

Live-ondertiteling met spraakherkenning: een case-analyse van reductie en fouten

Luuk Van Waes, Tijs Delbeke, Mariëlle Leijten & Aline Remael

Steeds meer televisie-uitzendingen worden live ondertiteld met gebruik van spraakherkenning. Dit gebeurt in de eerste plaats voor doven en slechthorenden. De delay waarmee een ondertitel in beeld verschijnt, is bepalend voor de perceptie van de kwaliteit van live ondertiteling. Twee factoren bepalen in hoge mate die vertraging: de parafrase van het live-commentaar om een gereduceerde ondertitel te maken en de correctie van fouten tijdens het productieproces (herstel van spraakherkenningsproblemen en inhoudelijke revisie).

Voorliggend onderzoek ontwikkelt een methode en een eerste classificatiemodel om de ondertitelreducties en -fouten (in proces en product) te identificeren en te beschrijven. Een eerste kwantitatieve analyse van een sample uit drie sportuitzendingen geeft inzicht in de gebruikelijke reductiepercentages en in het aantal gecorrigeerde fouten. De analyse toont ook een duidelijk verband tussen de hoeveelheid commentaar en het reductiepercentage in de ondertitel. De kwalitatief georiënteerde reductieclassificatie brengt de methodes in kaart die de ondertitelaar gebruikt bij het reduceren en leidt tot een voorlopige classificatie van de reductieoorzaken. De foutenanalyse ten slotte maakt het mogelijk om de hoofdcategorieën in de verschijningsvormen van fouten te onderscheiden.

Deze verkennende studie creëert een methodologische basis voor verder onderzoek over reductie- en correctieprocessen in live ondertiteling. Het uiteindelijke doel is bij te dragen tot een procedureverbetering van het ondertitelproces.

1 Inleiding

Steeds meer televisie-uitzendingen worden live ondertiteld. Heel wat openbare omroepen streven er zelfs naar om op relatief korte termijn bijna alle programma's van ondertitels te voorzien. Een complexe opdracht, te meer omdat ook de eisen die kijkers stellen, steeds hoger worden. Daarom willen we in dit onderzoek nagaan hoe we op basis van een accurate beschrijving van de huidige ondertitelpraktijk houvasten kunnen bieden om live-ondertiteling verder te optimaliseren en te professionaliseren. Het voorliggende onderzoek is exploratief van aard en betreft een kwantitatieve en kwalitatieve verkenning van drie uitzendingen die live ondertiteld werden. Het onderzoek gebeurde in situ bij de Vlaamse openbare omroep, de VRT.

In deze inleiding schetsen we kort een overzicht van de evolutie van live-ondertitelen en beschrijven we meer in detail het proces van live-ondertiteling zoals dat op dit ogenblik in de praktijk verloopt.

1.1 De evolutie van live-ondertiteling

Ondertiteling is een fenomeen waarmee we als Nederlandstaligen sterk vertrouwd zijn. In Vlaanderen en Nederland - net als in veel andere kleine taalgebieden - worden ondertitels vooral intralinguaal gebruikt om zo via vertaling de anderstalige uitzending toegankelijk te maken. Toch zijn we ook steeds meer vertrouwd geraakt met interlinguaal ondertitelen (Remael, De Houwer & Vandekerckhove, 2008), bijvoorbeeld bij uitzendingen waarin uitgesproken regionale taalvarianten voorkomen of ook in Vlaamse fictieprogramma's die in Nederland uitgezonden worden (en vice versa).

Naast deze vormen van *open* ondertiteling, bestaat er ook een vorm van *gesloten* ondertiteling. Daarbij heeft de kijker zelf de keuze om al dan niet de ondertitel te activeren. Teletextondertiteling is daar een voorbeeld van en is voornamelijk bedoeld voor doven en slechthorenden. In de praktijk blijkt echter dat ook allerlei andere doelgroepen gebruikmaken van deze ondertitels, voornamelijk allochtonen en bejaarden (Neves, 2005, 2008). In Groot-Brittannië bijvoorbeeld blijken 80% van de 7,5 miljoen gebruikers geen auditieve handicap te hebben (Ofcom, 2006). In Vlaanderen en Nederland is de situatie erg vergelijkbaar (Dewulf, 2005; SOAP, 2007). Op dit moment worden bij de VRT ongeveer 80% van alle uitzendingen ondertiteld; het streefdoel – door de overheid bepaald – is echter om over enkele jaren 95% te ondertitelen.

Een bijzondere - en erg complexe - vorm van *gesloten* ondertiteling is ‘live-ondertiteling’. Bij live-uitzendingen is het immers onmogelijk om van te voren de ondertitels voor te bereiden, maar kan er alleen in real-time ondertiteld worden (Remael, 2007). De omzetting van gesproken naar geschreven taal gebeurt dus simultaan met de uitzending van het programma. Omdat in de meeste gevallen niet geopteerd wordt om het programma vertraagd uit te zenden, veroorzaakt dit onvermijdelijk een *delay*. Een vertraging die tot gevolg heeft dat de gesproken uiting (en het betreffende beeld) niet volledig synchroon loopt met de overeenkomende ondertitel. Om die *delay* zoveel mogelijk te beperken, wordt gebruik gemaakt van allerlei hulpmiddelen: (a) gewoon toetsenbord (max. 90-110 woorden per minuut); (b) velotype toetsenbord (120-140 woorden na intensieve training van ongeveer 3 jaar; accuraatheid 95-98%); (c) stenografisch toetsenbord (tot 200 woorden per minuut na intensieve training; accuraatheid van 95-98%; Lambourne, 2006).¹

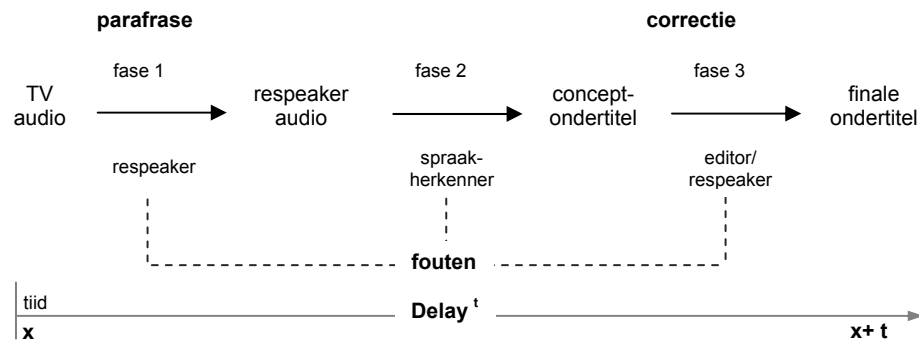
De laatste jaren maken ondertitelaars bij live-uitzendingen steeds meer gebruik van spraakherkenning (Lambourne et al., 2004). De VRT was hierbij een van de voorlopers. Het grote voordeel van spraakherkenning - zeker bij de huidige versies - is dat na een beperkte training (een dag) het mogelijk is om 150-160 woorden tekst per minuut te produceren met een optimale accuraatheid van 96 tot 98% (Honeycutt, 2003; Leijten, 2007). De VRT maakte op het moment van het onderzoek gebruik van het programma Dragon Naturally Speaking 9.5 (Nuance).

Een veel voorkomend misverstand is dat het commentaar van de televisiepresentator rechtstreeks als input dient voor het spraakherkenningsprogramma en zo automatisch naar tekst wordt omgezet. De huidige generatie van spraakherkenners is echter nog steeds in hoge mate sprekerafhankelijk en is gebaseerd op een sprekerprofiel voor elke individuele spreker. Bovendien werkt het programma beter als er vloeiend ‘gedicteerd’ wordt in een taalvorm die dicht aansluit bij geschreven taal. Lambourne et al. (2004) bevestigen dit en wijzen ook nog op drie andere redenen om respeakers in te schakelen bij live-ondertiteling: “interference from background music or noise; likelihood of multiple simultaneous speakers; the need for highly accurate speaker-independent recognition in real time; and the need to control style and position” (p. 270).

1.2 Het proces van live-ondertiteling met spraaktechnologie

Hoe verloopt de invoer van live-ondertitels met spraakherkenning precies? Figuur 1 geeft een visuele representatie van het proces in drie fasen. De figuur beschrijft de aanpak die op dit moment het meest gangbaar is bij de openbare omroep. De *respeaker* beluistert² via een koptelefoon het live-commentaar en parafraseert die in een dictaat (eerste fase). Dit is een proces dat in zekere mate vergelijkbaar is met simultaan tolken, maar zonder vertaalslag (Van der Veer, 2007). Het dictaat van de respeaker vormt

de audio-input voor de spraakherkenner (tweede fase). Die zet de gesproken taal om in digitale tekst en maakt die beschikbaar in het ondertitelprogramma (i.c. Swift). Het resultaat is een concept-ondertitel die - na eventuele correctie - op antenne gezet wordt via de Teletexttoepassing (derde fase). In de meeste gevallen is die laatste fase (bewerking en uitzending) in handen van een tweede persoon (editor) die op een gesynchroniseerd computerscherm werkt en de bewerking en uitzending voor zijn rekening neemt. Op die manier is het eenvoudiger voor de respeaker om zich op de hoofdtak te concentreren.



Figuur 1 Temporele representatie van het proces bij live ondertitelen met spraaktechnologie (VRT).

De voorstelling van het proces in Figuur 1 geeft duidelijk aan dat er een onvermijdelijke *delay* ontstaat bij de uitzending van live-ondertitels. Die *delay* (t) wordt cumulatief opgebouwd en vormt de som van:

1. de tijd die de respeaker moet wachten tot een commentaaruiting (TV audio) is afgerond;
2. de tijd om de parafrase te dicteren (fase 1);
3. de tijd die de spraakherkenner nodig heeft om de spraak naar tekst om te zetten (afhankelijk van de lengte en de complexiteit 1 tot 3 seconden; fase 2);
4. de tijd om eventueel correcties uit te voeren en de opmaak van de ondertitel aan te passen (fase 3).

De praktijk toont aan dat dit al snel leidt tot een gemiddelde *delay* van een tiental seconden.

Ook laat het schema in Figuur 1 zien dat er in de verschillende fasen fouten kunnen optreden: de respeaker kan het live-commentaar bij de parafrase (reductie) verkeerd interpreteren, of onzorgvuldig spreken; de spraakherkenner kan technisch falen bij de herkenning; de editor kan ten slotte bij de correctie onbedoeld een typefout maken. Er kan dus op heel wat momenten een fout in de uiteindelijke ondertitel sluipen.

Omdat beide factoren – reductie en fouten – cruciaal zijn in het ondertitelproces, hebben we ons in dit onderzoek voornamelijk hierop geconcentreerd. Het zijn immers deze factoren die in hoge mate bepalen hoe groot de *delay* is tussen geluid/beeld en ondertitel. Een factor die cruciaal is bij de perceptie van de kwaliteit van de ondertiteling.

2 Methode

In dit onderzoek streven we ernaar om een beter beeld te krijgen van de reductie in live-ondertiteling. Daarnaast willen we de accuraatheid bestuderen via een analysemodel dat fouten in het proces en in het eindproduct beschrijft. De ontwikkeling van een adequate methode en de beschrijving van verschillende vormen van reductie en fouten staan daarbij centraal. Daartoe hebben we data verzameld van drie verschillende uitzendingen die live ondertiteld werden. Het betreft telkens sportuitzendingen die in het voorjaar 2008 op de VRT werden uitgezonden.

2.1 Deelnemers en materiaal

De data werden verzameld op de VRT tijdens drie verschillende sessies. Het betrof telkens een live uitgezonden sportwedstrijd die live ondertiteld werd door getrainde respeakers van de openbare omroep. Zoals we in de inleiding reeds aangaven, vormen sportuitzendingen een groot aandeel van de live ondertitelde uitzendingen. We hebben geopteerd voor drie verschillende sporten omdat we zo niet alleen een variatie kregen in gesproken live-commentaren (telkens andere sportcommentatoren), maar ook omdat we het vermoeden hadden dat de hoeveelheid gesproken commentaar (per minuut) voor de verschillende disciplines sterk varieerde tussen tennis, voetbal en veldrijden. Dat vermoeden van variatie werd bevestigd in een verkennende analyse waarin we het aantal woorden in het gesproken commentaar berekenden per minuut (tennis: 73.4 woorden per min; voetbal 107.4; veldrijden: 169.5).

In totaal werden 3349 ondertitels verzameld. Omdat de uitzendingen sterk in lengte varieerden en de verkennende analyses erg arbeidsintensief zijn, opteerden we ervoor om telkens een subset uit het begin, het midden en het einde van de uitzending te selecteren. Voor de reductieanalyse bestonden die subsets telkens uit 100 ondertitels; voor de foutenanalyse uit 60 ondertitels.

2.2 Observatie en dataverzameling

Om een zo rijk mogelijke dataset op te bouwen met zowel product- als procesgegevens hebben we vier verschillende instrumenten ingezet:

1. Inputlog
2. Audio-opname respeakers
3. Softel Swift
4. Videopname live-uitzending

Inputlog is een softwareprogramma dat ontwikkeld werd voor schrijfonderzoek (Leijten & Van Waes, 2006; Van Waes & Leijten, 2006). Het loggingprogramma registreert alle toetsaanslagen, muishandelingen en gesproken input (via de loggingmodule van het programma Dragon Naturally Speaking, Nuance). Bovendien worden alle 'events' gekoppeld aan een tijdstipmoment (met een accuraatheid van ongeveer 20 ms). De logging zelf interfereert niet met andere programma's waardoor de observatie kan gebeuren zonder de ecologische validiteit van de context te verstoren. De observaties vonden plaats in de productiestudio van de VRT zelf tijdens een reële uitzending, de normale setting voor de respeakers dus. De audio-opnames werden als back-up en als controle van de output gebruikt.

Via merging van de basisfiles en een algemene data-analyse ('General analysis' in Inputlog) konden we een XML-bestand genereren dat in Excel en SPSS ingevoerd werd. Tabel 1 toont een voorbeeld van een fragment.

Tabel 1 Voorbeeld van een fragment uit de basisdata van Inputlog (fragment uit tennisverslag)

Modus	Productie	In-tijd	Out-tijd	Lengte actie	Pausetijd (ms)
3	een-heel-stuntelige-volley-van-Williams	817596	820399	2803	38566
1	NUM *	818347	818587	240	0
1	UP	822222	822403	181	3875
3	een-beetje-een-beginnersfout	822602	824497	1895	380
1	CTRL + DOWN	822793	822933	140	0
1	INS	823094	823234	140	301
1	INS	827079	827219	140	3985
3	het-was-niet-duidelijk-of-ze-met-één-of-twee-handen-ging-volleren	827459	831469	4010	380
1	UP	833218	833368	150	1749
1	HOME	833579	833669	90	361
1	DEL	833849	833939	90	270
3	het-was-een-makkelijke-volley	835221	836897	1676	1372
1	INS	835361	835511	150	0
3	maar-verslaat-met-één-hand-in-het-net	839066	841580	2514	3705
1	INS	839848	839998	150	0
1	UP	843212	843373	161	3364
1	HOME	843513	843593	80	301
1	DEL	843703	843813	110	190
1	CTRL + RIGHT	844014	844144	130	311
1	DEL	844514	844574	60	500
1	DEL	844665	844765	100	151
1	Z	845846	845996	150	1181
1	E	845896	846016	120	50
1	SPACE	845996	846137	141	100
1	DEL	846137	846247	110	141
1	CTRL + DOWN	846918	846948	30	781
1	CTRL + DOWN	847008	847138	130	90
3	een-cruciale-fase-in-de-tiebreak	847639	850123	2484	631

Legende: modus 1 = toetsenbord; modus 3 = spraaktechnologie

Het fragment in Tabel 1 omvat ongeveer een halve minuut en is de procesweergave van de productie van zes ondertitels (in totaal 43 woorden). De gedicteerde spraakherkenning (modus 3) verschijnt in de productiekolom en wordt meestal gevolgd door functietoetsen die het ondertitelprogramma aansturen. Bij de productie van de vijfde ondertitel treedt er echter een probleem op in de spraakherkenning die een revisie noodzakelijk maakt: 'maar-verslaat-met-één-hand-in-het-net'. 'Ze slaat' werd hier herkend als 'verslaat'. Na toetsnavigatie corrigeert de ondertitelaar (3 x DEL-toets) en

typt ‘ze’ in de plaats. Dit resulteert in de ondertitel: ‘Maar ze slaat met één hand in het net’. De revisie neemt ongeveer 5 seconden in beslag.

De gelogde gegevens uit Inputlog werden vervolgens aangevuld met de ondertitels die in het ondertitelprogramma Swift (www.softel.co.uk) werden geregistreerd. Een universele tijdsaanduiding in de verschillende loggingfiles en het gemeenschappelijke XML-formaat stelde ons in staat de inhoud van de verschillende dataverzamelingen aan elkaar te koppelen. Ten slotte werd het gesproken commentaar van de drie uitzendingen getranscribeerd aan de hand van de video-opname. Via het hulpprogramma LiveComment³ werden ook tijdsindicaties toegevoegd aan elke uiting. Uiteindelijk hebben we de drie datasets bij elkaar gevoegd in één Excelbestand (zie Tabel 2).

Deze geïntegreerde representatie van de data maakte het mogelijk om de reductie te berekenen die het gevolg is van parafrase alsook het tijdsverloop tussen het gesproken commentaar en het verschijnen van de ondertitel. Deze data vormden ook de basis voor de foutenanalyse.

3 Analyse

In deze paragraaf bespreken we respectievelijk de resultaten van de reductie- en de foutenanalyse. De besprekingen vertrekken telkens vanuit een explorerende, kwantitatieve beschrijving van de data. We presenteren telkens de basisberekeningen voor de geselecteerde fragmenten in de drie uitzendingen per ondertitel en per gecumuleerd tijdsinterval van één minuut. De berekeningen betroffen het aantal karakters en woorden per eenheid in het gesproken commentaar en in elke ondertitel, de effectieve reductie tussen commentaar en ondertitel (absoluut en relatief), het aantal gecorrigeerde fouten in het proces en in de uitgezonden ondertitels. Een correlatie-analyse had tot doel eventuele verbanden tussen de berekende variabelen na te gaan.

Om een beter beeld te geven van de onderliggende processen, illustreren we de kwantitatieve analyse telkens met een aantal voorbeelden. Die voorbeelden koppelen we aan een classificatiemodel. De modellen voor de reductie- en foutenanalyse kwamen tot stand via een iteratief proces: we vertrokken vanuit een theoretisch model en hebben dit model verder uitgebreid en gespecificeerd via een iteratief voortschrijdende kwalitatieve analyse van de datasets.

Tabel 2 Geïntegreerd rekenblad met de basisgegevens van de gesproken commentaren, het dictaat en de uiteindelijke ondertitel (inclusief telling aantal woorden en karakters).

Gesproken commentaar			Dictaat			Ondertitel				
starttijd	eindtijd	tekst	#k	#w	tekst	starttijd	eindtijd	tekst	#k	#w
0:05:01	0:05:05	Een wel heel stuntevolle volley daar van Venus.	46	8	een-heel-stuntevolle-volley-van-Williams	0:05:08	0:05:11	Een heel stuntevolle volley van Williams	40	6
0:05:06	0:05:08	Een beetje een beginnersfout als het ware.	42	7	een-beetje-een-beginnersfout	0:05:12	0:05:14	Een beetje een beginnersfout	28	4
0:05:09	0:05:13	Niet duidelijk of ze met een of twee handen ging volleren.	58	11	het-was-niet-duidelijk-of-ze-met-één-of-twee-handen	0:05:20	0:05:24	Het was niet duidelijk of ze met één of twee handen	65	13
0:05:15	0:05:16	En een makkelijke volley	25	4	het-was-een-makkelijke-volley	0:05:24	0:05:26	Het was een makkelijke volley	29	5
0:05:17	0:05:24	Maar helemaal beneden in het net, ja dat is iets om je hand van voor je mond te houden: wat heb ik nu gedaan?	109	24	maar-verslaat-met-één-hand-in-het-net	0:05:32	0:05:35	Maar ze slaat met een hand in het net	37	9
0:05:24	0:05:28	En dat in een cruciale fase, in de tiebreak van de eerste set.	62	13	een-cruciale-fase-in-de-tiebreak	0:05:37	0:05:40	Een cruciale fase in de tiebreak	32	6

#k = aantal karakters
#w = aantal woorden

3.1 Reductieanalyse

In een eerste analyse hebben we de reductie berekend voor elk van de segmenten in de drie uitzendingen. Tabel 3 geeft een overzicht van de resultaten. Het reductiepercenta-

ge varieert sterk per uitzending. Bij tennis – waar het aantal gesproken woorden in het commentaar per minuut het laagst is – zien we een reductiepercentage van gemiddeld 38.4%; bij veldrijden – waar de hoeveelheid gesproken commentaar meer dan dubbel zo hoog is – loopt het reductiepercentage op tot 57.8%. Ondanks de hoge reductiegraad daar, ligt het gemiddeld aantal woorden per minuut in de ondertitels ongeveer de helft zo hoog. Toch blijft dit nog beduidend beneden de maximumgrens van 120-125 woorden die doorgaans gehanteerd wordt bij ondertiteling in Vlaanderen en Nederland (Gielen & d'Ydewalle, 1989; Remael, 2007). Opmerkelijk is ook dat we kunnen vaststellen dat de hogere reductiegraad ervoor zorgt dat de vertraging (*delay*) bijna halveert, waardoor er een betere synchronisatie ontstaat tussen commentaar/beeld en ondertitel.

Tabel 3 Aantal woorden per minuut in de gesproken commentaren en de ondertitels

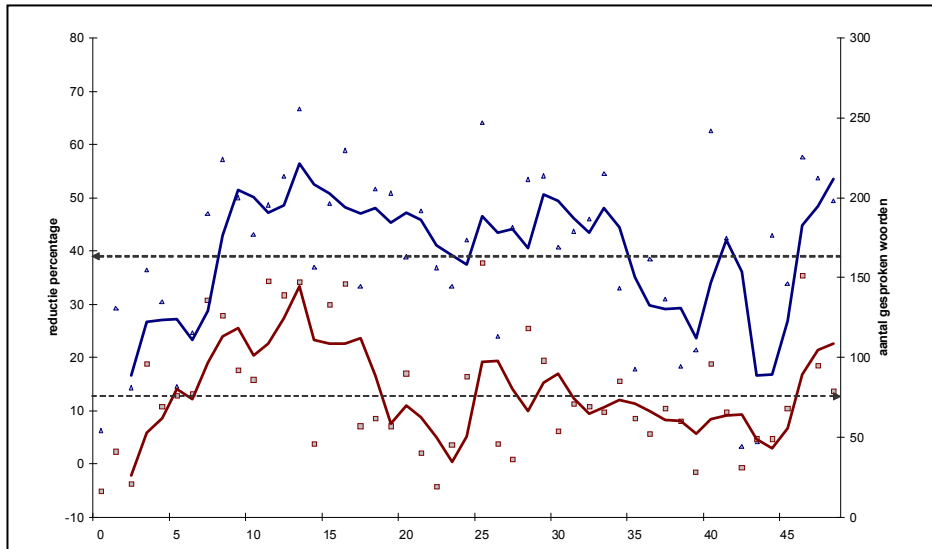
	gesproken commentaar (woorden per min)	ondertitel (woorden per min)	reductiepercentage %	vertraging (in sec)	
				Intijd ¹	Uittijd ²
tennis	76.2 (39.3)	42.1 (16.4)	38.4 (17.8)	12 (5)	11 (4)
voetbal	107.4 (41.2)	70.1 (26.7)	27.6 (11.8)	11 (4)	11 (4)
veldrijden	169.5 (56.4)	72.4 (31.4)	57.8 (24.6)	6 (2)	5 (2)

¹ De 'intijd' verwijst naar de periode (in sec) tussen de start van het gesproken commentaar en het moment waarop de ondertitel in beeld verschijnt.

² De 'uittijd' verwijst naar de periode (in sec) tussen het einde van het gesproken commentaar en het moment waarop de ondertitel in beeld verschijnt.

Voor de ondertiteling van de voetbalwedstrijd zien we een vergelijkbare ondertitelratio (70 woorden per minuut), maar die is wel het resultaat van een veel beperktere reductie (27.6%). Een mogelijke verklaring hiervoor is dat er bij tennis vaak stiltes vallen in het live-commentaar tijdens de balwisselingen en het commentaar relatief geconcentreerd aangeboden wordt tijdens de onderbrekingen van het spel. Bij voetbal en veldrijden is er – mede door de aard van de activiteit - sprake van een meer continue becommentariëring. Bij de bespreking van de verschillende soorten reductie komen we hierop terug.

In een aanvullende correlatieanalyse hebben we onderzocht of er een verband bestaat tussen het aantal woorden⁴ in de ondertitel en in het overeenstemmende live-commentaar. De resultaten wijzen op een significant verband (tennis: $r=.850$; $p<.001$ - voetbal: $r=.412$; $p<.001$ - veldrijden: $r=.388$; $p<.001$). In Figuur 2 stellen we dit verband grafisch voor aan de hand van de data van de tennisuitzending. De bovenste lijn toont het reductiepercentage (linkse y-as); de onderste lijn toont het aantal woorden per minuut van het gesproken commentaar (rechtse y-as). De punten geven de exacte data per minuut weer (interval); de lijngrafiek is gebaseerd op het voortschrijdende gemiddelde ($n=3$). De grafiek illustreert duidelijk de correlatie tussen beide variabelen. De lijnen zijn in hoge mate gekenmerkt door een gelijklopende verloop: als het aantal woorden in het gesproken commentaar toeneemt, stijgt ook het reductiepercentage, en vice versa.



Figuur 2 Aantal woorden per minuut van het gesproken commentaar en het overeenstemmende reductiepercentage van de uitgezonden ondertitel (intervallen van 1 minuut).

Hoewel de grafiek op intervalniveau al een grote variatie laat zien (van 3 tot 67%), is die variatie op segmentniveau (individuele ondertitels) uiteraard nog groter. Opvallend daarbij is ook dat er in een aantal gevallen een negatieve reductie plaatsvindt. M.a.w. er wordt soms tekst toegevoegd t.o.v. het oorspronkelijke gesproken commentaar. Dat komt vooral voor op plaatsen waar de spreeknelheid lager ligt. De uitbreiding van het commentaar kan verschillende vormen aannemen. Soms is het vooral een omzetting naar schrijftaal ('Niks.' wordt 'Hij doet niets verkeerd.'). Maar ook worden soms backward referenties toegevoegd om de achterlopende synchronisatie op te vangen: 'De coach van Ivanovic.' wordt 'We zagen even de coach van Ivanovic in beeld.' In een aantal gevallen krijgen we ook – conform de regels voor ondertitels voor doven en slechthorenden - verbalisaties van achtergrondgeluid of van een gebeurtenis die in het live-commentaar niet verwoord wordt (Ø wordt 'Hij windt zich op.' of 'Djokovic lijkt helemaal zeker van de overwinning.').

Om een beter beeld te krijgen van wat het reductieproces precies inhoudt, hebben we geprobeerd de verschillende vormen van reductie te beschrijven en te classificeren. Zo proberen we een beter inzicht te krijgen in veel voorkomende manieren waarop de respeaker het gesproken commentaar verkort, meestal met de bedoeling om zo weinig mogelijk betekenis verloren te laten gaan. Voor de classificatie van dit reductieproces hebben we enerzijds gekeken naar de *operatie* die de reductie kenmerkt en anderzijds naar de *oorzaken* die aan de basis liggen van de reductie (zie Tabel 8 en 9 in de bijlage).

Voor de classificatie van de reductie-analyse⁵ baseerden we ons op interviews met de respeakers van de VRT en op modellen uit gerelateerd onderzoek, voornamelijk Schilperoord, De Groot en Van Son (2005) en Lambourne (2006). Zo wijzen Schilperoord et al. (2005) in hun onderzoek naar de ondertitels van Nederlandstalige

soap-uitzendingen (n=3) erop dat er bij de parafrase voornamelijk een omzetting gebeurt van gesproken naar geschreven taal die gekenmerkt wordt door: (a) herstel van de canonieke woordvolgorde van de hoofdzin (SVO, subject – verb – object); (b) vereenvoudiging van complexe (samengestelde) zinnen; (c) reductie door de voorkeur te geven aan inhoudelijk minder relevante informatie; (d) het schrappen van conversatiemarkeerders.

Lambourne (2006) wijst in zijn onderzoek voornamelijk op een aantal praktische oorzaken voor de reductie tijdens de parafrase van de gesproken commentaren. Zo wijst hij op het verschil tussen lees- en spreeknelheid, de technische beperking van de spraakherkenners (waardoor bijvoorbeeld niet geoefende eigennamen anticipatief vervangen worden door voornaamwoorden) en de beperking dat in een ondertitel slechts één persoon tegelijkertijd kan worden ondertiteld (in die gevallen waar twee sprekers door elkaar spreken, zoals vaak gebeurt in debatten).⁶

De omvang van deze bijdrage maakt het niet mogelijk om het volledige model toe te lichten en te illustreren. Graag beperken we ons tot een paar voorbeelden uit het geanalyseerde corpus om zo de lezer een beter beeld te geven van de processen die we hebben beschreven (Tabel 4).

Tabel 4 Voorbeelden van reductieoperaties

	Commentaar	Ondertitel	Karakterreductie	Woordreductie
(1)	Maar dat kunnen ze vergeten, denk ik. <u>Dat kunnen ze vergeten.</u>	Maar dat kunnen ze vergeten, denk ik.	-24	-4
(2)	En <u>dit is verschrikkelijk</u> , dit is balen.	Dit is balen.	-28	-4
(3)	Maar waarschijnlijk heeft <u>Walter</u> Vromans het bij het rechte eind.	Maar waarschijnlijk heeft Vromans het juist	-22	-4
(4)	Score 30 beide.	∅	-13	-3

Veel van de reductievormen betreffen typische kenmerken van spreektaal in het livecommentaar die aanleiding geven tot reductie (deletie of beperkende substitutie). Een uitgesproken voorbeeld van deletie vormt de *herhaling*. Het weglaten van gespiegelde informatie in herhalingen is immers een zeer efficiënte vorm van reductie die meestal slechts beperkt impact heeft op de oorspronkelijke inhoud. Soms is er sprake van letterlijke herhaling (voorbeeld 1); soms vormt de herhaling een parafrase (2). Ook de weglating van voornamen is een vaak voorkomende vorm van reductie (3). Voor de verstaanbaarheid is het weglaten van een *voornaam* immers meestal niet problematisch.⁷ Voorbeeld (3) toont tegelijkertijd een andere vorm van reductie, namelijk de substitutie van beeldende taal (zegswijzen, metaforen) door een meer beknopte beschrijving.

Dat de respeakers ook rekening houden met de afstemming tekst en beeld, maakt voorbeeld (4) duidelijk. Daar ‘leest’ de commentator de tussenstand die in beeld is. Een typisch voorbeeld waar de ondertitel van de gesproken commentaar niets toevoegt aan het beeld, zeker als we rekening houden met de onvermijdelijke *delay*.

Bij de classificatie hebben we ook een tweede invalshoek gekozen, met name een indeling van de reducties op basis van de mogelijke oorzaken (Tabel 9 als bijlage). Uiteraard gebeurde dit soms op basis van inferenties; meestal echter maakt de combi-

natie van beeld, commentaar en ondertitel een eenduidige interpretatie mogelijk. De basisindeling maakt een onderscheid tussen drie hoofdoorzaken: commentaar, beeld en spraaktechnologie. Ook hier geven we een paar voorbeelden (Tabel 5).

Tabel 5 Voorbeelden van reductieoorzaken

	Commentaar	Ondertitel	Karakter-reductie	Woord-reductie
(5)	Die gaat weer deze namiddag meestrijden, maar dat wordt in principe een walk-over van Niels Albert. Hij heeft de betere loopstijl, zie je. Sagan heeft de kortere beentjes.	Bij de beloften zal Niels Albert te winnen	-129	-20
(6)	De Smet die met de bal aan de haal was. Hands gefloten nu van Oleg lachtchouk.	Handspel gefloten tegen lachtchouk	-48	-12
(7)	Nee, niet te geloven. Hij valt in de laatste ronde.	Ai!		
(8)	Maar met Sven Vermant in de ploeg, kan dit ook wel eens een indraaier worden. <u>En let op de positie van Sterchele. Nu nog even buitenspel maar die gaat zeker en vast beginnen bewegen.</u>	Maar met Sven Vermant in de ploeg, kan dit een indraaier worden	-119	-22
	Vermant met een bal die te veel snelheid meekreeg.	De bal van Vermant kreeg teveel snelheid mee	-6	-1
(9)	<u>lachtchouk. De Smet. Afleggen. Van Mol. De voorzet. De Sutter.</u> Ja. Cercle Brugge levensgevaarlijk via die patronen.	Cercle is levensgevaarlijk via die patronen	-72	-11

Voorbeeld (5) is een typisch illustratie van een reductie die veroorzaakt wordt door het gesproken *commentaar*. In dit voorbeeld ligt de spreeknelheid ontzettend hoog: 28 woorden die in minder dan 7 seconden uitgezonden werden (extrapolatie: 240 woorden per min). Dergelijke fragmenten leiden onvermijdelijk tot reductie, vaak met een meer arbitraire selectie dan in de andere gevallen. Betekenisverlies is hier in vele gevallen onvermijdelijk.

Anders is het als het beeld (of de actie) de directe aanleiding vormt van de reductie. Dat zien we in voorbeelden (6-7) waar een plotse gebeurtenis (*sudden event*) leidt tot een forse reductie. Immers, als er plots iets gebeurt tijdens een wedstrijd – bijvoorbeeld een ontsnapping, een valpartij, een overtreding of een doelpunt - is het belangrijk dat die gebeurtenis zo snel mogelijk ondertiteld wordt en de *delay* zo beperkt mogelijk gehouden wordt. In voorbeeld (7) wordt een commentaar gereduceerd tot een uitroep ('Ai!') om zo dicht mogelijk aan te sluiten bij de dramatische gebeurtenis. Als de 1-op-1 relatie tussen beeld en ondertitel minder dwingend is – bijvoorbeeld een achtergrondverhaal of algemene beschouwing – is de noodzaak om de *delay* zo sterk mogelijk te beperken minder dwingend.

Het kan ook gebeuren dat de respeaker anticipeert op een belangrijke gebeurtenis in de wedstrijd. In die gevallen wordt er in afwachting van die gebeurtenis vaak extra gereduceerd, om zo het commentaar over het *sudden event* zo volledig mogelijk

mee te kunnen geven. Dat zien we bijvoorbeeld in de beschrijving van een hoekschop in voorbeeld (8). Het commentaar over die hoekschop zelf wordt volledig meegegeven. De gedetailleerde tussencomentaar wordt gereduceerd om het mogelijke doelgevaar sneller te ondertitelen. Ook zogenaamde simultane beschrijvingen worden vaak gereduceerd of helemaal niet ondertiteld. Zeker bij voetbaluitzendingen komt dit vaak voor, bijvoorbeeld als de commentator de balwisselingen continu en simultaan beschrijft. Dergelijke beschrijvingen die een relatief kleine meerwaarde hebben omdat ze ook visueel te volgen zijn (voor wie de spelers kent), worden in hoge mate weggelaten om die reden (voorbeeld 9).

Deze voorbeelden tonen duidelijk aan dat de respeakers gebruik maken van verschillende reductietechnieken. De veelheid aan mogelijke oorzaken leidt ertoe dat er niet alleen sprake is van verkortende parafrases met (stilistische) omzettingen van gesproken naar geschreven taal, maar dat er soms ook meer ingrijpende keuzes gemaakt dienen te worden. Naast de hierboven beschreven oorzaken, vormt het gebruik van de spraaktechnologie en de fouten die daarbij kunnen ontstaan ook een belangrijke vorm van *delay* (zie Figuur 1). De correctie van fouten vergt immers tijd en vergroot zo de *delay*. Daarom concentreren we ons in de volgende paragraaf meer bepaald op de fouten die ongewild opduiken in het proces van de live-ondertiteling.

3.2 Foutenanalyse

In de foutenanalyse hebben we een onderscheid gemaakt tussen fouten die voorkwamen in de origineel gedicteerde ondertitel (proces) en fouten in de uiteindelijke ondertitel (eindproduct). Tabel 6 laat de resultaten zien van de analyse voor de verschillende uitzendingen. In totaal werden 360 ondertitels geanalyseerd (2364 woorden). Daarin troffen we tijdens het proces 126 fouten aan; 100 ervan werden gecorrigeerd voordat ze op het scherm verschenen.

Tabel 6 *Overzicht van de fouten en het correctheidspercentage voor de verschillende uitzendingen*

	aantallen		woorden	fouten		Correctheidspercentage		
	titels per minuut	ondertitels		proces (woord)	product (ondertitel)	per woord	per dict. ondertitel	per uitg. ondertitel
Tennis	7.6	120	738	33	3	95.5	80.9	97.5
Veldrijden	9.7	120	829	43	16	94.8	70.0	86.7
Voetbal	12.5	120	779	50	7	93.6	65.9	94.2
Totaal	9.8	360	2346	126	26	94.6	72.2	92.8

De respeakers bereiken tijdens hun ondertiteltaak een correctheidsniveau van ongeveer 95%. De accuraatheid is dus vergelijkbaar met de bodemreferentiewaarden voor spraakherkenning zoals die naar voren worden geschoven voor gebruik in optimale omstandigheden. We moeten er echter rekening mee houden dat de complexiteit van de respeaking context het uiterste vergt van de spraakherkenner en de dicteerder. De cognitieve inspanning bij live-ondertitelen ligt immers een stuk hoger dan bij ‘gewo-

ne' spraakherkenning, omdat hier luisteren (kijken) en parafraseren simultaan gebeuren. Bovendien spelen ook factoren als tijdsdruk en vermoeidheid ongetwijfeld een rol. Bekijken we de correctheid echter op het niveau van de ondertitels, dan zou met deze correctheidsratio – zonder correctie na het dictaat – 1 op de 4 ondertitels nog een fout bevatten. Dankzij de correctie⁸ daalt de foutenkans echter sterk en wordt de kijker per 5 minuten nog ongeveer 3 keer met een fout geconfronteerd. Daarbij dienen we nog op te merken dat niet alle fouten even storend zijn. Bijvoorbeeld samengestelde woorden die als twee woorden verschijnen (bijvoorbeeld: 'er van' of 'gebruik maken' i.p.v. 'ervan' en 'gebruikmaken') vormen strikt genomen een fout, maar zullen vaak niet als erg storend gepercipieerd worden (zie ook foutenclassificatie).

Ook zien we in Tabel 6 dat de variatie per uitzending vrij groot is. Bij de 'tragere' tennisuitzending slaagt de respeaker erin om ruim 80% van de titels foutloos te produceren. Na correctie zijn meer dan 97% van de titels correct. Bij veldrijden daalt dit niveau beduidend, nl. met 10%. Het aantal geproduceerde woorden lijkt hierbij een rol te spelen, maar is duidelijk niet de enige factor. Zo zien we dat ook in de voetbaluitzending het aantal woorden relatief hoog is – zowel in het commentaar als in de gerealiseerde ondertitels – en dat het dictaat soms vrij problematisch is, maar dat de meeste fouten hersteld worden. Toch slaagt men er nog in om ongeveer 3 ondertitels meer per minuut op het scherm te krijgen (9.5 versus 12.5).

Aanvullend op deze analyse hebben we geprobeerd veel voorkomende fouten te beschrijven. Het basismodel dat we hanteerden in de foutenanalyse is vooral gebaseerd op het classificatiemodel dat Leijten (2007, pp. 39-43) uitwerkte voor de analyse van foutcorrecties in schrijfprocessen waarin spraakherkenning als schrijfmodus gebruikt wordt. Haar model vormt een uitbreiding op het model voor spraakherkenningsfouten van Karat, Halverson, Horn et al. (1999), het model voor tekstverwerkingsfouten van Van Waes en Schellens (2003) en het klassieke model van Faigley & Witte (1981). In haar onderzoek maakt Leijten bijvoorbeeld een onderscheid tussen technische fouten (veroorzaakt door de spraakherkenner of de typemotoriek) en meer inhoudelijke revisies (veroorzaakt door de gebruiker). Ook wijst ze op het bereik en de niveauverschillen tussen fouten die typisch zijn voor spraakherkenning (bijvoorbeeld één woord wordt herkend als twee woorden: 'voor komen' i.p.v. 'voorkomen', of navigatie-instructies worden herkend als tekst: 'schrapp dat' resulteert niet in de gewenste verwijdering van tekst, maar in nieuwe tekst). Omdat er een duidelijke parallel is tussen de schrijfscontext die zij onderzocht en de ondertitelcontext, gebruikten we haar model als uitgangspunt. Via iteratieve cycli in de analyse werd het aangepast voor de ondertiteldata.

In Tabel 7 presenteren we het concept voor een classificatiemodel van fouten in ondertitels.⁹ In het model leggen we niet de nadruk op de oorzaak van de fout. Zo maken we bijvoorbeeld geen onderscheid tussen een misrecognitie die een technische oorzaak betreft (bijvoorbeeld omdat het woord niet in het lexicon voorkomt) dan wel een menselijke fout betreft (bijvoorbeeld een verspreking). Wel onderscheidt het model verschillende verschijningsvormen van fouten. We maken daarbij gebruik van drie hoofdcategorieën die we hieronder kort toelichten.

De eerste hoofdcategorie omvat verschillende soorten *misrecognities*. Het bereik van de verkeerde herkenning ('niets', één woord, meer woorden) en het type fout (persoonsnamen of andere eigennamen) vormen hier de belangrijkste verschillen. In tegenstelling tot verkeerd herkende woorden of woordgroepen, onderscheiden we in een tweede hoofdcategorie een aantal vormen van *misinterpretaties*. In die gevallen veroorzaakt niet de herkenning zelf het probleem, maar gebeurt er een onterechte tech-

nische transformatie, vaak op basis van verkeerd toegepaste parsingregels. In de laatste categorie hebben we de *spel- en typefouten* ondergebracht. Dit soort fouten komen in principe niet voor in teksten die via de spraakherkenner gedictieerd werden, maar omdat de ondertitelaars ook gebruik maken van hun toetsenbord, doken ze sporadisch op in ons datamateriaal. In een uitzonderlijk geval betrof het een woord dat verkeerd was toegevoegd aan het lexicon van de spraakherkenner.¹⁰

Tabel 7 *Categorisatiemodel voor fouten in ondertitels*

Misrecognitie	Fout	Correct
één woord als één ander woord	En dus niet het hele <u>bleek</u> Wit-Russen, Moldaviërs enzovoort	En dus niet het hele <u>blik</u> Wit-Russen, Moldaviërs enzovoort
meer woorden als meer andere woorden	En de <u>O'Grady zien</u> dat het goed is	En de <u>ogen die zien</u> dat het goed is
één woord als meer (andere) woorden	Hij kreeg niet overal <u>start terecht</u> .	Hij kreeg niet overal <u>startrecht</u> .
meer woorden als een ander woord	Hij nadert nu <u>opzij</u> seconden.	Hij nadert nu <u>op zes</u> seconden.
'niets' als een woord (<i>ghost-recognition</i>)	Maar dan komen er altijd weer <u>een</u> langere stroken.	Maar dan komen er altijd weer <u>_</u> langere stroken.
woord als 'niets' (<i>non-recognition</i>)	Sven Nys heeft <u>_</u> al drie keer gewonnen.	Sven Nys heeft <u>er</u> al drie keer gewonnen.
persoonsnamen	Vrije trap met <u>portier</u> Hij heeft ook te wapens om het <u>verder</u> lastig te maken.	Vrije trap met <u>Portier</u> Hij heeft ook te wapens om het <u>Federer</u> lastig te maken.
andere eigennamen en acroniemen	Hij gaat eraf, de sympathieke <u>zijn babbel aan</u>	Hij gaat eraf, de sympathieke <u>Zimbabweaan</u>
Misinterpretatie		
parsing-fout: één (samengesteld) woord als meer woorden	10 tegen 9, als we even <u>door redeneren..</u>	10 tegen 9, als we even <u>doorredeneren..</u>
parsing-fout: meer woorden als één (samengesteld) woord	De Sutter was <u>daarvoor</u> het eerst bij de bal.	De Sutter was <u>daarvoor</u> het eerst bij de bal.
grammaticale fout (bijvoorbeeld congruentie)	De lijnrechter <u>zitten</u> ertussen.	De lijnrechter <u>zit</u> ertussen.
commando als tekst	Hij had de 2 <u>een streepje</u> 0 aan de voet	Hij had de <u>2-0</u> aan de voet
tekst als commando	<u>_</u> uit!	<u>Punt</u> uit!
achtergrondgeluid/aarzeling/kuch als dictaat herkend	<u>Regent.</u>	<u>(keel schrapen)</u>
Overig		
spelfout	Gomb <u>an</u> i aan de bal.	Gomb <u>a</u> mi aan de bal.
typefout (alleen voor input via toetsenbord)	Dadelijk komt er een <u>moeilijkee</u> bocht.	Dadelijk komt er een <u>moeilijke</u> bocht.

De meest voorkomende fouten zijn te situeren in een beperkt aantal categorieën. Zo vallen ongeveer 40% van de fouten in de eerste drie categorieën. Fouten in persoons- en eigennamen vormen ongeveer 20% van de problemen.

4 Conclusie en discussie

Tot slot vatten we de belangrijkste bevindingen uit dit onderzoek kort samen. Daarnaast wijzen we in de discussie op enkele kritische aspecten van het onderzoek en presenteren we mogelijke perspectieven voor vervolgonderzoek.

Voor het verkennende onderzoek waarover hier gerapporteerd wordt, hebben we een methode en een eerste classificatiemodel ontwikkeld. Die stellen ons in staat om nauwgezet reducties en fouten te identificeren en te beschrijven in het ondertitelproces. Een eerste *kwantitatieve analyse* van een representatieve sample uit drie sportuitzendingen met variërende hoeveelheden commentaar en verschillende spreeknelheden, geeft inzicht in de gebruikelijke reductiepercentages en schetst een beeld van het aantal gecorrigeerde fouten in proces en product. Er blijkt ook een duidelijk verband tussen de hierboven vermelde variabelen.

Hoewel de reductiegraad varieert naargelang van het programma, blijft hij onder de waarden die gebruikelijk zijn bij traditionele open ondertiteling. Als het aantal woorden in de gesproken commentaren toeneemt, stijgt echter ook het reductiepercentage. Anderzijds blijkt ook dat de *delay* kleiner wordt naarmate de reductiegraad toeneemt.

De *kwalitatief georiënteerde reductieclassificatie* uitgevoerd in de tweede fase van het onderzoek, brengt de aanpak in kaart die de ondertitelaar gebruikt bij het reduceren. De verschillende reductieoperaties worden gekenmerkt door een (bewuste) omzetting van gesproken naar geschreven taal; er wordt rekening gehouden met het verschil tussen lees- en spreeknelheid en met de technische beperkingen van de spraakherkenner. In de voorlopige classificatie van de reductieoorzaken wijzen we op drie factoren: commentaar, beeld en spraaktechnologie.

De *foutenanalyse* is duidelijk gerelateerd aan deze laatste oorzaak. Ze toont aan dat de ondertitelaars een correctheidsniveau van ongeveer 95% bereiken tijdens het dicteren, maar na correctie stijgt dit percentage nog, deels afhankelijk van de spreeknelheid van commentator en respeaker. Een eerste concept voor een classificatiemodel van fouten in ondertitels toont de hoofdcategorieën in de verschijningsvormen van fouten: misrecognities, misinterpretaties en spel- en tyfefouten.

Dit verkennende onderzoek legt in de eerste plaats de basis voor verder onderzoek op korte en lange termijn. In vervolgonderzoek willen we verder het verband uitdiepen tussen de soort reductie en het voorkomen van bepaalde fouten. In een gecontroleerde, quasi-experimentele setting willen we bijvoorbeeld nagaan welke invloeden verschillende reductiegraden (*verbatim*, *licht gereduceerd* en *sterk gereduceerd*) hebben op de productie van fouten en de ondertitel-*delay* van eenzelfde programma. Het antwoord op die vraag is onder meer van belang als men ondertitelaars zonder correctors wil laten werken.

De resultaten van dit vervolgonderzoek zijn ook van belang als basis voor verder onderzoek op langere termijn, zowel vanuit het gezichtspunt van de ondertitelproductie als van het ondertitelgebruik - en zelfs vanuit didactisch oogpunt. Sommige misvattingen over ondertiteling voor doven en slechthorenden die Neves (2007) beschrijft, vormen daarvoor interessante uitgangspunten. Zo wijst ze onder meer op het feit dat de doelgroep van ondertiteling voor doven en slechthorenden niet homogeen is en argumenteert ze dat enkel *verbatim* ondertiteling de doelgroep echt gelijkwaardige mediatoegankelijkheid biedt. De mate waarin ondertitels al dan niet ingekort moeten worden om begrijpelijk te blijven, is het onderwerp van discussie onder onderzoekers, ondertitelaars, televisiezenders en gebruikers. Het is echter een complexe aangelegen-

heid precies omdat de doelgroep niet homogeen is en er spelen verschillende elementen een rol. Daarom is gericht en systematisch onderzoek nodig om te bepalen welke vormen van reductie de voorkeur genieten, zowel technisch als inhoudelijk. Daarvoor is naast meer analytisch onderzoek ook perceptieonderzoek nodig bij de verschillende doelgroepen.

Uiteindelijk is het de bedoeling om alle onderzoeksresultaten te gebruiken voor procedureverbeteringen in het live-ondertitelproces. Dit betekent ook dat we input willen leveren die nuttig is vanuit een didactisch oogpunt. Op dit ogenblik wordt live-ondertitelen voornamelijk gedoceerd binnen tolk- en vertaalopleidingen. Live ondertitelen vertoont immers duidelijke punten van overeenkomst met bijvoorbeeld simultaantolken (Anderson, 1994), maar hoe groot zijn de overeenkomsten en hoe groot zijn de verschillen? Ook die vragen moeten worden beantwoord om het leerproces van toekomstige ondertitelaars verder te verbeteren.

Dank

Graag danken we Erik de Snerck, Bernard Dewulf en de andere medewerkers van de VRT voor hun bereidheid om aan dit onderzoek deel te nemen. Alexander Lesser, Evelynne D'Milonas, Joyce Verschuere en Annelies den Haerincq, masterstudenten in de opleiding Meertalige Professionele Communicatie van de Universiteit Antwerpen, danken we voor hun hulp bij de transcripties. Stijn van Even en Hans Geuns (Nuance, Boston) maakten het mogelijk om de logging via Dragon Naturally Speaking in Inputlog te integreren. Veel dank aan hen voor de technische en inhoudelijke ondersteuning daarbij. Dit onderzoek werd mede mogelijk gemaakt dankzij een beurs uit het Bijzonder Onderzoeksfonds (BOF) van de Associatie Universiteit en Hogeschool Antwerpen (AUHA).

Noten

1. Ter informatie: De meeste omroepen in Europa hanteren de zogenaamde 6 seconde-regel (of een vergelijkbare richtlijn): die regel bepaalt dat een ondertitel van twee volledige regels (met 64 karakters) maximaal zes seconden op het scherm blijft staan; kortere regels blijven proportioneel minder lang zichtbaar (Gielen & d'Ydewalle, 1989). Het gebruik van die regel resulteert in een maximaal aanbod van 120 tot 125 woorden per minuut, ongeveer 50% van wat een gemiddelde volwassen lezer aankan (Schilperoord, De Groot, & Van Son, 2005). Op die manier kan de aandacht redelijk verdeeld worden tussen tekst en beeld. Of dit tempo geschikt is voor alle doelgroepen, staat echter geenszins vast, bovendien wordt de regel niet altijd gerespecteerd en is de leesnelheid van ondertitels vaak hoger.
2. In de ondertitelstudio staan ook televisieschermen opgesteld. Die bieden de mogelijkheid om simultaan de uitgezonden televisiebeelden te bekijken.
3. Het programma LiveComment is een VB.net programma ontwikkeld voor dit project. Het programma faciliteert de koppeling van tijdsaanduidingen aan transcripts van gesproken teksten. We danken Bart van de Velde die dit programma schreef.
4. We rapporteren hier enkel de resultaten voor de telling van de woorden. De resultaten voor de telling van de karakters waren echter in zeer hoge mate vergelijkbaar, wat erop wijst dat de reductie niet in de eerste plaats op woordniveau gebeurt (keuze voor kortere woorden), maar op zinsniveau (weglaten van woorden).
5. De reductieclassificatie is gedeeltelijk gebaseerd op het onderzoek van Joyce Verschuere. Ze rapporteert hierover in haar masterscriptie 'De oorzaken van reductie bij live-ondertiteling via spraakherkenning' (Master in de Meertalige Professionele Communicatie, Universiteit Antwerpen, 2008).

6. Dat laatste geldt vooral voor ‘scrolling’ ondertitels. Bij ‘block’-ondertiteling is het immers mogelijk om via kleuren en lay-out (twee regels) de uitingen van verschillende sprekers van elkaar te onderscheiden.
7. Soms zien we echter ook het tegenovergestelde proces en stellen we vast dat de respeaker ervoor kiest om de voornaam toe te voegen, ook al kwam die niet voor in het live-commentaar. De reden daarvoor is dat de combinatie van voornaam en familienaam de herkenning van persoonsnamen soms verbetert. Een voorbeeld: ‘Wellens’ wordt soms herkend als ‘wel eens’; ‘Bart Wellens’ levert meestal geen problemen op.
8. Uit een aanvullende analyse blijkt dat 78% van de fouten gecorrigeerd worden door de editor; de respeaker zelf neemt de overige 22% voor zijn rekening.
9. De foutenclassificatie is gedeeltelijk gebaseerd op het onderzoek van Alexander Lesser. Hij rapporteert hierover in zijn masterscriptie ‘Live-ondertiteling via spraakherkenning: Een foutenanalyse’ (Master in de Meertalige Professionele Communicatie, Universiteit Antwerpen, 2008).
10. Zo is bijvoorbeeld door een respeaker de naam van voetballer Gombani ingegeven als “Gombani”, waardoor de naam telkens weer fout wordt gespeld.

Literatuur

- Anderson, L. (1994). Simultaneous interpretation: Contextual and translation aspects. In S. Lambert & B. Moser-Mercer (Eds.), *Bridging the Gap. Empirical Research in Simultaneous Interpretation* (pp. 101-120). Amsterdam: Benjamins.
- Dewulf, B. (2005). VRT ondertiteling live-uitzendingen. Paper presented at the Congres Taal in Bedrijf (Nederlandse Taalunie).
- Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition and Communication*, 32, 400-414.
- Gielen, L., & d'Ydewalle, G. (1989). Hoe worden ondertitelde televisieprogramma's bekeken? *De Psycholoog*, 24, 425-431.
- Honeycutt, L. (2003). Researching the use of voice recognition writing software. *Computers and Composition*, 20, 77-95.
- Karat, C., Halverson, C., Horn, D., & Karat, J. (1999). *Patterns of entry and correction in large vocabulary continuous speech recognition systems*. Paper presented at the CHI 99, Pittsburgh.
- Lambourne, A. (2006). *Subtitle respeaking: A new skill for a new age*. [Electronic Version]. InTRAlinea 8 from http://www.intralinea.it/specials/respeaking/eng_more.php?id=447_0_41_0_M.
- Lambourne, A., Hewitt, J., Lyon, C., & Warren, S. (2004). Speech-based real-time subtitling services. *International Journal of Speech Technology*, 7, 269-279.
- Leijten, M. (2007). *Writing and Speech Recognition: Observing Error Correction Strategies of Professional Writers* (Vol. 160). Utrecht: LOT.
- Leijten, M., & Van Waes, L. (2003). *The writing processes and learning strategies of initial users of speech recognition: A case study on the adaption process of two professional writers* (Research Report No. 2003022). Antwerp: University of Antwerp.
- Leijten, M., & Van Waes, L. (2006). Inputlog: New perspectives on the logging of online writing. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer Key-Stroke Logging and Writing: Methods and Applications* (Vol. 18, pp. 73-94). Oxford: Elsevier.
- Neves, J. (2005). *Audiovisual Translation: Subtitling for the Deaf and Hard-of-Hearing*. University of Surrey-Roehampton: PhD Dissertation.

- Neves, J. (2008). 10 Fallacies about subtitling for the d/Deaf and the hard of hearing [Electronic Version]. *JosTrans: The Journal of Specialised Translations*, 10.
- Ofcom. (2006). *Television Access Services: Review of the Code and Guidance*. [Electronic Version]. Retrieved 24 February 2008 from <http://www.ofcom.org.uk/consult/condocs/accessservs/access.pdf>.
- Remael, A. (2007). Sampling subtitling for the deaf and the hard-of-hearing in Europe. In J. Díaz Cintas, A. Remael & P. Orero (Eds.), *Media for All* (pp. 23-52). Amsterdam: Rodopi.
- Remael, A., De Houwer, A., & Vandekerckhove, R. (2008). Intralingual open subtitling in Flanders: Audiovisual translation, linguistic variation and audience needs. *JosTrans: The Journal of Specialised Translations*, 10, 76-105.
- Schilperoord, J., De Groot, V., & Van Son, N. (2005). Nonverbatim captioning in Dutch television programs: A text linguistic approach. *Journal of Deaf Studies and Deaf Education*, 10(4), 402-416.
- SOAP. (2007). *Samenwerkingsverband Ondertitel Alle Programma's* [Electronic Version]. Retrieved 24 February 2008 from <http://www.leefwijzer.nl/default.asp?location=nieuwsartikel&layout=nieuws&id=3537>.
- Van der Veer, B. (2007). De tolk als respeaker: een kwestie van training. *Linguistica Antverpiensia*, 6, 315-328.
- Van Waes, L., & Leijten, M. (2006). Logging writing processes with Inputlog. In L. Van Waes, M. Leijten & C. Neuwirth (Eds.), *Writing and Digital Media* (Vol. 17, pp. 158-166). Oxford: Elsevier.

Bijlage

Tabel 8 Classificatie van reductie-operaties

Deletie	
1	Herhaling
2	Onbelangrijke en vanzelfsprekende gegevens
3	Voornaam
4	Voegwoord aan het begin van de zin
5	Spreektaalig tussenwerpsel, aarzeling, verspreking en 'verbal graffiti'
6	Aanspreking publiek, co-presentator of geïnterviewde
7	Numerieke gegevens in beeld
8	Rest
Substitutie	
9	Lange of moeilijke woorden door korte of eenvoudige woorden
10	Spreektaal of niet-Standaardtaal door (meer) formele Standaardtaal
11	Substantieven door voornaamwoorden en verwijzwoorden
12	Zegswijze, metafoor door beknopte omschrijving
13	Rest

Tabel 9 Classificatie van reductie-oorzaken

Commentaar	
1	Delay voorkomen
2	Te veel commentaar op korte tijd (spreeksnelheid > leessnelheid > schrijfsnelheid)
3	Onvolledige uiting en incorrecte grammaticale constructie
4	Onnodig complexe grammaticale constructie
5	Ingewikkelde of onduidelijke betekenis
6	Ingewikkelde interactie
Beeld	
7	Nieuwe (vooraf opgenomen) ondertitels staan klaar
8	Sudden event
9	Commentaar niet meer van toepassing
Spraaaktechnologie	
10	Anticipatie op probleem met spraakherkenner (niet in lexicon/vorige keer fout)
Extra	
11	Andere duidelijk aanwijsbare reden
12	Reden onduidelijk
