.

“kun je daar wel mee gaan werken?” ⇒ “kun je wel gaan werken met dat gereedschap?” ⇒ ***meewerken**, *geen koppeling*

“hij moet er wel echt wat mee gaan doen” ⇒ “hij moet wel echt wat gaan doen met die computer” ⇒ ***meedoen**, *geen koppeling*

“hier zit meer cichorei in” ⇒ “er zit meer cichorei in die poeder” ⇒ ***inzitten**, *geen koppeling*

“een paar huizen staan d’rtussen waar ze geen dak op hebben gezet” ⇒ “ze hebben geen dak gezet op een paar huizen” ⇒ ***opzetten**, *geen koppeling*

“je hebt nergens schuld aan” ⇒ “je hebt geen schuld aan iets” ⇒ ***aanhebben**, *geen koppeling*

“die soapster staat overal in” ⇒ “die soapster staat in elk tijdschrift” ⇒ ***instaan**, *geen koppeling*

“hij heeft hem er toen uit geknikkerd” ⇒ “hij heeft hem toen uit het team geknikkerd” ⇒ ***uitknikkeren**, *geen koppeling*

(*uitknikkeren* is daarmee alleen toegestaan in de zeldzame betekenis van “knikkers uit het knikkerputje schieten” of “iem. verslaan bij het knikkeren”).

Als de abverbiale voornaamwoorden in dergelijke contexten een locatief karakter hebben en vervangen kunnen worden door “op die plek” (of in het geval van *nergens* “niet op die plek”), dan kan het echter **wel** gaan om een scheidbaar samengesteld werkwoord.

Veel hangt hierbij af van de context en de klemtoon, die soms pas te achterhalen is, als je het gesproken fragment afluistert. Ook een incorrecte orthografische transcriptie kan hierbij de annotator parten spelen.

B.v.

“omdat je daar goed op moet pássen” ⇒ “omdat je daarop (op die tas) goed moet passen” ⇒ *geen koppeling*
vs.
“omdat je daar goed óp moet passen” ⇒ “omdat je op die plek goed moet oppassen” ⇒ *wel koppeling*

“omdat je hier goed op moet létten” ⇒ “omdat je hierop (op de verkeersborden) goed moet letten” ⇒ *geen koppeling*
vs.
“omdat je hier goed óp moet letten” ⇒ “omdat je op die plek goed moet opletten” ⇒ *wel koppeling*

“dan is ’t aan jou om daar een antwoord op te geven” ⇒ “daarop moet je antwoord geven” ⇒ *geen koppeling*
vs.
“dan is ’t aan jou om daar je voorsprong op te geven” ⇒ “op die plek moet je je voorsprong opgeven” ⇒ *wel koppeling*

Tot slot kan door weglating van elementen als *er* e.d. in lopende spraak ten onrechte de indruk ontstaan van een scheidbaar samengesteld werkwoord, als in de standaard telefoonfrase “ja, (daar) spreekt u mee” (**meespreken*). Ook hier wordt dus **niet** gekoppeld.

4.1.5 Constructies met kwantitatief ’er’

Als *er* in feite deel uitmaakt van een partitieve of kwantitatieve constructie, nl. uitbreidbaar is tot *ervan*, dan kan er ook sprake zijn van een scheidbaar samengesteld werkwoord.

> “ik heb er zes (van) over” ⇒ **overhebben** (*“erover hebben”)
>
> “ik vraag er twee (van) aan” ⇒ **aanvragen** (*“eraan vragen”)
>
> “dan blijft er vast nog wel iets (van) over” ⇒ **overblijven** (*“erover blijven”)

4.1.6 Constructie met ’om te’-beknopte bijzin

Let op de constructie *’zijn’ (PV) + ’om te’-beknopte bijzin*, waarbij sprake kan zijn van een scheidbaar samengesteld werkwoord, maar ook van een prepositioneel object-constructie zonder (adverbiaal) voornaamwoord. Een hoofdklemtoon op het niet-werkwoordelijk deel duidt altijd op een scheidbaar samengesteld werkwoord.

> “ik vertrouw haar zonder meer. Ze is echt iemand om op te bóuwen” (nl. “bouwen op iemand”) ⇒ *geen koppeling*
> “Lego, dat is echt leuk speelgoed om (dingen) óp te bouwen” (nl. “iets opbouwen”) ⇒ *wel koppeling*
>
> “dit is een dag om op te schíeten” (nl. “een vreselijke dag”, “schieten op de dag”) ⇒ *geen koppeling*
> “dit is een dag om óp te schieten” (nl. “een dag waarop je veel kunt doen”, “opschieten”) ⇒ *wel koppeling*

4.1.7 Constructies met ’PV – ’uit/mee’ – infinitief

Constructies als “Wij gaan/zijn/willen/(...) uit vissen” en “Wij gaan/komen/zijn/(...) mee eten” zijn zeer lastig te analyseren en lijken wellicht op het eerste gezicht scheidbaar samengestelde werkwoorden te bevatten, nl. *uitgaan* (vissen) en *meekomen/meegaan* (eten). Hoewel **uitzijn, *uitwillen* en **meezijn* uiteraard geen samenstellingen kunnen zijn, ligt het voor de hand hier weglating van het werkwoord *gaan* te postuleren (respectievelijk “wij zijn uit vissen gegaan”, “wij willen uit vissen gaan” en “wij zijn mee gaan eten”).

In het geval van *uit* is echter geen sprake van een scheidbaar samengesteld werkwoord, omdat aaneengeschreven vormen onmogelijk zijn (*“Wij zijn uitgegaan vissen”, “*het was al laat, toen wij uitgingen vissen”).

In het geval van *mee* zijn drie lezingen mogelijk, die elk een eigen codering vereisen:

1. De spreker plaatst een enkele hoofdklemtoon op *mee*, zoals in “ik wil méé werken”. In dit geval is de transcriptie fout, en moet “ik wil meewerken” worden getranscribeerd. In “ik wil wel méé gaan werken” duidt de hoofdklemtoon eveneens op een werkwoord *meewerken* (nl. een variant van “ik wil wel gaan meewerken”) en moeten “mee” en “werken” worden gekoppeld.
2. De spreker plaatst een dubbele hoofdklemtoon op *mee* en de infinitiefvorm, met een korte pauze tussen beide vormen en soms een lijdend voorwerp, zoals in: “ik ga méé (...) fíetsen” of “ik ga méé (...) de blóembakken buitenzetten”. Dit zijn gevallen van het werkwoord *meegaan*, die gekoppeld moeten worden.
3. De spreker plaatst een een enkele hoofdklemtoon op de infinitiefvorm, zoals in “ik ga wel mee fíetsen”. In dit geval is noch sprake van **meegaan*, noch van **meefietsen*, en wordt dus niets gekoppeld.

In de koppelingspraktijk zal het erop neerkomen, dat *mee* niet wordt gekoppeld, tenzij uit de context en/of beluistering duidelijk blijkt dat het hoort bij de infinitiefvorm, zoals in het geval van *meewerken* bij punt 1 hierboven.

Zie ook de ANS [2], paragraaf 18·5·6 (p. 1054-6).

Met dank aan Peter-Arno Coppen [1].

4.1.8 Gescheiden vormen en ellipsis

Vanwege de Vlaamse (in feite: Zuid-Nederlandse) tendens om een aantal adverbiale voornaamwoorden, voorzetsels, voorzetselcomplementen en andere zinselementen niet te realiseren t.o.v. de Standaardnederlandse norm, geven we uitsluitend bij twijfelachtige ellipsis toch een koppeling aan, mits de betekenis van het scheidbaar samengestelde werkwoord dat toelaat. Dit volgens de filosofie: wat niet gecodeerd is, zal later ook niet of nauwelijks meer te vinden zijn, m.a.w. liever te veel aangegeven dan te weinig.

Dit lijkt in tegenspraak met het punt over ellipsis onderaan paragraaf 4.1.4. Voor het koppelen bij een vermoeden van ellipsis moet dan ook een goede reden zijn, nl. structurele weglating, vooral in een Zuid-Nederlandse context, in samenhang met een toepasselijke betekenis.

Voorbeelden:

> "ik stuur (er) nog wel een foto bij" ⇒ **?bijsturen**
> "wij hebben alle spullen voor rotsklimmen bij (ons)" ⇒ **?bijhebben**
> "toen vloog Maarten weer (naar) buiten in onze klas" ⇒ **?buitenvliegen**
> "had jij al je boekentas mee(genomen)?" ⇒ **?meehebben**

Van een andere orde is de ellipsis van elementen die anders zouden worden herhaald in de zin, vergelijkbaar met meerledige samentrekkingen, zoals in:

> "je zou er wat vrijer mee om kunnen gaan en willen gaan"
> "ze keek niet op of om"
> "dat levert niet extra mankracht of levert niet extra uhm uh geld op"
> "'k probeer vaak foto's na te tekenen of te schilderen"

In de eerste zin wordt tweemaal het werkwoord *omgaan* bedoeld, met weglating van het tweede *om*, in de tweede zin zowel *opkijken* als *omkijken*, met weglating van de werkwoordelijke stam, in de derde zin tweemaal *opleveren* (eerste *op* weggelaten), en in de vierde zowel *natekenen* als *naschilderen*.

In al deze gevallen koppelen we de delen van het samengesteld werkwoord tweemaal, alsof alle delen wel twee keer gerealiseerd waren, als volgt:

```
dus             _                     _
'k              _                     _
probeer         _                     _
vaak            _                     _
foto's          _                     _
na              WW(inf,vrij,zonder)=> natekenen(1)
te              _                     _
tekenen         WW(inf,vrij,zonder)=> natekenen(1)
of              _                     _
te              _                     _
schilderen      WW(inf,vrij,zonder)=> naschilderen(2)
.               _                     _
```

De nabewerkingssoftware zorgt bij de omzetting naar XML-files voor de juiste dubbele koppeling.

4.1.9 Individuele lastige gevallen

aankomen "dan kom ik d'r zo aan" kan zowel slaan op "aankomen op die plek" ("arriveren"), wat altijd een samengesteld werkwoord is, als op "komen aan een object" ("aanraken", maar ook "verkrijgen"), wat meestal samengaat met een adverbiaal voornaamwoord. In geval van twijfel koppelen we de items aan elkaar. Idem voor *aanzitten* ("iets aanraken").

afweten Geen scheidbaar samengesteld werkwoord in constructies als "ik weet er niet veel van af". Alleen scheidbaar samengesteld werkwoord in "het laten afweten", zoals in "je kan meekomen, omdat Pieter het weer eens af heeft laten weten".

goedvinden Zeer spaarzaam als samenstelling gekoppeld, alleen zonder voorlopig object, om onderscheid te maken tussen "ik vind het goed/slecht/mooi dat het huis wordt gerenoveerd" etc. en "ik vind goed dat zij intrekt".

werkwoorden met 'heen-' *Heengaan* beschouwen we alleen als scheidbaar werkwoord in de (verouderde/formele) betekenis van "sterven", "weggaan":

> "Ondanks tekenen van herstel ging zij de dag erop van ons heen"
> "Gaat allen heen in vrede"

Constructies met discontinue voorzetsels als "om...heen" en "door...heen" zijn dus geen gevallen van het werkwoord *heengaan*.

> "Hij gaat om het hek heen" ⇒ ***heengaan** *(geen koppeling)*
> "De trein gaat door Roosendaal heen" ⇒ ***heengaan** *(geen koppeling)*

Ook geen koppelbaar item is de constructie: "heen en weer gaan".

Heenlopen is alleen een scheidbaar samengesteld werkwoord in de verouderde/formele betekenis van "weglopen", als in "verbouwereerd liep zij heen". Idem voor werkwoorden als *heenrijden*, *(zich) heenspoeden*, *heensnellen*, *heenvlieden* en *heenzenden*.

inhebben Alleen scheidbaar werkwoord in de betekenissen van:

- *bevatten, in zich hebben*: "hij had wel vijf pilsjes in", "hij heeft de smoor/pest in" (niet: "hij heeft er de smoor/pest in")
- *te betekenen hebben*: "wat heeft die glimlach in?", "dat heeft allemaal niets in"
- *in zijn macht hebben*: "Holland en Zeeland hadden bijna de gehele visserij in"

medeaanzitten Dit werkwoord, alleen vermeld in Van Dale, is uitzonderlijk, omdat in de gescheiden werkwoordsvormen de beide niet-werkwoordelijke delen niet aaneen worden geschreven, althans volgens Van Dale: "de notabelen zaten mede aan in de grote balzaal". Vergelijk dit met "de personenauto's reden achteraan (*achter aan) in de stoet" en 'het regiment liep voorop (*voor op) in de parade". Omdat Van Dale dit zo expliciet vermeldt en opname van een bijwoord *medeaan* onjuist lijkt, volgen we Van Dale in deze.

toelopen/toegaan Alleen scheidbaar samengesteld werkwoord indien niet deel uitmakend van constructie met prepositioneel complement *naar*:

> "ik liep op mijn broer toe" ⇒ **toelopen**
> "ik liep naar mijn broer toe" ⇒ ***toelopen** *(geen koppeling)*
>
> "ik ging naar huis toe" ⇒ ***toegaan** *(geen koppeling)*

uithebben Alleen scheidbaar werkwoord in de verouderde of zelden gebruikte betekenissen van:

- (onovergankelijk) *afgelopen zijn, afgedaan hebben*
- (onovergankelijk) *genoeg punten hebben om te winnen*
- (overgankelijk) *geld uit hebben staan*

Dus niet: "een boek *uithebben".

uitzien/eruitzien/d'ruitzien Er zijn twee afzonderlijke scheidbaar samengestelde werkwoorden *uitzien*, nl. één met de betekenissen:

- *uitkijken (op/over)*: "de kamer ziet uit op het zuiden", "de reiziger zag uit over het bergmassief"
- *verlangend wachten op*: "ik zie erg uit naar de familiereünie"
- *erop lijken, verwacht worden*: "het ziet ernaar uit dat zij niet meer komt'
- *afkijken*: "het is jammer dat ik de film niet uit kon zien"

en één met de betekenissen:

- *een bepaald voorkomen hebben*: "zij ziet er beeldig uit", "ik zie er niet uit vandaag", "wat ziet die kamer eruit"
- *erop lijken, verwacht worden*: "het ziet eruit of zij niet meer komt"

Het laatste werkwoord wordt volgens bronnen als Van Dale, het Groene Boekje en de Spellingwijzer Onze Taal geschreven als *eruitzien (ziet eruit, zag eruit, heeft eruitgezien)*. Ook het Taaladviesdossier van *Onze Taal* op het internet concludeert: "het lijkt ons het best de in het Groene Boekje genoemde spelling als één woord over te nemen".

Bij de koppeling moet dan ook een onderscheid worden gemaakt tussen **uitzien** aan de ene kant, en **eruitzien** en zijn volgens het CGN-protocol toegestane informele variant **d'ruitzien** aan de andere kant. Alle drie de werkwoorden zijn dan ook opgenomen in het CGN-meerwoordslexicon.

de kwestie 'weer' en 'terug' We moeten erop bedacht zijn, dat de bijwoorden *weer* en *terug* in combinatie met werkwoorden anders gebruikt kunnen worden in Nederland en Vlaanderen. Zo wordt *terug* in Vlaanderen vaak gebruikt, waar Noord-Nederlanders *weer* zouden zeggen. Het lijkt mij echter niet dat het hier in de meeste gevallen gaat om scheidbaar samengestelde werkwoorden:

fv600292.ort:

"door 't feit dat ge genoeg van hetzelfde gebruikt, krijgt de (= ge) terug een indrukwekkend en heel rijkelijk gegeven."

fv600215.ort:

"en zeker nu de SP terug een meer progressieve radicalere standpuntinname wil doen"

"dat was voor u een aanleiding om terug te gaan nadenken over"

Omgekeerd komen in het Vlaams werkwoorden met *weer* voor, waar in Nederland vormen met *terug* vaak gebruikelijker zijn en het gebruik van 'we(d)er' wordt opgevat als zeer formeel of archaïsch:

weerhebben Normaliter zijn combinaties met "hebben" en "zijn" geen scheidbaar samengestelde werkwoorden, maar worden ze los geschreven. In Vlaanderen zou dit werkwoord toegelaten kunnen worden in de betekenis van "terugkrijgen/terug hebben". De klemtoon moet dan wel vallen op het eerste lid:

> "de kilo bananen mag je me thuis eens komen afleveren, samen met een stempel die ik nog altijd moet weerhebben"

we(d)erkrijgen Vlaams voor "terugkrijgen". Equivalent aan de onscheidbare variant met klemtoon op de werkwoordelijke stam:

"hij weerkreeg zijn vrijheid" - "hij kreeg zijn vrijheid weer"
"Wie iets bijstuurt, krijgt zeker en vast iets weer"

we(d)erkomen Synoniem "terugkomen", "we(d)erkeren".

we(d)ervinden Synoniem "terugvinden".

weldoen Alleen scheidbaar werkwoord in de betekenis van "goede daden verrichten":

"hij gebruikte zijn rijkdom om wel te doen"
"alles heeft hij welgedaan, tot wie zou ik anders gaan"
"doe wel en zie niet om"

welgaan Het gaat hier niet om werkwoordscombinaties met het bevestigend bijwoord van modaliteit ("hij wil niet meegaan, maar ik ga wél"), of het onbeklemtoonde *wel* in "ik ga wel een keertje mee, wanneer ik tijd heb", maar om het bijwoord van hoedanigheid met vergrotende trap *beter* en overtreffende trap *best*, zoals in:

"alles gaat wel, dus ik mag niet klagen" ⇒ **welgaan**
"vaarwel en het ga u wel" ⇒ **welgaan**

Referenties

[1] P.A. Coppen. *Linguïstisch Miniatuurtje LXXXIII: Mee uit.* In: Neder-L: Elektronisch tijdschrift voor de neerlandistiek. Col: 0203.43, maart 2002, <http://baserv.uci.kun.nl/~salemans/archieven/miniatuurtjes/>, 2002.

[2] W. Haeseryn e.a. *Alegmene Nederlandse Spraakkunst (ANS).* Martinus Nijhoff/Wolters Plantyn, Groningen/Deurne, 1997.

[3] F. Van Eynde. *Part of Speech Tagging en Lemmatisering.* Openbaar CGN-document, KU Leuven, 2001.