Protocol voor manuele verificatie van automatisch gegenereerde woordsegmentaties

D. Binnenpoorte, 29-01-02 (+ aanvulling 15-01-04)

A. Algemeen

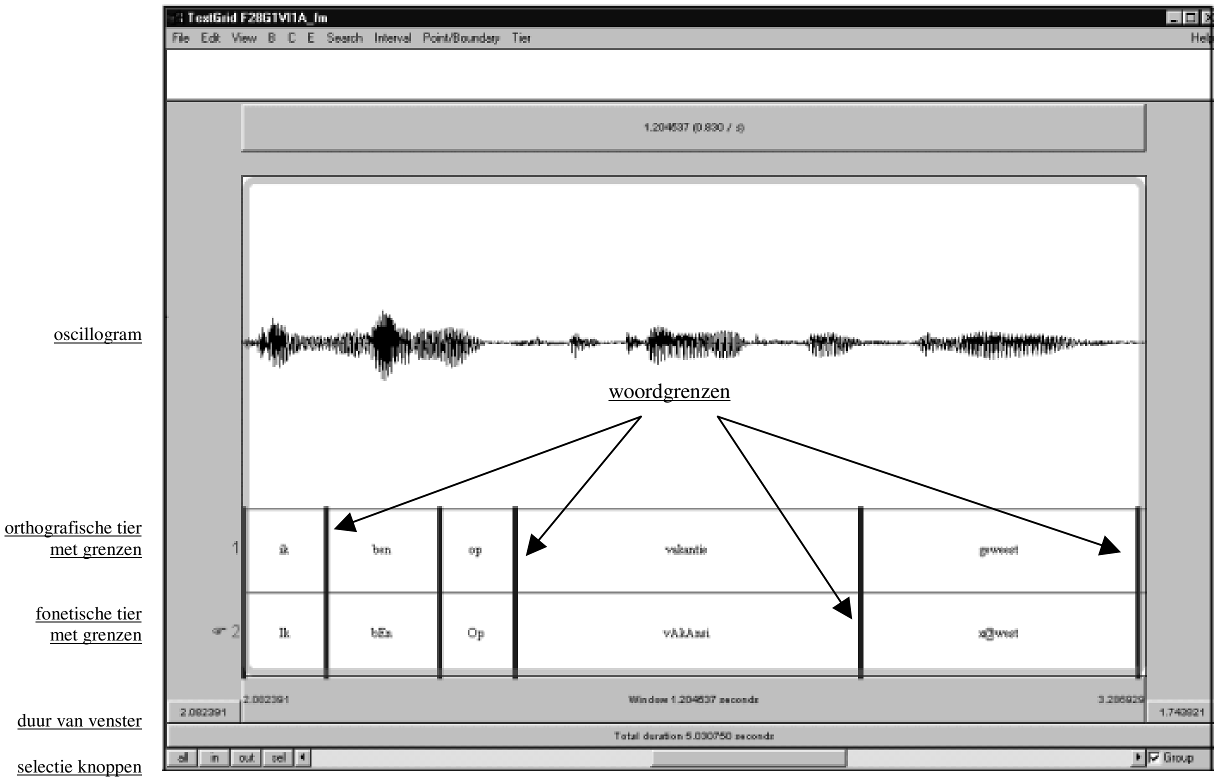
A.1 Doelstelling

Het doel dat nagestreefd wordt, is om in verbonden spraak woorden van elkaar te scheiden door grenzen te plaatsen, zodanig dat de afzonderlijke woorden akoestisch acceptabel klinken. Dit betekent dat de desbetreffende woorden als zodanig auditief herkenbaar zouden moeten zijn. De praktijk leert dat korte woorden zoals 'de', 'en', 'ik', 'met' etc geïsoleerd uit verbonden spraak nauwelijks herkenbaar zijn. Dit soort woorden zullen dan ook niet voldoen aan het 'akoestisch acceptabel' criterium, maar probeer de grenzen altijd zo goed mogelijk te zetten.
De functie van de te plaatsen woordgrenzen is om te dienen als ankerpunten in de tijd. De uiteindelijke gebruiker van het corpus moet een zoekopdracht kunnen geven waarbij bijvoorbeeld een woord wordt gespecificeerd. De ankerpunten geven dan het begin- en eindpunt van het signaal.
De taak van de manuele oplijner is om volgens onderstaande principes automatisch gegenereerde woordgrenzen te controleren en indien nodig te corrigeren.

A.2 Uitgangspunten

Al het spraakmateriaal is zowel orthografisch als fonetisch getranscribeerd. Beide transcripties zijn door mensen gemaakt en vormen de basis voor de reeds automatisch geplaatste woordgrenzen. Om deze te genereren is gebruik gemaakt van een automatische spraakherkenner. De grenzen die door de herkenner worden gevonden moeten dus handmatig worden gecontroleerd en zonodig worden aangepast. Het programma waarmee gewerkt wordt is Praat. In Figuur 1 is een voorbeeld van het werkvenster gegeven van het programma Praat met de volgende termen:

- *Orthografische transcriptie met grenzen* (in bovenste *tier*) geeft de gespelde vorm weer van de woorden die zijn geuit;
- *Fonetische transcriptie met grenzen* (tweede *tier*) laat zien hoe deze woorden zijn gerealiseerd met behulp van een set fonetische tekens die de uitspraak weergeven.
- De verticale lijnen tussen de woorden in de beide tiers zijn de *segmentgrenzen*, of woordgrenzen. Deze grenzen moeten indien nodig worden verschoven door de manuele oplijner.
- Boven de tiers staat het *oscillogram*, een weergave van het spraaksignaal.
- Door middel van de *selectieknoppen* kan de venstergrootte bepaald worden. In Figuur 1 is iets meer dan één seconde van een langere uiting geselecteerd.



Figuur 1 Praat-werkvenster voor de verificatie van de woordoplijning

A.3 Wat zijn woorden?

Woorden zijn alle entiteiten die in de orthografie, de spelling, van elkaar gescheiden zijn door een spatie. De orthografie van het CGN bevat ook entiteiten, codes, zoals ‘ggg’ voor sprekergeluiden, of ‘xxx’ voor onverstaanbare spraak. Ook dit zijn woorden, als deze gescheiden zijn van andere entiteiten middels een spatie. Deze codes kunnen echter ook binnen woorden voorkomen, maar moeten dan uiteraard niet beschouwd worden als woorden. Meer hierover is te lezen in C.5 en C.6.

De vuistregel voor de verificatie van de automatisch gevonden woordsegmentaties is dat <u>alle woorden</u> van elkaar gescheiden worden door een grens. Dus tússen alle woorden staat een grens. Er bestaan hierop twee uitzonderingen.

1. Ten eerste kan er een pauze of stilte tussen woorden optreden. Deze pauze wordt dan voorafgegaan en gevolgd door een grens, zie paragraaf C.4.
2. De tweede uitzondering treedt op als twee opeenvolgende woorden een plosiefklank delen op hun grens, dat wil zeggen het eerste woord eindigt op een plosiefklank en het daaropvolgende woord begint met dezelfde plosiefklank. In dit geval wordt de gedeelde plosief ook omgeven door grenzen. Hierover is meer te lezen in hoofdstuk D.

B. Types woordgrenzen

De automatische woordgrenzen zijn ontstaan doordat de spraakherkenner de klanksymbolen uit de fonetische transcriptie gekoppeld heeft aan de bijbehorende gedeelten van het geluidssignaal. Voor ieder foneem (klanksymbool) wordt zo een stukje signaal gevonden waar het bij hoort. Uit de foneemsegmentaties zijn de uiteindelijke woordgrenzen afgeleid. Deze woordgrenzen zijn zichtbaar gemaakt door middel van de verticale blauwe lijnen. Er moet gecontroleerd worden of deze grenzen op de juiste tijdstippen staan.

Perceptief en fonetisch kan men verschillende types woordgrenzen of plaatsen waar deze voorkomen onderscheiden:

- De meest voorkomende grenzen zijn de normale vloeiende overgangen tussen de woorden. De eindgrens van het eerste woord valt samen met de begingrens van het daarop volgende.
- Tussen twee woorden kan een pauze voorkomen. Deze zou best als een afzonderlijk segment voorgesteld kunnen worden, om te beletten dat een geselecteerd woord door een te lange pauze omringd is.
- Tussen woorden kunnen klanken (*fonemen*) zijn geïnserteerd ten behoeve van een vloeiende uitspraak. Zoals bijvoorbeeld in: *Daarom doe* ***[w]*** *ik het nu.* In de fonetische transcriptie is dit weergegeven als: *darOm du*-**w**-*Ik @t ny.*
  In paragraaf D.2 wordt hier dieper op ingegaan.
- Woorden kunnen fonemen delen op hun grenzen. Dit komt voor als woord1 eindigt op dezelfde klank als waarmee woord2 begint. Bijvoorbeeld: *…naa***r R***otterdam*…. In deze gevallen wordt vaak niet tweemaal de klank [r] uitgesproken. In de fonetische transcriptie zijn deze gevallen weergegeven als *na***r_r***Ot@rdAm.*
  In paragraaf D.1.2 worden meer voorbeelden gegeven en wordt beschreven waar de woordgrenzen in die gevallen moeten staan.
- Bij de overige woordgrenzen kunnen zich assimilatieverschijnselen voordoen. Dit betekent dat klanken elkaar over woordgrenzen heen beïnvloeden. Hier volgt een voorbeeld:
  Is d*it goed?* Hier wordt de [*s*] onder invloed van de [*d*] uitgesproken als een [*z*] en fonetisch wordt het als volgt weer gegeven: *Iz dIt Gut?* In de sectie over fonetische achtergrondinformatie kan meer uitleg over dit verschijnsel gevonden worden.

C. Basisprincipes voor de verificatie

Het belangrijkste uitgangspunt bij het verifiëren van de woordgrenzen is dat dit consistent gebeurt. Dit is te verwezenlijken door het volgen van enkele regels.

C.1 Consistentie met orthografische transcriptie

De woordoplijning moet consistent blijven met de orthografische transcriptie. Deze transcriptie staat in de bovenste transcriptie-*tier*, zie paragraaf A.2. Voor elk woord (inclusief ‘xxx’, ‘ggg’) in de orthografische transcriptie moet er in de oplijning een segment afgebakend zijn.

Hou de oplijning consistent met de orthografische transcriptie.

Aarzelingen, gestotter en alle andere tussenwerpsels zijn orthografisch getranscribeerd. Beschouw deze verschijnselen als afzonderlijke woorden, omgrensd door woordgrenzen. Aarzelingen binnen een woord geven nooit aanleiding tot extra woordgrenzen, aangezien ze het aantal orthografische eenheden, woorden, niet veranderen.

Aarzelingen, bijvoorbeeld *uh*, en andere tussenwerpselen dienen middels segmentgrenzen te worden afgezonderd. Indien deze spraakgeluiden zich BINNEN een woord voordoen worden ze niet afgezonderd met segmentgrenzen.

Het is niet toegestaan iets te veranderen aan de orthografische transcriptie. Eventuele fouten of vreemde verschijnselen dienen te worden gerapporteerd aan de leidinggevende.

Verander NOOIT iets aan de orthografische transcriptie.

In de orthografische transcriptie kunnen woorden zijn opgesierd met zogenoemde sterretje-codes. Deze codes worden door de orthografie gebruikt om de volgende mogelijke verschijnselen te kunnen onderscheiden.

*v → woorden uit vreemde taal
*d → dialectwoorden
*z → zware dialectwoorden
*n → 'nieuwe' woorden
*t → nieuwe tussenwerpsels
*a → afgebroken woord
*u → vreemde uitspraak of verspreking
*x → niet goed verstane woorden

Verander nooit wat aan deze codes, maar gebruik de extra informatie bij de controle van de woordgrens. De code vormt samen met het woord waarachter het staat één entiteit.

C.2 Consistentie met fonetische transcriptie

In de fonetische transcriptie-*tier*, zie paragraaf A.2, is zeer veel informatie weergegeven over de verschijnselen, vermeld in sectie B, die zich voor kunnen doen op woordgrenzen. Een voorbeeld:

(1) a. zE+ hAd@n nits [zij hadden niets]
b. zE+ hAd@n_nits [zij hadden niets]
c. zE+ hAd@ nits [zij hadden niets]

Deze voorbeelden zijn verschillende realisaties van de zin 'zij hadden niets', zoals deze zijn getranscribeerd door de expertluisteraars van de fonetische transcriptie groep. Voorbeeld (1)a geeft een uitspraak weer van de zin waarin zowel de eind-/n/ in 'hadden' als de begin-/n/ in 'niets' beide zijn uitgesproken. De twee /n/-en zijn prima te scheiden. Deze situatie zal zich in verbonden spraak niet vaak voordoen. In voorbeeld (1)b is sprake van een gedeeld foneem, *degeminatie*. Het is de bedoeling dat in beide woorden dat foneem hoorbaar is, zie hoofdstuk D. Voorbeeld (1)c laat zien dat de /n/ in 'hadden' helemaal niet is uitgesproken. De grens tussen de woorden ligt hier na de /@/ en voor de /n/ van 'niets'.
De gegeven fonetische en orthografische transcripties zijn doorslaggevend voor de auditieve eisen die men stelt aan de gesegmenteerde woorden. De transcripties zijn gemaakt door getrainde transcribenten die bovendien contextinformatie hebben gebruikt bij de generatie van de transcripties.

Hou de oplijning consistent met de fonetische transcriptie.

Evenals voor de orthografische transcriptie geldt ook hier dat het niet is toegestaan iets te veranderen aan de fonetische transcriptie. Merkwaardigheden of fouten dienen te worden gerapporteerd bij de leidinggevende.

Verander NOOIT iets aan de fonetische transcriptie.

C.3 Geen onnodige verplaatsing van grenzen

De woordgrenzen zoals gegenereerd door de automatische spraakherkenner moeten worden gecontroleerd. Elk woord dat is gesegmenteerd dient daartoe te worden beluisterd. Als op basis van de auditieve impressie een woordgrens verplaatst moet worden, moet dit altijd gebeuren door beluistering van het woord vóór de grens en het woord na de grens. Nogmaals, het criterium voor het verplaatsen van grenzen is altijd de auditieve impressie. Voor lange woorden geldt: als de woorden akoestisch acceptabel klinken, dan is de grens acceptabel. Voor korte woordjes geldt dat verplaatsing van de grens vaak niets meer oplevert dan de uitgangspositie. In die gevallen moet de grens blijven staan.
Het is mogelijk om één maal de 'ongedaan maken' functie in Praat te gebruiken om een grens terug te plaatsen naar zijn voorgaande positie. Let op: als een grens vanuit de oorspronkelijke positie twee maal wordt verplaatst, kan met de 'ongedaan maken' functie alleen de voorlaatste, dus niet de oorspronkelijke, positie worden teruggevonden.

Hou je bij de controle aan de pragmatische instelling: 'Goed is goed'.

Verschuif de genzen niet onnodig over 'kleine afstanden'.

C.4 Pauzes en andere niet getranscribeerde geluiden

De automatische segmentatie heeft ook 'stiltes' gemarkeerd als een afzonderlijk segment. Verzet de grenzen ervan alleen als het woord ervoor of erna niet goed klinkt, of wanneer de zogenaamde 'stilte' pertinent een stuk wel getranscribeerde spraak bevat. In uitzonderlijke gevallen kan het voorkomen dat een 'stilte' helemaal verwijderd moet worden. Als de automatische segmentatie hinderlijke 'stiltes' niet aangeduid heeft, voeg ze dan zelf in, met een lege transcriptie.

Verander zo weinig mogelijk aan de automatisch gevonden pauzes.

Indien storende niet-getranscribeerde geluiden nog niet automatisch zijn gesegmenteerd, voer dan grenzen in voor deze geluiden.

Figuur 2 illustreert hoe pauzes tussen woorden in kunnen staan en omgeven worden door segmentgrenzen.
Om te voorkomen dat er teveel pauze segmenten worden omgeven met grenzen is de regel gesteld dat een pauze korter dan 50 ms (0.050 s) niet als apart segment is weergegeven door de automatische segmentatie. Als een stilte of pauze door verschuiving van grenzen korter wordt dan 50 ms dient deze te worden verwijderd.

Pauzes, of stiltes, korter dan 50 ms worden niet gemarkeerd.

Een pauze of stilte begint als het voorafgaande woord volledig is.

C.5 Segmenteren van sprekergeluiden

Sprekergeluiden, zoals gelach, gegil, gekuch, zijn in de orthografische transcriptie-*tier* aangeven met 'ggg', en in de fonetische transcriptie met het teken '#'. Deze moeten worden beschouwd als woorden als ze zijn omringd door spaties. Uiteraard moeten deze ook geverifieerd worden. Bepaal de begingrens altijd door beluistering van het voorafgaande woord en verander zijn positie enkel als dit woord niet acceptabel klinkt. Voor de bepaling van de eindgrens van het sprekergeluid moet het daaropvolgende woord goed klinken. Het sprekergeluid, 'ggg', is dus van ondergeschikt belang ten opzichte van de omringende woorden. Verzet de grenzen om deze geluiden dan ook niet als dit niet nodig is.

Sprekergeluiden buiten de woorden zijn middels segmentgrenzen afgezonderd.

C.6 Segmenteren van 'xxx'

In de transcripties komen ook '*xxx*'-coderingen voor, in de fonetische transcriptie zijn deze weergegeven als '[]'. Ze wijzen op één of meerdere opeenvolgende woorden die niet goed verstaan zijn. Beschouw een dergelijk stuk als één woord af te bakenen door segmentgrenzen.

Onverstaanbare spraak segmenteren als één woord, omgeven door segmentgrenzen.

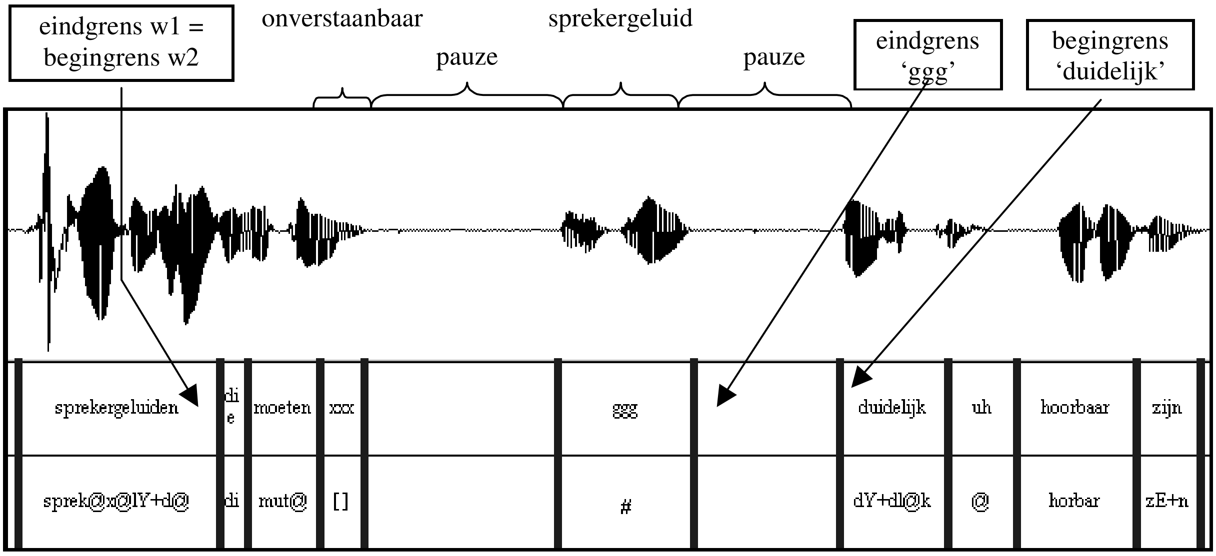
De '*xxx*'-codering kan ook binnen woorden voorkomen. Het woord waarin dit voorkomt moet worden behandeld als een gewoon woord en in zijn geheel worden omgeven met grenzen.

C.7 Voorbeeld

In Figuur 2 is een fragment genomen waarin de meeste van bovengenoemde verschijnselen zich voordoen. Belangrijke punten zijn:

- de eindgrens van woord1 valt samen met de begingrens van woord2, de grens staat tussen de twee woorden;
- onverstaanbare spraak ('xxx') is begrensd met woordgrenzen;
- sprekergeluiden ('ggg') zijn ook begrensd met woordgrenzen. Zowel de onverstaanbare spraak als de sprekergeluiden zijn getranscribeerde entiteiten;

- een pauze (ofwel stilte ofwel een storend achtergrondgeluid) begint waar een voorafgaand woord (hier: 'ggg') eindigt. Een pauze eindigt waar een volgend woord (hier: 'duidelijk') begint.



Figuur 2 Voorbeeld van pauzes en sprekergeluiden

D. Gedeelde fonemen op woordgrens

In hoofdstuk B is al kort aan de orde gekomen welke verschijnselen kunnen optreden op woordgrenzen in verbonden spraak. Nogmaals, in verbonden spraak zijn woorden niet van elkaar gescheiden door pauzes, zoals spaties in geschreven taal, maar lopen de klanken van de woorden in één stroom door. Er wordt uiteraard wel gepauzeerd voor ademhaling, of om andere redenen die bijdragen tot de verbetering van de communicatie. De woordgrenstypes die voor moeilijkheden zouden kunnen zorgen zijn die waarbij gedeelde (hoofdstuk D) of juist ingevoegde fonemen (hoofdstuk E) voorkomen. Hieronder staat per verschijnsel beschreven hoe de woordgrenzen moeten worden gecontroleerd. Elk soort kent ook zijn eigen notaties en codes die al automatisch zijn weergegeven, en daardoor makkelijker te vinden. Deze extra codes dienen als wegwijzers, ze mogen niet worden veranderd of toegevoegd.

D.1 Wat zijn gedeelde fonemen?

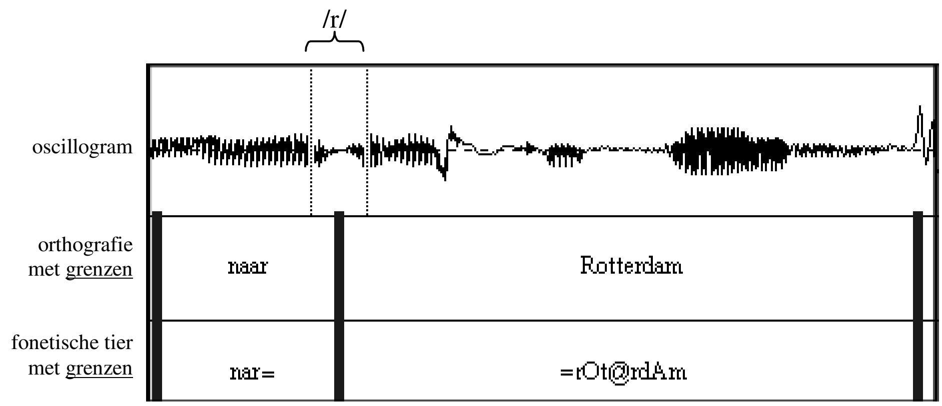
In verbonden spraak vindt *assimilatie*, of 'aanpassing', van de klanken plaats, zowel woordintern als op woordgrenzen. Een apart geval van assimilatie op woordgrenzen, heet *degeminatie*. Hiervan kan sprake zijn als de eindklank van het ene woord dezelfde is als de beginklank van het er opvolgende woord zodat die klank gedeeld wordt door beide woorden. Bijvoorbeeld in:

(2)	a. Ik wil naa**r R**otterdam.	[gedeelde sonorant]
	b. Hij i**s z**eker vandaa**g g**evallen.	[gedeelde fricatief]
	c. Hij kom**t t**erug.	[gedeelde **stemloze plosief**]
	d. Ik he**b b**elachelijk veel te doen.	[gedeelde **stemhebbende plosief**]
	e. Hij wil va**st st**oppen.	[**meervoudig gedeelde foneemsequentie**]
	f. Het lief**st st**a ik vroeg op.	[**meervoudig gedeelde foneemsequentie**]

Deze klanken worden vaak als één uitgesproken en zullen daarom wellicht moeilijk te scheiden zijn. Hieronder zijn de richtlijnen voor deze gevallen weergegeven.

D.2 Niet-plosieven als gedeeld foneem

In het geval van voorbeeldzin (2)b: ‘*Hij is zeker vandaa**g** **g**evallen*’ zal het goed mogelijk zijn om ongeveer het midden van de ‘g’ als grens tussen /*vandaa**g***/ en /***g**evallen*/ te nemen. Dit is eveneens het geval in zin (2)a: ‘*Ik wil naa**r** **R**otterdam*’. Dit soort fonemen, de niet-plosieven, lenen zich daar uitstekend voor, omdat ze niet, zoals bij plosieven wel het geval is, uit twee fases bestaan, zie hiervoor paragraaf D.3.
Doordat deze woordparen een foneem delen, zullen de woorden na het zetten van de grens niet klinken zoals dat het geval zou zijn als deze zijn gesproken in isolatie. De gedeelde fonemen van de twee woorden zullen niet langer zijn dan wanneer ze binnen een woord uitgesproken worden. Toch moeten ze gedeeld worden, waardoor de woorden niet ‘af’ zullen klinken, maar toch goed herkenbaar.
Indien er in het foneem meerdere energiemaxima optreden, zoals bij een /r/ kan voorkomen, in Figuur 3 geïllustreerd, dan moet men een energieminimum in de buurt van het midden selecteren.



Figuur 3 Niet-plosief als gedeeld foneem: ongeveer midden

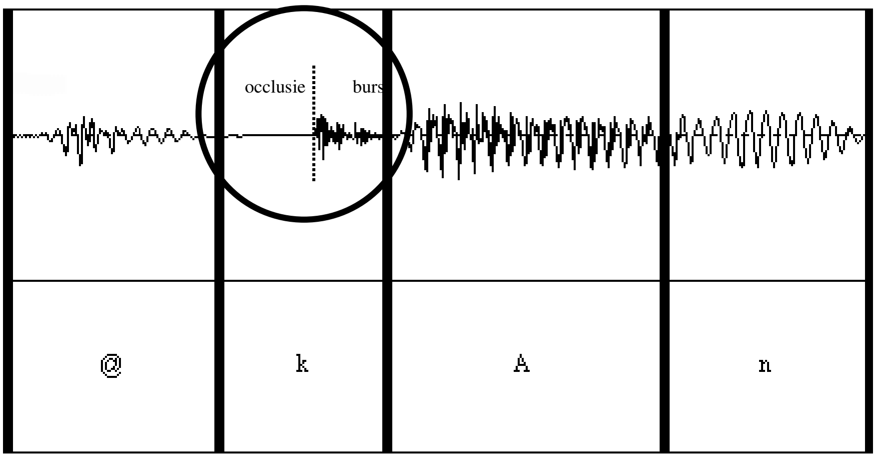
Figuur 3 geeft een voorbeeld, uit de voorbeeldzin (2)a. Er is duidelijk te zien dat ongeveer het midden van de /r/ een woordgrens kan zijn. In de tier met de fonetische transcriptie is deze grens gemarkeerd met een ‘**=**‘-teken om aan te geven dat de /r/ in dit geval een gedeeld foneem is waardoor de woorden die de /r/ delen niet volledig klinken.

Zorg in geval van gedeelde fonemen, die niet tot de plosieven behoren, dat het gedeelde foneem hoorbaar is in beide woorden.

D.3 Plosieven (/p/, /t/, /k/, /b/, /d/, /g/) als gedeeld foneem

De gedeelde fonemen in voorbeeld (2)c en (2)d zijn *plosieven*. De plosieven in het Nederlands zijn /p/, /t/, /k/ (de stemloze plosieven) en /b/, /d/, /g/ (de stemhebbende plosieven). Deze fonemen heten zo omdat bij de realisatie ervan een soort explosie plaatsvindt. Er wordt een afsluiting gemaakt door een van de *articulatoren*, zie paragraaf F.2, en deze wordt plotseling geopend, waardoor er een ontsnapping van lucht plaatsvindt. De plek waar de afsluiting zich bevindt en de aan- of afwezigheid van trilling van de stembanden, bepaalt welke plosief hoorbaar is gemaakt. Voordat er een ontsnapping van de lucht plaatsheeft, wordt er eerst een

luchtdruk opgebouwd achter de afsluiting. Dit wordt de *occlusie* genoemd. Daarna wordt de lucht losgelaten, de *burst*, waardoor de klank tot uiting komt. Als twee woorden een plosief op de grens delen, zoals in (2)c en (2)d, wordt niet tweemaal de luchtdruk opgebouwd en weer losgelaten in verbonden spraak.
Als twee woorden gescheiden moeten worden, en het midden van de duur van een plosief als leidraad wordt genomen als scheidingspunt, zal in het eerste woord meestal niets van de identiteit van de plosief hoorbaar zijn. In dat gedeelte wordt immers de luchtdruk opgebouwd, en de plosief zal dan enkel pas in het tweede woord deels te horen zijn. Figuur 4 verheldert het probleem. Hier is @kAn weergegeven, als deel van het woordpaar 'd**e kan**tine'. De plosief /k/ is omcirkeld. Er is te zien hoe de plofklank is opgebouwd uit een luchtdrukopbouwend gedeelte, de occlusie, en een deel waarin de occlusie wordt geopend en de lucht vrijkomt, de burst. Dit voorbeeld heeft betrekking op een stemloze plosief, de /k/. Een vergelijkbare situatie doet zich voor bij de andere stemloze plosieven en de stemhebbende plosieven, de /b/, de /d/ en de /g/.
Het verschil tussen de stemhebbende en de stemloze plosieven is de stembandtrilling die vaak zowel tijdens de luchtdrukopbouw als tijdens de burst aanwezig is bij de stemhebbende plosieven. Het occlusiegedeelte van de stemhebbende plosieven zal een trilling laten zien met een kleine amplitude, in plaats van een vlakke lijn. Echter, de overgang naar de burst is doorgaans nog steeds goed zichtbaar.

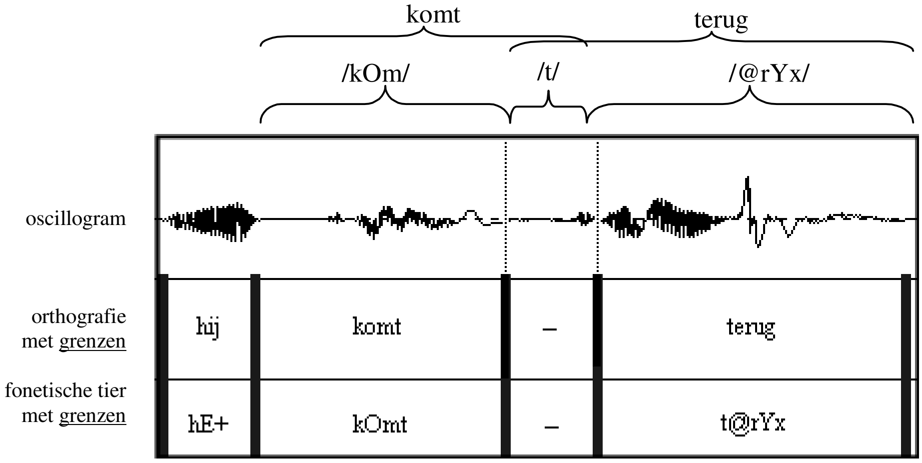


Figuur 4 Stemloze plosief, de /k/

Gedeelde plosieven zijn dus strikt genomen niet op te splitsen en we zullen dat dan ook niet proberen. In plaats daarvan wordt de gedeelde plosief als een zelfstandig segment gezien.

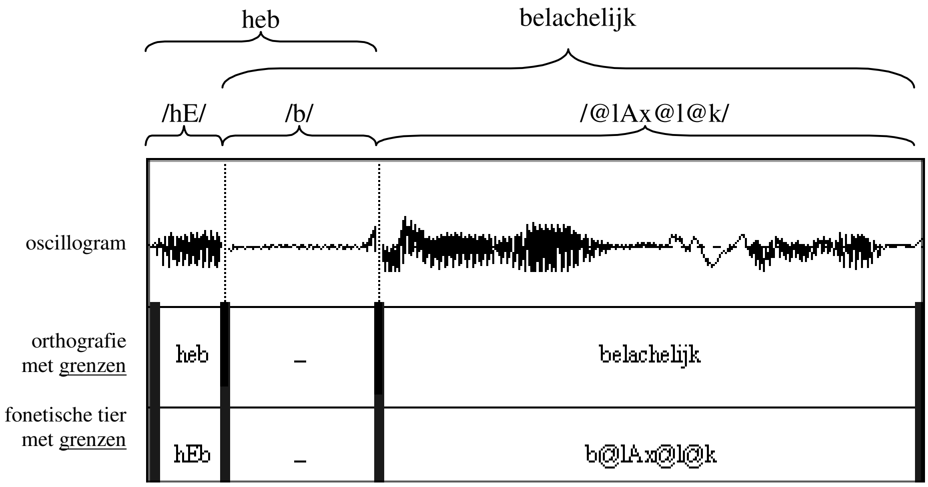
D.3.1. Enkelvoudige gedeelde plosieven

Figuur 5 is de weergave van voorbeeld (2)c waarin de /t/ als gedeelde plosief optreedt in de zin '*Hij komt terug*'. Zoals te zien is, wordt de /t/ als geheel als een apart segment genomen en is in het werkvenster weergegeven als een underscore, '_'. Bij beluistering zal ook slechts een /t/ te horen zijn. Om het woord 'komt' te beluisteren moeten twee segmenten worden geselecteerd, zowel het gedeelte /kOm/ als het segment dat de hele /t/ bevat, en achterelkaar worden beluisterd. Om het woord 'terug' te beluisteren, begint de selectie met het /t/-segment en daarbij het /@rYx/-segment.



Figuur 5 Gedeelde stemloze plosief

Figuur 6 is een voorbeeld van een stemhebbende plosief die gedeeld wordt. Het geeft een detail uit voorbeeld (2)d: '*Ik he**b** **b**elachelijk veel te doen*'. Ook hier geldt weer dat het woord 'heb' hoorbaar moet worden gemaakt door zowel het segment waarin /hE/ te horen is als het segment waarin alleen de /b/ zit te selecteren en achtereenvolgens te beluisteren.



Figuur 6 Gedeelde stemhebbende plosief

Hetzelfde geldt weer voor het woord 'belachelijk', eerst het segment met de /b/ en daarna het deel met /@lAx@l@k/.

Gedeelde plosieven worden in een apart segment gezet, weergegeven middels een '_' in zowel de fonetische als de orthografische tier en dienen bij beluistering samengevoegd te worden met het voorafgaande of het volgende woord.

D.3.2 Gedeeld plosief als woord

Door snelle uitspraak komt het regelmatig voor dat woorden worden gereduceerd. Bij korte woorden kan dit zelfs doorgaan tot één klank. Bijvoorbeeld: *het is* wordt gereduceerd tot *'t is.*

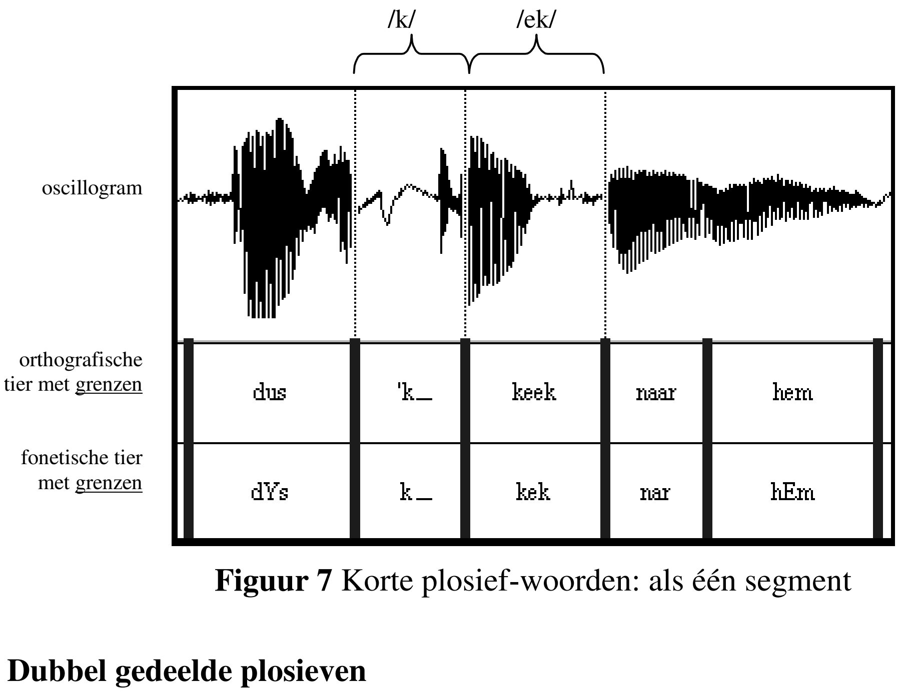
Het kan natuurlijk voorkomen dat een gereduceerd woord een plosief deelt met een ander woord, of wel zichzelf deelt met een ander woord zoals in onderstaande voorbeelden:

(3)	a. dat ’t niet uitmaakt	[orthografie]
	b. dAt_t nit Y+tmakt	[fonetische weergave]
	c. dus ‘k keek naar hem	[orthografie]
	d. dYs k_kek nar hEm	[fonetische weergave]

De woorden ‘’t’ (het) en ‘’k’ (ik) zijn in deze gevallen zodanig kort weergegeven dat de gehele duur samenvalt met de eerste (of laatste) plosief van een ander woord. Toch is er voor de orthografische en voor de fonetische transcribent voldoende evidentie om dit toch als een apart woord te zien.
Er is gekozen om dit soort gevallen een aparte behandeling te geven. De woorden die bestaan uit de plosief die gedeeld is, worden niet in twee delen gescheiden, woord en ‘_’-teken, maar behoudt de status van woord en blijft één segment. Om te weten of dit woord een ‘zelfstandig’ woord is of zoals in de voorbeelden in (3) tevens een gedeeld plosief, moet dit zowel in de orthografische als in de fonetische tier worden aangegeven met een ‘_’ teken aan de kant waar het foneem gedeeld wordt binnen het segment. De voorbeelden uit (3) worden als volgt in de orthografische tier weergegeven:

(4)	a. dat _’t niet uitmaakt	[_ aan linkerzijde van ‘t]
	b. ‘k_ keek naar hem	[_ aan rechterzijde van ‘k]

Figuur 7 verheldert de oplossing. In deze figuur is een deel van voorbeeld (3)c/d weergegeven. Om het woord ‘keek’ te kunnen beluisteren moet zowel het segment met ‘k_’ als het ‘keek’ segment worden beluisterd.



Figuur 7 Korte plosief-woorden: als één segment

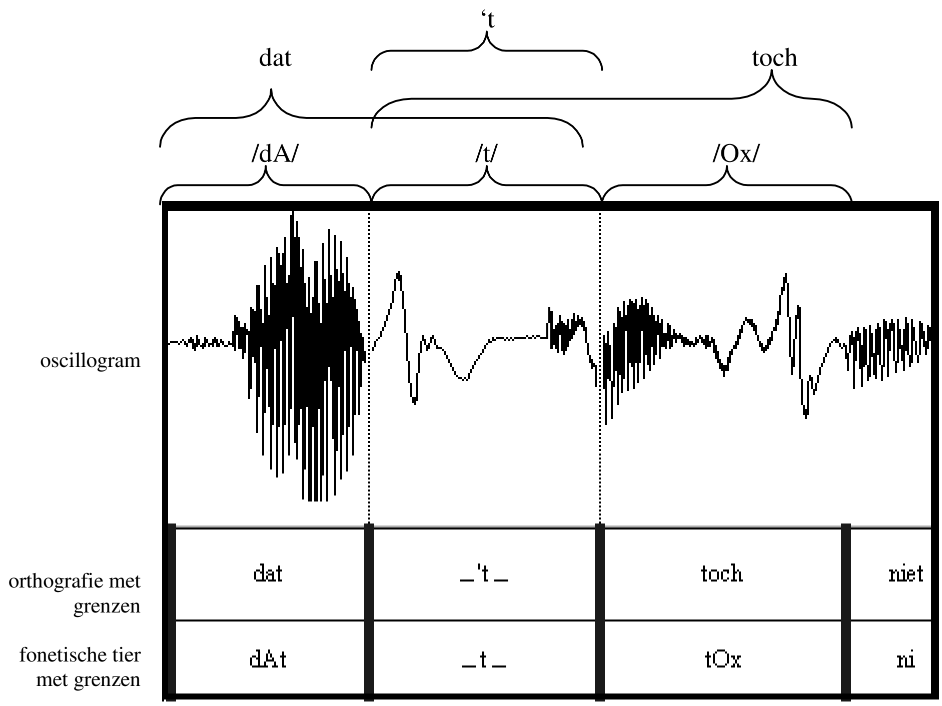
D.3.3 Dubbel gedeelde plosieven

We spreken van dubbel gedeelde plosieven wanneer een gedeeld plosief van twee woorden zelf ook nog een woord is. Een voorbeeld:

(5)	a. **dat ’t toch** niet kan	[orthografie]
	b. **dAt_t_tOx** nit kAn	[fonetische weergave]

c. dan **maak ‘k kip** met rijst [orthografie]
d. dAn **mak_k_kIp** mEt rE+st [fonetische weergave]

Zowel in voorbeeld (5)b als (5)d is te zien dat respectievelijk de woordjes ’t en ‘k dermate gereduceerd zijn uitgesproken dat ze samen vallen met het woord ervoor als erna. Er is dan sprake van een zogenoemd dubbel gedeeld plosief. Het plosief deelt zichzelf met het voorafgaande woord met het daaropvolgende.
Voor deze gevallen is een aparte notatie gekozen, ‘**_t_**’, om weer te geven dat het gaat om een dubbel gedeeld plosief in dat specifieke segment. In Figuur 8 is een afbeelding van het voorbeeld (5)a/b. Voor de duidelijkheid, om het woord ‘dat’ te kunnen beluisteren, moet ook het segment met het ‘_t_’-label worden geselecteerd. Voor het woord ‘’t’ alleen het uitzonderingssegment ‘_t_’ en tenslotte dient voor beluistering van het woord ‘toch’ zowel het segment met ‘_t_’ als het segment ‘toch’ te worden geselecteerd en afgespeeld.



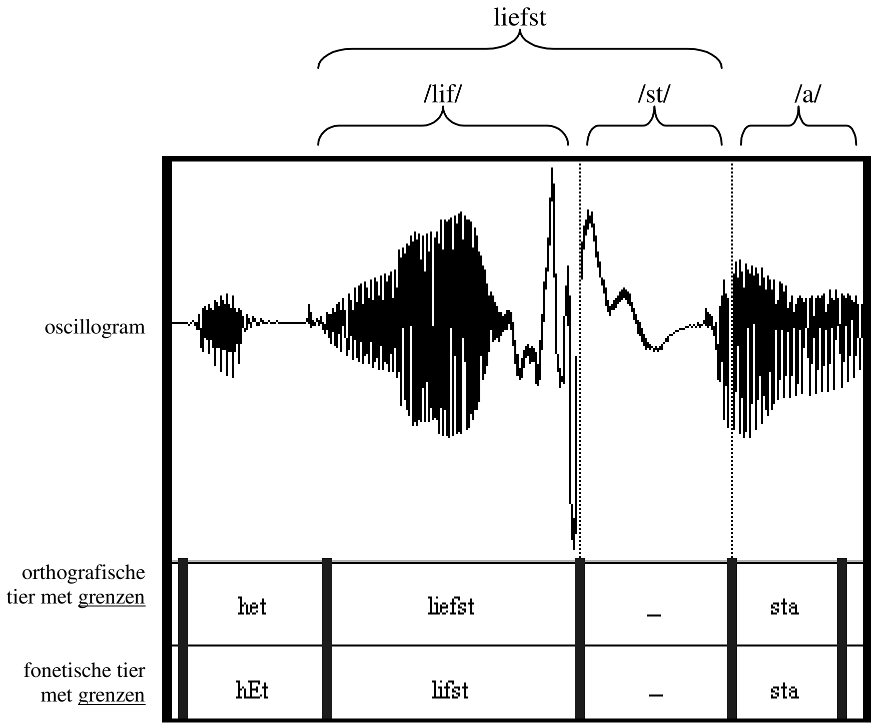
Figuur 8 Dubbel gedeeld plosief

D.3.4 Meervoudig gedeelde fonemen

In geval van meervoudig gedeelde fonemen worden op woordgrenzen een reeks van fonemen gedeeld. Voorbeelden (6)a en (6)c (uit hoofdstuk D.1), laten twee fonemen zien die gedeeld worden.

(6) a. Hij wil va**st st**oppen. [meervoudig **gedeelde foneemsequentie**]
b. hE+ wIl vA**st_st**Op@n [fonetische transcriptie]
c. Het lief**st st**a ik vroeg op. [meervoudig **gedeelde foneemsequentie**]
d. @t lif**st_st**a Ik vrux Op. [fonetische transcriptie]

Van deze gedeelde foneemreeks is niet met zekerheid te zeggen bij welk van de twee woorden ze worden uitgesproken. Dit verschijnsel moet conform de enkelvoudige gedeelde plosieven worden behandeld. De sequentie van gedeelde fonemen moet worden gezien als één entiteit en moet in een apart segment worden gezet. In zowel de orthografische als de fonetische tier wordt het weergegeven met het '_' teken. In Figuur 9 is de zin uit voorbeeld (6)c/d weergegeven met de juiste segmentgrenzen.



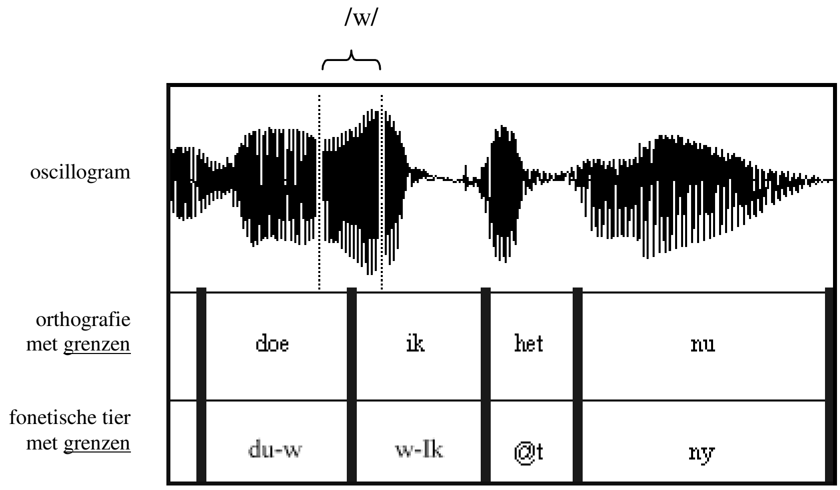
Figuur 9 Meervoudig gedeelde fonemen

E. Tussenklanken

Een ander verschijnsel is <u>insertie</u> van klanken. Bijvoorbeeld in:

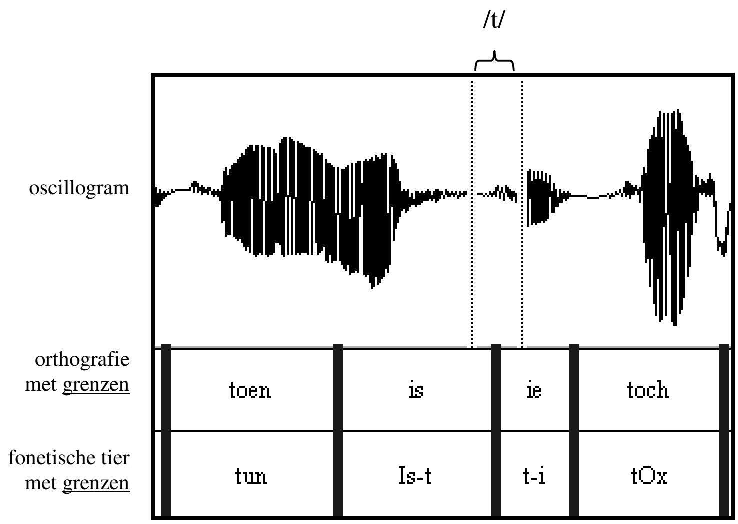
(7) a. Daarom doe **/w/** ik het nu.
b. Toen belde **/n/** ie naar huis.

In deze gevallen wordt een klank tussengevoegd ten behoeve van de vloeiendheid van de spraak. Het is een soort tussenklank, die is ontstaan doordat de articulatoren zich van de ene naar de andere uiterste positie moeten verplaatsen. Dit verschijnsel is optioneel.
Als hier twee verbonden woorden moeten worden gescheiden zal een deel van het geïnserteerde foneem, bijvoorbeeld in (7)a de /w/, hoorbaar blijven als beide woorden beluisterd worden. De richtlijn voor de positiebepaling van de grens in geval van een tussenklank is ongeveer het midden van die tussenklank. Anders dan bij gedeelde fonemen waar de hoorbaarheid van het gedeelde foneem erg belangrijk is, is dit bij geïnserteerde fonemen minder van belang. Het belangrijkste is dat de woorden aan beide zijden acceptabel klinken. Dit geldt ook als het geïnserteerde foneem een plosief is. In dit laatste geval wordt deze plosief niet als een apart woord behandeld.



Figuur 10 Geïnserteerde tussenklank: ongeveer midden

Figuur 10 geeft een illustratie van het verschijnsel waarbij de /w/ is geïnserteerd. Let op zowel de fonetische als de orthografische transcriptie. In de fonetische transcriptie worden geïnserteerde fonemen geschreven met een koppelteken, '-', tussen het woord en de tussenklank.
Figuur 11 geeft een voorbeeld waarbij het geïnserteerde foneem een plosief is, de /t/. Deze worden zoals gezegd op dezelfde wijze behandeld als niet-plosieven die zijn geïnserteerd. Er bestaat weliswaar geen midden van een plosief, daarom geldt hier dat de woorden aan beide zijden van het plosief acceptabel moeten klinken.



Figuur 11 Geïnserteerde plosief: beide woorden acceptabel

F. Werkwijze

Samengevat, de controle van de automatische woordgrenzen gaat als volgt:

- Verander niets aan de transcripties, maar rapporteer altijd eventuele fouten. Dit geldt zowel voor de orthografische als de fonetische transcripties;
- Gebruik het geluid, de orthografie en de fonetische transcriptie als leidraad;

- Beluister woord voor woord. Als iets niet acceptabel klinkt, beluister dan de woordparen aan beide zijden van de betreffende grens;
- Verzet geen grenzen als dit niet nodig is: goed is goed;
- Gebruik de grafische informatie van het oscillogram slechts als aanvulling bij de auditieve impressie;
- Kies voor een werkbare situatie een venstergrootte van 1 tot maximaal 2 seconden;
- Gebruik voorgenoemde regels voor het verzetten van de grenzen;

G. Achtergrondinformatie

G.1 Verbonden spraak

Het materiaal dat verzameld wordt in het Corpus Gesproken Nederlands (CGN) bestaat uit veel verschillende soorten spraak. Deze spraakmodi variëren van spontane dialogen tot voorgelezen spraak (monologen). De spraak is te karakteriseren als verbonden spraak. Dat wil zeggen dat er geen pauzes tussen woorden worden ingelast, zodat woorden niet geïsoleerd worden uitgesproken. In verbonden spraak worden de eindklank van het eerste en de beginklank van het tweede woord vaak met elkaar verbonden. Hieronder wordt kort, op een fonetische manier, uitgelegd welke verschijnselen zich rond woordgrenzen voordoen die relevant zijn voor het plaatsen van woordgrenzen.

G.2 Fonetiek

Als je spraak produceert gebruik je daarvoor je articulatoren. Dit zijn onder andere je lippen, je tong, je kaak, je stembanden, etc. Deze articulatoren werken samen voor de productie van spraak. Ze zijn op elkaar afgestemd met betrekking tot timing (op het juiste moment gaan je lippen van elkaar als je het woord /maar/ wil zeggen en van de /m/ naar de /a/-klank gaat). Maar ze zijn ook afgestemd op elkaar met betrekking tot plaats. Met plaats wordt bedoeld waar de vernauwing van de opening of de complete afsluiting tussen de tong en het verhemelte zich bevindt, voor of achter in de mondholte. Elke stand van de tong geeft een typische klank. Bijvoorbeeld, er ontstaat een /x/-klank of /G/-klank als de afsluiting zich achterin bevindt, en een /l/-klank kan worden gemaakt als de tongpunt zich tegen de binnenkant van de boventanden bevindt. Uiteraard is de stand van alle andere articulatoren op dat moment doorslaggevend voor de typische klank.
Naast de afwisseling van standen van de articulatoren is er ook afwisseling in het wel of niet laten trillen van de stembanden. Het verschil tussen een /s/ en een /z/ is de stembandtrilling die wel aan de orde is bij de /z/-klank maar niet bij de /s/. De stand van de articulatoren blijft in deze vergelijking gelijk. De /z/-klank is een stemhebbende klank, de /s/ is een stemloze klank.
Een uiting, een zin of een woord, bestaat uit een opeenvolging van verschillende klanken (fonemen) die ieder een eigen manier van produceren (standen van de articulatoren) hebben. De verandering van de standen van de articulatoren gaat in een rap tempo. Als klanken elkaar snel opvolgen zullen ze elkaar beïnvloeden, de articulatoren kunnen in korte tijd waarschijnlijk niet helemaal in de karakteristieke stand terecht komen als de volgende klank vraagt om een andere uiterste positie. Dit verschijnsel van beïnvloeding heet *coarticulatie*. Dit is ook hoorbaar. Een ver doorgevoerde vorm van coarticulatie, wanneer zelfs de foneemidentiteit verandert, is *assimilatie*.
Als er woorden worden geknipt uit geïsoleerde spraak, dus met een duidelijke pauze tussen de woorden, klinken deze anders dan woorden geknipt uit verbonden spraak. In het geval van de geïsoleerd uitgesproken woorden zal er geen assimilatie plaatsvinden aan de begin- of eindklanken, fonemen, van het woord. Dit is dus wel het geval bij woorden uit verbonden spraak. De begin- en eindfonemen hebben een hoorbare invloed ondergaan van de aangrenzende klanken.

Bijvoorbeeld:

(8)	a. is dit alles	[orthografie]
	b. **Iz** dIt Al@s	[fonetisch]

De stemhebbende /d/ aan het begin van *dit* beïnvloedt de realisatie van de /s/ in *is*. In plaats van stemloos wordt de /s/ ook stemhebbend uitgesproken, als een /z/.

(9)	a. dat die nog loopt	[orthografisch]
	b. dAt **t**i nOx lopt	[fonetisch]

In (9)b ondergaat de stemhebbende /d/ van *die* een verandering onder invloed van de stemloze /t/ aan het einde van *dat* en wordt daardoor ook stemloos.
In zowel (8) als (9) zijn voorbeelden gegeven waarbij de plosief van het ene woord zijn stemhebbendheid of stemloosheid doorgeeft aan het foneem van het ander woord.

(10) a. ik schreef veel in bad
b. Ik sxref **f**el i**m** bAt

Het eerste vetgedrukte voorbeeld in (10)b geeft een verandering weer van een stemhebbende /v/, naar een stemloos foneem, de /f/. Dit is gebeurd onder invloed van de daaraan voorafgaande fricatief /f/ uit *schreef*. Het tweede voorbeeld in (10)b laat zien dat de plaats van articulatie van de /n/ meer gaat lijken op die van de /b/ waardoor de /n/ verandert in een /m/. Bij realisatie van zowel de /m/ als de /b/ zijn de lippen gesloten.

H. Aanvulling

H.1 Dubbel gedeelde plosieven (3 september 2003 door W. Goedertier)

In paragraaf D.3.2 (Gedeeld plosief als woord) en paragraaf D.3.3 (Dubbel gedeelde plosieven) wordt uitgelegd hoe men omgaat met gedeelde plosieven in de woordoplijning. Ondertussen is er echter nog een speciaal geval opgedoken dat beide fenomenen combineert. Een praktijkvoorbeeld (11):

(11)	a. dan*z gaat de*d ’t zeker weten	[orthografisch]
	b. tA **Gat_t_t** sek@r wet@n	[fonetisch]

De woordoplijning zou dan als volgt moeten zijn:

ORT-tier:	dan*z	gaat	_de*d_	't	zeker	weten
FON-tier:	tA	Gat	_t_	t	sek@r	wet@n

De positie van de grenzen (tijdsinformatie van de segmenten) moet dan als volgt zijn:

2de segment (gaat/Gat) : bevat foneem /G/ en /a/
3de segment (_de*d_/_t_) : bevat foneem /t/
4de segment ('t/t) : bevat NIETS !!

Theoretisch zou het 4de segment dus een lengte NUL moeten hebben. Aangezien een segment van lengte NUL problemen geeft in Praat, zullen we dit segment een HEEL KORTE lengte geven (5 à 10 milliseconden). (In de automatische woordoplijning krijgen deze segmenten een lengte van exact 1 milliseconde.)

Een tweede praktijkvoorbeeld (12) van hetzelfde fenomeen dat analoog kan worden opgelost:

(12)	a. je hebt ’t de hele tijd zo gezegd	[orthografisch]
	b. j@ **Ept_t_t** hel tE+t so x@z@xt	[fonetisch]

H.2 Twee woorden als één (15 januari 2004 door W. Goedertier)

In paragraaf D.3.4 (Meervoudig gedeelde fonemen) wordt uitgelegd hoe te werk te gaan wanneer er meer dan één foneem gedeeld wordt over twee woorden. Ondertussen is gebleken dat het volgende geval regelmatig voorkomt, voorbeeld (13):

(13)	a. heb je je computer al geïnstalleerd?	[orthografisch]
	b. hEp **j@_j@** kOmpjut@r A G@Int@lert	[fonetisch]

De transcribent die de orthografiesche transcriptie maakte, schreef tweemaal "je". Dit is ook het enige dat grammaticaal gezien zinvol is. De transcribent die de brede fonetische transcriptie maakte, hoorde echter maar één keer /j@/ uitspreken. Volgens het protocol voor fonetische transcriptie kan dit genoteerd worden als /j@_j@/. Op deze manier blijft het ‘woord voor woord’-principe gerespecteerd. Na strikte toepassing van regel D.3.4 (en D.3.2) zou de woordoplijning er als volgt moeten uitzien:

ORT-tier:	heb	je	_	je	computer	al	geïnstalleerd?
FON-tier:	hEp	j@	_	j@	kOmpjut@r	A	G@Inst@lert

De positie van de grenzen (tijdsinformatie van de segmenten) zou dan als volgt zijn:

2de segment (je/j@) : bevat NIETS, lengte NUL (in praktijk dus 5 à 10 milliseconden)
3de segment (_/_) : bevat de fonemen /j@/
4de segment (je/j@) : bevat NIETS, lengte NUL (in praktijk dus 5 à 10 milliseconden)

Ook hier moeten we gebruik maken van segmenten met lengte NUL.

Er is echter een andere manier om deze situatie aan te pakken. In de fonetische transcriptie kan men in plaats van /j@_j@/ ook /j_j@/ schrijven. Volgens het protocol voor fonetische transcriptie is de interpretatie hetzelfde, er werd namelijk slechts één /j/ en slechts één /@/ gehoord. De woordoplijning ziet er nu als volgt uit:

ORT-tier:	heb	je	je	computer	al	geïnstalleerd?
FON-tier:	hEp	j=	=j@	kOmpjut@r	A	G@Inst@lert

De positie van de grenzen (tijdsinformatie) wordt nu:

2de segment (je/j=) : bevat de eerste helft van het foneem /j/
3de segment (je/=j@) : bevat de tweede helft van het foneem /j/ én het volledige foneem /@/

Bij deze oplossing hoeft men geen gebruik maken van segmenten met lengte nul en heeft conceptueel de voorkeur.

Een tweede voorbeeld (14) van hetzelfde fenomeen, dat analoog kan worden opgelost, dus van (14)b naar (14)c:

(14)	a. anders loop je jezelf een beetje…	[orthografisch]
	b. Anz lop **j@_j@zelv** @m betj@	[fonetisch]
	c. Anz lop **j_j@zelv** @m betj@	[zoals in woordoplijning]