Protocol voor Brede Fonetische Transcriptie

Versie 12 (24/09/2001)

A. Algemeen

Taak: Verificatie van een automatisch gegenereerde brede fonetische transcriptie (hierna: voorbeeldtranscriptie) aan de hand van het bijbehorende geluidsfragment.

Doel: Het verkrijgen van een geverifieerde brede fonetische transcriptie van het gesproken materiaal die binnen de gedefinieerde foneemset blijft, waarin inserties, deleties en substituties van fonemen tot uiting komen. Graduele processen, zoals mate van stemhebbendheid bij plosieven en fricatieven en monoftongering of diftongering bij vocalen, worden niet getranscribeerd. Ook nasalisering en verlenging of verkorting van vocalen worden niet getranscribeerd.

Foneemset: De transcripties worden vervaardigd met gebruikmaking van de onderstaande symbolen. De symbolen voor leenvocalen en nasale vocalen worden uitsluitend gebruikt in leenwoorden en niet om nasalisering en verlenging of verkorting van vocalen te transcriberen.

Klasse	**Voorbeeld**	**CGN-symbool**
Plosieven	put	p
	bad	b
	tak	t
	dak	d
	kat	k
	goal	g
Fricatieven	fiets	f
	vat	v
	sap	s
	zat	z
	sjaal	S
	ravage	Z
	licht	x
	regen	G
	geheel	h
Sonoranten	lang	N
	mat	m
	nat	n
	oranje	J
	lat	l
	rat	r
	wat	w
	jas	j
Korte vocalen	lip	I
	leg	E
	lat	A
	bom	O
	put	Y
Lange vocalen	liep	i
	buur	y
	leeg	e

	deuk	2
	laat	a
	boom	o
	boek	u
Sjwa	gelijk	@
Diftongen	wijs	E+
	huis	Y+
	koud	A+
Leenvocalen	scène	E:
	freule	Y:
	zone	O:
Nasale vocalen	vaccin	E~
	croissant	A~
	congé	O~
	parfum	Y~

B. Transcriptieregels

B.1. De gegeven voorbeeldtranscriptie is met behulp van de computer tot stand gekomen. Het is een automatische vertaling van de orthografische transcriptie (de gewone spelling met enkele aanpassingen) naar een brede fonetische transcriptie. Er is niet naar de fragmenten geluisterd bij de totstandkoming van de voorbeeldtranscriptie. Ga er dus nooit van uit dat 'het wel zal kloppen'.

B.2. Het is de bedoeling dat er een één-op-één-relatie op woordniveau blijft bestaan tussen de voorbeeldtranscriptie en de geverifieerde transcriptie. Eén woord (één of meer symbolen omgeven door spaties of underscores) in de voorbeeldtranscriptie moet ook één woord zijn in de geverifieerde transcriptie.

B.3. Pas de voorbeeldtranscriptie zo aan dat die overeenkomt met hoe iets daadwerkelijk is uitgesproken, uiteraard door uitsluitend de symbolen in de set te gebruiken. Doe dit door (herhaaldelijk) te luisteren naar het desbetreffende stukje. Luister naar stukjes van 0,5 tot 1 seconde. Zorg ervoor dat je ook altijd naar de overgang tussen twee woorden luistert.

B.4. Transcribeer wat je *hoort* en laat je zo min mogelijk beïnvloeden door wat er al in de voorbeeldtranscriptie staat.

B.5. Voeg geen spaties in en haal geen spaties weg. Punten en komma's maar ook apostrofs, koppeltekens en andere leestekens die tot de gewone spelling van woorden behoren, zijn weggelaten in de voorbeeldtranscriptie. Daardoor ontstaat soms een wat vreemd woordbeeld (zie onderstaande voorbeelden), maar het is niet de bedoeling dat je daar iets aan verandert.

Gewone spelling	**Voorbeeldtranscriptie**
'k zal er zijn.	k sAl @r zE+n
is 't echt?	Is t Ext
HEMA-trui	hematrY+
twee-onder-een-kap-woning	twejOnd@renkApwonIN
R&D-consultant	ErEndekOnsYlt@nt
varkens- en rundvlees	vArk@ns En rYntfles

B.6. Als je zeker weet dat er in de voorbeeldtranscriptie woorden ontbreken, voeg ze dan toe. Als je zeker weet dat er in de voorbeeldtranscriptie te veel woorden staan, schrap ze dan. Als je zeker weet dat in de voorbeeldtranscriptie een ander woord is getranscribeerd dan gezegd is, vervang die transcriptie dan door het goede woord. *Meldt deze fouten altijd aan degene die verantwoordelijk is voor de brede fonetische transcriptie.* Zorg te allen tijde dat de geverifieerde transcriptie overeenkomt met wat daadwerkelijk gezegd is.

B.7. Laat bij twijfel het symbool uit de voorbeeldtranscriptie staan:

Gewone spelling	**Twijfel tussen**		**Voorbeeldtranscriptie**
we **z**ijn weer thuis.	w@ **s**E+n wer tY+s	w@ **z**E+n wer tY+s	w@ **z**E+n wer tY+s
oma w**ordt** negentig.	oma w**Ot** nex@nt@x	oma w**Ort** nex@nt@x	oma w**Ort** nex@nt@x

B.8. Als je een klank hoort waarvoor volgens jou geen geschikt symbool te vinden is in de symboolset, laat ook in dat geval het symbool uit de voorbeeldtranscriptie staan:

Gewone spelling	**Je hoort**	**Je transcribeert**
we z**ij**n weer thuis.	w@ z**E~**n wer tY+s	w@ **z**E+n wer tY+s
margarine	m**A:**G@rin@	m**A**G@rin@

B.9. Gebruik de symbolen voor leenvocalen en nasale vocalen /E:, Y:, O:, E~, A~, O~, Y~/ *alleen* in leenwoorden.

Gewone spelling	**Je hoort**	**Je transcribeert**
z**o**ne	z**O:**n@	z**O:**n@
goeiem**or**gen	xuj@m**O:**x@	xuj@m**O**x@
d**oor**verwijzen	d**O:**v@wE+z@	d**o**v@wE+z@
parf**um**	pArf**Y~**	pArf**Y~**
uh even denken.	**Y~** ev@ dENk@	@ ev@ dENk@
vacc**in**	vAks**E~**	vAks**E~**
b**en** je daar nog?	b**E~** j@ dar nOx	b**E** j@ dar nOx
s**cè**ne	s**E:**n@	s**E:**n@
'k zal er z**ij**n.	k sAl @r z**E:**n	k sAl @r z**E+**n

B.10. Als twee opeenvolgende woorden een klank (of, in uitzonderingsgevallen, meer klanken) delen, zodat je niet weet bij welk woord die klank hoort, transcribeer je dat met behulp van een *underscore* (_) op de volgende manier:

Gewone spelling	**Brede fonetische transcriptie**
hij vraagt o**m m**eer.	hE+ vraxt o**m_m**er
al**s s**int komt.	Al**s_s**Int kOmt
vee**l l**iever.	ve**l_l**iv@r
ik vin**d d**at ook.	Ik fIn**d_d**Ad ok
o**p p**ad.	O**p_p**At
de naam die op de lij**st st**ond.	d@ nam di Ob d@ lE+**st_st**Ont
zij he**bb**en **be**reikt.	zE+ hE**b**@**_b**@rE+kt

B.11. Als de spreker een pauze laat vallen tussen twee woorden, kunnen die woorden nooit een klank delen. In dat geval gebruik je geen underscore.

Gewone spelling	**Brede fonetische transcriptie**
hij vraagt o**m (pauze) m**eer.	hE+ vraxt o**m m**er
al**s (pauze) s**int komt.	Al**s s**Int kOmt

B.12. Twee opeenvolgende woorden kunnen ook een klank delen als de laatste letter van een woord en de eerste letter van het daaropvolgende woord in de gewone spelling niet overeenkomen:

Gewone spelling	**Brede fonetische transcriptie**
o**p b**allet.	O**b_b**AlEt
ik he**b p**apier.	Ik hE**p_p**Apir
da**t d**oet pijn.	dA**d_d**ut pE+n
ik wor**d t**achtig.	Ik wOr**t_t**Axt@x
gee**f v**eel effect.	xe**f_f**el EfEkt
dat i**s z**eker.	dAt I**s_s**ek@r
al**s j**e wilt.	Al**S_S**@ wIlt
va**n m**'**n m**oeder.	vA**m_m**@**m_m**ud@r
ku**n j**e helpen?	kY**J_J**@ hElp@
sommig**e e**rvan.	sOm@x@_@rvAn

B.13. Vergelijk ook de volgende gevallen, waarin er geen sprake is van twee opeenvolgende woorden die een klank delen, maar van wegvallen (*deletie*) van een klank:

Gewone spelling	**Brede fonetische Transcriptie**
een sigaret.	@ **s**iGarEt
k**un j**e helpen?	k**Y j**@ hElp@
sommi**ge e**rvan.	sOm@**x E**rvAn

Transcribeer dus niet: /@**s_s**iGarEt/ of /k**Yj_j**@ hElp@/.

B.14. Soms wordt tussen twee woorden een klank ingevoegd. Transcribeer dat met behulp van koppeltekens, op de volgende manier:

Gewone spelling	**Brede fonetische Transcriptie**
speeld**e hij** toen al?	speld@**-n-E+** tun Al
ik d**oe aa**n sport.	Ig d**u-w-a**n spOrt

Gebruik de koppeltekens uitsluitend *tussen* woorden, niet *binnen* woorden (zie ook C.6.).

C. Belangrijke aandachtspunten

Let op: een brede fonetische transcriptie is meer dan een omzetting van de gewone spelling in fonetische symbolen. Zie bijvoorbeeld de volgende gevallen (waarbij niet steeds alle mogelijke uitspraakvarianten worden weergegeven):

C.1. Een *b*, *d* of *g* aan het einde van een woord of lettergreep wordt meestal uitgesproken als resp. *p, t* of *x*:

Gewone spelling	**Brede fonetische transcriptie**
kra**b**	krA**p**
bor**d**	bor**t**
vla**g**	vlA**x**
a**b**sur**d**	A**p**sYr**t**
a**d**miraal	A**t**miral

C.2. Als twee medeklinkers uit de categorieën "plosief" en "fricatief" (zie het schema onder A) elkaar opvolgen, hetzij binnen een woord, hetzij tussen twee woorden, is er vaak sprake van *assimilatie* (één van de twee klanken verandert of verdwijnt onder invloed van de andere klank):

Gewone spelling	**Brede fonetische transcriptie**
kra**bd**e	krA**bd**@
'k he**b d**e tijd.	k E**p t**@ tE+t, k E**b d**@ tE+t
kra**bt**	krA**pt**
kla**pb**and	klA**b**Ant
o**p b**allet.	O**b_b**AlEt
za**kd**oek	zA**gd**uk
o**p d**ansles.	O**b d**AnslEs
o**pv**allend	O**pf**Al@nt
o**p z**aterdag.	O**p s**at@rdAx
han**dv**at	hAn**tf**At
ik wor**d z**iek.	Ik wOr**t s**ik
kloo**fd**e	klo**vd**@
lie**f d**ing.	li**v d**IN
mi**sd**aad	mI**zd**at
vi**sv**ijver	vI**sf**E+v@r, vI**sf**E+f@r
a**fv**al	A**f**Al

C.3. Als een *n* gevolgd wordt door een medeklinker uit de categorieën "plosief" en "fricatief" of door een *m*, hetzij binnen een woord, hetzij tussen twee woorden, treedt ook vaak assimilatie op:

Gewone spelling	**Brede fonetische transcriptie**
o**nb**epaald	O**mb**@palt
i**n b**ad.	I**m b**At
o**ng**econtroleerd	O**Nx**@kOntrolertHYPERLINK
o**nm**iddellijk	O**m**Id@l@k
i**n m**emoriam.	I**m_m**emorijAm

C.4. Ook als een *s* of een *n* gevolgd wordt door een *j*, hetzij binnen een woord, hetzij tussen twee woorden, treedt vaak assimilatie op:

Gewone spelling	**Brede fonetische transcriptie**
ku**sj**e	kY**S**@
ik mi**s j**e.	Ik mI**S_S**@
ora**nj**e	orA**J**@
be**n j**e thuis?	bE**J_J**@ tY+s

C.5. Bij een woord dat eindigt op *p, t, k, s, f, g* treedt soms klankverandering op als het eropvolgende woord met een klinker begint:

Gewone spelling	**Brede fonetische transcriptie**
o**p** aanvraag.	O**b** anvrax
da**t** is zo.	dA**d** Is_so
i**k** ook.	I**g** ok
net al**s** iedereen.	nEt Al**z** id@ren
o**f** is dat niet zo?	O**v** Iz dAt nit so
ik la**ch** altijd.	Ik lA**G** AltE+t

C.6. Als twee klinkers elkaar opvolgen, hetzij binnen een woord, hetzij tussen twee woorden, hoor je vaak een tussenklank:

Gewone spelling	Brede fonetische transcriptie
b**io**	b**ijo**
televis**ie-i**nterview.	tIl@viz**ijI**nt@rvju
d**ue**t	d**ywE**t
d**oe he**t snel.	d**u-w-@**t snEl

C.7. Als een *r* of een *l* gevolgd wordt door een andere medeklinker, hoor je vaak de tussenklank /@/:

Gewone spelling	Brede fonetische transcriptie
mel**k**	mE**l@k**
we**rk**	wE**r@k**
e**rg**	E**r@x**
De**lf**t	dE**l@f**t
a**rm**	A**r@m**

C.8. De slot-*n* van woorden die eindigen op *-en* wordt vaak niet uitgesproken:

Gewone spelling	Brede fonetische transcriptie
tege**n**	teG@
kope**n**	kop@
mole**n**	mol@

C.9. Waar in de gewone spelling een dubbele medeklinker wordt geschreven, hoor je er doorgaans maar één:

Gewone spelling	Brede fonetische transcriptie
rus**tt**e	rYs**t**@
o**nn**odig	O**n**od@x

C.10. Hoe woorden worden uitgesproken hangt van veel factoren af. Zorgvuldigheid en snelheid waarmee iemand spreekt zijn van invloed, maar ook regionale herkomst en specifieke eigenaardigheden van de spreker. Kijk bijvoorbeeld naar de volgende mogelijke uitspraken:

Gewone spelling	Brede fonetische transcriptie
kop**eren**	kop**r@**, kop@**r@**, kop@**r@n**
ke**gelen**	ke**Gl@**, ke**G@l@**, ke**G@l@n**
b**a**naan	b@nan, b**A**nan, b**a**nan
f**ilos**ofie	f**Il@z**@fi, f**ilozo**fi, f**Il@s**@fi, f**ilos**ofi
m**i**nuut	m@nyt, m**i**nyt
p**ara**dijs	p**A**r@dE+s, p**ara**dE+s
t**ele**visie	t**Il**@vizi, t**el**@vizi, t**ele**vizi
k**o**lom	k**O**lOm, k**o**lOm
v**ill**a	v**Il**A, v**Il**a, v**il**a
st**ati**on	st**AS**On, st**Asj**On, st**AtS**On
lin**eaa**l	lin**ja**l, lin@**ja**l, lin**eja**l, lin**ija**l
auto	**A+**to, **o**to
pro**mpt**	prO**mpt**, prO**mt**
ma**rkt**	mA**rkt**, mA**rt**, mA**r@kt**
p**olit**ie	p**OlIs**i, p**olits**i, p**lis**i
ka**tj**e	kA**tj**@, kA**tS**@

D. Speciale gevallen

D.1. Gebruik vierkante haakjes [] voor materiaal dat niet getranscribeerd kan worden, bijvoorbeeld omdat het onverstaanbaar is of omdat de spraak overstemd wordt door achtergrondlawaai. Dat geldt voor woorden én voor delen van woorden. Zorg dat de één-op-één-relatie met de voorbeeldtranscriptie behouden blijft, dus als in de voorbeeldtranscriptie meerdere woorden staan die je niet verstaanbaar acht, gebruik je voor ieder woord haakjes.

Je hoort (? = onverstaanbaar)	**Brede fonetische transcriptie**
hij leest ?endertig bladzijden.	hE+ lest []@ndErt@x blAtsE+d@
hij leest ? bladzijden.	hE+ lest [] blAtsE+d@
hij leest achten?tig bladzijden.	hE+ lest Axt@n[]t@x blAtsE+d@
hij leest ? ?	hE+ lest [] []

D.2. Soms (voornamelijk bij snelle spraak) kun je van een woord geen enkel individueel foneem terugvinden, terwijl je toch kunt horen dat het woord wordt uitgesproken. Gebruik ook in dit soort gevallen vierkante haakjes [].

D.3. In gesproken taal wordt een opeenvolging van drie of meer identieke symbolen in de gewone spelling vaak gerealiseerd als één klank (zie ook B.10.). Omdat de één-op-één-relatie tussen voorbeeldtranscriptie en geverifieerde transcriptie moet blijven bestaan (zie B.2.) transcribeer je een dergelijke opeenvolging op de volgende manier:

Gewone spelling	**Je hoort**	**Brede fonetische transcriptie**
dat ’t toch	dAtOx	dAt_t_tOx

Met behulp van de underscores geef je aan dat het om een klank gaat die meerdere malen gedeeld is.

D.4. Een hekje ‘#’ wordt in de voorbeeldtranscriptie gebruikt voor zogenaamde niet-talige sprekergeluiden, bijvoorbeeld lachen, huilen, kuchen, niezen etc.

Voorbeeldtranscriptie
ev@ mE+n kel srap@ # zo dAt Iz bet@r

Laat het hekje uit de voorbeeldtranscriptie staan, tenzij je ervan overtuigd bent dat het echt onterecht is gebruikt. Vervang in dat geval het hekje door een fonetische trancriptie met gebruikmaking van de symbolenset. Voeg dit symbool **nooit** zelf in.

D.5. Wanneer een woord onderbroken wordt door ‘uh’ of een ander tussenwerpsel is dit in de voorbeeldtranscriptie als volgt weergegeven:

Er wordt gezegd:	**Brede fonetische transcriptie**
tussen-uh-oplossing	tYs@n@OplOsIN
heb jij geld over-ja-gemaakt?	hEp jE+ xElt ov@rjaG@makt

Voeg geen spaties in.

D.6. Eigennamen worden getranscribeerd zoals andere woorden. Denk eraan dat de één-op-één-relatie met de gewone spelling moet blijven bestaan, dus:

Gewone spelling	**Brede fonetische transcriptie**
New York	nuw jOr@k, nuw jOrk, nu jOk
Roeland Van Hout.	rulAnt vAn hA+t

D.7. Woorden die afgebroken zijn, versprekingen en woorden uit een vreemde taal worden zo goed als mogelijk is met de beschikbare symbolen getranscribeerd:

Er wordt gezegd:	**Brede fonetische transcriptie**
neem de link uh de rechterdeur.	nem d@ lINk @ d@ rExt@rd2r
ben je naar de kapper gewee-weest?	bEJ_J@ nar d@ kAp@r x@wewest
ken je ‘t lied La Vie En Rose?	kEJ_J@_@t lit lA vi A~ roz