Protocol voor de Prosodische Annotatie

Versie 2.1; 24 augustus 2001

1. Doelstelling

Een geschreven tekst bestaat uit woorden en leestekens (interpunctie). Insgelijks kan gesproken taal worden neergeschreven als een opeenvolging van woorden (de orthografische transcriptie) en prosodische tekens. Deze prosodische tekens geven aan welke lettergrepen extra werden beklemtoond, op welke plaatsen onderbrekingen (prosodische grenzen) voorkomen, en waar klanken abnormaal werden verlengd. De hier voorgestelde prosodische annotatie beoogt nu precies de **toevoeging van die prosodische elementen aan de orthografische transcriptie** van de spraak**.**

Sommige prosodische fenomenen zijn bewust door de spreker ingelast, maar andere zijn eerder toevallige aarzelingen of zijn het gevolg van versprekingen. Dit onderscheid is echter achteraf vaak niet meer met zekerheid te maken. Vandaar dat de hier voorgestelde prosodische annotatie enkel **een registratie en geen interpretatie** van de prosodische fenomenen beoogt.

2. De aangeleverde orthografische transcriptie

Alvorens uit te leggen hoe de prosodische annotatie in de orthografie aan te brengen, overlopen we eerst enkele belangrijke eigenschappen van de aangeleverde orthografische transcriptie.

- De orthografische transcriptie is opgebouwd uit zogenaamde **tiers** (lagen), en elke tier bevat de transcriptie van hetgeen door één van de sprekers werd gezegd. Normaal gesproken is er dus 1 tier per spreker, en elke tier is vergezeld van een **sprekercode**. Er kan echter nog een extra tier **unknown** voorkomen met de transcripties van niet-gekende sprekers.

- De orthografische **transcriptie van een spreker** bevat geen interpunctie, maar is wel gefragmenteerd in **spraakfragmenten** afgewisseld door **pauzefragmenten**. De transcriptie van een spraakfragment bevat woorden, deze van een pauzefragment is leeg of bestaat enkel uit *ggg* (dit staat voor een niet-talig sprekergeluid zoals een ademhaling, een gekuch, etc.: zie verder). Wanneer de transcriptie van een pauzefragment leeg is betekent dit niet noodzakelijk dat er in het fragment niets te horen is: er kan achtergrondlawaai aanwezig zijn, of er kan een andere spreker aan het woord zijn.

- De zonet vermelde fragmentatie van de tiers is **afgeleid van de manuele woordoplijning** die voor de te annoteren bestanden beschikbaar is. Deze bevat de begin- en eindpunten van alle woorden en (voldoende lange) pauzes. Bij de fragmentatie werd uitgegaan van de hypothese dat elke pauze van meer dan 0.5 seconden lang aanleiding geeft tot een pauzefragment, maar dat ook kortere pauzes als pauzefragmenten kunnen worden aanvaard daar waar dit nodig is om te verkrijgen dat alle spraakfragmenten korter zijn dan 10 seconden. *Dit heeft voor gevolg dat er in een spraakfragment pauzes kunnen voorkomen die langer zijn dan de kortste pauzefragmenten. Dit aspect komt later nog meer expliciet aan bod.*

- De orthografie is een woordelijke neerslag van hetgeen werd gezegd. Ze hoeft dus niet grammaticaal correct te zijn. Men moet dus niet verbaasd zijn als men zaken leest zoals *het paard die op stal staat ziet er ziek uit*.

- Hoofdletters zijn gereserveerd voor eigennamen, titels en letterwoorden.

- Sommige woorden zijn voorzien van een extra code (een * gevolgd door een letter). De codes zijn de volgende:

  *v : woorden uit een vreemde taal
  *d : dialectwoorden
  *z : woorden die met een sterk regionaal accent werden uitgesproken
  *a : afgebroken woorden die al dan niet worden gevolgd door een herhaling van het volledige woord
  *u : al dan niet opzettelijke versprekingen, of woorden die aarzelingen /herhalingen bevatten
  *x : woorden die onvolledig of niet met zekerheid werden verstaan

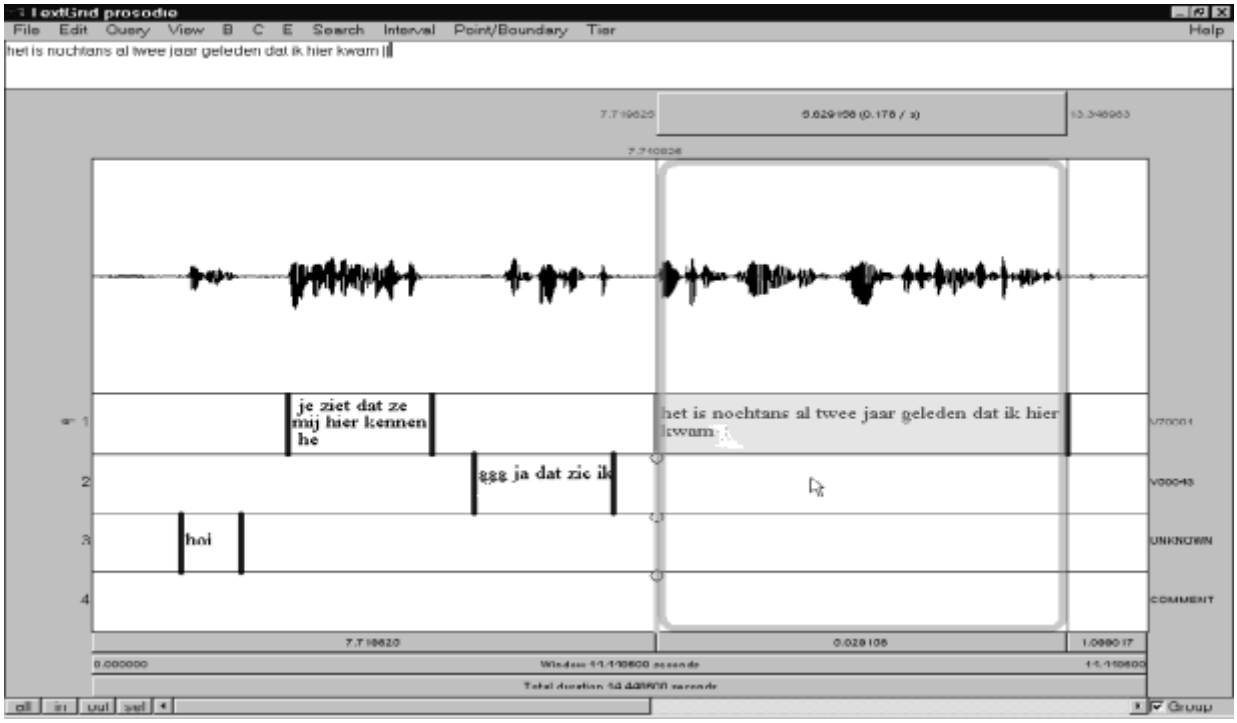
- Er kunnen ook nog codes *n en *t voorkomen. Zij duiden respectievelijk nieuwe woorden (*n) en tussenwerpsels (*t) aan die pas in een volgende release in het CGN-lexicon als normale woorden zullen worden opgenomen.

- Naast de speciaal gemarkeerde woorden komen in de orthografie ook nog de lettersequenties *ggg* en *xxx* voor die we in het vervolg ook als 'woorden' beschouwen. Met *ggg* duidt men goed hoorbare sprekergeluiden aan (b.v. gezucht, gekuch, gelach), met *xxx* onverstaanbare passages (eventueel bestaande uit meerdere woorden).

Belangrijk voor het vervolg is dat de transcripties spreker per spreker zijn georganiseerd, en dat elke spreker-tier is ingedeeld in spraak- en pauzefragmenten.

3. Visualisatie van de informatie

Met behulp van het programma **Praat** wordt op het scherm een visuele voorstelling verkregen van het spraaksignaal en zijn corresponderende orthografische transcriptie.



Figuur 1: Visuele voorstelling van de informatie

Een voorbeeld hiervan is weergegeven op figuur 1. Bovenaan herkent men het oscillogram van het spraaksignaal (in het geval van stereo-opnames zijn er twee spraakkanalen onder elkaar zichtbaar). Daaronder ziet men de verschillende spreker-tiers. Let in het bijzonder op de tijdsmarkers die de spraak- en de pauzefragmenten afbakenen.

De tier **unknown** bevat hier spraak van een toevallige voorbijganger.

Zoals je merkt is er ook nog een tier **comment** voorzien. Daarop kun je al je opmerkingen met betrekking tot dit bestand kwijt (b.v. over de moeilijkheden die je hebt ondervonden bij de prosodische annotatie).

Het is met behulp van het programma **Praat** dat je de prosodische annotatie zult aanbrengen. Een handleiding van **Praat** is te vinden op de website van het CGN en op de annotatie-CD van elke release.

4. Basisprincipes van de prosodische annotatie

1. De prosodische annotatie beoogt de aanduiding van **drie soorten fenomenen**: prosodische grenzen, prominente lettergrepen en abnormale klankverlengingen.
2. De prosodische annotatie beoogt enkel de registratie van **fenomenen die voor een onervaren luisteraar gemakkelijk auditief herkenbaar zijn**.
3. De prosodische annotatie wordt **aangebracht in de spraakfragmenten van de orthografische tiers** (inclusief in de unknown-tier) van de met elkaar in gesprek zijnde sprekers.
4. **De orthografie (inclusief de fragmentatie) mag in principe niet gewijzigd worden**. Mocht je er toch flagrante fouten in ontdekken, meld die dan aan je supervisor alvorens verder te doen.
5. **Grenzen tussen woorden** worden voorgesteld door symbolen die **omgeven** zijn **door spaties, grenzen binnen woorden** door symbolen die **niet** omgeven zijn door spaties (dit om het aantal woorden in de orthografie niet te wijzigen).
6. **Tijdens de interactie tussen sprekers let men op wat de individuele sprekers doen**. Dit betekent niet dat men even dikwijls door het bestand moet navigeren als er sprekers zijn, maar wel dat men tijdens passages waarin een spreker bijvoorbeeld door een andere wordt onderbroken, moet letten op wat elke spreker afzonderlijk doet.
7. Het is niet de bedoeling nutteloze tijd te investeren in de annotatie van fragmentjes waarin sprekers door elkaar heen praten. In die fragmentjes brengt men best alleen die informatie aan die snel en trefzeker kan worden aangebracht.

In wat volgt worden de meeste van de zonet genoemde principes verder uitgewerkt

5. Aanduiden van prosodische grenzen (symbolen || en |)

Een prosodische grens treedt op **daar waar** men de gehoorsindruk heeft dat **de spraakstroom** (de opeenvolging van woorden) even **onderbroken is**. Er wordt een onderscheid gemaakt tussen sterke en zwakke grenzen. Volg in geval van twijfel tussen sterke en zwakke grens de onderstaande regels.

5.1. Sterke grenzen (symbool ||)

Een sterke grens is **ernstige onderbreking van de spraakstroom**. Ze wordt typisch gerealiseerd door de insertie van een duidelijke (lange) stilte of adempauze tussen twee woorden of woorddelen.

Een groot aantal sterke grenzen werden reeds automatisch gedetecteerd. Het zijn de **grenzen van de spraakfragmenten** die zichtbaar zijn als dikke blauwe verticale strepen in de verschillende tiers. Zij **hoeven hier dus niet meer** verder **gemarkeerd te worden**. Het wordt wel aangeraden om, alvorens de annotatie te starten, even naar spraakgedeeltes rond pauzefragmenten te luisteren. Op die manier krijg je een idee van pauzelengtes die al zeker als sterke grenzen werden aanzien.

Een aantal **andere sterke grenzen** – nemen we aan – werden niet automatisch gedetecteerd en dienen nu dus **binnen de spraakfragmenten te worden toegevoegd** (m.b.v. ||-symbolen). Dat niet alle sterke grenzen automatisch werden gedetecteerd kan vele oorzaken hebben. Enkele daarvan (ter informatie!) zijn:

(1) De pauze is niet zo heel lang, maar de onderbrekingsindruk die ze opwekt is toch sterk – en wordt dan (mede) veroorzaakt door andere factoren (b.v. het toonhoogteverloop).

(2) De pauze treedt op binnen een woord, en kon bijgevolg niet automatisch worden gevonden, aangezien tijdens de woordoplijning enkel woordgrenzen werden aangeduid.

(3) De pauzegrens werd tijdens de woordoplijning niet correct geplaatst, waardoor de pauze korter lijkt dan ze in werkelijkheid is.

Vraag je niet teveel af waarom een sterke grens niet automatisch werd aangeduid, maar duid ze hier gewoon aan!

Veronderstel nu de volgende orthografie (inclusief een suggestie van de positie van de spraak-fragmenten in de tijd):

Dialoog: spreker A : ik was hier al eens eerder twee jaar geleden.
spreker B : ja dat weet ik maar wanneer

Monoloog : ik ben bij de politiecommissaris ontboden

En neem aan dat er duidelijke pauzes voorkomen na *weet ik* (dialoog, spreker B) en na *politie* (monoloog). De prosodische annotatie wordt dan:

Dialoog: spreker A : ik was hier al eens eerder twee jaar geleden
spreker B : ja dat weet ik ‖ maar wanneer

Monoloog : ik ben bij de politie‖commissaris ontboden.

Een vaak voorkomend verschijnsel is dat een spreker een duidelijke pauze laat nadat hij door zijn gesprekspartner werd onderbroken.

spreker A: hij heeft negen medailles ‖ zeven medailles gewonnen
spreker B: zeven

Spreker A wordt hier door B verbeterd, A onderbreekt zijn spraak, en gaat dan correct verder.

Toe te passen regels:

A.1: Zoek vooraleer je met de annotatie begint naar enkele kortere pauzefragmenten, en luister naar stukjes spraak rondom deze fragmenten. Op die manier krijg je een indruk van wat nog als een sterke grens kan worden beschouwd.
A.2: Duid ernstige onderbrekingen van de spraakstroom *binnen spraakfragmenten* aan als sterke grenzen (symbool ‖).
A.3: Omgeef de ‖-symbolen *tussen* woorden door spaties, deze *binnen* woorden niet.

5.2. Zwakke grenzen (symbool |)

Een zwakke grens is een **zwakke maar toch duidelijk hoorbare onderbreking van de spraakstroom.** Men hoort duidelijk dat de woorddelen niet op de normale (vlotte, 'verbonden') manier bij elkaar aansluiten, terwijl er in het signaal toch geen duidelijke stilte of adempauze wordt waargenomen. Het is niet ondenkbaar dat in het volgende voorbeeld:

spreker A : jan en ook piet waren hier al eerder twee jaar geleden
spreker B : ja dat weet ik ‖ maar wanneer

spreker A de woorden *jan* en *en*, en spreker B de woorden *ja* en *dat* niet vlot op elkaar laat volgen. Dit leidt dan uiteindelijk tot het volgende resultaat:

spreker A : jan | en ook piet waren hier al eerder twee jaar geleden
spreker B : ja | dat weet ik ‖ maar wanneer

Zwakke grenzen kunnen uiteraard ook binnen woorden optreden, b.v. om een sterke emotie uit te drukken.

Dit is werkelijk on|ge|looflijk

Verder mag men b.v. zwakke grenzen verwachten in de buurt van afbrekingen, herhalingen, hernemingen en gestotter. Beschouw volgende voorbeelden:

afbreking: ik kwam hem tegen in het cent*a centraal station van Amsterdam
herhaling: hij is in de in de fout gegaan
herneming: hij is in die op die bijeenkomst aanwezig geweest
gestotter: hij is noodgedwongen te*a te*a teruggekeerd

Men kan verwachten dat de prosodische annotatie daarvan het volgende resultaat oplevert:

ik kwam hem tegen in het cent*a | centraal station van Amsterdam
hij is in de | in de fout gegaan
hij is in die | op die bijeenkomst aanwezig geweest
hij is noodgedwongen te*a te*a | teruggekeerd

Maar het kan evengoed zijn dat er b.v. geen grens optreedt na de *cent*a* in voorbeeld 1, of na de *in de* in voorbeeld 2. **Plaats enkel grenzen waar je ze echt hoort**!

Men kan vaak ook zwakke grenzen horen voor of na een *uh*. Volgende voorbeelden zijn zeker niet uitgesloten:

er was uh | gisteren een grote grens-|uh|-controle*u op de E17

Zoals reeds gezegd komt het vaak voor dat een spreker zijn spraak even onderbreekt wanneer hij door zijn gesprekspartner wordt onderbroken. Deze onderbreking hoeft niet altijd een sterke grens op te leveren; zo'n grens kan ook zwak zijn.

Laat je niet leiden door de geschapen verwachtingen (volgend uit de syntax of uit de aanwezigheid van aarzelingen of *uh*'s), maar **laat je leiden door het geluidsignaal**! Annoteer dus een | of een || op basis van wat je hoort!

Toe te passen regels:

B.1: Duid zwakke doch duidelijk hoorbare onderbrekingen aan als zwakke grenzen (symbool |).
B.2: Omgeef de |-symbolen tussen woorden door spaties, deze binnen woorden niet.
B.3: Plaats de grenzen op de plaatsen waar je ze hoort, zelfs al zijn dit niet de plaatsen waar je ze op basis van de syntax of de aanwezigheid van *uh's* had verwacht.
B.4: Ingeval je twijfelt of er al dan niet een grens is tussen twee woorden (woorddelen), probeer die dan eens zonder grens na elkaar uit te spreken, en plaats slechts een grens wanneer dit duidelijk anders klinkt dan in het te annoteren fragment.
B.5: Ingeval je twijfelt tussen een zwakke en een sterke grens, kies dan voor een zwakke.
B.6: Bij twijfel tussen een zwakke grens en geen grens, kies dan voor 'geen grens'.

6. Aanduiden van prominente lettergrepen (symbool ^)

In een spraakfragment kunnen **lettergrepen extra** worden **beklemtoond** (benadrukt) om een woord belangrijk (prominent) te maken, of om het in contrast te plaatsen met een ander woord (waardoor het eigenlijk ook belangrijk wordt). Een beklemtoonde lettergreep noemt men een **prominente lettergreep**. Een woord met een prominente lettergreep noemt men een **prominent woord**. **Let goed op welke lettergreep er precies wordt beklemtoond**!

Prominentie wordt doorgaans gerealiseerd door een toonhoogtebeweging, zij het vaak in combinatie met een volume(luidheids)toename en een klinkerverlenging. Normaal gesproken bevat een prominent woord slechts 1 prominente lettergreep, maar in uitzonderlijke gevallen kunnen dat er ook twee zijn.

Een prominent woord lijkt dikwijls luider te klinken dan de andere woorden. Probeer in geval van twijfel het woord in zijn context een keer mét en een keer zónder nadruk uit te spreken, en op basis daarvan een beslissing te nemen. Twijfelt men b.v. aan de prominentie van de woorden *ik* en *thuis* in het fragment

> ik ben thuis weggegaan omwille van mijn broer

dan kan men zelf de uitspraken van de volgende paren met elkaar en met het te annoteren fragment vergelijken:

voor klemtoon op “ik”:	^i^k ben thuis	←→	ik ben th^ui^s
voor klemtoon op “thuis”:	thuis w^e^ggegaan	←→	th^ui^s weggegaan

Prominente lettergrepen worden aangeduid door de karakters die het **klinkergedeelte** van die lettergreep vormen, te **omgeven door ‘^’-symbolen**. Hier volgen enkele voorbeelden:

1) ik stond niet v^oo^r | maar ^a^chter hem
2) ik doe alles z^e^lf | tenzij het sl^e^cht bet^aa^ld wordt
3) ik doe ^a^lles zelf | tenzij het sl^e^cht betaald wordt
4) ik kom pas m^o^rgenavond naar huis
5) deze courg^e^tte is d^o^nkergr^oe^n
6) het is ^eeu^wen gel^e^den | dat ik hem nog eens gespr^o^ken heb
7) ik heb dit niet g^e^kocht | maar v^e^rkocht
8) d^a^t had ik n^ooi^t verwacht zie.

De voorbeelden 2 en 3 zijn orthografisch identiek, maar prosodisch verschillend. Dit om te benadrukken dat wat je hoort de doorslag geeft, en niet wat je verwacht. Voorbeeld 4 is er een waarbij de klemtoon in het prominente woord niet op de normale plaats ligt (de spreker contrasteert nu **morgen**avond met **van**avond of **overmorgen**avond). Voorbeeld 5 presenteert een prominent woord met twee prominente lettergrepen. Voorbeeld 6 toont aan dat het klinkergedeelte van de prominente lettergreep tot 3 karakters lang kan zijn. Het laatste voorbeeld illustreert het gebruik van beklemtoning voor het aanbrengen van een contrast tussen lettergrepen die normaliter geen klemtoon dragen.

Speciale gevallen

- Indien een spraakfragment slechts 1 woord bevat, dan is het veel moeilijker om van prominentie te spreken. Men dient hier enkel prominentie aan te duiden wanneer men er van overtuigd is dat de spreker bewust een grote nadruk op het woord heeft gelegd.
- Indien er in de tekst een benadrukt letterwoord optreedt (b.v. SMS), dan dient men de prominentiesymbolen rond de beklemtoonde letter te plaatsen (in de uitspraak vormt iedere letter immers een eigen lettergreep). Het volgende voorbeeld is dus mogelijk: "ik kreeg pas een SM^S^-bericht binnen"

Het gebeurt wel eens dat het laatste woord voor een ||-grens een bijzondere toonhoogteverloop vertoont (een zogenaamde grenstoon) die b.v. signaleert dat de spreker nog zal verdergaan, of dat hij een vraag stelt. Een dergelijk woord mag slechts als prominent worden aangeduid als het ook effectief een beklemtoonde lettergreep bevat.

Toe te passen regels:

C.1: Een lettergreep is prominent als hij wordt beklemtoond en daardoor het woord belangrijk maakt binnen het fragment waartoe het behoort.
C.2: Elk spraakfragment bevat normalerwijze een of meer prominente woorden, maar vooral bij korte spraakfragmenten (enkele woorden) kan het wel eens voorkomen dat ze geen enkel prominent woord bevatten.
C.3: Duid slechts twee of meer prominente lettergrepen in hetzelfde woord aan wanneer die bedoeld zijn om verschillende delen van het woord belangrijk te maken (alsof elk van die woorddelen aparte prominente woorden waren).
C.4: Indien je twijfelt aan de aanwezigheid van prominentie van een woord, probeer dan zelf het woord eens met en zonder nadruk uit te spreken (zie tekst), en duid slechts prominentie aan wanneer de benadrukte uitspraak duidelijk het best op deze van het te annoteren woord lijkt.
C.5: Duid geen prominentie aan in onbekende woorden (xxx).

7. Aanduiding van abnormale klankverlengingen (symbool %)

In spontane spraak kunnen **klanken** (zowel klinkers als medeklinkers) **abnormaal verlengd** zijn **zonder dat zij** daarom een **lettergreep prominent maken**. Hierna volgen enkele voorbeelden om dit te illustreren (de verlenging werd hier aangegeven door het repeteren van dezelfde letter):

1) ik heb hemmmmmm negen maal ontvangen denk ik
2) hij is al deeeeeee negende die dat vraagt
3) dit is nogal sssssstout gesproken
4) ik heb dat ongeeeeeedisciplineerd genoemd
5) hij is pas vierennnnndertig
6) jaaaaaa || maar dat is eigenlijk verkeerd hoor

In de voorbeelden 2 en 4 lijken de klankverlenging op gecamoufleerde *uh's.* Het voorlaatste voorbeeld kan erop wijzen dat de spreker b.v. heeft getwijfeld tussen 24 en 34. In het laatste voorbeeld zal de verlenging van de *a* waarschijnlijk ook gepaard gaan met een specifieke toonhoogtebeweging, maar is het woord *ja* daardoor niet echt beklemtoond.

Wat ook hun bedoeling moge geweest zijn, alle abnormaal verlengde klanken die niet tot prominentie leiden dienen te worden omgeven door %-symbolen. Voor de zonet gegeven voorbeelden leidt dit tot de volgende resultaten:

1) ik heb he%m% | n^e^gen maal ontvangen denk ik
2) hij is al d%e% | negende die dat vraagt
3) dit is nogal %s%tout gesproken
4) ik heb dat ong%e%disciplin^ee^rd genoemd
5) hij is pas viere%n%d^e^rtig
6) j%a% || maar dat is eigenlijk verk^ee^rd hoor

Niet aan te duiden verlengingsfenomenen

(1) De meeste *uh*'s en *mmm*'s zijn eigenlijk ook abnormaal verlengde klanken. Ze zijn echter op zichzelf herkenbaar, en dienen hier niet meer van %-symbolen te worden voorzien.

(2) Het is normaal dat de eindmedeklinkers van woorden voor een grens langer zijn dan deze in andere woorden. Zolang de verlenging niet echt abnormaal is hoeft ze niet te worden aangeduid, maar dient er een sterke of zwakke grens te worden geplaatst.

Toe te passen regels

D.1: Omgeef abnormaal verlengde klanken die niet tot prominentie leiden door %-symbolen.
D.2: Een klinker mag niet tegelijkertijd als verlengd en prominent worden aangeduid. Kies enkel voor verlenging als je zeker bent dat ze geen prominentie veroorzaakt.
D.3: Indien je twijfelt of de verlenging wel abnormaal lang is, duid ze dan niet aan, maar ga na of er na de verlenging geen grens dient te worden aangeduid.
D.4. Een duidelijk abnormaal verlengde klank die niet tot geen prominentie leidt mag ook voor een grens als verlengd worden aangeduid.
D.5: Markeer geen abnormale klankverlengingen in onbekende woorden (xxx), en in *uh's* en *mmm's*.

9. Voorgestelde werkwijze

Bij het aanmaken van de prosodische annotatie wordt de volgende werkwijze aangeraden:

1. Kies in *Praat* voor een tijdschaal van om 15..20 seconden per volledig scherm, zodat een spraakfragment nooit meer dan tweederden van het scherm in beslag neemt.

2. Luister alvorens met de annotatie te beginnen eerst eens naar de spraak rondom de pauzefragmenten (dit zijn intervallen tussen blauwe grenzen waarin niets werd getranscribeerd) teneinde een idee te krijgen over de lengtes van deze per definitie sterke grenzen. Doe dit voor de verschillende sprekers in het bestand.

3. Ga van links naar rechts door het bestand en annoteer de opeenvolgende spraakfragmenten uit de aangeleverde orthografie.

4. Gebruik de *Alt*-toets in combinatie met de pijltjes naar links en rechts om een spraakfragment naar links of rechts te bewegen (op dezelfde tier), en met de pijltjes naar onder en boven om van tier te verspringen.

5. Wanneer op een bepaald ogenblik meerdere sprekers tegelijk aan het woord zijn, werk dan die passages spreker per spreker af, maar steek geen onnodige tijd in het proberen annoteren van fragmenten waar sprekers echt door elkaar heen spreken, en waar geen kwaliteit kan worden gegarandeerd.

6. Controleer bij de annotatie van een spraakfragment eerst of er sterke grenzen (‖) dienen te worden ingevoegd, duid vervolgens de zwakke grenzen (|) aan, en duid dan pas de prominente lettergrepen (^) en de klankverlengingen (%) aan.

7. Beluister bij het annoteren van een spraaksegment (vooral als het kort is) een langer stukje spraak dat zich uitstrekt over de beide fragmentgrenzen heen.