

## **De tolk als respeaker: een kwestie van training**

**Bart van der Veer**

University College Antwerp

*As a result of political decisions, all Dutch-spoken television programmes that are broadcast by the Flemish Broadcasting Company (VRT) should be provided with subtitles for the deaf and hard of hearing by the year 2010. In order to meet these high expectations, the VRT is constantly improving and changing its subtitling techniques, as are other broadcasting companies worldwide. One of the main areas of change concerns the technique of live subtitling, i.e. real time subtitling of live television programmes. This type of subtitling has definitely benefited from the use of modern speech recognition software. Live subtitling, therefore, requires not only technical skills but also excellent 'respeaking' skills that are reminiscent of the skills of conference interpreters. The central question in the first part of this paper is to what extent 're-speaking' is related to simultaneous (and other forms of) interpreting: is a good interpreter automatically a good respeaker? In the second part, I adopt a didactic point of view in order to investigate the teaching aspects of real time subtitling skills: the conclusion is that it is best included in an education programme for conference interpreters.*

### **1. Inleiding: live ondertitelen**

In de beheersovereenkomst van de Vlaamse openbare omroep (VRT) voor de periode 2007-2011 is vastgelegd dat in 2010 nagenoeg alle Nederlands-talige programma's moeten worden ondertiteld voor doven en slechthorenden. Ook de Nederlandse Publieke Omroep (NPO) moet dit streefdoel in dezelfde periode zien te halen volgens het Tussentijds Concessiebeleidsplan 2006-2010. (De genoemde beleidsdocumenten zijn te raadplegen op de websites van de omroepen, respectievelijk [www.vrt.be](http://www.vrt.be) en [www.omroep.nl](http://www.omroep.nl)). In andere Europese landen wordt eveneens volop gewerkt aan een ruimere toegankelijkheid van televisieprogramma's voor doven en slechthorenden. Zo zal de Britse BBC reeds vanaf april 2008 het volledige programma-aanbod ondertitelen.

In Europa worden de ondertitels voor doven en slechthorenden aangeboden in de vorm van 'closed captions', ondertitels die optioneel opgeroepen kunnen worden door de gebruiker. Sinds de vroege jaren tachtig gebeurt dat via teletekst, terwijl door de opkomst van digitale televisie tegenwoordig de toepassingsmogelijkheden worden verruimd. Zo kunnen televisieprogramma's met teletekstondertiteling eenvoudig worden opgenomen op een video- of dvd-recorder.

In 1987 werd er in het Verenigd Koninkrijk gestart met het ondertitelen van live nieuwsuitzendingen, een praktijk die al gauw ingang vond in andere Europese landen (zie Lambourne *et al.* 2004). Er bestaan verschillende systemen om ondertitels voor live uitzendingen te generen. In de VS wordt nog vaak gewerkt met stenografen. Een andere mogelijkheid is het inzetten van velotypisten. Een velotypist kan eventueel samenwerken met een respeaker<sup>1</sup> die op basis van de input de ondertitels dicteert, zoals dat bijvoorbeeld gebeurt bij het Nederlandse bedrijf NOB Cross media facilities voor live interlinguaal ondertitelen. Recente ontwikkelingen in spraaktechnologie hebben het mogelijk gemaakt om spraakherkenning te gebruiken voor het simultaan ondertitelen. Hierbij dicteert de respeaker de ondertitels aan een geautomatiseerde spraakherkenner.

Lambourne (2006) toont aan dat de techniek van het respeaken met spraakherkenning inmiddels zodanig is geëvolueerd dat er ondertitels kunnen worden geproduceerd met een snelheid van 140 tot 160 woorden per minuut en een nauwkeurigheid van 95-98%, afhankelijk van de gebruikte taal. Een bijkomend voordeel is dat de training van respeakers aanzienlijk minder tijd vergt dan het opleiden van stenografen en velotypisten: drie maanden versus één tot twee jaar. Spraakherkenning wordt tegenwoordig gebruikt door omroepen in verschillende landen waaronder BBC, NPO, VRT en omroepen in Denemarken, Duitsland, Australië en Nieuw-Zeeland.

De gunstige resultaten laten onverlet dat er uitdagingen zijn om de techniek te verbeteren. Een nauwkeurigheid van 95-98% resulteert gemiddeld in één fout per vier ondertitels. Bij de VRT vangt men dit probleem momenteel op door naast de respeaker ook een corrector/uitzender in te zetten die een snelle controle uitvoert op de gesproken tekst voordat die uitgezonden wordt. Door die 'filter' verloopt de synchronisatie tussen beeld en tekst soms problematisch en kan de vertraging van de ondertiteling ten opzichte van het beeld oplopen tot ongeveer acht seconden. Algemeen wordt aangenomen dat een vertraging van meer dan vijf à zes seconden problematisch is voor de slechthorende kijker (Lambourne *et al.*, 2004).

Een factor die een grote rol speelt bij de kwaliteit van de live ondertiteling is de wijze waarop de titels op het scherm worden weergegeven. Twee verschillende basissystemen worden gebruikt. Traditioneel worden ondertitels consecutief in blokjes van één tot maximaal drie regels gedurende een aantal seconden in beeld gebracht. Het is ook mogelijk de tekst continu over het scherm te laten lopen zoals in een lichtkrant. Deze laatste methode wordt gehanteerd in o.a. het Verenigd Koninkrijk, Italië en Denemarken, omdat de vertraging op deze manier aanzienlijk beperkt wordt (Baaring 2006). Dat de VRT en de NPO (vooralsnog) voor de blokmethode blijven opteren, is vooral ingegeven door taalgebonden softwarematige beperkingen.

De weergavemethode heeft een directe impact op de respeaking-techniek. Deze techniek kan worden opgevat als een continuüm zoals weergegeven in figuur 1.

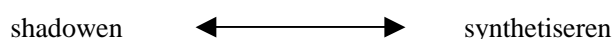


Fig. 1. Techniek van het respeaken: een continuüm

Het spreekt vanzelf dat de lichtkrantmethode (*scrolling mode*) het in principe mogelijk maakt de brontekst te ‘shadowen’, de techniek waarbij spraak – die via een koptelefoon wordt aangeleverd – vrijwel simultaan wordt herhaald. Een woordelijke weergave van de originele spraak geeft echter niet per definitie de beste resultaten, noch voor de kijker noch voor de respeaker. Voor de kijker kan er verwarring optreden, omdat de leesnelheid soms onaanvaardbaar hoog wordt of de inhoud van de tekst niet logisch of onbegrijpelijk is (Chen 2006). De respeaker zou volgens Kurz (1992) het risico kunnen lopen dat hij als een soort papegaai spraak weergeeft, zonder de inhoud ervan actief te analyseren en te interpreteren. Op het gebruik en de voordelen van de shadowtechniek bij de training van tolken en respeakers kom ik terug in paragraaf 2.3.

Van de respeaker wordt daarom verwacht het origineel enigszins te bewerken zonder afbreuk te doen aan de boodschap die moet worden overgedragen. Deze bewerking kan het karakter aannemen van het andere uiterste van het respeaking-continuüm, waar de brontekst gesynthetiseerd wordt weergegeven. Het gebruik van korte zinnen die de kern van de boodschap bevatten, is aangewezen bij de blokmethode, om de vertraging zo veel mogelijk te beperken. Een nadeel van deze techniek is dat er veel informatie verloren kan gaan. Voor zover ik kan nagaan, is nog niet onderzocht welke mate van synthese een ideale verhouding oplevert tussen de kwaliteit van de blokondertiteling en de synchronisatie ervan met het beeld. Hopelijk kan een onlangs opgestart onderzoeksproject aan de Universiteit Antwerpen en het Hoger Instituut voor Vertalers en Tolken van de Hogeschool Antwerpen hierin meer verheldering brengen.<sup>2</sup>

In dit artikel staat de techniek van het (meer of minder) synthetiserend respeaken met spraakherkenning centraal. Mijn uitgangspunt is dat de vaardigheden waarover respeakers moeten beschikken, grote gelijkenissen vertonen met de vaardigheden die nodig zijn voor het simultaan tolken (en andere vormen van tolken) (paragraaf 1), wat ook zijn weerslag heeft op de training (paragraaf 2).

## 2. Respeaken en tolken

Voor het live interlinguaal ondertitelen gebruikt het Nederlandse NOB bij voorkeur vertalers-ondertitelaars en geen conferentietolken. Volgens De Korte (2006) is dat een logische keuze omdat de NOB niet wil dat elk woord van de brontaal wordt vertaald. De velotypisten en de kijkers zouden dat tempo niet kunnen volgen. Bovendien zouden de ondertitelaars altijd in huis aanwezig zijn, in tegenstelling tot tolken. De teletekstafdeling van de

VRT ondervindt sinds kort juist dat er goede resultaten bereikt worden met afgestudeerde conferentietolken. In deze paragraaf zal ik uiteenzetten dat de techniek van het respeaken inderdaad zeer grote gelijkenissen vertoont met die van het simultaan tolken en andere vormen van tolken, al zijn er natuurlijk ook verschillen. Mijn verwachting is dat het inbouwen van een speciale training als respeaker in een tolkopleiding een meerwaarde is, niet alleen voor de studenten maar ook voor de omroepen.

### 2.1. Respeaken en simultaan tolken

Jones (2002) definieert vier basisprincipes van het tolken: luisteren, begrijpen, analyseren en herformuleren. Alvorens de ideeën van een ander te kunnen herformuleren en communicatie tussen twee of meer individuen mogelijk te maken, moet de tolk eerst aandachtig luisteren, achterhalen welke ideeën er uitgedrukt worden en de structuur en stijl van het betoog kunnen analyseren. Deze principes gelden voor alle vormen van tolken en uiteraard ook voor het respeaken. Voor het simultaan tolken signaleert Jones twee bijkomende moeilijkheden. Ten eerste is er de akoestische moeilijkheid van het spreken en luisteren tegelijk (*split attention*). Ten tweede is simultaan tolken in intellectueel opzicht een moeilijke activiteit omdat de tolk niet weet wat hem te wachten staat, noch op macroniveau (waar gaat het betoog als geheel heen?) noch op microniveau (hoe loopt de zin af?), factoren die eveneens een rol spelen bij het respeaken.

In het zojuist genoemde artikel van De Korte (2006) gaat de auteur voorbij aan het feit dat tolken zeker niet elk woord van de brontaal vertalen. Jones (2002) gaat uitvoerig in op het vierde basisprincipe van het tolkproces: het herformuleren. Lange, moeilijke zinnen worden bij voorkeur in korte zinnen weergegeven; ondergeschikte zinnen kunnen van plaats veranderen of worden zoveel mogelijk vermeden; passieve zinnen kunnen actief gemaakt worden; etc. Vanwege de tijdsdruk is efficiënt formuleren van groot belang. Waar nodig en mogelijk kunnen elementen worden weggelaten. Vlot en bondig kunnen herformuleren zijn dus troeven die tolken en respeakers gemeen hebben.

Zowel de simultaantolk als de respeaker werkt op basis van een eenmalige gesproken input die wordt omgezet in een gesproken output. De output van de tolk kan gewoonlijk worden beschouwd als een eindproduct dat niet gereproduceerd wordt. De output van een respeaker daarentegen fungeert op zijn beurt als input voor het genereren van geschreven ondertitels. Dankzij deze gemengde spraak-schriftdimensie situeert het respeaken zich op het raakvlak van tolken en transcriptiepraktijken in ruime zin, waarbij de transcribent als filter optreedt (bijv. transcripties in rechtbanken en vergaderingen en elke vorm van audiovisuele vertaling). Voor een interessante analyse van de overgang van gesproken naar geschreven taal verwijs ik graag naar Mack (2006).

De geschreven dimensie van het respeaken is bijzonder omdat het schrijfproces verloopt via spraakherkenningssoftware. Het gebruik hiervan

bepaalt het proces en de productie van de live ondertiteling (zie Leijten & Van Waes 2005). Dit zijn factoren waar een respeaker – in tegenstelling tot een tolk – zich zeer bewust van moet zijn bij zijn werk. Elk leesteken<sup>3</sup> moet worden uitgesproken en een goede articulatie primeert boven een aangename presentatie. Het herstellen van een pertinente fout, ook al niet eenvoudig bij het simultaan tolken, is bij het respeaken helemaal een netelige zaak, omdat alles wat gezegd wordt, ook in tekst en dus in ondertitels wordt omgezet. Werken met een corrector kan hier uitkomst bieden.

Een ander opvallend verschil tussen respeaken en simultaan tolken wordt veroorzaakt door de situationele context. Live ondertiteling is een vorm van communicatie waarbij de betrokken partijen zich alle op verschillende locaties bevinden: de spreker(s) in de televisiestudio of elders in de wereld, de kijkers thuis en de respeaker/ondertitelaar als bemiddelaar daartussenin op de teletekstafdeling. Simultaantolken werken doorgaans in dezelfde ruimte als de sprekers en de toehoorders en kunnen dus eventueel feedback krijgen (bij verwarring, technische problemen etc.). De laatste decennia vonden echter verschillende innovatieve vormen van tolken ingang. Daarvan worden er enkele besproken in de volgende paragrafen.

## 2.2. Respeaken en tolken op afstand

Tolken op afstand ('remote interpreting') kan verschillende vormen aannemen, van simultaan tolken op afstand tot telefoontolken (Niska 1998). Doorgaans worden bij deze manieren van tolken nieuwe technologieën gebruikt, bijvoorbeeld de toepassing van videoconferenties, maar vaak zijn de tolken dan nog wel aanwezig in de vergaderzaal en volgen ze een deel van de debatten via een beeldscherm. Mouzourakis (2006: 46) reserveert de term 'remote interpreting' voor

situations in which interpreters are no longer present in the meeting room, but work from a screen and earphones without a direct view of the meeting room or the speaker.

AIIC, de internationale beroepsorganisatie voor conferentietolken, vindt deze manier van tolken onaanvaardbaar (zie AIIC 2000), maar ondertussen gaan de experimenten verder. De tolken fysiek scheiden van de vergaderzaal heeft duidelijk een aantal voordelen. De keuze om op deze manier te werken is eerst en vooral ingegeven door financiële en logistieke overwegingen. Internationale organisaties kunnen bezuinigen op reis- en verblijfskosten van de tolken en er kan efficiënter worden ingespeeld op de specifieke noden van elke vergadering ('exotische' talen, specifieke talencombinaties etc.). De motieven kunnen ook heel praktisch van aard zijn, bijv. het ontbreken van tolkencabines in historische vergaderzalen of zalen die te klein zijn om er voldoende cabines te plaatsen.

De overeenkomsten qua technische set-up van het tolken op afstand en respeaken zijn opvallend. In beide gevallen wordt de audiovisuele input

gegeven via een beeldscherm en een koptelefoon en is er geen direct contact mogelijk met de spreker noch met de gebruiker. Beide activiteiten spelen zich af in een virtuele omgeving, waar 'virtueel' geïnterpreteerd moet worden als een predikaat voor teams waarvan de leden tot resultaten komen zonder dat ze zich op dezelfde locatie bevinden, wat mogelijk gemaakt wordt door de moderne communicatiemiddelen.

Onderzoek heeft aangetoond dat een dergelijke werkomgeving bij tolken kan leiden tot een gevoel van vervreemding – of het ontbreken van een participatiegevoel (Mouzourakis 2003). Het zou interessant zijn na te gaan welke rol het concept 'presence' (zie Kalawsky 2000) speelt in de activiteit van respeakers. Een groot deel van het werkcomfort en dus ook het gevoel 'erbij' te zijn, wordt bepaald door de kwaliteit van de beelden die op het scherm worden weergegeven. Kwaliteitsvolle televisiebeelden zullen in dit opzicht vermoedelijk een ander effect hebben op degene die ermee werkt dan de eerder statische beelden van de webcams die in vergaderzalen staan opgesteld. Mouzourakis (2006) pleit ervoor de kwaliteit van het tolken op afstand te vergroten door te investeren in innovatieve beeldcaptatie- en visualisatiemethodes.

Wordt er voor het tolken op afstand voor geijverd om op zijn minst de spreker op het beeldscherm te kunnen zien (Moser-Mercer 2005), dan is dat niet zo vanzelfsprekend voor respeakers. Een groot deel van het te ondertitelen programma-aanbod bestaat uit live uitgezonden sportwedstrijden, waar het commentaar buiten beeld wordt gegeven. De respeakers werken dan voornamelijk op het gehoor, al worden zij voor de interpretatie ervan dikwijls gesteund door de televisiebeelden.

### 2.3. Respeaken en simultaan filmtolken

Simultaan filmtolken wordt door Russo (2005: 1) gedefinieerd als "an oral mode of screen language transfer often required of professional interpreters for international film festivals." Voor filmtolken is een speciale competentie nodig die raakvlakken vertoont met andere vormen van vertalen:

- literaire vertaling: de tolk moet beschikken over een zeer fijn taalgevoel en stilistische vaardigheden, gezien de grote variatie in het aanbod aan filmteksten;
- conferentietolken: de tolk beschikt over de cognitieve vaardigheid om te spreken en tegelijk te luisteren naar de brontekst en kan waar nodig tekstreductiestrategieën toepassen;
- audiovisuele vertaling (nasynchronisatie en ondertiteling): de tolk heeft te maken met een grote tijdsdruk, aangezien de vertolking toch zo synchron mogelijk met de filmdialogen moet verlopen; de tekstreductiestrategieën zijn ook hier vergelijkbaar.

Over de vergelijking met ondertiteling merkt Gambier (2003: 178) het volgende op:

[...] subtitling is a kind of written simultaneous interpreting: both are constrained by temporal factors [...]; both are conditioned by a considerable density of information [...]; both are caught in the relationship between the written and the oral [...]; finally, both have to be conscious of special issues of reception [...].

Uit het voorgaande mag blijken, dat respeaken grote gelijkenissen vertoont met filmtolken. Evenals filmtolken worden respeakers geconfronteerd met zeer uiteenlopende inputgenres (variërend van nieuwsuitzendingen tot sport en variëte). Ook zij zien zich genoodzaakt tekstreductie toe te passen en te werken onder tijdsdruk, zoals reeds in de vorige paragrafen duidelijk is geworden. Tot slot zijn beide activiteiten gerelateerd aan audiovisuele vertaling, met dit verschil dat filmtolken gekenmerkt wordt door een 'written-to-be-spoken' tekstvariant en respeaken juist door een omgekeerd scenario, 'spoken-to-be-written'. De filmtolk vertaalt in de meeste gevallen filmdialogen vanuit een script (een soort "'feigned' fresh talk", Russo 2005:2; over de taal van filmdialogen, zie Remael 2001), terwijl de respeaker als het ware leert in ondertitels te formuleren.

Uit receptieonderzoek (Russo 2005) blijkt dat vooral de volgende kwaliteiten doorslaggevend zijn voor de kwaliteit van filmtolken: stemgebruik, correct taalgebruik, synchronisatie en algemene inhoud. Om de kwaliteit van live ondertiteling met spraakherkenning in kaart te brengen, is, voor zover ik kan nagaan, nog geen receptieonderzoek verricht. Naar verwachting zullen synchronisatie, taalgebruik en inhoud (compleetheid) hier ook een belangrijke rol spelen, terwijl stemgebruik uiteraard alleen van belang is voor de spraakherkenningssoftware. Een belangrijk verschil tussen respeaken en filmtolken is immers dat de filmtolk zich direct richt tot een publiek, dat, zo blijkt uit Russo's enquête, een aangename stem zeer belangrijk vindt.

#### **2.4. Respeaken en live televisietolken**

In zekere zin verhoudt live televisietolken (of 'live voice-over') zich tot respeaken/live ondertiteling als nasynchronisatie tot ondertiteling, zoals blijkt uit de onderstaande tabel:

Tabel 1: Taaltransfermogelijkheden voor televisie

a1. live voice-over	taalinput = spraak taaloutput = spraak
a2. nasynchronisatie	
b1. live ondertiteling	taalinput = spraak taaloutput = schrift
b2. voorbereide ondertiteling	

Hiermee is meteen een duidelijk verschil tussen de output van het respeaken en die van het live televisietolken aangestipt. Maar er zijn ook significante overeenkomsten.

Respeakers als televisietolken werken in een vergelijkbare setting met een 'off-screen cast' (bijv. de televisiekijkers en de opdrachtgever van het programma) en een 'on-screen cast' (bijv. de interviewer en de geïnterviewde) waarbij de televisietolk van één van beide casts deel kan uitmaken (Mack 2001) terwijl de respeaker uitsluitend off-screen werkt. Voor beiden is de live situatie een bijzonder grote stressfactor die het werk extra moeilijk maakt. Mack (2001) wijst op een aantal pijnpunten die zonder meer ook van toepassing zijn op het werk van de respeaker:

- er is nauwelijks tot geen voorbereidingstijd;
- de prestatie kan niet worden overgedaan;
- er is geen direct contact met de spreker(s);
- er moet snel en onder tijdsdruk worden gewerkt om vertraging te voorkomen;
- korte interventies maken het onmogelijk aan de stem en de stijl van de spreker(s) te wennen;
- inadequate beelden of een slechte geluidskwaliteit kunnen tot foutieve interpretaties leiden.

Mack (2001: 130) concludeert dat in het geval van televisietolken "working for TV is often frustrating because of the impossibility to meet (...) interiorised standards, mainly related to sense consistency and completeness". Waar het gaat om respeaken, besluit Den Boer (2001: 172) het volgende:

The ideal person for the job would be someone who is a qualified interpreter and a professional subtitler. These are two different fields of expertise. People who combine these qualities are extremely thin on the ground. The chances of being able to deliver a quality product are therefore quite small.

Inmiddels zijn we zes jaar verder en wordt er nog steeds volop door omroepen geëxperimenteerd om de kwaliteit van live ondertiteling te verbeteren. Er is recent ook veel belangstelling voor wetenschappelijk onderzoek op het gebied van respeaken, getuige het 'First international seminar on new tech-



nologies in real time intralingual subtitling' dat in het najaar van 2006 in Forlì werd georganiseerd. Met name bij academische instellingen die een tolkopleiding verzorgen, is de interesse gewekt en sinds kort worden dan ook respeakerstrainingen aangeboden aan tolkstudenten, o.a. in Italië en Vlaanderen. Over de didactische aspecten van het respeaken gaat de volgende paragraaf.

### 3. Didactiek van het respeaken

Op basis van de analyse van het respeaken die ik in de vorige paragraaf heb gemaakt, kunnen we concluderen dat een trainingsprogramma voor respeakers een aantal onderdelen bevat die ook aan bod komen in een opleiding tot simultaantolk. Hoewel het respeaken voornamelijk gericht is op intralinguale communicatie, is het zeker niet uitgesloten dat respeakers ook voor interlinguale transfer worden ingezet, bijv. voor het vertalend ondertitelen van grote live evenementen.<sup>4</sup> In zo'n geval is de aansluiting met simultaan (interlinguaal) tolken uiteraard nog groter.

Momenteel worden professionele (inter- en intralinguale) ondertitelaars om- of bijgeschoold tot respeaker door de omroepen zelf, bijv. de VRT. Een voordeel is uiteraard dat zij getraind zijn in segmentatietechnieken en op de hoogte zijn van de eisen van de klant (bijv. de doven en slechthorenden). Het is echter niet vanzelfsprekend dat zij ook live via spraakherkenning in staat zijn keurig gesegmenteerde ondertitels te produceren. Het is wellicht eenvoudiger voor een simultaantolk om te leren ondertitelen dan voor een ondertitelaar om te leren tolken. Een omroep als de VRT heeft dit ook ingezien en werft sinds kort bij voorkeur afgestudeerde conferentietolken als respeakers aan.

Het is voor tolkopleidingen interessant op deze ontwikkelingen in te springen. De behoefte aan (live) ondertiteling voor doven en slechthorenden neemt toe terwijl voor kleinere talen in Europa de behoefte aan conferentietolken eerder afneemt, mede door de opmars van het Engels als lingua franca, ook bij de instellingen van de Europese Unie. In deze paragraaf ga ik in op de vraag hoe tolkstudenten getraind kunnen worden als respeaker.

#### 3.1. Ondertiteling: segmentatietechnieken

Het volgen van een degelijke inleiding tot de theorie en de praktijk van het ondertitelen lijkt onontbeerlijk. Zo verwerven de studenten een goed inzicht in de noden van de klant/kijker, met name de doven en slechthorenden, en hebben ze kennis gemaakt met de zeer specifieke vertaalproblematiek van het ondertitelen (de technische entourage, spot- en segmentatietechnieken, reductiestrategieën etc.).

Na deze inleiding kunnen de studenten hun ondertitelingsintuïties uittesten aan de hand van een videofragment, bijv. een live interview tijdens een journaaluitzending. Ze transcriberen het fragment en bepalen, gewoon

op papier, hoe ze het fragment in ondertitels van acceptabele lengte zouden respeaken. Zo leren ze op gevoel segmenteren. Spotten is immers niet mogelijk bij live ondertiteling.

### 3.2. Spraakherkenning

In een tweede fase leren de studenten werken met spraakherkenningssoftware. Voor het Nederlandstalig taalgebied is dat Dragon Naturally Speaking van de firma Nuance. Na het aanmaken van een stemprofiel kunnen ze bijvoorbeeld de teksten (ondertitels) die ze in een eerdere oefening hebben gemaakt, inspreken in een tekstverwerkingsprogramma. Daarbij moeten ze letten op hun articulatie en spreektempo, versprekingen zo veel mogelijk leren vermijden en eraan wennen dat ze ook alle leestekens oplezen. Deze oefening is tevens nuttig om studenten te leren hoe ze nieuwe woorden aan het lexicon kunnen toevoegen en hoe ze met macro's kunnen werken.<sup>5</sup>

Een stap verder is natuurlijk het spontaan produceren van spraak en dan nog zodanig dat de spraakherkenner ermee overweg kan. Om de overgang van voorbereide naar spontane spraak te vergemakkelijken, kunnen de studenten eerst semispontane spraak produceren door te shadowen.

### 3.3. Shadowen

Lambert (1988: 381) geeft de volgende definitie van shadowen:

Technically speaking, shadowing is a paced auditory tracking task which involves the immediate vocalization of auditorily presented stimuli, i.e. word-for-word repetition, in the same language, parrot-style, of a message presented through headphones.

Het nut van shadowen in een respeakersprogramma (en een tolkopleiding) is tweeledig. Ten eerste is het een goede oefening om te wennen aan *split attention*: de cognitieve vaardigheid van het luisteren en spreken tegelijk. In de tweede plaats biedt shadowen de mogelijkheid tot het produceren van semispontane spraak, hetgeen nuttig is voor beginnende respeakers die hun spraak niet alleen moeten monitoren op inhoud maar ook op articulatie en spreektempo ten behoeve van de spraakherkenningssoftware.

Omdat shadowen het mogelijk maakt semispontane spraak te ontlokken, wordt de techniek trouwens ook toegepast in de experimentele taalkunde, niet zelden met gemanipuleerd bronmateriaal (zie bijv. Van der Veer 2006). De resultaten die hiermee bereikt worden, alsmede geheugentests en foutenanalyses uitgevoerd door Marslen-Wilson (1973 en 1975) en Marslen-Wilson & Welsh (1978), tonen aan dat shadowers toegang hebben tot syntactische en semantische informatie, ongeacht de decalage waarmee zij shadowen. Tot soortgelijke bevindingen komen Tonelli en Riccardi (1995). Hiermee lijkt de bezorgdheid van Kurz (1992) dat “[t]he advocates of sha-

dowing should be aware that a crucial element is missing in those exercises: the active analysis of the speech input” ongegrond.

### 3.4. Respeaking

In een laatste fase leren de studenten respeaken. Op de website <http://interpreters.free.fr> (“the only dedicated site for students of conference interpreting”) zijn zeer veel zinvolle oefeningen te vinden die ook voor respeakers van groot nut zijn. De oefeningen hebben o.a. betrekking op algemene kennis, *split attention*, decalage, herformuleren en stressbeheersing. Absolute beginners kunnen bijvoorbeeld starten met ‘cheating’: ze blijven hetzelfde fragment respeaken tot ze tevreden zijn. Er kan gevarieerd worden met de moeilijkheidsgraad en de lengte van de fragmenten. Sportwedstrijden zullen in het algemeen als makkelijker worden ervaren dan talkshows; een interview in het jeugdjournaal zal eenvoudiger zijn dan een interview in het gewone journaal.

### 2.5. De ‘salamitechniek’

Er is één techniek die bij simultaan tolken vaak wordt toegepast en die zeer zeker aandacht verdient bij het respeaken: de ‘salamitechniek’. Jones (2002) gebruikt deze term om aan te duiden hoe een tolk lange, ingewikkelde zinnen van een spreker opdeelt in kortere betekenisgehelen en die dan adequaat met elkaar verbindt. De lange zinnen worden als het ware als een worst in plakjes gesneden. Hoewel deze techniek vooral te pas komt op conferenties met een hoog stijlregister, is het voor respeakers aangewezen zich deze techniek ook eigen te maken omdat zij, zoals reeds aangegeven in paragraaf 1.3, ‘in ondertitels’ moeten spreken. Door korte, eenvoudige zinnen te gebruiken, kan vermeden worden dat de synchronisatie van de ondertitels met het beeld te problematisch verloopt en wordt de leesbaarheid vergroot. Oefeningen in de salamitechniek zijn dus bij uitstek geschikt om de kunst van het ‘op gevoel segmenteren’ verder te verfijnen (zie § 3.1).

Het nu volgende voorbeeld van de salamitechniek is geïnspireerd op een voorbeeld dat Jones in zijn boek geeft. Neem de volgende zin:

In het licht van de uitspraak van het internationale panel en in reactie op de niet-betaling van de compensatie door Amerikaanse staalexporteurs, die door de Amerikaanse overheid niet tot betaling werden gedwongen, ondanks haar wettelijke verplichtingen en de garanties die ze had gegeven, heeft Japan besloten unilateraal te handelen, waartoe het zeker gerechtigd is in het geval van niet-naleving van een uitspraak van een internationaal panel – en dat is hier het geval – door bij wijze van straf extra invoerrechten te heffen op de import van bepaalde platstaalproducten, hoewel die maatregelen voor de nabije toekomst niet gelden voor lange producten.

Voor iemand die direct naar deze zin luistert, in een bepaalde context, zal het niet moeilijk zijn de inhoud ervan te begrijpen. De tolk of de respeaker echter moet de zin analyseren en tegelijk hardop vertalen en bovendien zal hij waarschijnlijk minder ingewijd zijn in het onderwerp. De beste strategie is de toepassing van de salamitechniek. Als we ervan uitgaan dat de tolk/respeaker begint met spreken op het moment dat hij het woord 'panel' hoort, zou hij kunnen komen tot een resultaat dat lijkt op het volgende:

Het internationale panel deed een uitspraak. De Amerikaanse staal-exporteurs hebben geen compensatie betaald. De Amerikaanse overheid heeft hen niet tot betaling gedwongen. Toch had ze wettelijke verplichtingen en heeft ze garanties gegeven. Dus heeft Japan besloten unilateraal te handelen. Het is daartoe zeker gerechtigd omdat de uitspraak van het internationaal panel niet is nageleefd. Japan straft met extra invoerrechten op de import van bepaalde platstaalproducten. De maatregelen gelden nog niet meteen voor lange producten.

### **3.6. Respeaken en ondertiteling**

Tot slot kunnen de studenten overgaan tot het daadwerkelijk live ondertitelen van filmbeelden. De spraakherkenningssoftware wordt daarbij dus als inputmiddel gebruikt voor de ondertitelingssoftware. Op die manier kunnen de studenten de kwaliteit van hun vaardigheden evalueren: hoeveel fouten bevatten de ondertitels, zijn de ondertitels prettig leesbaar, hoeveel vertraging is er in de synchronisatie met het beeld? Er kan gewerkt worden in teams van twee studenten, waarbij de een respeakt en de ander de rol van corrector op zich neemt (waarvoor een grondige kennis van de spelling van de moedertaal vereist is). Het verdient zeker aanbeveling een bezoek aan de teletekstafdeling van een omroep te organiseren, zodat de studenten inzicht krijgen in hoe het er in de praktijk aan toegaat.

## **4. Conclusie**

Zoals het er nu naar uitziet zullen televisieomroepen (publieke en commerciële) in Europa (en daarbuiten) een steeds groter deel van hun programma-aanbod toegankelijk maken voor mensen met gehoorproblemen. Een gevolg daarvan is dat ook steeds meer live programma's via teletekst ondertiteld zullen moeten worden. Dankzij de ontwikkelingen in de spraaktechnologie is het live ondertitelen steeds efficiënter geworden, al is de techniek zeker nog voor verbetering vatbaar. Een essentieel onderdeel van de live ondertiteling, het respeaken, vertoont grote gelijkenissen met diverse vormen van tolken, met name simultaan tolken. Het ligt daarom voor de hand tolkstudenten te trainen als respeakers, vooropgesteld dat een tolk waarschijnlijk sneller leert ondertitelen dan dat een ondertitelaar leert tolken/respeaken.

Nader onderzoek moet uitwijzen of het inzetten van tolken als respeakers een positieve impact heeft op zowel de productie als de receptie van live ondertiteling.

## Bibliografie

- AIIC (2000). "Code for the use of new technologies in conference interpretation". *Communicate!* Online at [www.aiic.net](http://www.aiic.net).
- Baaring, Inge (2006). "Respeaking-based online subtitling in Denmark". *inTRAlinea Special issue: Respeaking*. Online at [www.intralinea.it](http://www.intralinea.it).
- Chen, Sheng-Jie (2006). "Real-time subtitling in Taiwan". *inTRAlinea Special issue: Respeaking*. Online at [www.intralinea.it](http://www.intralinea.it).
- De Korte, Thijs (2006). "Live inter-lingual subtitling in the Netherlands". *inTRAlinea Special issue: Respeaking*. Online at [www.intralinea.it](http://www.intralinea.it).
- Den Boer, Corien (2001). "Live interlingual subtitling". Gambier & Gottlieb (2001), 167-172.
- Eugeni, Carlo (2006). "Introduzione al rispeaking televisivo". *inTRAlinea Special issue: Respeaking*. Online at [www.intralinea.it](http://www.intralinea.it).
- Gambier, Yves (2003). "Introduction. Screen transadaptation: Perception and reception". *The Translator* 9(2), 171-189.
- Gambier, Yves & Henrik Gottlieb (eds) (2001). *(Multi)Media Translation. Concepts, Practices, and Research*. Amsterdam/Philadelphia: John Benjamins.
- Jones, Roderick (2002). *Conference interpreting explained*. Manchester: St. Jerome.
- Kalawsky, Roy S. (2000). "The validity of presence as a reliable human performance metric in immersive environments". 3<sup>rd</sup> International Workshop on Presence, Delft, Netherlands. Online at [www.presence-research.org/Kalawsky.pdf](http://www.presence-research.org/Kalawsky.pdf).
- Kurz, Ingrid (1992). "Shadowing Exercises in Interpreter Training". C. Dollerup & A. Loddegaard (eds). *Teaching translation and interpreting 1: Training, talent and experience*. Amsterdam/Philadelphia: John Benjamins, 245-250.
- Lambert, Sylvie (1988). "A Human Information Processing and Cognitive Approach to the Training of Simultaneous Interpreters". D.L. Hammond (ed.). *Languages at crossroads. Proceedings of the 29th Annual Conference of the American Translators Association ATA*. Medford, NJ: Learned Information Inc., 379-387.
- Lambourne, Andrew (2006). "Subtitle respeaking". *inTRAlinea Special issue: Respeaking*. Online at [www.intralinea.it](http://www.intralinea.it).
- Lambourne, Andrew, Jill Hewitt, Caroline Lyon & Sandra Warren (2004). "Speech-based real-time subtitling services". *International Journal of speech technology* 7, 269-279.
- Leijten, Mariëlle & Luuk Van Waes (2005). "Writing with speech recognition: The adaptation process of professional writers with and without dictating experience". *Interacting with Computers* 17(6), 736-772.
- Mack, Gabriele (2001). "Conference interpreters on the air: live simultaneous interpreting on Italian television". Gambier & Gottlieb (2001), 125-132.
- Mack, Gabriele (2006). "Detto scritto: un fenomeno, tanti nomi". *inTRAlinea Special issue: Respeaking*. Online at [www.intralinea.it](http://www.intralinea.it).
- Marslen-Wilson, William D. (1973). "Linguistic structure and speech shadowing at very short latencies". *Nature* 244, 522-523.
- Marslen-Wilson, William D. (1975). "Sentence perception as an interactive parallel process". *Science* 189, 226-228.
- Marslen-Wilson, William D. & A. Welsh (1978). "Processing interactions and lexical access during word recognition in continuous speech". *Cognitive Psychology* 10, 29-63.
- Moser-Mercer, Barbara (2005). "Remote interpreting: issues of multi-sensory integration in a multilingual task". *Meta* 50(2), 727-738.
- Mouzourakis, Panayotis (2003). "That feeling of being there: vision and presence in Remote Interpreting". *Communicate!* Online at [www.aiic.net](http://www.aiic.net).
- Mouzourakis, Panayotis (2006). "Remote interpreting: a technical perspective on recent experiments". *Interpreting* 8(1), 45-66.

- Niska, Helge (1998). "What is remote interpreting?" Handout presented at the Conference *Anवादores de vos, anavadores de nos*, Vigo, Galicia, 25-27 March. Online at <http://lisa.tolk.su.se/remote-niska.html>.
- Remael, Aline (2001). "Some thoughts on the study of multimodal and multimedia translation". Gambier & Gottlieb (2001), 13 – 22.
- Russo, Mariachiara (2005). "Simultaneous film interpreting and users' feedback". *Interpreting* 7(1), 1-26.
- Tonelli, Livia & Alessandra Riccardi (1995). "Speech Errors, Shadowing and Simultaneous Interpretation". *The Interpreters' Newsletter* 6, 67-74.
- Van der Veer, Bart (2006). *The Italian 'mobile diphthongs': a test case for experimental phonetics and phonological theory*. Utrecht: LOT (doctoraal proefschrift).

- 
- <sup>1</sup> Uiteraard zijn 'respeaker' en 'respeaken' neologismen, die gebaseerd zijn op het Engelse werkwoord 'to respeak' dat op zijn beurt een nieuwvorming is. In het Italiaans is onlangs de term 'respeakeraggio' geïntroduceerd (zie Eugeni 2006).
- <sup>2</sup> De titel van het project luidt: Live ondertitelen met spraakherkenning en wordt begeleid door Mariëlle Leijten en Luuk Van Waes van de Universiteit Antwerpen en door Aline Remael en Bart van der Veer van het Hoger Instituut voor Vertalers en Tolken, Hogeschool Antwerpen.
- <sup>3</sup> In recentere versies van een applicatie als Dragon Naturally Speaking hoeven punten aan het eind van een zin niet meer gedictieerd te worden.
- <sup>4</sup> Interlinguaal respeaken gebeurt ook al, zoals o.a. blijkt uit De Korte (2006), zij het dan niet altijd met spraakherkenning.
- <sup>5</sup> De spraaktechnologie maakt het mogelijk om voor veelgebruikte en problematisch herkenbare woorden of uitdrukkingen macro's of ideolectische items te maken. De VRT gebruikt daarvoor bijvoorbeeld het 'tox'-item. Voorbeeld: 'tox-komop' resulteert na herkenning in 'Kom op tegen Kanker'; 'tox-èmèr' resulteert in 'MR'.