

Master in het tolken

**Evaluatie van de kwaliteit van een tolkprestatie:
een vergelijking tussen de peerevaluatie en de zelfevaluatie**

Masterproef aangeboden door
Claire NARDON
tot het behalen van de graad van
Master in het tolken

Promotor: Franciska VANOVERBERGHE

Academiejaar 2016 – 2017

Voorwoord

Deze masterproef zou nooit tot stand zijn gekomen zonder de hulp van mijn promotor prof. Franciska Vanoverberghe. Dankzij haar terechte opmerkingen en raadgevingen over zowel het onderzoek als het schrijfproces, heb ik dit werk tot een goed einde kunnen brengen.

Verder wil ik alle zestien deelnemers aan mijn experiment bedanken voor hun tijd en inzet. Ook hun aanmoediging hebben zeker hun effect niet gemist. Dankzij Dina Marinus en Lisa Ernst heb ik de nodige postgraduaatstudenten kunnen contacteren om mijn experiment te kunnen verwezenlijken, waarvoor mijn oprechte dank. Om dezelfde reden ben ik dhr. Dirk Reunbrouck, coördinator van het postgraduaat Conferentietolken aan de KU Leuven Campus Antwerpen, en de verantwoordelijke van de ombudsdienst, dhr. Wim Schramme, zeer dankbaar om mij in contact te hebben gebracht met een laatste respondent in het postgraduaat Conferentietolken met Engels in zijn talencombinatie.

Daarnaast hebben mijn docente van het vak ‘Methodologie van het onderzoek in het tolken’, dr. Emmanuelle Gallez, en twee buitenlandse studenten, Sarah Thuderoz en Libby Morris, mij heel wat moeite bespaard door mij enkele online of schriftelijke bronnen te bezorgen die ik via mijn eigen instelling KU Leuven of via andere databanken helaas niet kon bemachtigen.

Ten slotte gaat mijn uiterste dank uit naar mijn meter Kristine Coenen. Haar redactionele hulp heeft mijn werk ongetwijfeld naar een hoger niveau getild. De ondersteuning van zowel mijn meter als mijn andere familieleden en hun vertrouwen in mijn kunnen hebben me steeds opnieuw in mijn doorzettingsvermogen gesterkt.

Inhoudstafel

Voorwoord

Résumé

1	Inleiding	7
2	Inzicht in kwaliteit en evaluatie	11
2.1	<i>Problematiek rond het concept ‘kwaliteit’</i>	11
2.1.1	Methodologische en praktische problemen	11
2.1.2	Gebrek aan consensus over het concept ‘kwaliteit’	14
2.1.3	Relatief karakter van kwaliteit	18
2.1.4	Mogelijke beoordelaars van de kwaliteit van een vertolking	23
2.2	<i>Kwaliteitscriteria</i>	28
2.2.1	Linguïstische criteria.....	29
2.2.2	Niet-linguïstische criteria.....	39
2.3	<i>Zelf- en peerevaluatie als beoordelingsmiddel</i>	43
2.3.1	Zelfevaluatie	43
2.3.2	Peerevaluatie	45
2.3.3	Vergelijking tussen de zelf- en de peerevaluatie	47
3	Methode	49
3.1	<i>Onderzoeksvraag</i>	49
3.2	<i>Onderzoeksopzet</i>	49
3.2.1	Deelnemers	49
3.2.2	Verloop van het experiment en het survey-onderzoek	50
3.2.3	Evaluatieformulier en glossarium	53
3.3	<i>Dataverzameling en –verwerking</i>	55
4	Resultaten	57
4.1	<i>Gemiddelde score van de zelfevaluaties</i>	57
4.2	<i>Afzonderlijke criteria</i>	58
4.2.1	Inhoudelijke overeenkomst met de brontekst: omissies	58
4.2.2	Inhoudelijke overeenkomst met de brontekst: toevoegingen.....	60
4.2.3	Inhoudelijke overeenkomst met de brontekst: substituties	62

4.2.4	Stopwoorden/stopgeluiden.....	64
4.2.5	Zelfverbeteringen	65
4.2.6	Pauzes	67
4.2.7	Correct gebruik van terminologie	69
4.2.8	Correct gebruik van grammatica.....	71
4.2.9	Tongval	73
4.2.10	Logische samenhang.....	75
4.2.11	Geschikte stijl	77
4.2.12	Aangename stem.....	79
4.2.13	Objectiviteit/onpartijdigheid.....	81
4.2.14	Spreektempo	83
4.2.15	Levendige intonatie	85
4.2.16	Zelfvertrouwen	87
4.2.17	Articulatie	89
4.2.18	Algemene beoordeling.....	91
4.3	<i>Algemene vergelijking tussen de zelf- en peerevaluaties.....</i>	93
4.4	<i>Totaalresultaat: gemiddelde scores per criterium.....</i>	95
5	Discussie en aanbevelingen	97
5.1	<i>Bemerkingen bij de resultaten</i>	97
5.2	<i>Methodologische beperkingen</i>	107
6	Conclusie.....	113
	Referentielijst.....	115
	Bijlage A.....	123
	Bijlage B	125
	Bijlage C	127
	Bijlage D.....	128
	Bijlage E	129
	Bijlage F	132

Résumé

Malgré une multitude de recherches d'une part sur la qualité d'une interprétation, d'autre part sur l'auto-évaluation et l'évaluation par les pairs, une comparaison entre les évaluations par les interprètes et celles de leurs collègues a rarement fait l'objet d'étude. Or, l'interprétation étant considérée comme un service, le contrôle de sa qualité reste indispensable afin de satisfaire aux attentes des clients.

Le but de cette thèse est de vérifier si l'hypothèse que l'auto-critique des étudiants est plus sévère ou non que la critique de leurs évaluateurs est bien fondée. Par le biais d'un formulaire d'évaluation présentant treize critères de qualité basés sur les études réalisées par Bühler (1986), Barik (1994), Gile (1990), Pöchhacker et Zwischenberger (2010) et Riccardi (2002), il a été demandé à des étudiants en master d'interprétation d'évaluer leur propre interprétation consécutive (du français ou de l'anglais vers le néerlandais). Ces mêmes interprétations ont ensuite été notées et annotées par leurs condisciples et des étudiants du master spécialisé en interprétation de conférence. Afin d'éviter les différences de compréhension sur les critères de qualité, un glossaire les spécifiant clairement a été ajouté en début de formulaire.

Contre toute attente, la comparaison individuelle des auto-évaluations et des évaluations par les pairs démontre que, contrairement à leurs évaluateurs, la plupart des étudiants s'attribuent un score plus élevé sur une majorité des critères de qualité. L'analyse des scores moyens par critère, en revanche, met en évidence que la majorité des paramètres est évaluée de façon plus rigoureuse par les étudiants en interprétation, confirmant l'hypothèse de cette thèse. Cette constatation contradictoire s'explique par certaines notes aberrantes des étudiants généralement affectant une attitude plutôt sévère. Lors de l'évaluation générale des prestations, les étudiants dans cette étude s'avèrent moins austères pour leur travail que leurs condisciples et que les étudiants du master spécialisé.

Par ailleurs, les annotations des étudiants en interprétation ainsi que des évaluateurs indiquent que les critères de qualité sont confondus, malgré le glossaire. Une répartition en catégories, comme l'a introduit Bühler (1986), est donc bien utile afin de pouvoir évaluer la qualité d'une interprétation, mais celle-ci s'avère insuffisante. Une plus ample description des critères est donc nécessaire.

Ces divers résultats confirment le caractère subjectif du concept « qualité » et montrent les visions hétérogènes des participants. Cependant, l'échantillon de recherche étant restreint, il reste nécessaire de mener des études complémentaires. L'évaluation mérite également un intérêt particulier de la part des professeurs, par exemple au moyen d'exercices pour familiariser les étudiants avec le système d'évaluation. En effet, un tel questionnaire comme dans cette étude montre aux étudiants que la critique et l'auto-critique sont nécessaires à une prise de conscience de leurs propres compétences et capacités.

1 Inleiding

In de laatste drie decennia groeide de interesse van onderzoekers in de kwaliteit van een vertolking (vb. Bartłomiejczyk, 2007; Chiaro & Nocella, 2004; Christensen, 2011; Grbíc, 2008; Kalina, 2005). Het beroep van tolken heeft een zekere status, wat dit fenomeen zou kunnen verklaren. Hoewel een vertolking misschien niet onmiddellijk als een verhandelbaar ‘product’ hoort te worden gecatalogeerd, wordt tolken toch algemeen als een van de vele diensten beschouwd die ter beschikking staan van consumenten. Net als bij andere ‘producten en diensten’ moet er daarom voldaan worden aan de verwachtingen van de gebruiker (Moser-Mercer, 1996). Volgens Grbíc (2008) is het bijgevolg zeer gemakkelijk om kwaliteit te garanderen: zodra de verwachtingen vervuld zijn, is de kwaliteit bereikt. Dat lijkt een relatief eenvoudige opdracht, alleen is kwaliteit een zeer subjectief concept. Slechts zelden slagen onderzoekers erin objectief onderzoek uit te voeren, omdat ze nauwelijks rekening houden met de werkomstandigheden van de tolk (Kahane, 2000).

We leven in “an era obsessed with quality control” (Kahana, 2000, par. 3); waarin klanten waar voor hun geld willen (Grbíc, 2008). Ongeacht de betrokken dienst, de kwaliteit speelt een belangrijke rol in de moderne maatschappij om algemene tevredenheid te waarborgen. Dat geldt ook voor de tolkenwereld, zeker omdat deze dienst, in de ogen van de gebruikers, behoorlijk wat geld kost en het om een discipline gaat die als doel heeft een efficiënte en duidelijke communicatie tussen de spreker en zijn toehoorders tot stand te brengen (Grbíc, 2008; Marrone, 1993; Moser-Mercer, 1996). Volgens Kalina (2002) mag het belang van kwaliteit niet onderschat worden: een kwalitatief slechte vertolking kan leiden tot ernstige misverstanden, die op hun buurt aanleiding kunnen geven tot spanningen en zelfs vijandigheid tussen verschillende partijen of gemeenschappen. Kalina (2005) benadrukt dat ook de brontekst kwalitatief aanvaardbaar moet zijn: sprekers zijn er vaak van overtuigd dat zij de Engelse taal, die zich in onze huidige maatschappij wijdverspreid tot *lingua franca* heeft ontwikkeld, voldoende beheersen. Toch is hun gesproken woord voor heel wat luisteraars – en dus ook voor tolken – niet altijd even verstaanbaar (Kalina, 2005). Het kan ook de andere kant op: aangezien Engels als *lingua franca* fungeert, is het volgens Shlesinger (1997) een cruciale vereiste geworden om een vertolking van een uitstekende kwaliteit te bieden. Een vertolking wordt immers steeds meer beschouwd als een luxediens. Dit is des te meer het geval bij Europese vergaderingen, waarbij afgevaardigden vaak een of meerdere vreemde talen voldoende

beheersen om de grote lijnen van de brontekst te begrijpen en bijgevolg enkel naar de tolk luisteren om meer inzicht te krijgen in nuances en details (Kurz, 1993).

De manier waarop kwaliteit beoordeeld hoort te worden, vormde de centrale focus van de meeste studies die zich toespitsen op kwaliteit (Christensen, 2011). Zo brachten de vele survey-onderzoeken van onder meer Bühler (1986), Kurz (1993), Chiaro en Nocella (2004) en Pöchhacker en Zwischenberger (2010), die de verwachtingen van de tolken enerzijds en van de gebruikers van de dienst anderzijds probeerden te achterhalen, belangrijke criteria aan het licht waaraan de vertolking zou moeten voldoen (zie 2.2). Ondanks deze verscheidene studies, blijft kwaliteit tot op heden een heikel thema. Voornamelijk over de invulling van het concept 'kwaliteit' (zie 2.1.2) en de personen die het best oordelen over de kwaliteit van een vertolking (zie 2.1.4) bestaat nog steeds enige mate van onenigheid (zie o.a. Gile, 1991a; Garzone, 2002; Moser-Mercer, 2008; Seleskovitch, 1986). In het merendeel van de studies werden de vertolkingen geëvalueerd door de personen die om een tolk hadden gevraagd. Deze masterproef probeert bovenstaand aspect te belichten vanuit een ander oogpunt door de zelfevaluaties van de tolkstudenten te vergelijken met de beoordelingen van medestudenten. Collega's beschikken namelijk over vergelijkbare kennis en zijn meer vertrouwd met de praktijk en de normen voor een vertolking, in tegenstelling tot het publiek (zie ook Gile, 1983). Deze masterproef bouwt voort op de masterthesis van DeBilde (2016), die voor haar onderzoek de evaluaties van de collega-tolk vergeleek met die van de *pure customer* om te achterhalen wie de tolk het meest kritisch beoordeelt. In tegenstelling tot de collega-tolk, kregen de *pure customers* de oorspronkelijke toespraak niet te horen.

Net als naar de kwaliteit van een vertolking, werd ook veelvuldig onderzoek gedaan naar het belang en de voor- en nadelen van zowel zelf- als peerevaluatie (o.a. Bartłomiejczyk, 2007; Iaroslavski, 2011; Ross, 2006; Sadeghi, Hassani, & Mohammadloo, 2015; Sadler & Good, 2006). De meeste empirische studies lieten echter vergelijkingen tussen de evaluaties van tolken en hun collega's over de kwaliteit van hun tolkprestatie achterwege. Kurz (1993, 2003) bracht aan het licht dat tolken wel zégggen dat ze strenger zijn voor zichzelf en hogere eisen stellen aan hun vertolking dan hun toehoorders, maar het gebeurt wel vaker dat een individu iets anders doet dan hij zelf te kennen geeft. De vraag kan dus gesteld worden of een dergelijke kritische houding zich ook vertaalt bij de beoordeling van de eigen tolkprestatie. In het onderwijs wordt geregeld een onderscheid gemaakt tussen 'beoordeling', waaronder een informele beoordeling over de eigen prestatie of die van een medestudent/collega wordt verstaan, en 'evaluatie', een

beoordeling die wordt verrekend in het eindcijfer (Gregory, Cameron, & Davies, 2000, geciteerd in Ross, 2006). Omdat dat verschil voor het onderzoek niet van toepassing is, krijgen de termen ‘beoordeling’ en ‘evaluatie’ in deze masterproef dezelfde betekenis, namelijk de informatie die door de student zelf of door een gelijke verstrekt wordt over een specifieke prestatie met als doel de eigen vaardigheden te verbeteren. Gebaseerd op onder meer de bovengenoemde bevinding van Kurz (1993, 2003), vertrekt deze masterproef vanuit de hypothese dat mede- en postgraduaatstudenten die gevraagd worden om de tolkprestaties van tolkstudenten te evalueren, minder streng zullen oordelen dan de tolkstudenten zelf.

Aangezien enkele vragen rond de positie van de beoordelaars van kwaliteit en de verschillende visies van tolken ten opzichte van hun prestatie in tegenstelling tot hun collega’s onbeantwoord blijven, beoogt deze masterproef meer inzicht te bieden in de evaluatie van een vertolking. De literatuurstudie (hoofdstuk 2) is tweeledig en bevat enerzijds een overzicht van de literatuur over het begrip kwaliteit en anderzijds een overzicht van publicaties over die evaluatie. Het derde hoofdstuk beschrijft het experiment en het survey-onderzoek. De resultaten van dit onderzoek worden belicht in hoofdstuk 4. Ze zullen al dan niet de hypothese bevestigen dat tolken een strengere zelfevaluatie zullen voorleggen dan hun medestudenten en bieden ruimte om na te gaan of de verwachte discrepanties tussen beide types evaluatieformulieren daadwerkelijk zullen optreden. Ten slotte volgt in hoofdstuk 5 de discussie met enkele aanbevelingen en in hoofdstuk 6 worden de conclusies uit het onderzoek getrokken.

2 Inzicht in kwaliteit en evaluatie

Ondanks de bewering van Moser-Mercer (1996) dat kwaliteit wel een veelbesproken maar nauwelijks onderzocht onderwerp is, getuigen vele tolkwetenschappelijke studies (vb. Bartłomiejczyk, 2007; Christensen, 2011; Garzone, 2002; Korpala, 2012; Pöchhacker, 2001, 2005; Shlesinger, 1997) van de aanzienlijke aandacht die reeds in het verleden aan kwaliteit besteed werd. Grbíc (2008) merkt wel op dat eerdere studies de theoretische aspecten vaak achterwege lieten: onderzoekers gingen niet op zoek naar de betekenis of een definitie van het begrip ‘kwaliteit’ (theoretisch onderzoek), maar richtten zich voornamelijk op de vertolkingen zelf en de evaluatie van de kwaliteit van die vertolkingen (empirisch onderzoek). Naast de lange ‘traditie’ van praktijkgerichte onderzoeken, zouden twee andere factoren dit gebrek aan theoretische conclusies kunnen verklaren: enerzijds de complexiteit van het concept, anderzijds de nadruk op het resultaat en de praktijk tijdens de les om zo hooggeschoolde tolken op de markt te brengen (Grbíc, 2008, p. 233). Ook al bestaat er ondertussen eensgezindheid over de kwaliteitscriteria waaraan vertolkingen getoetst horen te worden (Chiaro & Nocella, 2004; Déjean le Féal, 1990; Kahane, 2000; Gile, 1991a; Pöchhacker, 1994), de problematiek rond andere aspecten van kwaliteit is nog niet volledig opgelost.

2.1 Problematiek rond het concept ‘kwaliteit’

2.1.1 Methodologische en praktische problemen

Uit de vele survey-onderzoeken (vb. Amini, Ibrahim-González, Ayob, & Amini, 2015; Bühler, 1986; Chiaro & Nocella, 2004; Kopuszyński, 1994; Kurz, 1993; Pöchhacker, 2000; Pöchhacker & Zwischenberger, 2010; Shlesinger, 1997) blijken vragenlijsten het meest gebruikte meetinstrument voor de kwaliteitscontrole van een vertolking. Gile (1991b, p. 163) beschrijft de aanpak zelfs als “[t]he most straightforward scientific way of collecting data on actual quality perception by delegates”. Shlesinger (1997) voegt hieraan toe dat het invullen van de vragenlijsten zich niet hoeft te beperken tot enkel de tolken of de gebruikers van de tolkdienst, maar dat alle aanwezigen van de bijeenkomst die kunnen beantwoorden. Moser-Mercer (2008) echter is van mening dat er niet voldoende richtlijnen voorhanden zijn om deze survey-onderzoeken op een correcte en efficiënte wijze uit te voeren. Om die reden formuleert de auteur meerdere richtlijnen om toekomstige onderzoekers op het juiste pad te

zetten. Zo raadt ze bijvoorbeeld aan om voornamelijk gesloten vragen (vb. meerkeuzevragen) te stellen, omdat die, in tegenstelling tot open vragen, makkelijker te verwerken zijn. De criteria laten rangschikken van belangrijk naar onbelangrijk zou een nog geschiktere methode zijn om de kwaliteit te beoordelen. Verder kan het zinvol zijn een algemene, gevarieerde doelgroep uit te kiezen die een representatief beeld vormt van de werkelijkheid. De betrouwbaarheid van het onderzoek zou anders in het gedrang kunnen komen. Eenzelfde of een gelijkaardige studie moet namelijk overeenkomstige resultaten weergeven in uiteenlopende omstandigheden en bij andere respondenten – een voorwaarde die Hale en Napier (2013, p. 165) onder de term “external validity” plaatsen. Daar wringt het schoentje meestal niet, wél blijft er discussie bestaan over de “population validity” – wat Moser-Mercer (2008, p. 145) beschrijft als de nood aan een representatieve steekproef om de geldigheid van het onderzoek te verzekeren. Zo uiten Chiaro en Nocella (2004) openlijke kritiek op het kleine aantal respondenten (namelijk 124 verspreid over drie verschillende conferenties) binnen de studie van Kurz (1993), een opmerking die op verzet stuit bij Pöchhacker (2005). Op dezelfde manier hebben Pöchhacker en Zwischenberger (2010) en Kurz (1993) bedenkingen bij het (te) bescheiden aantal van zevenenveertig proefpersonen in het survey-onderzoek van Bühler (1986). Toch valt op te merken dat latere (replicatie)studies frequent verwijzen naar haar werk en dat de parameters die door deze onderzoeker werden gehanteerd regelmatig opnieuw opduiken in verscheidene vervolgstudies. Zelfs al zijn er redenen om de mening van Chiaro en Nocella (2004) te delen, dan nog mag niet vergeten worden dat een representatieve doelgroep verzamelen geen sinecure is. In tegenstelling tot het grote aantal beschikbare vertalers, kan er namelijk niet altijd een beroep gedaan worden op een ruime groep van tolken, die algemeen genomen beperkter zijn in aantal. Elke tolk biedt ook een eigen talencombinatie en de beschikbare groep verspreidt zich over verschillende continenten (Shlesinger, 1997; Moser-Mercer, 2008). Bovendien zijn sommige conferenties vertrouwelijk en kunnen onderzoekers er niet toegelaten worden (Kahane, 2000).

Verder komen meermaals de ongelijke variabelen ter sprake wanneer de voltooide survey-onderzoeken worden bestudeerd; studies omtrent kwaliteit streven vaak diverse doelen na. De vragenlijsten, doorgaans gebaseerd op het model van Bühler (1986), behandelen verschillende aspecten van kwaliteit, met een gebrek aan homogeniteit als gevolg. Zodoende ligt een betrouwbare vergelijking tussen de resultaten niet voor de hand (Kahane, 2000; Moser-Mercer, 2008; Shlesinger, 1997). Amini et al. (2015), Bühler (1986), Chiaro en Nocella (2004), Christensen (2011), Kurz (1993), Pöchhacker en Zwischenberger (2010) leggen zich toe op een

reeks linguïstische (semantische) criteria, soms aangevuld met niet-linguïstische (prosodische) kenmerken zoals bij de studie van Bühler (1986). Barik (1994) en Korpál (2012) richten zich dan weer op omissies (weglatingen) en Holub (2010), Pöchhacker en Zwischenberger (2010) en Shlesinger (1994) halen het belang van intonatie aan. Bij Vančura (2013) wordt aandacht besteed aan het spreektempo. Ook het aantal beoordeelde criteria stemt niet altijd overeen: de bekende studie van Bühler (1986) omvat bijvoorbeeld zestien linguïstische en niet-linguïstische kenmerken terwijl Marrone (1993) onderzoek doet naar het belang van drie kwaliteitscriteria (*volledigheid van de vertolking, stijl en correct gebruik van terminologie*) en drie parameters die een negatief element in de vertolking aanhalen (*incorrect gebruik van terminologie, onaangename stem en de fouten van de spreker overnemen*).

Niet enkel de methodologie vormt een knelpunt binnen onderzoek naar kwaliteit, ook de tolken die als proefpersoon optreden maken het de onderzoekers soms lastig. Wegens de harde economische concurrentie waarmee tolken te maken krijgen, staan professionele tolken vaak terughoudend tegenover de evaluatie van hun vertolking en tegenover een onderlinge vergelijking met het werk van hun collega's – uitzonderingen zijn te vinden in de studies van Ahrens (2005), Bühler (1986), Chiaro & Nocella (2004), Kurz (1993, 2003), Pöchhacker (2000) en de casestudy uitgevoerd door Christensen (2011). Bovendien zijn de toehoorders, als deelnemers aan een dergelijk onderzoek, voornamelijk geïnteresseerd in de inhoud van de toespraken en niet zozeer in de eigenlijke observatie en studie van tolken en kwaliteitsparameters (Chiaro & Nocella, 2004; Gile, 1983, 1995a; Kahane, 2000). Bijgevolg doen onderzoekers dikwijls een beroep op studenten of zelfs gewoonweg tweetalige personen, de zogenaamde 'natuurlijke tolken', die bij gelegenheid dienst doen als tolk hoewel ze geen tolkopleiding gevolgd hebben (Chiaro & Nocella, 2004; Pöchhacker, 2016, p. 23). Daarom spoort Gile (1990) onderzoekers aan om replicatiestudies uit te voeren die voorgaande resultaten moeten bevestigen of ontcrachten. Maar Gile (2003) spreekt zijn eerdere stelling wat tegen door te vermelden dat er desondanks geen probleem gemaakt hoeft te worden van studenten als proefpersonen, omdat klassen van een zekere omvang nu eenmaal meer mogelijkheden bieden om experimenteel onderzoek te verrichten. Evaluatie maakt bovendien deel uit van het leerproces en is daarom bij elk stadium – of het nu om een student of een professionele tolk gaat – een zinvolle toetsing (Gile, 2003). Ook bij de eigen prestatie zijn tolken er niet erg op gebrand om hun vertolking te toetsen aan de criteria die zij hoog in het vaandel dragen, omdat ze niet graag geconfronteerd worden met hun zwaktes. Daar zou echter verandering in moeten komen, aangezien professionele tolken herhaaldelijk klagen over de lage

kwaliteit van de prestaties van pas afgestudeerde – en daardoor nog te onervaren – tolkstudenten (Gile, 1990; Kalina, 2005).

Ten slotte bestaat er discussie over wie nu eigenlijk de verantwoordelijkheid draagt voor de kwaliteit: wie garandeert de kwaliteit van de tolkprestatie? (zie 2.1.4) Zowel Garzone (2002, 2003) als Chiaro en Nocella (2004) zijn het erover eens dat die verantwoordelijkheid volledig bij de tolk ligt. Hij heeft immers als enige een zekere mate van controle over de inhoudelijke overeenkomst met de oorspronkelijke tekst, de coherentie van de boodschap en het vlotte verloop van het spreektempo. Dat is inderdaad het geval bij freelancetolken, maar hetzelfde geldt niet noodzakelijk voor het vaste personeel dat wordt aangeworven bij internationale organisaties, zoals de EU of de VN, waar bepaalde standaardnormen voor kwaliteit en bijkomende opleidingen worden opgelegd (Kalina, 2005, p. 778). Volgens Vuorikoski (2004, geciteerd in Kalina, 2005) is het echter de nauwe samenwerking tussen de tolk en de spreker die garant staat voor de kwaliteit van de vertolking.

2.1.2 Gebrek aan consensus over het concept 'kwaliteit'

De AIIC (The International Association of Conference Interpreters) hecht uitzonderlijk veel belang aan de kwaliteit die haar leden kunnen waarborgen. Daarom onderwerpt de organisatie nieuwe leden aan zeer strenge selectiecriteria en kijkt het CACL ('Committee on Admissions and Language Classification'), het statutaire orgaan van de AIIC, toe op die kwaliteit (Admissions Committee, 2000, par. 1). Voor kandidaat-leden gelden strenge voorwaarden: toekomstige tolken moeten uiterlijk drie jaar vóór het indienen van hun kandidatuur reeds honderdvijftig dagen actief bij de instelling gewerkt hebben, waarbij ze in authentieke werkomstandigheden geobserveerd en geëvalueerd werden door minstens drie ervaren professionele AIIC-tolken (Admissions Committee, 2000; Bühler, 1986; Kahane, 2000; Kurz, 1993). Kalina (2005) haalt aan dat, nu er bij internationale organisaties steeds meer vraag is naar tolken, ook andere organisaties het voorbeeld van de AIIC zijn gevolgd door toelatingstesten in te voeren. Er worden klaarblijkelijk hoge eisen gesteld aan de tolken en de kwaliteit die ze moeten leveren, maar dan stelt zich de vraag: welke eisen worden er juist gesteld en hoe definieert men concreet die kwaliteit?

Ondanks de normen die de AIIC heeft opgesteld, bestaat er nog steeds geen consensus over de definitie van kwaliteit, noch over een universeel model voor de evaluatie van kwaliteit (Kahane,

2000; Kalina, 2005; Garzone, 2002, 2003; Gile, 1983; Grbíc, 2008; Moser-Mercer, 2008; Shlesinger, 1997). De moeilijkheden bij de opstelling van een definitie – die volgens Grbíc (2008, p. 236) voor het eerst buiten de tolk- en vertaalwetenschap is ontstaan – komen duidelijk naar voren wanneer de uiteenlopende voorstellen van verscheidene auteurs onder de loep genomen worden. Volgens Garzone (2002, p. 110) staat kwaliteit voor “a construct embodying the norms which are deemed appropriate to guarantee the intrinsic and extrinsic properties considered ideal for an interpretation performance in a given social, cultural and historical situation”. Deze definitie stemt echter niet tot gehele tevredenheid, aangezien Chiaro en Nocella (2004) erop wijzen dat onderzoekers zelfs in het begin van de 21^{ste} eeuw nog steeds niet kunnen steunen op bevredigende, algemene standaardnormen – en deze term mag begrepen worden als “general values or ideas shared by a certain community – as to what is right and wrong, adequate and inadequate – into specific performance-instructions appropriate for and applicable to specific situations” (Toury, 1980, p. 50) – die de fundamenteen leggen voor een optimale kwaliteitscontrole. Dit staat in scherp contrast met de beroepscategorie van vertalers, voor wie een kwaliteitscontrole gemakkelijker wordt geacht wegens het onvergankelijke karakter van een vertaling: wat geschreven staat is immers makkelijker te beoordelen dan het gesproken woord. Toch betwist Shlesinger (1997) deze uitspraak door te stellen dat er wel degelijk normen worden aangebracht tijdens universitaire tolkopleidingen – weliswaar een beperkt aantal –, maar het blijft te onduidelijk in welke mate die universiteiten die normen ook zelf hanteren en of ze dézelfde normen toepassen. Ook Déjean le Féal (1990) bespreekt het bestaan van deze professionele standaardnormen, naast de persoonlijke eisen die de tolk zichzelf oplegt. De auteur vat de normen – en in wezen dus ook het begrip ‘kwaliteit’ (niet te verwarren met de term ‘kwaliteitscriteria’, zie onder) – als volgt samen:

What our listeners receive through their earphones should produce the same effect on them as the original speech does on the speaker’s audience. It should have the same cognitive content and language. Its language and oratory quality should be at least on the same level as that of the original speech, if not better, given that we are professional communicators [...]. (Déjean le Féal, 1990, p. 155)

De auteur doelt met andere woorden op de vertaaltechniek ‘equivalentie’: in plaats van de taaluitingen in de brontaal letterlijk om te zetten naar de doeltaal, moet de tolk op zoek gaan naar een equivalent waarbij de structuur en de stijl van de zin mogelijk veranderen, maar de oorspronkelijke betekenis wel behouden blijft (Molina & Hurtado Albir, 2002; Vinay &

Darbelnet, 1995). Zo zegt de zin “Il pleut des chats et des chiens” (een letterlijke vertaling van de Engelse uitdrukking “It is raining cats and dogs”) een Franstalige toehoorder helemaal niets. “Il pleut des cordes” zal hem daarentegen wél bekend in de oren klinken. (Vinay & Darbelnet, 1995, p. 38)

Al deze maatstaven zouden doorheen de tijd opgesteld zijn door onder meer de toehoorders, collega’s, professionele instellingen of de onderzoekers zelf (Grbíc, 2008, p. 239). Gile (1983, p. 239) schuift een andere omschrijving naar voren: “la qualité de l’interprétation est une somme pondérée de la fidélité informationnelle et de la qualité de la présentation du discours de l’interprète, dont les coefficients sont fonction de la nature du message et de la situation.” Deze definitie lijkt zowel inhoudelijke kwaliteitscriteria (“la fidélité informationnelle”) als vormelijke parameters (“la présentation du discours”) op te nemen en houdt rekening met het relatieve karakter van kwaliteit (“la nature [...] de la situation”), maar wanneer kan een tolk of luisteraar spreken van een kwalitatief aanvaardbare inhoud en presentatie? Moser-Mercer hanteert een gelijkaardige definitie, waarin ze een “optimum quality” beoogt:

Optimum quality in professional interpreting implies that an interpreter provides a complete and accurate rendition of the original that does not distort the original message and tries to capture any and all extralinguistic information that the speaker might have provided subject to the constraints imposed by certain external conditions. (Moser-Mercer, 1996, p. 44)

De status van een perfecte tolk – de zogenaamde ‘elite’ van de tolken waarnaar Europa op zoek is (Chiaro & Nocella, 2004) – wordt inderdaad nagestreefd door het invoeren van strenge toelatingsvoorwaarden, al moet in het achterhoofd gehouden worden dat kwaliteit afhankelijk is van diverse factoren, zoals “certain external conditions” (zie ook 2.1.3), en bijgevolg zeer moeilijk haalbaar is. Deze overweging raakt echter bij sommige beoordelaars af en toe in vergetelheid (Grbíc, 2008; Moser-Mercer, 1996). Ondanks alles is het volgens Chiaro en Nocella (2004) geen onhaalbare opdracht, omdat het onbereikbare karakter van de perfectie niet noodzakelijk het leveren van goede kwaliteit uitsluit.

Om allerhande facetten van kwaliteit te belichten (streven naar perfectie vertegenwoordigt namelijk slechts een aspect) onderscheidt Grbíc (2008) drie manieren om naar kwaliteit te kijken: kwaliteit als een uitzondering, kwaliteit als de perfectie, of als voldoende geschikt voor

het beoogde doel. Kwaliteit wordt volgens haar herhaaldelijk beschouwd als een uitzondering of exclusiviteit, met als gevolg dat personen die kwaliteit kunnen waarborgen een hoge status bekleden. Dat gold zeker in een ver verleden waar simultane tolken als ‘miraculeuze talenten’ beschouwd werden en de eerste werken over het beroep tolken meermaals melding maakten van “intuitive, and even self-congratulatory references to an impossible job well done” (Shlesinger, 1997, p. 124). Vanuit dit standpunt is een optimale kwaliteit volgens Grbíc (2008) bijgevolg onhaalbaar. Enkel wanneer ze voldoet aan opgelegde criteria, is er geen sprake meer van exclusiviteit, maar van een luxe. Kwaliteit als perfectie kan bereikt worden door doorzettingsvermogen, persoonlijke inzet en regelmatige oefening en gaat gepaard met de “optimum quality” van Moser-Mercer (zie hoger). In dat geval ligt bij de tolk de verantwoordelijkheid om geen enkel gebrek te vertonen en een coherent resultaat te leveren. Deze perfectie wordt niet alleen nagestreefd door de AIIC; ook de CIUTI (Conférence des directeurs d’instituts universitaires d’enseignement de la traduction et de l’interprétation) vraagt een certificaat alvorens nieuwe leden toe te laten. Een laatste door Grbíc (2008) geopperde mogelijkheid, waarbij de kwaliteit geschikt wordt geacht voor het beoogde doel, heeft betrekking tot de dienst als product. Zoals eerder vermeld, kan kwaliteit volgens deze visie bereikt worden wanneer aan de verwachtingen van de consumenten voldaan is. Aangezien deze dienst echter vanuit verschillende perspectieven bekeken kan worden, rijst de terechte vraag: “For whom is the product or service useful?” (Grbíc, 2008, p. 247)

Ook al bestaat er een sterke behoefte aan een definitie van kwaliteit, studies streven slechts zelden naar het creëren van een dergelijke beschrijving en laten de theoretische aspecten veelal achterwege (Gile, 1990; Grbíc, 2008). Slechts nu en dan wordt dankzij onderzoek duidelijk hoe verschillende auteurs die het besproken onderwerp behandelen over het concept ‘kwaliteit’ denken (Moser-Mercer, 2008), maar hun vage formuleringen maken het voor tolken onduidelijk wat er precies van hen verwacht wordt (Shlesinger, 1997). Gile (1995a) probeert nochtans een zeer concrete beschrijving voor te leggen:

Interpretation quality can be viewed as a subjectively weighted sum of a number of components: the fidelity of the target-language speech, the quality of the interpreter’s linguistic output, the quality of his or her voice, the prosodic characteristics of his or her delivery, the quality of his or her terminological usage, all of them as perceived by the assessor. (Gile, 1995a, p. 151)

Dat heeft dan weer als gevolg dat tolken onderling moeite hebben om een duidelijk onderscheid te maken tussen de kwaliteitscriteria, namelijk elementen waaraan elk afzonderlijk individu een zekere mate van belang hecht, en standaardnormen, d.i. door instellingen (zoals de AIIC) opgelegde regels die reeds volledig in de tolkwereld geïntegreerd zijn en als vanzelfsprekend beschouwd worden (Shlesinger, 1997). De verschillende definities aangeboden door de auteur wekken immers de indruk dat de termen ‘normen’ en ‘criteria’ eenzelfde begrip weergeven, namelijk kwaliteit. Het is evenwel meer aangewezen om kwaliteit in categorieën en criteria op te delen en specifiekere variabelen te bestuderen, want het concept is te uitgebreid om zeer grondig en tot in de details te analyseren (Moser-Mercer, 2008). Bijgevolg kunnen bepaalde studies slechts een gedeeltelijk beeld weergeven van het uitgebreide concept ‘kwaliteit’ (Garzone, 2003). Wel komen in de omschrijving van Gile (1995a) – net als bij de definitie van Moser-Mercer (1996) – de twee aspecten aan bod die volgens Kopczyński (1994) onontbeerlijk zijn om kwaliteit te benoemen: het linguïstische en het pragmatische aspect. Gile (1990) vult aan met een correct gebruik van de microfoon (vb. geen storend geluid van ritselend papier, geen luidruchtige bewegingen), alsook probleemoplossende strategieën van de tolken.

Uit wat voorafgaat blijkt dat kwaliteit moeilijk te vatten is onder één noemer (Kurz, 2003; Shlesinger, 1997). Toch acht Gile (1983) een duidelijke definitie noodzakelijk, omdat ze de nauwkeurige analyse van een tolkprestatie mogelijk maakt en de resultaten betere opleidingen en werkmethoden kunnen opleveren voor toekomstige tolken. Geen enkele definitie met betrekking tot kwaliteit geeft tot op heden volledige voldoening of beantwoordt uitsluitend aan objectieve criteria.

2.1.3 Relatief karakter van kwaliteit

2.1.3.1. *Invloed van externe omstandigheden*

Kwaliteit is geen intrinsiek kenmerk van een vertolking, maar hangt in grote mate af van diverse externe factoren waarop je als tolk slechts weinig greep hebt (Chiaro & Nocella, 2004; Garzone, 2002, 2003; Gile, 1983, 1991a, 1995a; Grbíc, 2008; Kahane, 2000; Kalina, 2002, 2005; Kopczyński, 1994; Kurz, 1993; Moser-Mercer, 1996; Vančura, 2013). De problemen waarmee tolken zich geconfronteerd zien, worden bovendien door heel wat toehoorders tijdens de evaluatie van een vertolking zwaar onderschat (Kahane, 2000). De context waarin de vertolking plaatsvindt, is vaak van doorslaggevend belang voor het resultaat van de tolkprestatie en is

veranderlijk van aard (Gile, 1983; Grbíc, 2008). Bovendien komt die context niet altijd overeen met de situatie waar de tolk zich op voorbereid had. Niet alleen de omvang van de bijeenkomst, maar ook het type conferentie speelt bijvoorbeeld vaak een rol: op een wetenschappelijk congres zal bij een vertolking wellicht de inhoud vooral van tel zijn, terwijl een bijeenkomst van literaire en diplomatieke tolken of een politieke vergadering eerder ‘stijl’ als belangrijk aandachtspunt zullen meekrijgen (Pöchhacker & Zwischenberger, 2010; Kopczyński, 1994). Bij een televisie-uitzending krijgen de stemkwaliteit, de tongval of de vlotheid van de getolkte doelttekst volgens Kahane (2000) meer aandacht dan tijdens andere activiteiten waar een tolk vereist is. Andere factoren zoals een slechte geluidskwaliteit, het ontbreken van de nodige documenten die vooraf ter voorbereiding aan de tolk worden overhandigd, de technische apparatuur, het aantal werkuren of de onmogelijkheid om de visuele hulpmiddelen of de spreker zelf goed te zien, kunnen een voortreffelijke vertolking in de weg staan (Garzone, 2002; Gile, 1983; Kahane, 2000; Kalina, 2002, 2005; Moser-Mercer, 1996; Pöchhacker, 2005; Vančura, 2013). Kalina (2002) is echter van mening dat bij meertalige evenementen onvoldoende rekening gehouden wordt met al deze factoren. Daarom pleit ze voor een betere samenwerking tussen de tolken en de organisatoren van dergelijke evenementen en spoort ze aan om bijkomend onderzoek te verrichten naar de invloed van de werkomstandigheden op een vertolking. Verder haalt Kalina (2005) in haar studie het belang aan van het aantal werktalen en de richting waarin met die talen gewerkt wordt. In West-Europa heerst de overtuiging dat hoe meer talen je beheerst, hoe minder kennis je kan verwerven over de linguïstische aspecten van de verschillende talen en de cultuur en de bijhorende tradities van de betrokken landen. Hoge kwaliteit kan dus beter gewaarborgd worden wanneer de tolk vertaalt naar zijn moedertaal. Toch is er een verschuiving merkbaar binnen deze visie; in Canada en Italië, bijvoorbeeld, worden er steeds meer retourtolken ingezet (Garzone, 2002).

2.1.3.2. Invloed van de spreker

Een vertolking is een momentopname en wordt beïnvloed door een groot aantal factoren die volledig bepaald worden door de spreker. Die spreker is zich echter niet altijd bewust van zijn invloed op de vertolking. Gile (1983), Kahane (2000), Kalina (2005) en Vančura (2013) stellen vast dat zowel het spreektempo als de intonatie en de tongval (moedertaalspreker of niet) van de woordvoerder bepalen op welke manier de tolk de oorspronkelijke boodschap interpreteert. Ook wanneer de spreker hele stukken citeert, zich verspreekt of zich onzorgvuldig in vage of ambigue bewoordingen uitdrukt, kan de kwaliteit van de vertolking in het gedrang komen. Ten

slotte kunnen metaforen, figuren of eigennamen waarover de tolk niet werd ingelicht of woordspelingen en beeldtaal waarvoor de tolk geen passend equivalent voorhanden heeft de vertolking negatief beïnvloeden.

2.1.3.3. Invloed van de tolk

Uiteraard draagt de tolk zijn deel van de verantwoordelijkheid. Interne factoren zoals een goede achtergrondkennis over het onderwerp en een uitstekende voorbereiding, kunnen een vertolking al heel wat professioneler doen klinken (Kalina, 2005). Stress kan dan weer het tegenovergestelde effect hebben (Vančura, 2013). Toch bieden deze elementen geen garantie voor kwaliteit. Gile (1983, 1991a) en Déjean le Féal (1990) wijzen erop dat, zelfs wanneer de tolk vooraf de nodige informatie heeft kunnen doornemen, hij doorgaans toch over minder uitgebreide kennis en woordenschat beschikt dan de spreker en de toehoorders die wél thuis zijn in het onderwerp en dat is al helemaal zo wanneer het gaat om technische of wetenschappelijke onderwerpen. Achtergrondkennis zou dus niet de enige vereiste mogen vormen waaraan de kwaliteit van een vertolking getoetst wordt. De tolk moet ook aandacht besteden aan het doel dat de spreker met zijn brontekst wil bereiken. Om die reden is het voor een tolk belangrijk om – in de mate van het mogelijke – zowel de spreker als het doelpubliek te leren kennen om hun verwachtingen te achterhalen (Garzone, 2003; Gile, 1991a; Iaroslavski, 2011). Een goede voeling met de participanten maakt het voor de tolk makkelijker om bij essentiële keuzes de knoop door te hakken: trouw blijven aan de brontekst of de inhoud voor de toehoorders begrijpelijker maken door een eenvoudiger taal te hanteren, toezien op een correct taalgebruik of precisie nastreven (Kalina, 2005)? Op dezelfde wijze is het voor de tolk belangrijk om op de hoogte te zijn van eventuele cultuurverschillen: bepaalde formuleringen of uitdrukkingen kunnen in de brontaal erg kwetsend overkomen wanneer deze ondoordacht vertaald worden naar de doeltaal. Zo wordt het beruchte Amerikaanse vierletterwoord door Poolse luisteraars als zeer schokkend ervaren wanneer het vertaald wordt naar hun moedertaal (Kopczyński, 1994). Kahane (2000) haalt een voorbeeld aan van een internationale delegatie waar de Noord-Europeanen veelal een letterlijke vertaling verwachtten, terwijl de bewoners van de mediterrane landen rekenden op een ‘interpretatie’ van de boodschap door de tolken. Anderzijds zou een negatieve attitude van de tolk een goede communicatie in de weg kunnen staan. Ervaring en voortreffelijke linguïstische vaardigheden kunnen naar de mening van Garzone (2002, 2003) in dergelijke situaties een grote hulp bieden. Het werkgeheugen kan echter niet onophoudelijk actief blijven en de inspanningen om alle

informatie op te nemen en om de tekst om te zetten naar de doeltaal eisen na verloop van tijd hun tol. Deze vermoeidheid vertaalt zich in een kwalitatief slechtere prestatie waarbij weglatingen, een mindere taalkwaliteit of zelfs een foutieve weergave van de inhoud, kunnen opduiken (Gile, 1983; Marrone, 1993). Tolken zijn zich wel degelijk bewust van de mogelijke problemen en trachten hierop te anticiperen. Volgens Kalina (2005) moeten onderzoekers daarom niet alleen de effectieve vertolking analyseren, maar ook aandacht hebben voor het proces dat zich net daarvoor afspeelt.

2.1.3.4. Invloed van de toehoorder

Een derde niet te onderschatten groep van actoren binnen de evaluatie van kwaliteit, zijn de toehoorders van de tolk. Het relatieve karakter van kwaliteit komt bij deze groep duidelijk naar voren als gevolg van de verscheidenheid aan verwachtingen (Garzone, 2002; Gile, 1983; Grbíc, 2008; Kahane, 2000; Kalina, 2005). Shlesinger maakt binnen dit kader een terechte opmerking: “Do our clients know what’s good for them?” (1997, p. 126) Ook de perceptie van het doelpubliek kan gestuurd worden door een gebrek of, integendeel, juist een overvloed aan kennis over het onderwerp of een bepaalde attitude (Kalina, 2005). Het doelpubliek laat zich overigens leiden door het honorarium dat aan de tolkdienst verbonden is: prijzige en goedkopere diensten zullen niet op dezelfde manier geëvalueerd worden (Gile, 1991a).

Opnieuw variëren de verwachtingen van de toehoorders al naargelang de cultuur en de gewoontes die in die cultuur vervat zitten, wat een bijkomende uitdaging vormt voor de tolk (Kopczyński, 1994). Die moeilijke situatie wordt uitstekend geschetst door Garzone:

The basic problem is that quality is the sum of several different, heterogeneous aspects, some of which involve different subjects — interpreters, clients, users, speakers — each with a different view and perception of quality [...]. There are variations even within the same group: among users, quality expectations tend to vary, but also individual tastes and criteria, often related to socio-linguistic factors. (Garzone 2002, p. 107)

Die verwachtingen worden onder meer gebaseerd op de eigen ervaring met tolken of de gedeelde ondervinding van collega’s en kunnen zowel een positief als een negatief beeld van de vertolking nalaten. Zo merkt Gile (1995a) in zijn studie een discrepantie op tussen de

perspectieven van de tolk enerzijds en zijn toehoorders anderzijds: een tolk kan uitbundig gefeliciteerd worden door zijn cliënt over de tolkprestatie, terwijl hij zelf ontevreden is over het geleverde werk en er zich bovendien van bewust is dat hij enkele inhoudelijke elementen achterwege liet. Dat kan tot grote ergernis leiden bij de tolk. Deze tegenstrijdige evaluatie zou verklaard kunnen worden door van de veronderstelling uit te gaan dat de toehoorders doorgaans niet opmerken wanneer bepaalde informatie foutief wordt weergegeven of eenvoudigweg niet vertaald wordt (Déjean le Féal, 1990; Gile, 1991a, 1995a). Een goede evaluatie staat dus niet noodzakelijk gelijk aan een kwalitatieve vertolking. Dit toont aan dat er wel degelijk een onderscheid gemaakt moet worden tussen de kwaliteit van een vertolking enerzijds en de evaluatie ervan anderzijds. Binnen hetzelfde kader verduidelijkt Shlesinger (1997) dat bepaalde toehoorders in een specifieke situatie een bijna letterlijke vertaling van de brontekst verkiezen, maar dat geldt zeker niet voor elk individu. Hetzelfde gaat op voor de kwaliteitscriteria die op verschillende manieren begrepen kunnen worden al naargelang de verwoording (Garzone, 2002; Iglesias Fernández, 2013). Dat probleem duikt helaas regelmatig op bij de meest gehanteerde methode binnen onderzoek naar kwaliteit, het survey-onderzoek, waarbij gebruik wordt gemaakt van vragenlijsten (Gile, 2003). Om aan deze problematiek zoveel mogelijk te remediëren, zouden bestaande definities verfijnd en criteria duidelijker geformuleerd moeten worden, hoewel Gile (2003) ervan overtuigd is dat variatie deel uitmaakt van het onderzoeksproces. Toch stelde Garzone onderzoekers gerust dat de besproken variatie tot op het moment van haar gepubliceerde werk in 2002 nog geen significant effect had vertoond op de resultaten van eerder uitgevoerde onderzoeken en experimenten.

Gile (1983) richt de aandacht op de grote verscheidenheid aan persoonlijke opvattingen: wat door de ene persoon als een onaangename intonatie of regionaal accent ervaren wordt, hoeft niet noodzakelijk op dezelfde manier beoordeeld te worden door een andere toehoorder van dezelfde spreker. Uitgesproken tics wekken evenmin bij iedereen dezelfde ergernis op. Sommige gebruikers van de tolkdienst hechten meer belang aan andere criteria, zoals de volledigheid van de vertolking (Garzone, 2002). Zelfs de fysieke verschijning van de tolk – niemand ontsnapt aan ‘de eerste indruk’ – of diens gedrag, zowel binnen als buiten de cabine, kunnen de gevoelsmatige ervaring van de toehoorder een bepaalde richting uit sturen (Gile, 1995b). Dat gaat gepaard met de vooroordelen (over zowel de tolk als het hele gebeuren) en emoties waarmee een individu naar een bijeenkomst vertrekt (Garzone, 2002; Grbić, 2008).

Om alle voornoemde redenen meent Bühler (1986) dat ideale kwaliteit enkel mogelijk is in een welbepaalde situatie. Ook Gile (1983) betuigt dat er meer aandacht besteed zou moeten worden aan het resultaat van de vertolking dan aan de eigenlijke tolkprestatie. Kalina (2005) voegt hier zelfs aan toe dat toehoorders moeten aanvaarden dat kwaliteit nooit een garantie kan zijn, maar het blijft wel de opdracht van de tolk om de linguïstische elementen die door de spreker verwaarloosd werden, zoals een correcte grammatica of terminologie, te verbeteren.

2.1.4 Mogelijke beoordelaars van de kwaliteit van een vertolking

Sommige tolken en docenten zijn ervan overtuigd dat ze de kwaliteit van een vertolking kunnen beoordelen op basis van hun intuïtie, die betrouwbaar is vanwege hun professionele ervaring (Kalina, 2005). Aangezien kwaliteit subjectieve eigenschappen heeft, blijft echter de vraag: wie is bevoegd om te beslissen over de opvatting en de invulling van een kwalitatief aanvaardbare tolkprestatie? Zo kan kwaliteit beoordeeld worden vanuit verschillende perspectieven: de docent, de professionele organisatie van geaccrediteerde tolken, de organisator van de conferentie (werkgever), de spreker (zender), de toehoorder (ontvanger), de ervaren collega-tolk en de tolk zelf (Bühler, 1986; Chiaro & Nocella, 2004; Christensen, 2011; Garzone, 2002; Gile, 1983, 1991a). Onderzoekers lijken hierover nog niet tot overeenstemming te zijn gekomen. Zelfs bij de toegepaste termen ontstaat er verwarring. Zo maakt onder meer Garzone (2002) een onderscheid tussen de ‘klant’ (d.i. de werkgever die de tolkdienst verzekert en geschikte tolken inzet voor de opdracht) en de ‘gebruiker’ van de dienst (in deze masterproef beschreven als ‘toehoorder’ of ‘publiek’), terwijl Kurz (2003) beide termen door elkaar gebruikt. Christensen (2011) spreekt uitsluitend over de ‘gebruiker’ om de toehoorder te benoemen, in tegenstelling tot Shlesinger (1997) en Chiaro en Nocella (2004) die voor diezelfde persoon de term ‘klant’ verkiezen.

2.1.4.1. De docent, de professionele organisatie en de werkgever

Niet alle actoren bevinden zich in de juiste context om een vertolking te beoordelen. Bühler (1986) vreest dat docenten en organisaties van geaccrediteerde tolken tijdens de opleiding mogelijk te weinig rekening houden met de verwachtingen van de toehoorders. Werkgevers zouden volgens Bühler (1986) beschouwd kunnen worden als competente beoordelaars, al laten zij zich sterk beïnvloeden door budgettaire factoren: de prijs blijft belangrijker dan de kwaliteit die de tolk moet waarborgen (Moser-Mercer, 1996). Deze stelling

druist in tegen het standpunt van Seleskovitch (1986), die Bühler (1986) erop wijst dat werkgevers slechts incidenteel tolken moeten voorzien en de aanstelling van een competente tolk meestal overlaten aan secretarissen. Daarenboven kunnen de werkgevers noch tot de groep van sprekers, noch tot de groep van tolken gerekend worden, met als gevolg dat ze dikwijls een ontoereikende kennis hebben over het onderwerp (Gile, 1991a).

2.1.4.2. De spreker en de toehoorder

De spreker, de zender van de brontekst, spreekt bij een simultane vertolking op hetzelfde moment als de tolk en kan bijgevolg de eigenlijke doeltekst niet horen – en dus ook niet evalueren (Gile, 1991a, 1995b). Enkel wanneer de vertolking in de consecutieve tolkmodus verloopt, de spreker de doeltaal voldoende beheerst en over de nodige achtergrondkennis beschikt, acht Gile (1991a) hem gekwalificeerd als volwaardige beoordelaar. Hetzelfde geldt voor de toehoorder, die als beoordelaar heel wat controversie uitlokt bij onderzoekers. Seleskovitch (1986) gelooft dat toehoorders zich wel degelijk in een geschikte positie bevinden om kwaliteit te evalueren, op voorwaarde dat zij niet vooraf werden ingelicht over hun rol als evaluator. Volgens de auteur zou de kans immers kunnen bestaan dat zij zich te veel concentreren op de onderzochte variabelen in plaats van op een gangbare manier gebruik te maken van de tolkdienst. Daarom stelt zij voor om een vertolking slechts te evalueren ná de prestatie. Op die manier kunnen de toehoorders optreden als *pure customers*; ze krijgen dus louter de doeltekst te horen en kunnen bijgevolg geen vergelijking maken met de oorspronkelijke toespraak. Bühler (1986) stemt in met deze stelling, al voegt ze toe dat toehoorders voor zichzelf een compromis moeten vinden tussen moeilijk haalbare verwachtingen en datgene waar ze werkelijk nood aan hebben. Drie paragrafen verder echter, stelt diezelfde auteur in haar studie de toehoorder als mogelijke beoordelaar in vraag: “Can a delegate reasonably be expected to judge the quality of interpretation since he/she does not know the original?” (Bühler, 1986, p. 231) Ook verscheidene meer recente studies beweren het tegendeel: de toehoorder zou inderdaad het meest bevoegd zijn om een vergelijking te maken tussen de bron- en de doeltekst aangezien hij beide versies te horen krijgt, maar hij zou geen tolk hoeven in te schakelen indien hij ook in staat was de oorspronkelijke boodschap te begrijpen (Garzone, 2002, 2003; Iglesias Fernández, 2013; Kahane, 2000). Het begrip van de brontekst is echter een cruciale vereiste om de kwaliteit van een vertolking te mogen beoordelen (Déjean le Féal, 1990). Verder is de toehoorder meestal vertrouwd met het onderwerp, maar niet met de taal. Hierdoor kan hij enkel oordelen over niet-linguïstische en prosodische

kenmerken, zoals de vlotheid van de voordracht of de intonatie van de tolk, maar niet over inhoudelijke elementen (Bühler, 1986; Gile, 1995b). Om die reden trekken Chiaro en Nocella (2004) de resultaten van eerder onderzoek uitgevoerd bij toehoorders in twijfel. Kalina (2005) en Déjean le Féal (1990) merken op dat het voor de toehoorder haast onmogelijk is om te achterhalen of bepaalde (storende) elementen te wijten zijn aan de vaardigheden van de spreker of van de tolk. Een enkele keer doet zich een situatie voor waarbij de toehoorder weliswaar erkent dat de tolk goed werk geleverd heeft, maar hem of haar toch een mindere beoordeling van de vertolking toeschrijft, vanwege een desinteresse in het vertolkte onderwerp (Gile, 1991a).

Opmerkelijk is de soms lage betrouwbaarheid van de beoordelingen van toehoorders: tijdens een (experimenteel) onderzoek komt de rangschikking van kwaliteitscriteria niet altijd overeen met de parameters die toehoorders daadwerkelijk toepassen om een vertolking in een authentieke werksituatie te evalueren (Garzone, 2003). Pöchhacker beschrijft het probleem als volgt:

We seem to know what the product should be like, but we are less sure about a method for establishing what a particular product is like in a given situation. Quite obviously, researchers, teachers and trainees need a method for looking at the product. (Pöchhacker, 1994, p. 235)

Een aangename stem kan bijvoorbeeld als ‘weinig belangrijk’ aangeduid worden op een vragenlijst, maar kan wel degelijk een heuse impact hebben op de perceptie van een vertolking (Garzone, 2002; Gile, 2003). Wie luistert er graag meerdere uren naar een onaangename stem? Omgekeerd kunnen een goede stemkwaliteit en intonatie de indruk wekken dat de vertolking van een zeer goede kwaliteit is, terwijl de inhoudelijke informatie verscheidene tekortkomingen vertoont (Déjean le Féal, 1990). Een tolk hoort bijgevolg altijd rekening te houden met de voortdurende wisselwerking tussen de “content”, de inhoud van de boodschap, en de “package”, de manier waarop die boodschap gepresenteerd wordt (Gile, 1991a, p. 194). Om dat contrast te minimaliseren, sluit Gile (2003) zich aan bij het eerdere voorstel van Seleskovitch (1986): wanneer toehoorders te horen krijgen dat ze deel uitmaken van een experiment, besteden ze mogelijk meer aandacht aan wat daadwerkelijk onderzocht wordt en gaat bijgevolg alle spontaneïteit verloren. Onderzoek waarbij een authentieke toespraak gemanipuleerd wordt, zou daarom betrouwbaardere resultaten kunnen opleveren (Gile, 2003).

In tegenstelling tot de bewering van Gile (1983, 1991a), luisteren toehoorders niet uitsluitend naar de spreker óf de tolk in een simultane tolksetting. Onderzoek van Amini et al. (2015) wijst uit dat bij een steekproef van 256 toehoorders slechts 19,6% voortdurend gebruik maakt van de tolkdienst en hetzelfde percentage nooit een beroep op de tolk doet, maar zich volledig tot de spreker richt. De helft van de respondenten gaf tijdens het onderzoek aan af en toe naar de tolk te luisteren uit pure interesse of nieuwsgierigheid, ook al hadden ze de boodschap van de spreker begrepen. Niet alle toehoorders handelen echter op die manier uit positieve redenen: sommigen menen dat tolken niet in staat zijn om een tekst adequaat over te brengen en beschouwen hen als een onbetrouwbare bron die foutieve inhoudelijke elementen kan verkondigen wegens een gebrek aan kennis. De spreker blijft in hun ogen de wezenlijke bron en de tolk biedt enkel hulp bij een moeilijke zin. (Gile, 1983) Toch mag volgens Gile (1983) het belang van een tolk niet onderschat worden, aangezien hij de tekst duidelijker en begrijpelijker weergeeft dan de spreker, beter articuleert en dankzij een klare taal de boodschap overtuigender kan brengen. De afwisseling tussen tolk en spreker wordt tot op zekere hoogte bepaald door de motivatie van de toehoorders. Zij zijn namelijk niet altijd even geïnteresseerd in de volledige toespraak en selecteren daarom welbepaalde delen van de vertolking. Hierdoor heeft hun evaluatie enkel betrekking op een fractie van de vertolking en wordt de resterende niet-beluisterde prestatie slechts beoordeeld op basis van de “package” (Gile, 1991a, p. 198).

2.1.4.3. De tolk

Ook de tolk wordt over het algemeen niet als een geschikte beoordelaar beschouwd. Kalina (2005) beweert dat tolken niet kunnen instaan voor de beoordeling van de kwaliteit, aangezien zij zelf geen controle kunnen uitoefenen over hun tolkprestatie. Toch zijn de zelfverbeteringen waar tolken bijna haast zonder uitzondering naar teruggrijpen het bewijs dat ze wel degelijk luisteren naar hun eigen reproductie (zie bv. de transcriptie van Gerver, 1975, p. 122). Hoewel de tolk zowel de bron- als de doeltaal beheerst en zich tegelijkertijd als ontvanger én zender profileert, beschikt hij dikwijls over te weinig achtergrondkennis om zijn vertolking integraal te evalueren (Gile, 1991a). Bovendien houdt hij zich te intensief bezig met de activiteit van het tolken en kan hij dus niet voldoende afstand nemen van de inhoud om de eigen prestatie objectief te beoordelen (Garzone, 2003; Gile, 1983). Seleskovitch (1986) oordeelt als een van de weinigen dat professionele tolken van instellingen voldoende competent zijn om de kwaliteit van hun eigen vertolking te evalueren.

2.1.4.4. *De collega-tolk*

Zowel Gile (1983) als Bühler (1986) zijn van mening dat de passieve, ervaren collega-tolk het meest bevoegd is voor een bevredigende kwaliteitscontrole. Ook al worden tolken niet graag beoordeeld door collega's vanwege de onderlinge concurrentie, toch kunnen deze laatsten, dankzij hun nodige vakkennis, het best nagaan hoe tolken met bepaalde moeilijkheden omgaan en kunnen zij elkaar helpen om deze ongemakken het hoofd te bieden. Gile (1983) waarschuwt dat een controle van de nauwkeurigheid van de beoordeling desondanks noodzakelijk blijft en dat de criteria die worden toegepast door de collega-tolken afgetoetst moeten worden aan de kwaliteitscriteria van de toehoorders, om eventuele discrepanties in beschouwing te nemen. Ondanks de overtuigende argumenten van de auteurs, blijft het voor zowat elke mogelijke beoordelaar een uiterst moeilijke taak om alle elementen van de brontekst te vergelijken met de vertolking, omdat deze opgave te veel inspanning vergt van het kortetermijngeheugen (Gile, 1995a).

Een beoordelaar die bij de collega-tolk aansluit is de tolkstudent, met als enige verschil dat hij de tolkopleiding nog niet afgerond heeft. Toch raadt Gile (1995a) af een beroep te doen op studenten om de kwaliteit van een vertolking te evalueren. In zijn experiment werden twaalf studenten, waaronder de spreker, gevraagd foutieve elementen uit de vertolking te halen. Hij kwam tot de conclusie dat geen enkele deelnemer erin slaagde alle fouten waar te nemen in de vertolking. Sterker nog, de studenten gaven kennis van zogenaamde 'valse fouten', elementen die als incorrect werden aangeduid, maar in werkelijkheid correct werden vertolkt. Ondanks de door de studenten ontwaarde fouten, gaf een grote meerderheid (83%) een zeer goede score. Volgens Gile (1995a) zou dit een gevolg kunnen zijn van de milde houding die studenten aannemen tegenover de beoordeling van een medestudent, omdat zij zich mogelijk laten beïnvloeden door de klassikale sfeer en omdat enkele studenten beseffen dat het risico bestaat dat ze nooit als tolk zullen afstuderen. Gile (1995a) maakt echter veelvuldig gebruik van de term 'fouten', een term die met de nodige voorzichtigheid benaderd moet worden. De term 'fout' legt namelijk een normatief karakter op en is bovendien zeer subjectief, aangezien niet iedereen dezelfde taalcomponenten als foutief ervaart. Ook dient opgemerkt te worden dat meer dan de helft van de studenten in Gile's steekproef boven de dertig jaar oud waren en al professionele ervaring hadden opgedaan, in de vertaalsector of in een ander domein.

2.2 Kwaliteitscriteria

In tegenstelling tot het gebrek aan een eenduidige definitie over de kwaliteit van een vertolking, bestaat er wel een algemene consensus over de kwaliteitscriteria die worden toegepast bij de evaluatie van een vertolking (Chiaro & Nocella, 2004; Gile, 1990; Déjean le Féal, 1990; Kahane, 2000). De kwaliteitscriteria opgesteld door Bühler (1986), de pionier binnen het onderzoek naar de kwaliteit van een vertolking vanuit het perspectief van de tolk, vormden een belangrijke inspiratiebron voor vele latere onderzoeken en replicastudies. Moser-Mercer (2008) wijst er echter op dat de meeste survey-onderzoeken geen pilotstudie uitvoerden. Die studie is nochtans noodzakelijk om na te gaan of het onderzoek en de beoogde steekproefgrootte realistische doelstellingen zijn en om de geldigheid en de betrouwbaarheid van het onderzoek te waarborgen. Het probleem dat zich voordeed tijdens de pilotstudie zal namelijk naar alle waarschijnlijkheid ook opduiken in het eigenlijke experiment of survey-onderzoek en moet dus vooraf zoveel mogelijk vermeden worden (Moser-Mercer, 2008). Ook Pöchhacker (2005) heeft bedenkingen bij de methodologie die eerdere onderzoeken toepasten: kwaliteitscriteria rangschikken in plaats van een cijferschaal toe te passen blijft volgens hem een valabele onderzoeksmethode, maar vergt behoorlijk wat concentratie en vereist veel inspanning van het werkgeheugen, dat extra zwaar belast wordt wegens de vele vergelijkingen die gemaakt moeten worden alvorens de criteria naar behoren te kunnen ordenen. Daarom pleit Pöchhacker (2005) ervoor beide onderzoeksmethodes te combineren, wanneer er gepolst wordt naar dergelijke parameters.

De verwachtingen en voorkeuren van zowel tolken als toehoorders werden in de afgelopen dertig jaar hoe langer hoe duidelijker, dankzij de resultaten van vragenlijsten die opgesteld waren op basis van kwaliteitscriteria (Pöchhacker, 2005). Professionele tolken nemen hierbij voor zowat elk criterium een strengere houding aan dan hun toehoorders (Kurz, 1993). De bovenstaande stelling druist evenwel in tegen het standpunt van Garzone (2002), die aangeeft dat onderzoekers onderling nog steeds niet tot overeenstemming zijn gekomen over dit onderwerp. Enkel over de nauwkeurigheid, de totstandkoming van succesvolle communicatie en de inhoudelijke overeenkomst met de brontekst zou eensgezindheid bestaan. Het laatste element werd volgens Gile (1995a) immers vaak als eerste criterium aangeleerd aan studenten, aangevuld met een logische samenhang, correct gebruik van grammatica, de volledigheid van de vertolking en de vlotheid van de voordracht.

De volgende paragrafen gaan dieper in op linguïstische criteria, die grotendeels overeenkomen met die van Bühler (1986), en niet-linguïstische criteria. Verschillende studies legden zich toe op uiteenlopende criteria voor simultane vertolkingen (zie hoger). Bühler (1986) voerde onderzoek bij professionele AIC-tolken en nam hiervoor acht linguïstische en negen niet-linguïstische parameters in aanmerking. Zowel Pöchhacker en Zwischenberger (2010) als Chiaro en Nocella (2004) namen alle criteria die Bühler identificeerde als linguïstisch integraal over. Pöchhacker en Zwischenberger (2010) voegden hier nog enkele criteria aan toe, zonder onderscheid te maken tussen linguïstisch en niet-linguïstisch. Ook Kurz (1993) nam, op één uitzondering na, alle linguïstische parameters over. Daarnaast vergeleek Kurz (1993) haar bevindingen met de evaluaties van drie verschillende groepen respondenten, verzameld op drie soorten bijeenkomsten: die van de afgevaardigden van de Raad van Europa, van artsen en van ingenieurs. Chiaro en Nocella (2004) verkozen, net als Bühler, een onderverdeling in linguïstische en niet-linguïstische kwaliteitscriteria. Opvallend is dat deze auteurs, in tegenstelling tot Bühler, het criterium *aangename stem* als een linguïstisch criterium beschouwen. Dit wijst op de vage scheidingslijn tussen linguïstische en niet-linguïstische criteria. De meeste onderzoekers echter (Amini et al., 2015; Christensen, 2011; Garzone, 2003; Moser, 1995), richtten hun aandacht op de toehoorders, eventueel in vergelijking met de sprekers (Kopczyński, 1994). Christensen (2011) voerde haar onderzoek uit in een juridische context, waarbij haar steekproef bestond uit alle aanwezigen in de rechtbank. Bühler (1986), Kurz (1993), Pöchhacker en Zwischenberger (2010), Amini et al. (2015) en Christensen (2011) vroegen in hun survey-onderzoek de respondenten om het belang van de criteria aan te duiden op basis van een cijferschaal, waarbij het hoogste cijfer het standpunt ‘heel belangrijk’ weergaf. Moser (1995) liet de puntenschaal achterwege en richtte zich rechtstreeks op een schaal van ‘onbelangrijk’ naar ‘heel belangrijk’. In de studies van Chiaro en Nocella (2004), Garzone (2003) en Kopczyński (1994) moesten de respondenten het gegeven aantal kwaliteitscriteria rangschikken.

2.2.1 Linguïstische criteria

Zoals blijkt uit de studie van Bühler (1986), kunnen linguïstische en niet-linguïstische criteria makkelijk van elkaar onderscheiden worden op basis van hun directe verband met de taal. Zo vormen de semantiek, woordenschat en grammatica van een taal belangrijke componenten binnen de linguïstiek (Aitchison, 2010). Bij de evaluatie van een vertolking worden deze criteria

over het algemeen als belangrijker ervaren dan de vorm van een vertolking, omdat er meer belang wordt gehecht aan inhoudelijke kenmerken (Kalina, 2002).

2.2.1.1. Inhoudelijke overeenkomst met de brontekst

Over één criterium zijn haast alle onderzoekers het eens: de oorspronkelijke boodschap moet zo getrouw en nauwkeurig mogelijk weergegeven worden. Deze parameter belandt in vrijwel elke studie op de eerste plaats in de rangschikking en dit zowel bij tolken als bij toehoorders, ongeacht het type bijeenkomst of de functie van de toehoorder, zij het dokter, ingenieur of afgevaardigde (Amini et al., 2015; Bühler, 1986; Chiaro & Nocella, 2004; Garzone, 2003; Kahane, 2000; Kurz, 1993; Pöchhacker & Zwischenberger, 2010). In Bühlers studie (1986), een van de eerste onderzoeken naar de kwaliteit van een vertolking vanuit het perspectief van de tolken, beschouwt 96% van de participerende AIIC-tolken de inhoudelijke overeenkomst met de brontekst als belangrijkste criterium. Ook in de replicastudie van Pöchhacker en Zwischenberger in 2010 – met een steekproef van wel vijftien maal groter – deelt 88,3% van de professionele tolken diezelfde mening. Toch blijven Chiaro en Nocella (2004) sceptisch over de genoemde percentages; Bühlers resultaten (1986) vertonen voor de inhoudelijke overeenkomst immers geen significante verschillen met de overige kwaliteitscriteria. De meeste linguïstische criteria worden als ‘belangrijk’ of ‘heel belangrijk’ gewaardeerd. En toch blijkt ook uit de studie van Chiaro en Nocella (2004), waarbij de bijna driehonderd respondenten uit de vijf continenten gevraagd werden om de kwaliteitscriteria te rangschikken, dat juist deze parameter bovenaan geplaatst wordt. Ook de sprekers vertrouwen op een nauwkeurige weergave van de oorspronkelijke tekst en ergeren zich wanneer de tolk deze taak niet naar behoren volbrengt (Kopczyński, 1994). Slechts twee onderzoekers komen tot andere bevindingen. Christensen (2011) concludeert dat de toehoorders in haar studie de volledigheid van de vertolking verkiezen boven de verstrekking van essentiële informatie. Kurz (1993) merkt een verschil op bij de toehoorders onderling: ook al ligt de evaluatie van de kwaliteit door de tolken, de artsen en de ingenieurs in dezelfde lijn als de bevindingen van de andere besproken onderzoeken, de afgevaardigden van de Raad van Europa beschouwen een correct gebruik van terminologie als het belangrijkste criterium. De inhoudelijke overeenkomst met de brontekst schuift binnen deze groep, als criterium, naar de vierde plaats.

Volgens Gile (1983) is de inhoudelijke overeenkomst met de brontekst het makkelijkst te evalueren criterium. Kahane (2000) deelt deze mening niet: zoals eerder vermeld, beperkt de kennis van de toehoorder zich doorgaans tot slechts één van de twee gehanteerde talen. En hoe, vraagt de auteur zich af, zou deze toehoorder dan in staat kunnen zijn om een volledige tolkprestatie te beoordelen, wanneer de evaluatie van de vertolking op basis van slechts één kwaliteitscriterium al een hele uitdaging vormt?

Het criterium *inhoudelijke overeenkomst met de brontekst* mag niet verward worden met een letterlijke vertaling. De weergave van het idee en de betekenis die achter de boodschap schuilt, wordt veelal belangrijker gevonden dan een woord-voor-woordvertaling – een opmerkelijk onderscheid tussen tolken en vertalen (Moser, 1995). Dit is des te meer het geval bij bijeenkomsten waar spontaneïteit en interactie centraal staan. Kahane (2000) meent echter dat toehoorders eerder een letterlijke vertaling verkiezen, een waarbij zelfs de ‘fouten’ van de spreker trouw worden weergegeven, dan een vertolking die de boodschap globaal weergeeft. De sprekers daarentegen zouden een tolerante houding aannemen tegenover de verbeteringen door de tolk. Ook de tolk in de juridische context, zoals beschreven in de studie van Christensen (2011), gelooft dat zijn taak erin bestaat een letterlijke vertaling te bezorgen.

2.2.1.2. *Omissies*

Enkele gedetailleerde studies (Barik, 1994; Korpala, 2012) behandelen omissies als een apart criterium. Het verband met de inhoudelijke overeenkomst met de brontekst en de volledigheid van de vertolking – criteria waarnaar veel aandacht uitgaat tijdens de beoordeling van de kwaliteit van een vertolking (zie hoger) – is echter moeilijk weg te denken. Barik (1994, p. 122-124) beschrijft omissies als elementen die aanwezig zijn in de brontekst, maar niet in de doelttekst, en maakt een onderscheid tussen vier verschillende weglatingen: “skipping omission”, “comprehension omission”, “delay omission” en “compounding omission”. Het eerste type weglating wordt gezien als een woord of een kort zinsdeel – vaak gaat het om een kwalificerend adjectief, een weggelaten voorzetsel of conjunctie – dat niet vertaald wordt door de tolk, maar geen verandering teweegbrengt in de grammaticale structuur en bijgevolg geen invloed heeft op de betekenis van de boodschap. Bij simultaan tolken wordt deze ‘fout’ als aanvaardbaar beschouwd. “Comprehension omission” is het gevolg van een verkeerd begrip door de tolk. Aangezien deze weglatingen betrekking hebben op grotere tekstdelen, gaat de betekenis van de brontekst hier toch wel deels verloren. Bij de “delay omission” loopt de tolk

achter op de spreker en kan hij ervoor kiezen om te wachten op een nieuw betekenisvol element of hij kan bepaalde zinsdelen niet vertolken met als doel de tekst verder weer op te pikken en zo weer aan te sluiten bij het ritme van de spreker. De laatst genoemde weglating, de “compounding omission”, beschrijft de weglating van meerdere (bij)zinnen en de samenstelling van nieuwe zinnen op basis van de resterende zinslementen, met een aanzienlijk betekenisverschil tot gevolg. Verder kunnen ook nog omissies van lidwoorden, overbodige of onvertaaltbare zinslementen of het nevenschikkend voegwoord ‘en’ voorkomen. Barik (1994) toont aan dat deze tendens vaker voorkomt bij beginners of tweetalige deelnemers zonder tolkopleiding dan bij professionele en meer ervaren tolken, omdat deze laatsten meer getraind zijn om te luisteren.

Korpala (2012) stelt dat niet alle omissies als even zwaarwichtig worden ervaren; volgens hem hangt de ernst van de weglatingen in hoge mate af van de subjectieve visie van de auteur. Bartłomiejczyk (2007) oordeelt bijvoorbeeld dat de tolk minder zwaar in de fout gaat wanneer hij een voorzetsel weglaat, dan wanneer hij het verandert. Barik (1994), van zijn kant, heeft geen bezwaar tegen de omissie van connectoren of lidwoorden. Volgens Pöchhacker (1994) bestaat de voornaamste opgave erin na te gaan of de toehoorders de boodschap begrepen hebben. Wanneer aan toehoorders de vraag werd voorgelegd of zij tolken zouden toestaan langere stukken tekst samen te vatten, volgden er haast uitsluitend negatieve reacties (Christensen, 2011; Kopczyński, 1994). Indien de toehoorders toch dienden aan te duiden in welke situatie ze de meeste omissies in de vertolking verwachtten, kwamen onder meer ‘herhaling van dezelfde woorden’, ‘irrelevante informatie’, ‘culturele verwijzingen’ en ‘subjectieve beoordelingen van de spreker’ naar voren (Korpala, 2012, p. 107).

Opnieuw wordt er, weliswaar in minder recente studies, de zeer subjectieve term ‘fout’ gehanteerd (zie o.a. Barik, 1994), maar wanneer mag een handeling onder ‘fout’ gecategoriseerd worden? Enkele auteurs (vb. Bartłomiejczyk, 2007; Garzone, 2002; Gile, 2003; Hale & Napier, 2013; Korpala, 2012; Napier, 2004) spreken in specifieke omstandigheden namelijk van een welbewuste keuze van de tolk om een beroep te doen op omissies. Alleen dan zou de kwaliteit van de vertolking gegarandeerd kunnen worden. Zo verklaart Jones (2002, p. 124-125): “There are cases when the interpreter is unfortunately not in a position to provide a totally complete and accurate interpretation. [...] [T]he interpreter omits in order to preserve as much of the essential message as possible.” Simultaan tolken overwegen bijvoorbeeld nu en dan bepaalde inhoudelijke of vormelijke elementen weg te laten wegens tijdgebrek of wanneer

ze dat stukje informatie als irrelevant beschouwen en zo overtollige afwijkingen wensen te vermijden (Bartłomiejczyk, 2007; Korpala, 2012). Ook kunnen omissies hulp bieden om de coherentie en de nauwkeurige weergave van de oorspronkelijke boodschap te bewaren en zo de communicatie vlotter te laten verlopen (Hale & Napier, 2013; Korpala, 2012; Napier, 2004; Riccardi, 2003). Zeker wanneer de informatie in de oren van de tolken foutief klinkt, verkiezen ze deze achterwege te laten. Enkelen beweren zelfs dat ze de brontekst onmogelijk vertaald zouden kunnen hebben, indien ze zich verplicht voelden om ieder element om te zetten naar de doeltaal (Korpala, 2012). Elke tolk beslist echter zelf over de informatie die hij wil vermelden, met als gevolg dat er een opmerkelijke variatie kan optreden binnen de vertolking van eenzelfde tekst (Gile, 2003).

Bij een consecutieve vertolking kunnen tolken er eveneens voor kiezen om niet alle informatie te noteren. Gile (2001) legt uit dat woorden waarvan de schriftelijke neerslag tijdrovend is bijvoorbeeld makkelijker weggelaten worden om zo te vermijden dat de tolk te veel achterloopt op de spreker. Deze techniek duikt niet op bij simultaan tolken aangezien de tolk diezelfde taaluitingen aan een hoog tempo kan herformuleren. Op dezelfde manier zullen bepaalde opsommingen bij consecutief tolken vaker onvolledig zijn. In zijn studie concludeert Gile (2001) dat de simultane vertolkingen veelal nauwkeuriger zijn dan de consecutieve. De resultaten van Russells onderzoek (2005) spreken die bevinding echter tegen: 95% van de oorspronkelijke boodschap wordt correct vertaald bij consecutief tolken, tegenover 87% bij simultaan tolken. De vraag blijft in welke mate omissies mogen voorkomen tijdens een consecutieve vertolking: aangezien de consecutieve vertolking per definitie volgt op de oorspronkelijke tekst en het geheel dus meer tijd in beslag neemt, is het zeker belangrijk om de nodige aandacht aan omissies te besteden. In tegenstelling tot de bevindingen van Christensen (2011) en Kopczyński (1994), zou volgens Gile (1983) een synthese van de brontekst zelfs de voorkeur krijgen, zeker wanneer de toehoorders in staat zijn de spreker grotendeels te begrijpen.

Weglatingen kunnen ook het gevolg zijn van externe factoren (zie 2.1.3. voor bijkomende voorbeelden) en kunnen dus onbewust voorkomen tijdens de vertolking. De lengte van de vertolking speelt bijvoorbeeld een belangrijke rol: naar het einde toe laat de tolk steeds meer elementen weg (Barik, 1994), mogelijk door mentale moeheid. Tegen de verwachtingen in, speelt het spreektempo echter geen belangrijke rol bij het aantal omissies (Korpala, 2012). Zo lijkt een laag spreektempo niet te leiden tot omissies om pragmatische redenen en wordt

irrelevante informatie niet noodzakelijk geschrapt vanwege een snelle voordracht door de spreker. Bij een vergelijking tussen tolkstudenten en professionele tolken blijken de vertolkingen van de professionele tolken iets minder inhoudelijke omissies te vertonen bij een hoog spreektempo vanwege hun ervaring en deskundigheid, al is er opnieuw geen sprake van een significant verschil. Aangezien het experiment van Korpál (2012) slechts een steekproef van zeventien respondenten behelsde en de onnatuurlijke, experimentele context mogelijk een invloed zou kunnen hebben uitgeoefend op het aantal omissies in vergelijking met een authentieke tolksetting, zou replicatieonderzoek naar omissies aangewezen zijn.

2.2.1.3. Toevoegingen

Ook toevoegingen aan de brontekst kunnen leiden tot een afwijking in de inhoudelijke overeenkomst met de brontekst en worden volgens Barik (1994) beschouwd als ‘fouten in de vertaling’. Opnieuw maakt Barik (1994, p. 125-127) een onderscheid tussen verschillende soorten toevoegingen: “qualifier addition”, “elaboration addition”, “relationship addition” en “closure addition”. Zoals de naam reeds aangeeft, staat een “qualifier addition” voor het bijvoegen van een bepaling. De “elaboration addition” zou in theorie kunnen bijdragen tot een beter begrip aangezien er bijkomende uitleg gegeven wordt bij een bepaald zinsdeel, maar ook deze toevoeging wordt volgens Barik (1994) niet algemeen aanvaard. Het derde type wordt volledig afgekeurd, omdat de tolk door een verandering van het verbindingswoord een andere relatie tussen de zinnen creëert. De “closure addition” wordt gebruikt om de zin af te ronden, wanneer de tolk een tekstdeel bijvoorbeeld niet volkomen begrepen heeft. Verder zou de bijkomstige vermelding van het voegwoord ‘en’, een voorzetsel of commentaar door de toehoorders niet op prijs gesteld worden. Enkel wanneer het toegevoegde element geen afbreuk doet aan de betekenis van de oorspronkelijke boodschap, zou er geen bezwaar tegen zijn. Barik (1994) toont aan dat deze toevoegingen in mindere mate optreden wanneer de tolk goed voorbereid was op de brontekst. De tolken geven in bepaalde omstandigheden de voorkeur aan toevoegingen, omdat ze kunnen bijdragen tot een logische samenhang van de boodschap (Gile, 2003; Riccardi, 2003).

2.2.1.4. Substituties

Barik (1994) beschrijft substituties of vervangingen als een combinatie van een omissie en een toevoeging, die betrekking kunnen hebben op één woord of op een volledige bijzin van de

spreker. Een voegwoord komt bijvoorbeeld soms in een andere bijzin te staan dan vermeld door de spreker. Deze verschuiving zou ondanks alles het begrip van de toehoorder niet verstoren, aangezien hij makkelijk zelf tot de conclusie kan komen dat het om een vergissing ging. Dezelfde redenering gaat op voor andere onbelangrijke substituties die de boodschap van de brontekst niet aanpassen, zoals het gebruik van een meervoudig in plaats van een enkelvoudig onderwerp. Foutieve vertalingen van lexicale elementen kunnen een vreemde vertaling opleveren en slechts een verwaarloosbaar betekenisverschil teweegbrengen, maar wanneer de tolk naar een homoniem grijpt, kan de semantiek geheel verloren gaan. Ten slotte kan ook erbij verzonden informatie of een verkeerd begrip van de oorspronkelijke boodschap leiden tot niet te onderschatten betekenisverschillen.

2.2.1.5. *Vlotheid van de voordracht*

Bühler (1986) en Pöchhacker en Zwischenberger (2010) tonen aan dat tolken veel belang hechten aan de vlotheid van de vertolking. In beide studies wordt het criterium namelijk op de derde plaats gerangschikt. Al naargelang de studie, krijgt de parameter *vlotheid van de voordracht* een andere score van de toehoorders: tussen de lijst van elf toegepaste criteria krijgt het criterium bij de studie van Amini et al. (2015) een derde plaats, de toehoorders in de studie van Chiaro en Nocella (2004) rangschikken het als vierde belangrijkste criterium binnen de negen linguïstische criteria en bij Kurz (1993) belandt het zelfs op de vijfde plaats. Toch besteden toehoorders heel wat aandacht aan deze parameter (Christensen, 2011). Zo krijgen tolken met een meer aarzelende vertolking een slechtere beoordeling dan hun collega's, zelfs al valt er op de inhoud van de vertolking verder niets aan te merken. Een aarzelende tolkprestatie overtuigt minder en is lastiger te vatten (Garzone, 2003). Met name de afgevaardigden van de Raad van Europa zijn hier behoorlijk streng op in vergelijking met ingenieurs en artsen, omdat zij op internationale congressen meer spontaneïteit verlangen (Kurz, 1993).

De vlotheid van de voordracht wordt beïnvloed door pauzes en zelfverbeteringen. Opvullingen als “ums” en “ahs” werden door ongeveer 34% van de toehoorders in de studie van Moser (1995) als behoorlijk irritant ervaren en bijna 41% stoorde er zich zelfs mateloos aan. Ook lange ingelaste pauzes die leiden tot het achterlopen van de tolk op de spreker tijdens een simultane vertolking kan (voornamelijk vrouwelijke) toehoorders behoorlijk ergeren. Zelfverbeteringen daarentegen, d.i. wanneer de tolk zichzelf verbetert terwijl hij reeds een (deel van een) woord of een zin gedeeltelijk heeft uitgesproken, worden door de toehoorders volledig

aanvaard (Christensen, 2011). Sprekers en toehoorders – hoewel deze laatsten in iets mindere mate – staan zelfs verbeteringen van de tolk in de doeltekst toe; de fouten van de spreker herhalen wordt daarentegen niet erg op prijs gesteld (Kahane, 2000; Kopczyński, 1994; Marrone, 1993).

2.2.1.6. *Correct gebruik van terminologie*

Chiaro en Nocella (2004) geven aan dat de rangschikking voor een *correct gebruik van terminologie* lager ligt bij de toehoorders dan bij de tolken. De studies van Amini et al. (2015), Bühler (1986), Pöchhacker en Zwischenberger (2010) en Kurz (1993) komen tot een ander besluit: zowel tolken, toehoorders als sprekers keuren een foutief gebruik van terminologie af. Opnieuw levert Seleskovitch (1986) commentaar op de verwoording van Bühler (1986): een ‘correcte terminologie’ kan op uiteenlopende manieren geïnterpreteerd worden. Zo zijn Engelse technische termen vaak begrijpelijker dan nieuwe Franse samenstellingen.

Over het algemeen wordt een foutief gehanteerde terminologie als ontzettend ergerlijk ervaren door zowel sprekers als toehoorders (Kopczyński, 1994; Marrone, 1993). Kurz (1993) stelt vast dat voor de afgevaardigden van de Raad van Europa een correct gebruik van grammatica zelfs als het belangrijkste criterium geldt en ook bij de ingenieurs belandt het op de tweede plaats. De overhandiging van informatieve documenten voorafgaand aan de internationale conferentie zou aan de basis liggen van de strenge houding van de afgevaardigden. Omdat zij de informatie al eens hebben kunnen doornemen, verwachten ze ook minimaal een overeenstemmende terminologie. Toch is Pöchhacker (2005) ervan overtuigd dat de verwachtingen van de afgevaardigden deels bepaald worden door het onderwerp van de conferentie. Ook de deelnemers aan technische bijeenkomsten zijn vertrouwd met een bepaald jargon en vertrouwen erop dat tolken daar gebruik van maken (Kurz, 1993).

2.2.1.7. *Correct gebruik van grammatica*

Correcte grammaticale constructies worden door de tolken hoger aangeschreven dan door de toehoorders (Bühler, 1986; Chiaro & Nocella, 2004; Kahane, 2000; Kurz, 1993; Pöchhacker & Zwischenberger, 2010). Volgens Kahane (2000) houden tolken meer vast aan grammatica, de tongval en de stem en zijn ze doorgaans veeleisender voor zichzelf. Seleskovitch (1986) wijst

er echter op dat de parameter *correct gebruik van grammatica* in Böhlers studie (1986) – en dus ook in andere onderzoeken – duidelijker geformuleerd dient te worden:

What is *correct grammatical usage* and how significant are minor grammatical errors in an interpretation. [...] Errors of gender, slight mispronunciations and incorrect association of words in a flowing, coherent speech are not even noticed by those who are listening [...]. (Seleskovitch, 1986, p. 236)

Toehoorders lijken toch niet ongevoelig te zijn voor grammaticale fouten, ook al brengen deze niet noodzakelijk een communicatiestoornis teweeg (Chiaro & Nocella, 2004; Christensen, 2011; Kurz, 1993). Uit de studie van Amini et al. (2015) kan zelfs afgeleid worden dat toehoorders binnen de elf opgesomde criteria *correct gebruik van grammatica* een vijfde plaats toekennen. Zeker binnen de functies van de toehoorders onderling, zijn er opvallende verschillen waar te nemen: de afgevaardigden van de Raad van Europa besteden relatief weinig aandacht aan een correct gebruik van grammatica en de ingenieurs plaatsten het criterium zelfs volledig achteraan in de lijst (Kurz, 1993).

2.2.1.8. *Tongval*

Een sterk gekleurd regionaal of vreemd accent blijkt niet als storend ervaren te worden. Hoewel Kahane (2000) beweert dat tolken ook tegenover dit expressieve criterium een strengere houding aannemen dan hun publiek, blijkt het criterium *tongval* volledig onderaan de lijst te belanden en dit zowel in de onderzoeken van Chiaro en Nocella (2004) als in die van Pöschhacker en Zwischenberger (2010) en Kurz (1993). Bij Bühler (1986) betreft het een voorlaatste plaats. Hetzelfde geldt voor de toehoorders: meer dan de helft erkent dat zij geen rekening houden met de tongval van de tolk (Amini et al., 2015; Kurz, 1993; Moser, 1995). Enkel wanneer toehoorders als *pure customer* optreden, besteden zij voornamelijk aandacht aan vormelijke kenmerken om de kwaliteit van de vertolking te beoordelen (Bühler, 1986).

2.2.1.9. *Logische samenhang van de vertolking*

Met slechts onbeduidende nuances binnen de verschillende studies, wordt over het algemeen de *logische samenhang van de vertolking* door zowel tolken als toehoorders als tweede belangrijkste criterium naar voren geschoven (Amini et al., 2015; Bühler, 1986; Garzone, 2003;

Kurz, 1993; Pöchhacker & Zwischenberger, 2010). Een deelnemende AIIC-tolk in de studie van Pöchhacker en Zwischenberger (2010, par. 26) maakte hierover echter een terechte opmerking: “Logical cohesion can only be a criterion when the original speech is coherent!” Uit deze opmerking komt opnieuw het relatieve karakter van kwaliteit naar voren.

Bij het onderzoek van Kurz (1993) treedt er variatie op al naargelang de functie van de toehoorder: ingenieurs zijn gewoonlijk toleranter tegenover de logische samenhang van de vertolking, omdat zij gedurende de presentaties visuele ondersteuning krijgen van bijvoorbeeld illustraties en tabellen, en bijgevolg minder inspanningen moeten leveren om de leidraad te volgen. De afgevaardigden van de Raad van Europa geven deze parameter een nóg lagere plaats in de rangschikking. Zij kunnen de cohesie van de boodschap tijdens een internationale vergadering bewaren, gesteund door hun eigen achtergrondkennis en extra geholpen door de documentatie die zij vooraf kregen.

2.2.1.10. *Geschikte stijl*

Een goed evenwicht vinden in het toepassen van de meest geschikte stijl, bijvoorbeeld door een register te kiezen dat niet té formeel is, maar ook niet te informeel, kan, volgens de resultaten van Bühler (1986) en Pöchhacker en Zwischenberger (2010), als een niet al te belangrijke parameter beschouwd worden (*geschikte stijl* komt respectievelijk op een laatste en een zevende plaats te staan binnen de lijst van linguïstische criteria). Toehoorders zouden hier evenmin een probleem van maken en kennen dezelfde plaats toe als de deelnemende AIIC-tolken aan het onderzoek van Bühler (Amini et al., 2015). Christensen (2011) kwam echter tot andere bevindingen: zowel de meerderheid van de toehoorders als de tolken zouden een geschikte stijl heel belangrijk achten, al zou dit te maken kunnen hebben met de juridische tolksetting.

2.2.1.11. *Volledigheid van de vertolking*

Een opvallend resultaat is terug te vinden in de studie van Chiaro en Nocella (2004), waarbij de parameter *volledigheid van een vertolking* naar een tweede plaats stijgt binnen de rangschikking. Zowel de tolken in de studie van Bühler (1986) en Pöchhacker en Zwischenberger (2010) als de deelnemende toehoorders bij het onderzoek van Amini et al. (2015) lijken echter niet erg zwaar te tillen aan de volledigheid van de vertolking: het criterium wordt slechts op een zesde plaats gerangschikt. Ook al komt het criterium bij de studie

van Kurz (1993) op de vierde plaats te staan, artsen en ingenieurs geven opnieuw blijk van een mildere houding. Zij selecteren tijdens zeer informatieve wetenschappelijke vergaderingen namelijk vaak slechts de interessantere delen uit de volledige toespraak en hebben dus meer baat bij een beknopte samenvatting. Ook de afgevaardigden van de Raad van Europa hebben niet voortdurend oor voor de tolk (Kurz, 1993). Aanwezig bij een rechtszaak kennen daarentegen meer waarde toe aan een volledige vertolking (zie hoger). Moser (1995) ondervond eveneens dat bijna de helft van de geïnterviewde respondenten het criterium als zeer belangrijk beschouwden (voornamelijk op bijeenkomsten over technische onderwerpen). Wanneer hen echter de keuze werd geboden, verkozen ze toch de weglating van details en de nadruk op essentiële elementen boven een volledige weergave.

2.2.2 Niet-linguïstische criteria

De meeste niet-linguïstische criteria worden gelinkt aan prosodie, oftewel “*phenomena that involve the acoustic parameters of pitch, duration, and intensity* [oorspronkelijke nadruk]” (Ladd & Cutler, 1983, p. 1). Prosodie is erg belangrijk bij mondelinge voordrachten – en dus ook bij vertolkingen – omdat ze de structuur aangeeft, helpt om nadruk te leggen op bepaalde zinslelementen en de aandacht van de toehoorder vasthoudt (Ahrens, 2005; Cutler, 1983). Toch hebben slechts weinig onderzoekers zich reeds over dit onderwerp gebogen. Met name intonatie genoot maar een matige belangstelling, vanwege de moeilijkheid om het begrip helder te definiëren (Ahrens, 2005; Holub, 2010). Bovendien werden er nog geen eenduidige regels opgesteld voor de transcriptie van prosodische kenmerken (Ahrens, 2005).

Replicatie- en vervolgstudies die voortbouwden op de criteria van Bühler (1986) richtten zich voornamelijk op de linguïstische criteria en lieten veelal de niet-linguïstische criteria achterwege. Zodoende wordt de vergelijking tussen de niet-linguïstische criteria bemoeilijkt (Pöschhacker, 2005). Omdat auteurs verschillende termen toepassen om eenzelfde criterium aan te duiden, is het ook hier onduidelijk waar de kwaliteitscriteria precies op doelen. Zo is de parameter *intonatie* niet terug te vinden in de vragenlijsten van Bühler (1986) en Kurz (1993), want die zit vervat in het kwaliteitscriterium *aangename stem* (Iglesias Fernández, 2013).

2.2.2.1. *Aangename stem*

In tegenstelling tot wat Déjean le Féal (1990) en Marrone (1993) te kennen geven, kan uit de studies van Chiaro en Nocella (2004), Amini et al. (2010), Iglesias Fernández (2013) en Kurz (1993) afgeleid worden dat toehoorders minder zwaar tillen aan een onaangename stem: bij Chiaro en Nocella (2004) krijgt het criterium *aangename stem* een voorlaatste plaats in de rangschikking van de kwaliteitscriteria, in de drie laatste studies een op twee na laatste. Die milde houding tegenover de stem van de tolk is mogelijk deels te danken aan de cultuur waartoe de evaluerende toehoorders behoren: Franstaligen zouden zich bijvoorbeeld minder tolerant opstellen tegenover een onaangename stem dan Engelstalige toehoorders (Gile, 1990). Toch komen deze resultaten niet altijd overeen met de werkelijkheid. Garzone (2003) ontdekte dat toehoorders bij de rangschikking van kwaliteitscriteria verzekeren dat ze in veel mindere mate rekening houden met prosodische en vormelijke kenmerken. Wanneer ze echter verzocht worden om een authentieke vertolking en een gemanipuleerde versie (d.i. met aarzelingen en zonder een bepaald aangename stem) te beoordelen, blijkt de aangepaste vertolking een veel slechtere score te behalen op basis van het criterium *aangename stem*. De evaluatie van kwaliteit wordt dus wel degelijk beïnvloed door prosodische parameters zoals intonatie.

Wat betreft de mening van de tolken spreken de bestaande studies elkaar behoorlijk tegen. Volgens Kurz (1993) en Pöchhacker en Zwischenberger (2010) blijken de tolken niet veel gevoeliger te zijn voor een aangename stem (net als bij de toehoorders, komt bij deze studies de parameter op de voorlaatste en een op twee na laatste plaats te staan); Kahane (2000) beweert het tegendeel. Ook in de studie van Bühler (1986) schuift het criterium een plaats naar boven. Opnieuw blijken er verschillen op te treden tussen de meningen van vrouwelijke en mannelijke tolken: vrouwen zouden meer aandacht besteden aan linguïstische kenmerken in vergelijking met hun mannelijke collega's. Binnen een lijst van tien kwaliteitscriteria, rangschikken de vrouwelijke deelnemers een aangename stem op de vierde plaats, in tegenstelling tot hun mannelijke collega's die het criterium op de zevende plaats rangschikken (Iglesias Fernández, 2013).

2.2.2.2. *Spreektempo*

Het spreektempo van de tolk ligt meestal lager lag dan dat van de spreker van de oorspronkelijke brontekst (Vančura, 2013). Wat betreft de articulatiesnelheid, d.i. in tegenstelling tot de

spreeksnellheid “een meting van het aantal taaleenheden” waarbij “de stille intervallen verwijderd en niet meegerekend” worden (Rosink, van Heeswijk, Kroon, & Schüppert, 2014, p. 197), vergeleek Ahrens (2005) deze articulatiesnelheid van een spreker enerzijds met die van drie tolken anderzijds en concludeerde dat deze snelheid over de hele lijn lager lag bij de tolken.

Een te hoog spreektempo zou bij simultane vertolkingen kunnen leiden tot een kwalitatief slechtere tolkprestatie (Gile, 1983; Kalina, 2005). Toch wijzen de studies van Shlesinger (2003) en Vančura (2013) uit dat tolkprestaties van een sneller voorgedragen brontekst (140 tot meer dan 160 woorden per minuut) beter geëvalueerd worden dan vertolkingen van tragere voordrachten (120 woorden per minuut of minder). Hoewel de spreker bij een lager spreektempo meer pauzes inlast en de simultaan tolken bijgevolg de kans krijgen om hun aandacht evenwichtiger te verdelen over het luister- en tolkproces, zou een hoger spreektempo ervoor zorgen dat het werkgeheugen minder belast wordt. Dat zou een positief effect uitoefenen op de tolkprestatie. De voorkeur voor een tragere of een snellere voordracht varieert bij elk individu. Toen Marrone tijdens zijn onderzoek in 1993 aan de toehoorders vroeg of ze een voorkeur hadden voor een snelle of een trage voordracht, bleken de meningen verdeeld: van de 87 respondenten verkozen 58 toehoorders een traag spreektempo, terwijl 29 respondenten opteerden voor een snelle vertolking.

2.2.2.3. Intonatie

Intonatie is onontbeerlijk voor een goed begrip van de tekst, omdat ze de spreker in staat stelt heldere signalen over te brengen (Holub, 2010). Deze stelling wordt bekrachtigd door Shlesinger (1994), die enkele passages van een tekst luidop liet voorlezen door een tolk en drie andere passages liet tolken. Op die manier verkreeg ze een opname met een standaardintonatie en een versie waar de iets minder natuurlijke intonatie van de tolk te horen was. Nadat de respondenten werden gevraagd om, na het beluisteren van de opnames, negen inhoudelijke vragen te beantwoorden, bleek de getolkte versie 20% minder correcte antwoorden te genereren dan wanneer er geluisterd werd naar de voordracht met standaardintonatie. Een abnormale intonatie zou dus een nefaste invloed kunnen hebben op het begrip van de tekst. Holub (2010) en Pöchhacker en Zwischenberger (2010) bevestigden deze bevinding door aan de proefpersonen een levendige en een eentonige uitvoering van eenzelfde vertolking voor te leggen. Alweer bleken de toehoorders van de eentonige versie slechtere resultaten te behalen op de vragenlijst. Een monotone intonatie leidt overigens over het algemeen tot een negatievere

beoordeling over de vertolking, omdat tolken dan als minder professioneel worden ervaren, ook al zetten ze inhoudelijk een correcte tolkprestatie neer (Holub, 2010).

Ondanks de bovenstaande resultaten menen de toehoorders dat hun aandacht nauwelijks uitgaat naar intonatie bij de beoordeling van de kwaliteit van een vertolking. Het criterium komt namelijk op de voorlaatste plaats terecht in de criterialijst van Amini et al. (2015) en bijna 70% van de door Moser (1995) verzamelde respondenten duidde de parameter als ‘zeer ergerlijk’ of ‘behoorlijk ergerlijk’ aan. Tolken zouden iets meer belang hechten aan een levendige voordracht, maar nemen doorgaans toch een overwegend neutraal standpunt in (Chiaro & Nocella, 2004; Pöchhacker & Zwischenberger, 2010). Binnen hun evaluaties duiken wel gender- en leeftijdsverschillen op. Zo blijken vrouwen gevoeliger te zijn voor een levendige intonatie dan mannen. Tijdens het experiment van Pöchhacker en Zwischenberger (2010) bleken mannen zelfs een lichte voorkeur te hebben voor de monotone versie. Diezelfde tendens is waar te nemen bij de jongere toehoorders (30-47 jaar) en de minder ervaren tolken binnen de studie.

2.2.2.4. Andere niet-linguïstische criteria

In een klein aantal van de geraadpleegde studies worden, naast de bovenstaande linguïstische en niet-linguïstische criteria, ook een aantal andere parameters aangehaald waar zowel toehoorders als tolken belang aan zouden hechten. Toehoorders storen zich bijvoorbeeld enorm aan een gebrek aan simultaneïteit; dit criterium wordt juist als een van de belangrijkste opdrachten van de simultane tolk ervaren en geldt dus ook als een van de grootste verwachtingen van de toehoorders (Amini et al., 2015; Moser, 1995). Tolken schenken er daarentegen niet veel aandacht aan en plaatsen het pas voorlaatste in de rangschikking van kwaliteitscriteria (Pöchhacker & Zwischenberger, 2010). Verder worden een grondige voorbereiding van de vooraf bezorgde documenten en het vermogen om in team te werken als belangrijke parameters beschouwd; zelfvertrouwen en afwezigheid van stress belanden eerder achteraan in de lijst van verwachtingen (Bühler, 1986; Chiaro & Nocella, 2004). Sommige onderzoekers breiden de criteria zelfs uit met vragen over de rol van de tolk (Christensen, 2011; Pöchhacker, 2000; Pöchhacker & Zwischenberger, 2010), maar aangezien deze te ver afwijken van het te onderzoeken onderwerp in deze masterproef, wordt hier niet dieper op ingegaan.

2.3 Zelf- en peerevaluatie als beoordelingsmiddel

Ondanks de algemene consensus over de rijke hoeveelheid studies omtrent kwaliteit, werden volgens Bartłomiejczyk (2007) toch niet alle aspecten voldoende onder de aandacht gebracht. De verwachtingen van de luisteraars én die van de tolken werden wel degelijk aangehoord, maar het antwoord op de vraag hoe tolken naar zichzelf kijken en welke criteria bij zelfevaluatie in aanmerking genomen moeten worden, bleef uit.

In een maatschappij waar studenten steeds meer onder druk staan om goed te presteren en een aanvaardbare kwaliteit te leveren, speelt feedback niet alleen een grote rol bij het leerproces, maar ook bij de ontwikkeling van het verdere (professionele) leven (Davies, 2000; Iaroslavschi, 2011; Sadeghi, Hassani, & Mohammedloo, 2015). Zowel zelf- als peerevaluatie verdienen daarom de nodige aandacht. Twee evaluatietypes worden hierbij onderscheiden: de summatieve en de formatieve evaluatie (Csapó et al., 2012). Terwijl de summatieve evaluatie eerder gericht is op het ‘verleden’ en nagaat over welke kennis de student op dat moment beschikt of hoe ver hij al gevorderd is, legt de formatieve evaluatie nadruk op de ‘toekomst’ en op de werkpunten van de student om zijn niveau op te trekken (Binkley et al., 2012; Csapó et al., 2012). De volgende paragrafen bespreken de formatieve evaluatie, aangezien zowel de beoordelaars als de beoordeelde studenten zich nog in een leerproces bevinden.

2.3.1 Zelfevaluatie

Naar de mening van Russo (1995, p. 75) is zelfevaluatie in het beroep van tolken bedoeld “to encourage SI [simultaneous interpretation] students to analyse their performance, discover their weaknesses and strengths and channel their resources during the training period accordingly”. Ook al ontvangen tolken voldoende opbouwende kritiek of complimenten van zowel hun collega’s als van het publiek, het gebrek aan zelfevaluatie blijft een knelpunt binnen dit beroep (Déjean le Féal, 1990). Zelfevaluatie is nochtans essentieel voor tolkstudenten die hun tolkvaardigheden verder wensen te ontwikkelen en te verbeteren (Déjean le Féal, 1990; Iaroslavschi, 2011). Volgens Déjean le Féal (1990) zouden professionele tolken, om de kwaliteit van hun vertolking te garanderen, hun eigen vertolkingen moeten beoordelen door bijvoorbeeld gebruik te maken van opnameapparatuur – al behoort deze methode niet tot de gebruikelijke gang van zaken. Opnames laten namelijk toe om voldoende afstand te nemen van de eigen tolkprestatie, op voorwaarde dat ze pas enkele uren na de opdracht beluisterd worden.

Zelfevaluatie wordt doorgaans niet door iedereen als een geldig controlemiddel aanvaard omwille van het subjectieve karakter. Nochtans wordt deze methode, uitgezonderd in de tolkopleiding, wél veelvuldig toegepast (Déjean le Féal, 1990; Ross, 2006). Er rijzen onder meer twijfels over de betrouwbaarheid van de evaluaties: bepaalde studenten onderschatten hun talenten en beoordelen zichzelf eerder ondermaats, terwijl anderen een heel ander beeld hebben van een kwalitatief aanvaardbare prestatie en zich bijgevolg een te positieve zelfbeoordeling geven (Ross, 2006). Deze kwestie vereist zeker de nodige voorzichtigheid indien de beoordelingen mee het eindcijfer bepalen. Enkele studenten zouden in dat geval namelijk uit opportunisme kunnen handelen. Deze stelling wordt bevestigd door enkele studenten die zichzelf niet in staat achten om hun eigen prestatie te beoordelen. Wat zij als ‘een goede prestatie’ ervaren, wordt niet noodzakelijk op dezelfde manier beoordeeld door een medestudent of de docent (Ross, Rolheiser, & Hogaboam-Gray, 1998). Verder merkt Ross (2006) op dat de scores van de zelfevaluaties meestal hoger liggen dan die van de docenten. Studenten ondervinden soms moeilijkheden om op basis van de opgelegde criteria hun prestatie te evalueren. Daarom is het belangrijk dat docenten en studenten onderling bespreken welke criteria toegepast worden tijdens de zelfbeoordeling. Uit een studie van Bartłomiejczyk (2007) blijkt echter dat tolkstudenten zichzelf over het algemeen een negatieve beoordeling toekennen wanneer ze gevraagd worden om hun eigen tolkprestatie te evalueren. Voornamelijk de volledigheid en een correct gebruik van grammatica werden door de studenten negatief beoordeeld, al kreeg dit laatste criterium bijzonder weinig aandacht. Enkel over de logische samenhang van hun vertolkingen klonken de respondenten onmiskenbaar lovend. De overige criteria *trouw aan de oorspronkelijke tekst* (vergelijkbaar met *inhoudelijke overeenkomst met de oorspronkelijke tekst*), *stijl*, *woordenschat* en *presentatie* kregen overwegend hetzelfde aantal negatieve als positieve commentaren.

Déjean le Féal (1990, p. 157) tracht bovenstaande kritiek te weerleggen door verscheidene pluspunten naar voren te schuiven. Vaak komen bij het beluisteren van de eigen prestatie verscheidene elementen aan het licht waarvan de tolk geen besef had vanwege zijn concentratie op de output. Hij kan zich onder meer bewust worden van een overmatig gebruik van zinloze stopwoorden of opvullingen (vb. ‘euh’), een ongelukkige woordkeuze of een onjuiste syntaxis. Op die manier zou zelfevaluatie wel degelijk tot verbeterde tolkvaardigheden leiden en de motivatie en het zelfvertrouwen van de studenten stimuleren (Déjean le Féal, 1990; Postigo Pinazo, 2008; Ross, 2006), al bleek uit een interview met studenten dat enkele uitzonderingen de confrontatie met hun fouten juist als demotiverend ervoeren (Ross, Rolheiser, & Hogaboam-

Gray, 1998). Studenten halen zelf ook enkele voordelen aan waarom zij de voorkeur geven aan zelfevaluatie: dankzij een zelfevaluatie realiseren ze zich bijvoorbeeld sneller aan welke elementen ze moeten werken en krijgen ze een vollediger beeld van zowel de negatieve als de positieve elementen in hun prestatie (Ross, Rolheiser, & Hogaboam-Gray, 1998). Deze voordelen gelden echter niet wanneer tolken werken naar hun vreemde taal ('retourtolken'). Om van een betrouwbare evaluatie te spreken, dient de prestatie door een moedertaalspreker beoordeeld te worden, aangezien tolken hun vreemde taal meestal minder goed beheersen (Déjean le Féal, 1990).

2.3.2 Peerevaluatie

Docenten worden vaak geconfronteerd met tijdgebrek (Sadler & Good, 2006). Bijgevolg zou peerevaluatie een ideaal hulpmiddel vormen om het probleem deels te verhelpen. Topping (2009, p. 20) definieert peerevaluatie als volgt: "an arrangement for learners to consider and specify the level, value, or quality of a product or performance of other equal-status learners." Peerevaluatie beperkt zich niet tot de academische wereld, maar wordt ook veelvuldig toegepast in het dagelijkse leven onder de vorm van feedback, een zeer belangrijk fenomeen in om het even welk leerproces (Iaroslavschi, 2011; Topping, 2009).

Slechts weinig onderzoek naar peerevaluatie richt zich op de tolkwetenschap (Lee, 2016). Hoewel studenten hier een makkelijkere doelgroep vormen voor onderzoekers dan professionele tolken (Gile, 2003), zijn ook studenten er niet altijd even op gebrand om zich te laten evalueren door hun medestudenten. In een experiment uitgevoerd door Iaroslavschi in 2011, verkozen 27% van de tolkstudenten feedback van hun docent boven de peerevaluatie. Positieve peerevaluaties zouden het zelfvertrouwen van de studenten kunnen opkrikken, maar de negatieve feedback van docenten weegt toch zwaarder door. Sommigen zijn namelijk van mening dat de docent de enige bevoegde persoon is om studenten te evalueren vanwege zijn professionele kennis en ervaring in het toekennen van scores (Davies, 2000; Iaroslavschi, 2011). In landen zoals Zweden is de evaluatie door de docent zelfs een teken van autoriteit (Sadeghi, Hassani, & Mohammadloo, 2015). Deze opmerkingen sluiten aan bij de kritiek van Gile (2003) op studenten als incompetent beoordelaars (zie hoger). Door hun gebrek aan ervaring zouden studenten vaker een 'gemiddeld' cijfer geven, zijn ze niet in staat om alle foutieve elementen te detecteren of spreken de door hen toegekende scores de bijkomende nuancerende uitleg volledig tegen (Davies, 2000; Iaroslavschi, 2011; Topping, 2005, 2009).

Om die reden zou deze methode volgens Topping (2009) betrouwbaardere resultaten kunnen opleveren indien de studenten begeleid worden door de docent, die hen aanleert hoe een correcte peerevaluatie in zijn werk gaat en feedback geeft over de uitgevoerde evaluaties. Dit vergt echter de nodige tijd en inspanning. Evaluatieformulieren zouden ook hulp kunnen bieden bij het bewaren van de nodige objectiviteit (Iaroslavschi, 2011). Verder bleken sommige studenten in het experiment geen eerlijke feedback te willen geven uit angst om een conflict te veroorzaken. Zo trachtte 34% van de studenten die deelnamen aan het onderzoek van Iaroslavschi (2011) hun opmerkingen minder streng of op een beleefde manier te verwoorden of kozen ze ervoor om enkel de belangrijkste ‘fouten’ uit de tolkprestaties te halen. Peerevaluaties geven nu en dan inderdaad aanleiding tot conflictsituaties en gekwetste gevoelens, maar de oorzaak ligt in dat geval meestal bij de scherpe manier waarop de commentaren werden geformuleerd (Davies, 2000). Om dit zoveel mogelijk te vermijden, raadt Iaroslavschi (2011) aan om gestandaardiseerde evaluatieformulieren te gebruiken. Deze zorgen niet alleen voor meer consistentie aangezien alle studenten verplicht zijn om de vertolkingen van hun medestudenten op dezelfde manier te beoordelen, maar ze vermijden ook uitgeschreven, beledigende commentaren. Hoewel Iaroslavschi (2011) dit als een voordeel beschouwt, gelooft Topping (2005) echter dat kwaliteitsvolle feedback nuttiger is dan enkel scores te geven. Ook de respondenten betrokken bij het experiment van Sluijsmans, Moerkerke, van Merriënboer en Dochy (2001) gaven aan dat ze het louter gebruik van een score afkeurden en bijkomende uitleg wel degelijk als een noodzaak beschouwden. Iaroslavschi (2011) merkt op dat conflictsituaties zich minder vaak voordoen dan verwacht en dat studenten veelal een volwassen gedrag vertonen wanneer ze negatieve commentaar ontvangen. Enkele respondenten meenden zelfs dat negatieve feedback ook een constructieve invloed kan hebben en kan aanzetten tot verandering en meer inzet. Een andere manier om de angst bij peerevaluatie te verminderen bestaat erin om eerst positieve feedback te geven, waardoor het makkelijker wordt om nadien negatieve feedback te aanvaarden (Topping, 2009).

Aan peerevaluatie zouden echter meer voordelen dan nadelen gebonden zijn. Zelf beoordeeld worden en werk van medestudenten evalueren helpt niet enkel om inzicht te krijgen in de eigen prestatie, maar leert volgens Gillies (2004), Topping (2005, 2009) en Sadeghi, Hassani, en Mohammadloo (2015) studenten ook hun eigen werk beter te beoordelen (zelfevaluatie). Door de evaluaties van hun medestudenten krijgen studenten meer inzicht in hun eigen vaardigheden en kunnen ze bijgevolg werken aan hun tekortkomingen (Sadeghi, Hassani, & Mohammadloo, 2015; Topping, 2009). Door naar andermans prestaties te kijken, leren studenten bovendien hoe

anderen te werk gaan en worden ze kritischer voor zichzelf. Anderen blijken simpelweg voldoening te putten uit de fouten van anderen. (Davies, 2000) Een ander voordeel dat werd aangehaald door studenten in het onderzoek van Iaroslavschi (2011), en bevestigd wordt door Topping (2009), luidt dat medestudenten vaak op een andere wijze commentaar leveren en meer feedback voorzien in vergelijking met docenten, al zou dit een gevolg kunnen zijn van tijdgebrek. Net als bij zelfevaluatie, wordt het zelfvertrouwen dankzij peerevaluatie gestimuleerd. Sadeghi, Hassani en Mohammadloo (2015) kwamen tot de vaststelling dat het zelfvertrouwen bij de studenten veel meer ontwikkeld was na zes keer beoordeeld te worden door hun medestudenten. Wanneer dezelfde handelingen werden uitgevoerd door de docenten, was er nauwelijks verbetering te zien. Evaluatie door de docenten zou namelijk meer stress teweegbrengen bij de studenten (Iaroslavschi, 2011; Sadeghi, Hassani, & Mohammadloo, 2015).

Peerevaluatie heeft niet alleen voordelen voor de persoon die geëvalueerd wordt: de beoordelaar krijgt tegelijkertijd een algemeen beeld over het prestatieniveau van andere medestudenten en kan zich op die manier spiegelen aan hun prestaties (Davies, 2000; Topping, 2009). Topping (2009) geeft zelfs aan dat de evaluaties hen in staat stellen om hun leermethodes bij te sturen.

2.3.3 Vergelijking tussen de zelf- en de peerevaluatie

Slechts in een kleine minderheid van reeds uitgevoerde onderzoeken komt een vergelijking tussen de zelfevaluatie en de peerevaluatie aan bod, een kwestie die deze masterproef aanhaalt. Uit de studies van Iaroslavschi (2011) en Sadler en Good (2006) blijken studenten toleranter te zijn tegenover hun eigen prestaties dan tegenover die van hun medestudenten. Volgens Iaroslavschi (2011) valt dit te verklaren door een verkeerde aanpak van de zelfevaluatie. In vergelijking met de toegekende scores van de docent, gaven de studenten zichzelf tijdens het onderzoek van Sadler en Good (2006) gemiddeld 1,9 punten meer, terwijl het cijfer van de medestudenten 3,3 punten onder de score van de docent lag. Voornamelijk zwakkere studenten hadden de neiging om zichzelf minder streng te beoordelen. Sterkere studenten kregen dan weer een minder goede evaluatie van hun medestudenten. Bij deze studie was er echter sprake van een zelfevaluatie, met als gevolg dat de eigen score in het eindresultaat verrekend werd en in die zin een zekere invloed had op de manier waarop studenten zichzelf beoordeelden (Sadler & Good, 2006).

De bovenstaande bevindingen druisen in tegen de stellingen van Bartłomiejczyk (2007) en Gile (1995a). Bartłomiejczyk (2007) concludeert dat studenten zichzelf over het algemeen een negatieve zelfbeoordeling toekennen; Gile (1995a) stelt vast dat medestudenten veelal een milde houding aannemen tijdens de peerevaluatie.

3 Methode

3.1 Onderzoeksvraag

Tot op heden werd er nauwelijks onderzoek uitgevoerd naar hoe de perceptie van de tolk over zijn eigen tolkprestatie zich verhoudt tot de beoordeling van diezelfde prestatie door een collega of een medestudent. Omdat enkele vragen omtrent de zelf-en peerevaluatie nog onbeantwoord blijven, tracht deze masterproef een antwoord te bieden op de volgende onderzoeksvraag: verschilt de zelfevaluatie van de peerevaluatie bij de beoordeling van een consecutieve tolkprestatie naar het Nederlands? Dit kwalitatieve, vergelijkende onderzoek bij master- en postgraduaatstudenten is gebaseerd op een selectie van dertien kwaliteitscriteria van Bühler (1986), Barik (1994), Gile (1990), Pöchhacker en Zwischenberger (2010) en Riccardi (2002).

Deze empirische studie vertrekt van de hypothese dat de masterstudenten zichzelf strenger zullen beoordelen dan hun mede- en postgraduaatstudenten. Deze hypothese steunt op de studie van Gile (1995a) over de mildheid van studenten bij de evaluatie van de vertolking van een medestudent: ze beseffen dat ze allen nog in de leer zijn. Verder berust deze veronderstelling op de bevindingen van Bartłomiejczyk (2007), die aantoont dat studenten zichzelf over het algemeen een relatief negatieve zelfbeoordeling voorleggen.

3.2 Onderzoeksopzet

3.2.1 Deelnemers

De steekproef bestaat uit zestien respondenten, onder wie acht studenten van de master in het tolken (medestudenten van de auteur van deze masterproef en de auteur inbegrepen) en acht studenten van het postgraduaat Conferentietolken. De keuze van de respondenten hoort een (gedeeltelijk) antwoord te bieden op de problematiek omtrent de competentie van de beoordelaar (zie 2.1.4). Volgens Gile (1983) en Bühler (1986) behoort de passieve collega-tolk tot de meest gekwalificeerde en competente beoordelaars van een vertolking, aangezien de beoordelaars in staat moeten zijn zowel de bron- als de doelttekst te begrijpen om de oorspronkelijke tekst en de vertolking met elkaar te kunnen vergelijken (Déjean le Féal, 1990). Omdat dit onderzoek zich richt op de vergelijking tussen de zelfevaluatie en de

peerevaluatie – en bijgevolg een beroep doet op medestudenten –, wordt de bovenstaande suggestie alvast in acht genomen.

De selectie van de respondenten gebeurde op basis van de makkelijke bereikbaarheid, een beslissing die gebaseerd is op “convenience sampling” (Gile, 1998, p. 80). Met uitzondering van één postgraduaatstudent die aan KU Leuven Campus Antwerpen studeert, werden alle respondenten gerekruteerd aan KU Leuven Campus Brussel. De masterstudenten werden tijdens een gemeenschappelijke les aangesproken en geïnformeerd over het opzet en de methodologie van het survey-onderzoek. Alle aanwezigen waren onmiddellijk bereid om deel te nemen. Door de iets moeilijkere bereikbaarheid van de postgraduaatstudenten, werden zij uitsluitend per mail gecontacteerd.

Om de homogeniteit binnen de steekproef te bewaren, werden specifieke criteria toegepast voor de selectie van de respondenten. Zo zijn, op één uitzondering na, alle deelnemers moedertaalsprekers van het Nederlands en hebben evenveel studenten óf Engels, óf Frans in hun talencombinatie. De postgraduaatstudenten hebben minstens één jaar consecutieve tolkervaring en hebben allemaal de toelatingsproef tot het postgraduaat succesvol afgelegd. Ook de masterstudenten beschikken over een gelijkwaardige ervaring in het tolken. Hoewel Moser-Mercer (2008) pleit voor een gevarieerde steekproef voor de representativiteit van het onderzoek, vormt de homogeniteit binnen de groep van medestudenten juist een sterkte van deze masterproef. De gelijkaardige eigenschappen beperken namelijk het aantal variabelen, die een invloed kunnen uitoefenen op de resultaten.

3.2.2 Verloop van het experiment en het survey-onderzoek

Het onderzoek ving aan met een experiment waarbij een vertolking geëvalueerd werd op basis van een evaluatieformulier met een cijferschaal van één tot vijf. Vervolgens moesten de respondenten bijkomende uitleg geven bij hun score (survey-onderzoek). Omdat de beoordelingen bedoeld zijn om zwaktes en sterktes in de vertolkingen vast te stellen, is er sprake van formatieve evaluatie.

Voor het experiment en het survey-onderzoek werden twee aparte instructiefiches opgesteld: een voor de masterstudenten die optraden als tolk en als beoordelaar, een tweede voor de postgraduaatstudenten die enkel de rol als beoordelaar op zich namen (zie respectievelijk

bijlagen A en B). De fiches beschreven alle stappen van het experiment en vermeldde de deadlines voor de opnames van de tolkprestaties en het terugsturen van de evaluatieformulieren. Via deze instructiefiche werd aan alle participanten de anonimiteit van het onderzoek verzekerd. De masterstudenten kregen de fiches bijna twee maanden vóór de deadline per mail opgestuurd, zodat de respondenten voldoende tijd zouden hebben om de opdracht in hun agenda in te plannen. Ook de postgraduaatstudenten werden meer dan twee en een halve maand vooraf ingelicht over het onderzoek. Zij dienden de nodige informatie door te sturen over hun talige achtergrond en het opleidingstraject dat ze reeds gevolgd hadden, om zo hun ervaring te bepalen. De medewerking van alle respondenten zou acht vertolkingen, acht zelfevaluaties en tweeëndertig peerevaluaties opleveren.

3.2.2.1. Consecutieve vertolking van het Frans/Engels naar het Nederlands

Via e-mail kregen de masterstudenten een speech toegewezen uit de Speech Repository (d.i. een online databank waarin honderden speeches voor niet-commerciële doeleinden beschikbaar worden gesteld aan studenten, docenten en professionele tolken (Europese Commissie, z.d.)). De beschikbare speeches waren hetzij in het Engels, hetzij in de Franse taal, al naargelang de talencombinatie van de respondenten (zie bijlage C). De door Speech Repository aangeboden audiovisuele fragmenten werden reeds beoordeeld en verbeterd door professionele tolken en aangevuld met relevante informatie over de taal, het niveau, de tolkmodus, het domein en het type (Europese Commissie, z.d.). Om de variabelen te beperken en geen enkele respondent te benadelen, werden de speeches zodanig uitgekozen dat ze in lengte (ongeveer vier minuten) en niveau (beginner) op elkaar afgestemd waren. De geringe ervaring van de masterstudenten in consecutief tolken op het ogenblik van de afname van het experiment verklaart de keuze van de toespraken.

De meerderheid van de toespraken (zes) behoorden tot de categorie ‘Gezondheid’, al dan niet aangevuld met een verwant domein (vb. ‘Voedsel’ of ‘Consumenten’). De overige twee toespraken hadden betrekking op het milieu, de landbouw en/of consumenten. Verdere informatie over de toegewezen speeches werd vermeld in de persoonlijke bijbehorende mail ter attentie van de masterstudenten.

De speech deed dienst als brontekst voor de consecutieve vertolking van de B-taal (Frans of Engels) naar de A-taal (Nederlands). Er werd bewust gekozen voor een consecutieve tolkmodus, omdat de masterstudenten op het ogenblik van het experiment op nog geen enkele wijze in contact waren gekomen met simultaan tolken. Bovendien leent de consecutieve tolkmodus zich makkelijker tot experimenteel onderzoek, aangezien de vertolkingen in een niet-natuurlijke tolksetting kunnen plaatsvinden en er geen nood is aan specifiek tolkmateriaal. Ook kunnen de studenten de eigen prestatie sneller vergelijken met de oorspronkelijke speech dankzij hun notities. De studenten namen hun vertolkingen op in een niet-natuurlijke tolksetting (namelijk thuis of op een andere werkplek) in WMA- of M4A-formaat. In de instructiefiche werd vermeld dat de voorkeur uitging naar WAV-formaat, aangezien dit formaat, naar verluidt, een hogere kwaliteit oplevert, maar na verificatie van de opnames, vormde de kwaliteit absoluut geen probleem. Gemakshalve stuurden de respondenten hun opnames via e-mail of via WeTransfer door naar de auteur van deze masterproef.

3.2.2.2. Zelfevaluatie van de vertolking

Verder kregen de masterstudenten hetzij persoonlijk, of – indien persoonlijk contact niet mogelijk bleek – via e-mail een evaluatieformulier (zie bijlage E) waarin dertien criteria (zeven linguïstische en zes niet-linguïstische) werden opgesomd, voorafgegaan door een glossarium met daarin een ruimere omschrijving van de geselecteerde criteria (zie bijlage D). Op basis van dit formulier (zie voor verdere informatie hoofdstuk 3.2.3), werden de masterstudenten gevraagd om hun eigen tolkprestatie te beoordelen aan de hand van een cijferschaal. Om de taak te vergemakkelijken, konden zij een beroep doen op hun eigen notities die ter ondersteuning dienden bij de vertolking (zie bijgevoegde cd).

3.2.2.3. Peerevaluatie van de vertolking

De titels van de oorspronkelijke speeches uit de Speech Repository werden vervolgens gelijktijdig met de door de masterstudenten getolkte speeches via e-mail aan vier andere respondenten (twee medestudenten uit de master in het tolken en twee studenten uit het postgraduaat Conferentietolken) bezorgd. In principe hadden meer beoordelaars uitgenodigd kunnen worden om deel te nemen aan de peerevaluatie. Omdat het beluisteren en vergelijken van de toespraken en de vertolkingen enerzijds en het evalueren van de tolkstudenten anderzijds echter al redelijk wat tijd in beslag nam, werd het evaluatiewerk beperkt tot slechts twee

vertolkingen per respondent. Vanwege de korte tijdspanne voor het onderzoek, ontvingen zij geen uitgeschreven transcripties van de oorspronkelijke speeches. Wel konden ze deze, net als de masterstudenten, beluisteren via de hyperlinks die in de e-mail werden ingevoegd. Op net dezelfde wijze als de masterstudenten kregen de respondenten hetzelfde evaluatieformulier en glossarium. Elke respondent beluisterde twee bronteksten (d.i. de oorspronkelijke speeches) en beoordeelde aansluitend de tolkprestaties op een cijferschaal van één tot vijf. De postgraduaatstudenten werden gevraagd om eerst de oorspronkelijke speech te beluisteren, om de variabele als *pure customer*, oftewel de toehoorder die enkel het product te horen krijgt, uit te sluiten. De tolkprestaties kunnen namelijk enkel geëvalueerd worden op basis van de gekozen criteria indien de beoordelaars op de hoogte zijn van de inhoud en van de presentatie van de brontekst (Déjean le Féal, 1990).

3.2.3 Evaluatieformulier en glossarium

Globaal genomen moesten de respondenten de tolkprestaties beoordelen op basis van zeven linguïstische en zes niet-linguïstische criteria. Net als bij Bühler (1986) werd ervoor gekozen om de semantische (linguïstische) kenmerken te scheiden van de prosodische (niet-linguïstische) kenmerken om een duidelijk onderscheid te maken tussen inhoud en vorm van de vertolking. De linguïstische criteria die in het evaluatieformulier werden opgenomen (*inhoudelijke overeenkomst met de brontekst, vlotheid van de vertolking, correct gebruik van terminologie, correct gebruik van grammatica, tongval, logische samenhang en geschikte stijl*), zijn voornamelijk gebaseerd op die van Bühler (1986). Omdat het eerste linguïstische criterium (namelijk *inhoudelijke overeenkomst met de brontekst*) ruim opgevat kan worden, werden enkele belangrijke subcriteria van Barik (1994) – *omissions, toevoegingen en substituties* – toegevoegd met de bedoeling om de antwoorden verder te verfijnen. Zoals omschreven in het glossarium, betreffen omissies ‘belangrijke informatie die weggelaten wordt’ en wordt er niet zozeer gedoeld op details of elementen die geen afbreuk doen aan de inhoud wanneer de tolkstudent deze achterwege laat. Het gebruik van omissies hangt namelijk in grote mate af van de methodologische opvatting van wat een consecutieve vertolking hoort te zijn: hoeveel omissies in een consecutieve vertolking leiden tot een mindere kwaliteit? Een langere consecutieve versie dan de oorspronkelijke toespraak is vanzelfsprekend uit den boze. Daarom opteert Gile (1983), zoals eerder vermeld, voor een synthese van de brontekst en zijn omissies met andere woorden gewenst. Vanwege de beperkte geheugencapaciteit, kan een consecutieve tolk hoe dan ook niet in staat zijn om elk detail weer te geven (zie 2.2.1). Hetzelfde geldt voor

substituties ('belangrijke informatie die gewijzigd wordt'), die in deze masterproef worden beschouwd als gewijzigde elementen die foutieve informatie weergeven. Geïnspireerd door Riccardi's bevindingen (2002), werden de subcriteria *stopwoorden*, *zelfverbeteringen* en *pauzes* aan het formulier toegevoegd om op eenzelfde wijze een nuance aan te brengen bij het algemene criterium *vlotheid van de vertolking*. Bij ingelaste pauzes speelt voornamelijk de tijdsduur en het uitgekozen moment een belangrijke rol: na een betekenisdragende eenheid is een korte pauze zeker gewenst om de structuur van de tekst vlot te laten overkomen (deze methode is gekend onder de term "chunking", Piccaluga, Nespoulous, & Harmegnies, 2007, p. 1317), maar in het midden van een zin komt dit eerder onnatuurlijk over. De parameter *volledigheid van de vertolking* die in verschillende studies wordt toegevoegd aan de lijst (zo ook bij Bühler (1986)) werd in deze masterproef niet apart opgenomen, omdat het criterium al wordt weergegeven dankzij het subcriterium *omissies*. Een apart criterium om de volledigheid afzonderlijk te beoordelen werd dus als overbodig beschouwd. Voor de niet-linguïstische criteria werd opnieuw de studie van Bühler (1986) geraadpleegd, aangevuld met elementen die naar voren worden geschoven door Pöchhacker en Zwischenberger (2010) en Gile (1990): uit Bühlers studie (1986) werden de parameters *aangename stem* en *zelfvertrouwen* overgenomen, Pöchhacker en Zwischenberger (2010) vormden inspiratie voor *levendige intonatie*, Gile (1990) voor het criterium *spreektempo*. De criteria *objectiviteit/onpartijdigheid* en *articulatie* bleken interessant en zinvol tijdens de eigen tolklessen en werden daarom uit eigen beweging aan het formulier toegevoegd.

In tegenstelling tot vele voorgaande studies, werd er in deze masterproef geen rekening gehouden met hoe zwaar de kwaliteitscriteria doorwegen volgens de tolken en beoordelaars (en dus ook luisteraars). Wel werden de respondenten gevraagd de tolkprestaties te beoordelen op basis van de voornoemde criteria via een cijferschaal van één tot vijf (1 = slecht, 2 = onvoldoende, 3 = aanvaardbaar, 4 = goed, 5 = zeer goed). Bij elk afzonderlijk criterium moesten ze hun specifieke antwoord motiveren; zodoende werd tegemoet gekomen aan de kritische bemerkingen van Gile (1983): deze auteur beschouwt het gebruik van louter cijferschalen als onvoldoende om de kwaliteit van een vertolking te beoordelen en houdt vast aan bijkomende, diepgaandere vragen om de score te nuanceren. Onderaan het formulier konden de respondenten een algemene beoordeling van de vertolking geven, opnieuw door middel van een cijfer en bijhorende motivatie. Ze hoefden hierbij geen gemiddelde score te berekenen. Een extreem verschil tussen de score van de algemene beoordeling enerzijds en de gemiddelde score van alle afzonderlijke criteria anderzijds zou echter betekenen dat de

beoordelaars inconsequent zijn in hun beoordeling, wat niet verwacht wordt in deze masterproef. Aangezien de vertolkingen als audio-opnames werden opgestuurd om de anonimiteit van de respondenten te garanderen, waren kenmerken met betrekking tot de presentatie (vb. oogcontact, gebaren, houding) niet van toepassing en werden ze bijgevolg niet vermeld in de niet-linguïstische criteria. Zoals Iglesias Fernández (2013) en Garzone (2002) terecht opmerken, schuilt er een zekere subjectiviteit in alle opgenomen elementen, aangezien elke luisteraar of beoordelaar de term een andere invulling kan geven. Zoals eerder vermeld, pleit ook Gile (2003) voor een duidelijkere formulering van de kwaliteitscriteria. Om dit probleem in de mate van het mogelijke te beperken, werd bij het evaluatieformulier een glossarium gevoegd (zie bijlage D). Het betreffende glossarium werd vooraan in het document gevoegd, met als doel de respondenten aan te sporen dit aandachtig door te nemen alvorens over te gaan tot de evaluatie. De respondenten bezorgden de zelfevaluatieformulieren samen met de notities persoonlijk of via e-mail aan de onderzoeker.

De mate van begrijpelijkheid, de presentatie en de gebruiksvriendelijkheid van zowel het glossarium als het evaluatieformulier werden voor aanvang van het experiment en het survey-onderzoek uitgetest in een zeer kleinschalig pilotonderzoek (zoals door Moser-Mercer (2008) wordt aangeraden, zie 2.2). Drie respondenten (waarvan slechts één betrokken bij het uiteindelijke onderzoek) evalueerden een consecutieve vertolking naar keuze in het kader van een van hun tolkvakken. Zo bleek uit de pilotstudie dat er meer ruimte voorzien moest worden om de antwoorden te staven bij de criteria *inhoudelijke overeenkomst met de brontekst*, *vlotheid van de vertolking* en *algemene beoordeling*. Hoewel eerdere studies frequent gebruik maakten van een vijfpuntenschaal, beweert Pöchhacker (2005) dat een zevenpuntenschaal toch de voorkeur verdient. Daarom werd aan de respondenten van de pilotstudie gevraagd of zij een zevenpuntenschaal als genuanceerder en preciezer ervoeren dan een vijfpuntenschaal en of zij bijgevolg de zevenpuntenschaal wensten te gebruiken. De respondenten antwoordden echter negatief, omdat volgens hen het gebruik van een vijfpuntenschaal het evaluatieformulier gebruiksvriendelijker maakte.

3.3 Dataverzameling en –verwerking

Op vier exemplaren na, bereikten alle vertolkingen en zelf- en peerevaluatieformulieren tijdig de onderzoeker (voor de ingevulde formulieren, zie bijgevoegde cd). De scores werden per speech anoniem ingevoerd in Excel. In het volgende hoofdstuk volgt een kort overzicht van de

gemiddelde scores van de zelfevaluaties, in vergelijking met de algemene beoordelingen. Vervolgens worden de beoordelingen van elke parameter afzonderlijk geanalyseerd. Per criterium werd ervoor gekozen om elke score van de zelfevaluatie te vergelijken met het gemiddelde van de vier peerevaluaties en deze resultaten weer te geven in een staafdiagram om een visueel beeld te krijgen. Nadien volgen de gemiddelde scores van alle zelf- en peerevaluaties voor een algemeen overzicht. Door de grotere hoeveelheid aan data is een weergave in tabelvorm hierbij meer aangewezen.

Bij de verwerking van de resultaten wordt zowel rekening gehouden met mogelijke tendensen binnen de beoordelingen van de respondenten als met de absolute cijfers. Om te verwijzen naar de uitgeschreven beoordelingen, wordt gebruik gemaakt van de code 'S' voor de tolkstudenten ('self') en van de code 'P' voor de beoordelaars ('peer'). Tolkstudenten 1 tot en met 4 (S1 → S4) vertolkten een Franse toespraak naar hun moedertaal Nederlands, tolkstudenten 5 tot en met 8 (S5 → S8) luisterden naar een Engelse speech en zetten deze eveneens om naar het Nederlands. Hoewel de taal van de oorspronkelijke brontekst (Frans of Engels) een bijkomende variabele vormt, werd er in deze masterproef geen rekening mee gehouden. De studenten en de talencombinaties werden immers uitgekozen in functie van hun toegankelijkheid. Voor de absolute cijfers gelden 'hoge' of 'hogere' scores vanaf 4 punten. Een 'lage' of 'lagere' score verwijst naar 1 of 2 punten. Een score van 3 punten wordt beschouwd als een 'middelmatische score'. Bij de vergelijking tussen de zelfevaluaties en de gemiddelde score van de peerevaluaties, worden de termen 'groot' of 'groter' verschil gehanteerd. In deze masterproef wijzen deze op een verschil van minstens 1 punt.

Omdat de formulieren van twee respondenten onvolledig waren (een of twee parameters werden niet beoordeeld), werd er nagegaan of dit al dan niet per ongeluk gebeurde. Beide studenten bleken deze elementen inderdaad over het hoofd gezien te hebben en stuurden daarom een tweede bericht om de ontbrekende scores verder aan te vullen. Om te vermijden dat de respondenten hun werk opnieuw moesten bezorgen, werden de onvolledige scores door de onderzoeker persoonlijk in kleur aangeduid (zie bijlage F). Het ging hierbij om de beoordelingen van de toespraken 'All about diets', 'Food crisis' en 'Sunburn and skincancer'.

4 Resultaten

4.1 Gemiddelde score van de zelfevaluaties

Tabel 1

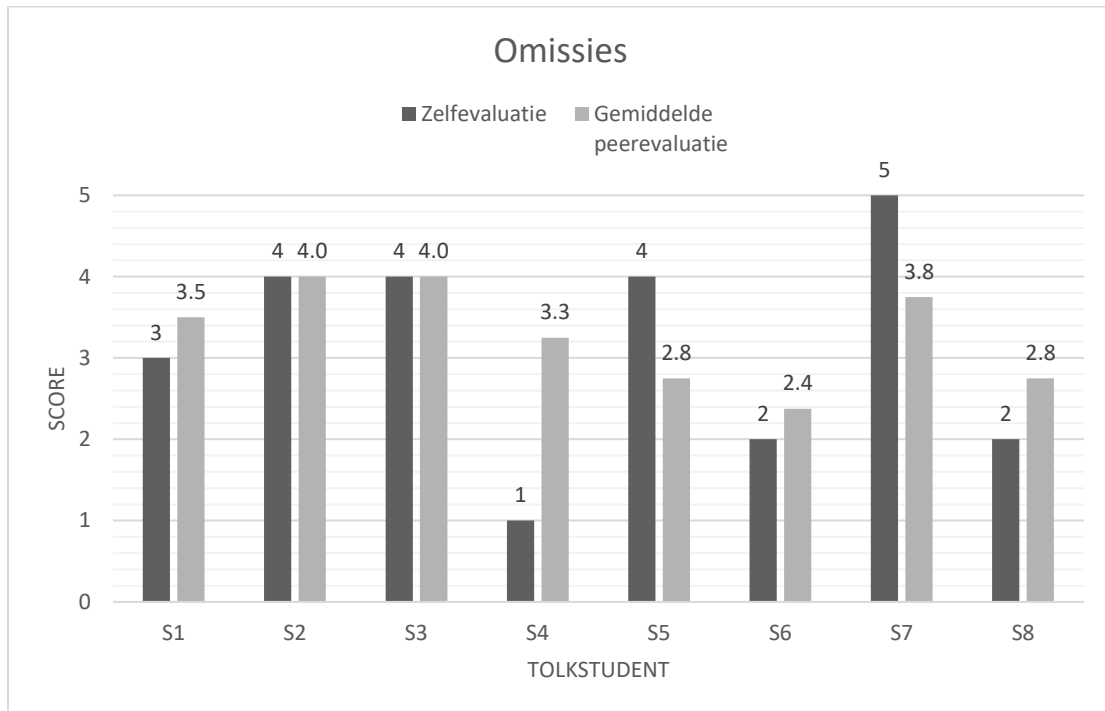
Vergelijking tussen de gemiddelde score van de zelfevaluaties en de algemene beoordeling

Tolkstudent	<i>S1</i>	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>S5</i>	<i>S6</i>	<i>S7</i>	<i>S8</i>
Gemiddelde van alle criteria	4	4	5	2	4	3	4	3
Algemene beoordeling	4	4	5	2	4	3	4	3

Uit tabel 1 kan worden afgeleid dat tolkstudent S4 zichzelf over de hele lijn zeer kritisch beoordeelt tijdens de zelfevaluatie, aangezien zijn gemiddelde score niet verder reikt dan 2 punten. Tolkstudent S3 daarentegen heeft een gemiddelde score van 5 punten, wat wijst op een milde houding. De cijfers die de tolkstudenten toekennen voor de algemene beoordeling komen precies overeen met de berekende gemiddelde scores. Deze gelijke resultaten voldoen dus aan de verwachtingen (zie 5.1 voor een diepgaande analyse).

4.2 Afzonderlijke criteria

4.2.1 Inhoudelijke overeenkomst met de brontekst: omissies



Figuur 1

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *omissies*

Data

Voor het criterium *omissies* kent de helft van de tolkstudenten zichzelf een score 4 of hoger toe; een enkele student beloont zichzelf met de maximumscore van 5 punten. Toch tonen ook drie tolkstudenten zich erg kritisch voor zichzelf en kennen ze zich een eerder lage score van 2 punten toe, of zelfs slechts 1 punt in het geval van tolkstudent S4. Bij diezelfde student wordt het grootste verschil waargenomen (2.3 punten verschil), het laagste bij student S6 (een verschil van 0.4 punten).

Open vragen

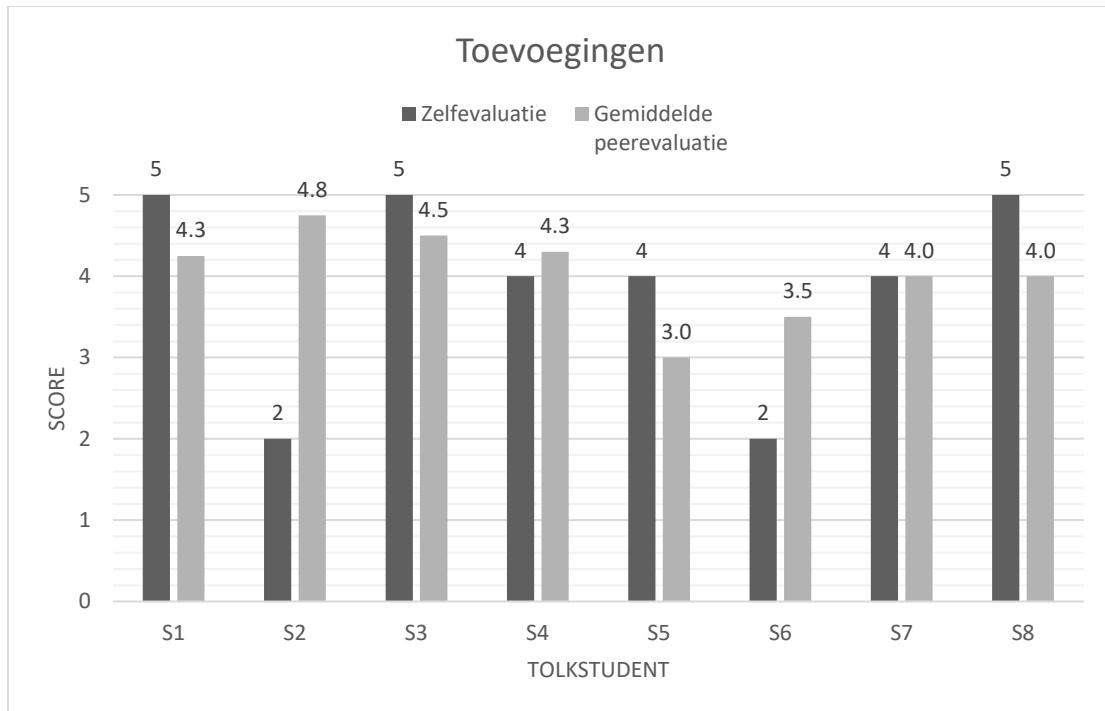
Opvallend is het verschil tussen de analyse van tolkstudent 4 en de beoordelaars. De student vindt bijvoorbeeld van zichzelf dat hij slechts de laagste score verdient voor zijn vertolking, terwijl een van de beoordelaars een score van 4 aanduidt: “De inhoud komt vrij goed over, geen van deze elementen is storend” (P2).

Ook het omgekeerde doet zich voor. Zo stelt tolkstudent S7 geen enkele ommissie in zijn vertolking vast, terwijl alle beoordelaars minstens één weglating aangeven. Dezelfde vaststelling kan gemaakt worden bij tolkstudent S5, die zichzelf een hoge score van 4 punten toekent. Ondanks een middelmatige score en eenmaal een hogere score van 4 punten, geven twee beoordelaars slechts een 2: “Tolk moet opletten voor te veel weglatingen” (P10).

Ten slotte blijkt er verwarring te zijn ontstaan over de introductie van de oorspronkelijke spreker. Twee tolkstudenten (S3 en S6) vertaalden de korte inleiding over het onderwerp van de speech niet en begonnen pas te tolken vanaf het ogenblik dat de oorspronkelijke spreker het begin van zijn speech aankondigde (door bv. “Je commence” te zeggen). Voor beoordelaars P7 en P14 behoorde dit deel echter nog tot de oorspronkelijke speech en moest deze informatie dus vertolkt worden. Bijgevolg beschouwden zij deze omissies als belangrijke weggelaten elementen.¹

¹ In deze masterproef worden deze commentaren als geldig aanvaard, omdat de cijfers en de aanvullende uitleg van de respondenten zo getrouw mogelijk worden overgenomen en de gegevens zo goed mogelijk vanuit hún standpunt worden verwerkt.

4.2.2 Inhoudelijke overeenkomst met de brontekst: toevoegingen



Figuur 2

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *toevoegingen*

Data

Opvallend zijn de hoge zelfevaluaties van de tolkstudenten: drie studenten geven zichzelf de maximumscore en eenzelfde aantal kent zich de op één na hoogste score toe. De resultaten van het criterium *toevoegingen* zijn overwegend homogeen, met uitzondering van twee duidelijk negatieve uitschieters (S2 en S6). Opmerkelijk is de grote discrepantie (2.8 punten) tussen de zelfevaluatie en de peerevaluatie bij S2: anders dan de tolkstudent, die zichzelf een lage score van 2 punten toekent, komt de gemiddelde score van de beoordelaars aardig in de buurt van het maximum. Bij de tolkstudenten die zichzelf een maximumscore geven voor dit criterium, wijkt het gemiddelde van de medestudenten niet sterk af van de zelfevaluatie – al ligt dit laatste cijfer wel hoger dan de peerevaluatie.

Open vraag

Een duidelijk contrast is te vinden bij S2: geen enkele beoordelaar maakt een opmerking over toevoegingen door de tolkstudent, terwijl de student er zelf wél enkele vaststelt. Om zijn gebruik van toevoegingen echter te verantwoorden, nuanceert de tolkstudent zijn antwoord: “Misschien een aantal toevoegingen om het duidelijk te kunnen tolken”² (S2).

Ook bij tolkstudent S5 is een discrepantie merkbaar: twee van de vier beoordelaars kennen dezelfde score toe als de student (score 4) en vermelden niets over eventuele toevoegingen. Beoordelaars P10 en P15 geven daarentegen een lagere score (2 punten), maar in hun persoonlijke uitleg komen de standpunten niet overeen: in tegenstelling tot P15 die de toevoegingen slechts als “kleine toevoegingen” beschouwt, lijkt P10 zwaarder te tillen aan dit criterium (“Tolk moet opletten met eigen toevoegingen”).

De toegekende scores en de aanvullende uitleg spreken elkaar soms tegen. Hoewel de volgende uitspraken overwegend positief klinken, krijgt geen enkele tolkstudent de maximumscore:

P2 over S1 (score 4): “goede vertolking zonder storende elementen”³

P3 over S3 (score 4): “er zijn me geen grote fouten opgevallen”⁴

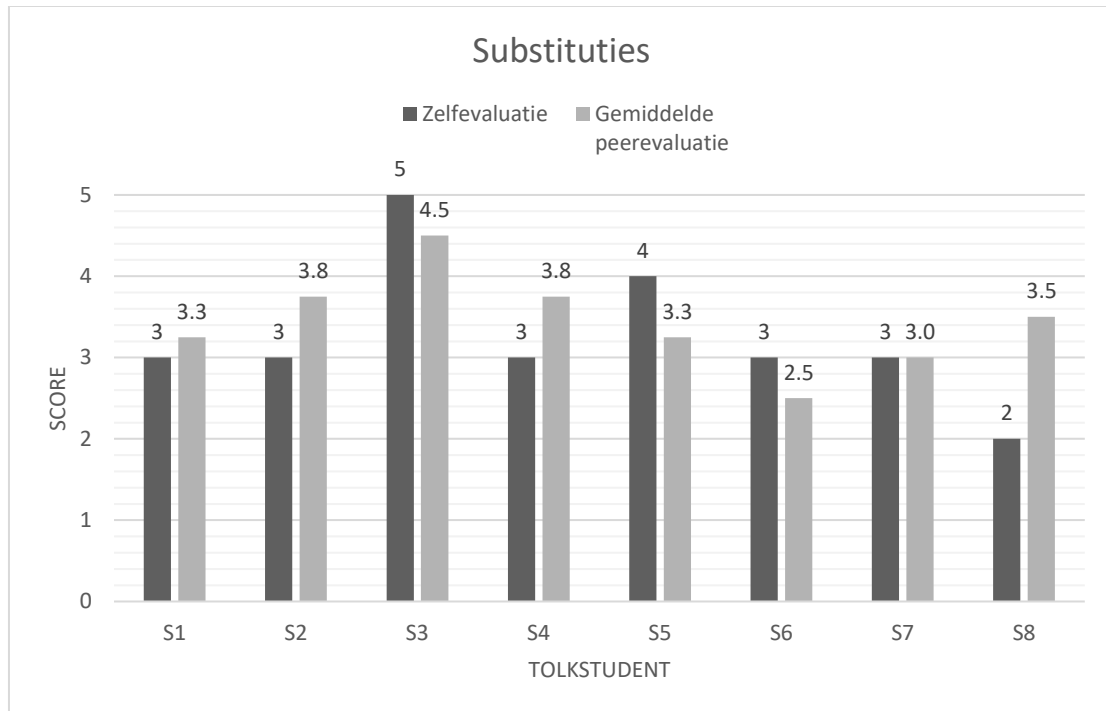
P2 over S4 (score 4): “De inhoud komt vrij goed over, geen van deze elementen is storend.”

² In eerste instantie lijkt tolkstudent S2 te doelen op ‘explicitaties’, een vertaaltechniek waarbij impliciete informatie van de brontekst expliciet wordt gemaakt in de doeltaal. Het doel van deze methode is om onbegrip door een gebrek aan achtergrondkennis van de toehoorder (bv. vanwege cultuurverschillen) te vermijden. (Mesa-Lao, 2011, p. 45-46) Na een grondige vergelijking tussen de oorspronkelijke toespraak en de vertolking, bleek hier echter geen sprake van te zijn, waardoor het onduidelijk blijft wat de tolkstudent precies bedoelt met zijn opmerking.

³ De commentaren van beoordelaars P2 en P3 wekken de indruk dat de respondenten het criterium *toevoegingen* niet volledig begrepen hebben. “Storende elementen” blijft immers een zeer vage omschrijving.

⁴ Zie voetnoot 3

4.2.3 Inhoudelijke overeenkomst met de brontekst: substituties



Figuur 3

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *substituties*

Data

De resultaten voor het criterium *substituties* liggen niet erg hoog: slechts twee zelfevaluaties bereiken een score van 4 of 5 punten, vijf van de acht scores voor de eigen prestatie reiken niet verder dan een middelmatige score. De meeste scores van de zelfevaluaties en de gemiddeldes liggen niet ver uit elkaar (het grootste verschil van 1.5 punten komt slechts eenmaal voor). Drie van de vier beoordelaars kennen bijvoorbeeld hetzelfde cijfer toe als tolkstudent S1 in zijn zelfevaluatie (namelijk een score van 3 punten). Een grote tegenstelling valt enkel op bij de peerevaluaties van student S7: twee beoordelaars duiden een zeer lage score aan (scores 1 en 2) terwijl de twee andere respondenten de hoogste scores toekennen (scores 4 en 5). Hierdoor komt de gemiddelde score wel gelijk te staan met de zelfevaluatie van de tolkstudent.

Open vraag

Hoewel tolkstudent S2 voor zichzelf een gemiddelde score van 3 punten vaststelt, wekt zijn bijkomende uitleg voor *substituties* een positievere indruk op: “niet dat ik heb gemerkt”. Ook bij de beoordelaars die werden vermeld bij de rubriek *toevoegingen* (P3 en P2), valt opnieuw de score van ‘slechts’ 4 punten op tegenover een zeer positieve commentaar.

Opvallend zijn de uiteenlopende evaluaties voor student S5, waarbij elke respondent een ander aantal substituties opmerkte:

S5 (score 4): “1 fout: ‘de ene’ en ‘de 2^e groep’, terwijl het niet ging over de groepen, maar individuen”

P10 (score 5): “Geen vervangingen gehoord”

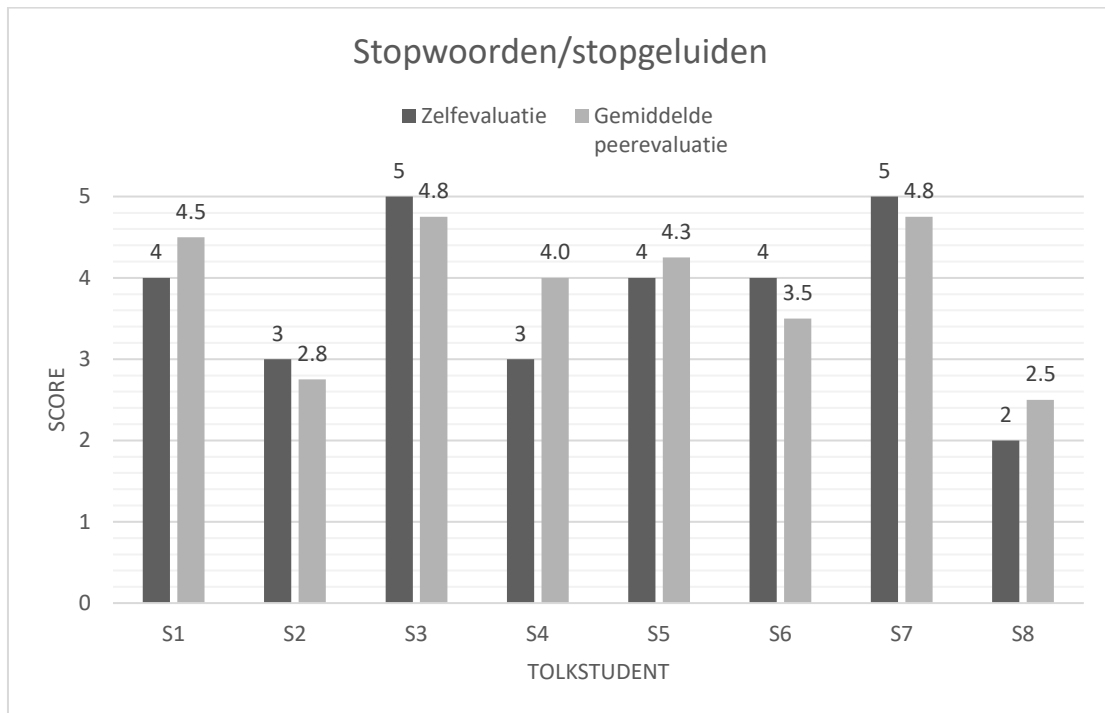
P11 (score 2): “Substituties echter waren wel aanwezig, soms was het gewoon fout zoals een stukje in het begin, en de connectoren waren ook niet altijd juist.”

Hetzelfde fenomeen duikt op bij tolkstudent S7:

P10 (score 1): “er vallen te veel verschuivingen op: de tolk maakt er een geheel van maar geeft niet de correcte inhoud”

P12 (score 4): “zeer weinig en niet bij belangrijke stukken”

4.2.4 Stopwoorden/stopgeluiden



Figuur 4

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *stopwoorden/stopgeluiden*

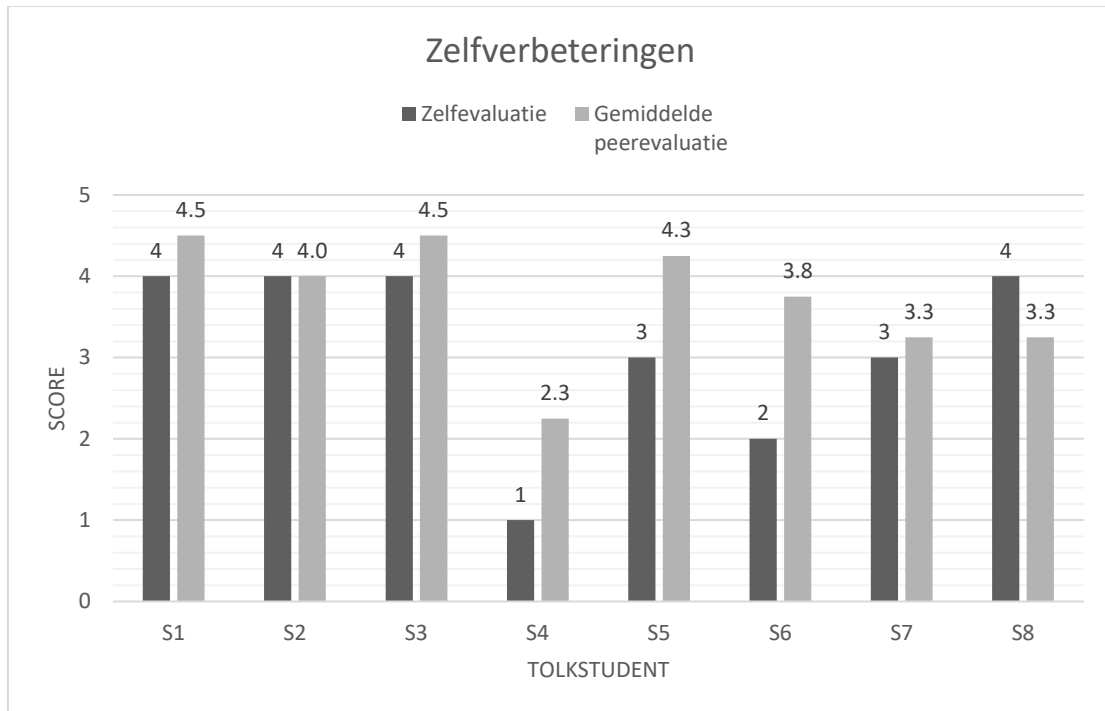
Data

Opnieuw liggen de meeste resultaten van de zelfevaluaties hoog (namelijk vijf scores van 4 of 5 punten). Slechts één uitschieter is te vinden bij student S8, van wie de eigen beoordeling 2 punten bedraagt. Net als bij het criterium *substituties* liggen de zelf- en peerevaluaties min of meer in elkaars verlengde. Het grootste verschil bedraagt 1 punt bij S4. Ook wanneer de cijfers van de tolkstudenten hoger uitvallen dan die van de evaluerende respondenten, wijken de scores niet erg van elkaar af (in drie van de vier gevallen betreft het een verschil van slechts 0.2 punten).

Open vraag

De tendens in de grafiek komt terug in de aanvullende motivering. Student S7 is bijvoorbeeld zeer tevreden over de afwezigheid van stopwoorden en dit wordt bevestigd door de beoordelaars (“Geen stopwoorden/geluiden, amper ‘euh’” (P10), “Wat betreft stopwoorden en pauzes heb ik geen commentaar, dit was heel goed.” (P11)) Tolkstudent S2 heeft zelf iets aan te merken op zijn frequente gebruik van het stopwoord/stopgeluid ‘euh(m)’. Deze opmerking komt tevens terug in alle peerevaluaties. Hetzelfde geldt voor S8, bij wie twee beoordelaars zelfs dezelfde omschrijving “hakkelig” gebruiken.

4.2.5 Zelfverbeteringen



Figuur 5

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *zelfverbeteringen*

Data

Geen enkele tolkstudent geeft zichzelf een maximumscore voor het criterium *zelfverbeteringen*. Wel bereikt de helft van de zelfevaluaties een hoge score van 4 punten. Net als bij omissies, zorgen ook hier tolkstudenten S4 en S6 voor grote uitschieters. De verschillen tussen de zelf- en peerevaluaties lopen op van 0.3 tot 1.8 punten.

Open vraag

Bij tolkstudenten met een lagere score, speelt het standpunt van de beoordelaar tegenover deze zelfverbeteringen een rol. Wanneer de verbeteringen als overbodig ervaren worden, treden de beoordelaars strenger op.

P2 over S4 (score 3): “Een paar keer jezelf hernomen waar je door kon, 2x een hele zin opnieuw gestart dat viel op”

P4 over S4 (score 1): “veel onnodig”

P12 over S7 (score 3): “De Meerderheid [*sic*] van de verbeteringen zijn onnodig, daarom bestraf ik dat meer.”

Indien de zelfverbeteringen de beoordelaars niet hinderen, ligt de toegekende score over het algemeen tamelijk hoog.

P8 over S2 (score 4): “Zelfverbeteringen [...] zaten er ook in maar waren niet te storend.”

P13 over S8 (score 4): “[...] zelfverbeteringen waren oké, niet storend.”

Dezelfde waar te nemen tendens als bij *substituties*, doet zich ook voor bij het criterium *zelfverbeteringen*. Opnieuw neemt niet elke beoordelaar hetzelfde aantal zelfverbeteringen waar.

S7 (score 3): “verschillende keren zichzelf verbeterd”

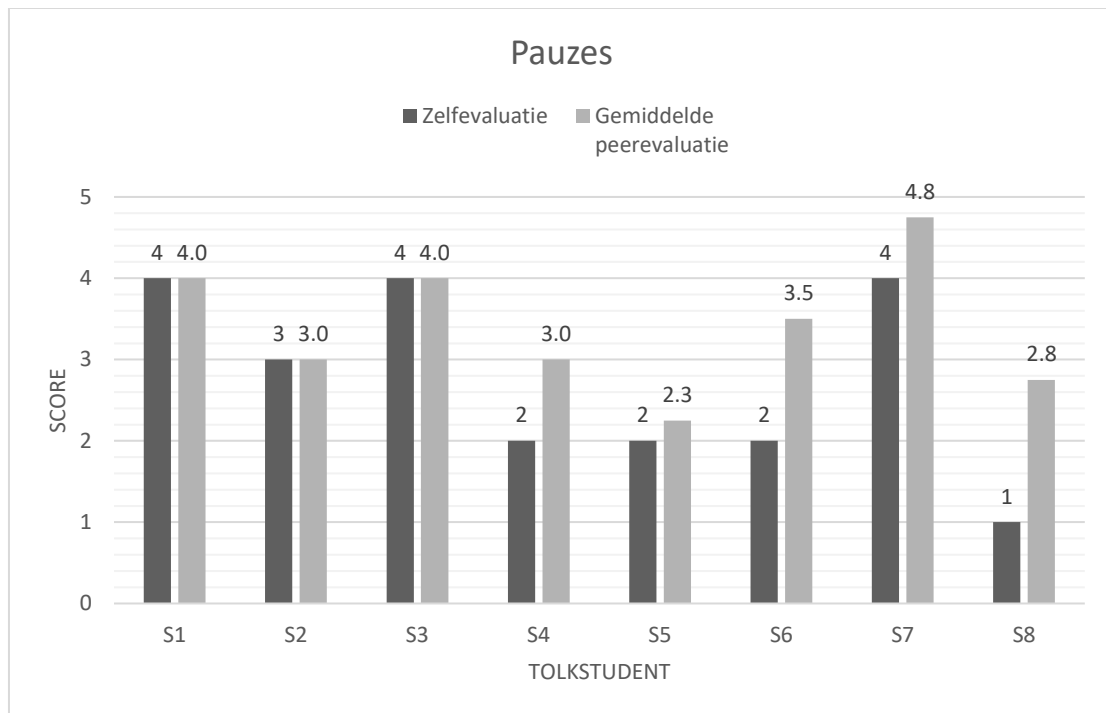
P10 over S7 (score 4): “herneemt zichzelf niet echt”

P14 over S8 (score 4): “Ze werkt wel haar zinnen af dus geen zelfverbeteringen.”⁵

P15 over S8 (score 2): “Erg veel hernemingen.”

⁵ De opmerking van P14 wekt echter vragen op: een tolk kan perfect zijn zinnen afwerken, maar zich in het begin een paar keer hernemen. De ene opmerking (zinnen afwerken) sluit met andere woorden de andere (zelfverbeteringen) niet uit. Het lijkt erop dat deze beoordelaar de twee kenmerken met elkaar verward. Het criterium *zelfverbeteringen* dient dus nog duidelijker omschreven te worden in het glossarium.

4.2.6 Pauzes



Figuur 6

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *pauzes*

Data

Ook al geven drie beoordelaars zichzelf de op één na hoogste score, toch vallen ook de vier lage scores onmiddellijk op: drie studenten (waaronder opnieuw studenten S4 en S6) kennen zichzelf een score van 2 punten toe, tolkstudent 8 kiest voor de laagste score. Bij die laatste student treedt overigens een groot contrast op tussen de zelf- en peerevaluaties: de laagst mogelijke score die de tolkstudent zichzelf toekent, wijst op een zeer strenge houding tegenover zijn gebruik van pauzes. Drie beoordelaars geven daarentegen een score van 3 punten. Niemand van de respondenten gunt zich de maximumscore. Opnieuw variëren de verschillen tussen de ongelijke zelf- en de peerevaluaties van 0.3 tot 1.8 punten.

Open vraag

Ook bij de aanvullende uitleg komt die kritische houding van de tolkstudent naar voren. Student S6 geeft zichzelf bijvoorbeeld een lage score (score 2) en schrijft: “Ik pauzeer toch wel regelmatig maar wel niet altijd té lang.” Beoordelaar P12 is het met deze stelling eens (“Een paar pauzes, maar bijna nooit lang”), maar geeft wel dubbel zoveel punten. Twee andere beoordelaars lijken dan weer helemaal geen pauzes op te merken:

P13 over S6 (score 4): “Niets storends, aangenaam om naar te luisteren.”

P14 over S6 (score 4): “Ze sprak vlot in die zin dat ze amper of niet aarzelde of haperde.”

Student S2 en zijn beoordelaars lijken het gebruik van pauzes erg verschillend te ervaren. Zo gelooft de tolkstudent dat hij “niet genoeg tijd laat tussen zinnen” (S2) en dus meer pauzes zou moeten inlassen om het spreektempo natuurlijker en aangenamer te maken. Twee van zijn beoordelaars echter vinden dat de student juist spaarzamer mag omgaan met stiltes:

P5 (score 2): “Sommige pauzes waren misschien wat lang en op de verkeerde momenten geplaatst wat het begrip van de boodschap licht ondermijnde.”

P6 (score 3): “[...] afgewisseld door (soms langere) pauzes.”

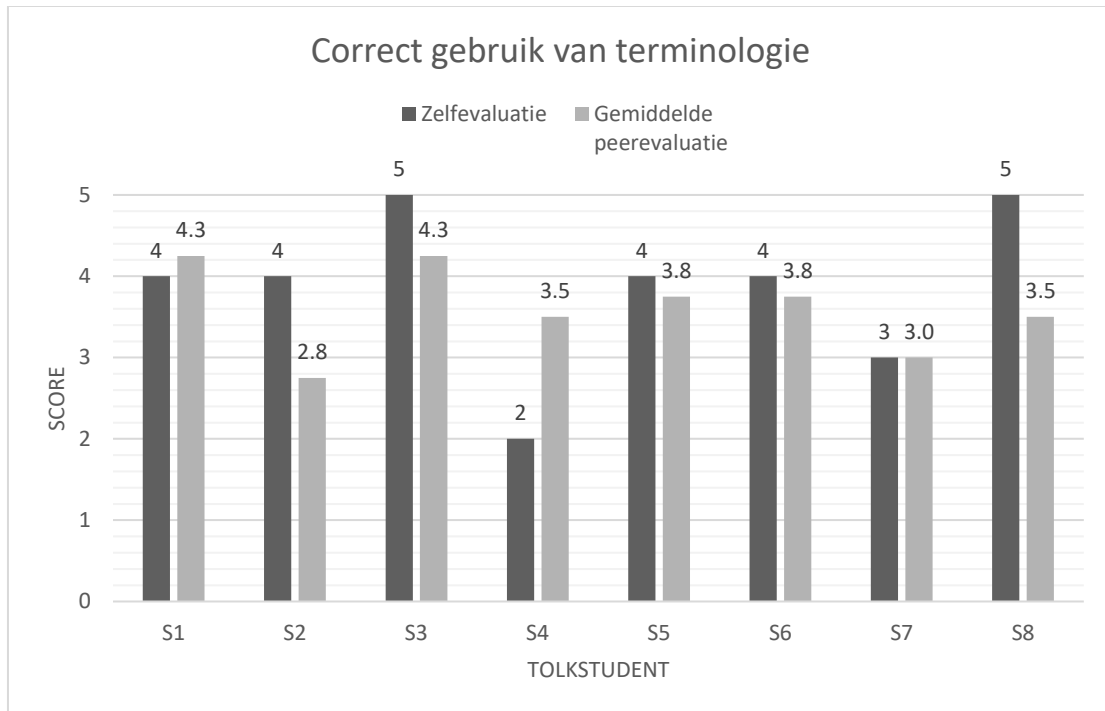
Pauzes kunnen in bepaalde omstandigheden wenselijk zijn, bijvoorbeeld om de gedachten en de structuur van de tekst te kunnen ordenen (zie 3.2.3). Onder meer de bovenstaande opmerking van beoordelaar P5 staft de stelling dat lagere of middelmatige scores in nauw verband staan met de lengte en de plaats van de pauzes in de zin.

P3 over S1 (score 3): “Soms vreemde pauzes in het midden van een zin.”

P15 over S5 (score 2): “Doordat de tolk probeert zich niet te hernemen vallen er vaak vrij lange pauzes. Dit geeft de tolkprestatie een onnatuurlijk effect.”

Ten slotte blijkt opnieuw het standpunt over het criterium van belang te zijn. Een derde respondent die student S2 hoorde te evalueren sluit zich aan bij de mening van de beoordelaars P5 en P6, maar voegt aan zijn commentaar toe: “[...] lange pauzes zaten er ook in maar waren niet te storend.” (P8) Bij het criterium *pauzes* geeft hij een relatief hoge score van 4 punten.

4.2.7 Correct gebruik van terminologie



Figuur 7

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *correct gebruik van terminologie*

Data

Bij een globale analyse van de resultaten vallen onmiddellijk de hoge cijfers op: zes tolkstudenten beoordelen zichzelf met een score van 4 of 5 punten. Enkel de score van S4 strandt op slechts 2 punten en vertoont een duidelijk verschil met de peerevaluatie (1.5 punten).

Open vraag

Opvallend is het kleine aantal foutieve collocaties en woorden dat wordt opgemerkt door enkele tolkstudenten in vergelijking met hun beoordelaars. Student S2 (met als zelfevaluatie score 4) geeft als eigen bemerking: “behalve “dingue” denk ik dat het goed was”. De andere beoordelaars merken echter nog andere elementen op:

P5 (score 3): “[...] dengue, ‘ordinaire mug’ ipv [*sic*] ‘gewone mug’, ‘vallen van de avond’ is in NL meer ‘zonsondergang’ of ‘schemering’, ‘performant’ ipv [*sic*] presterend”

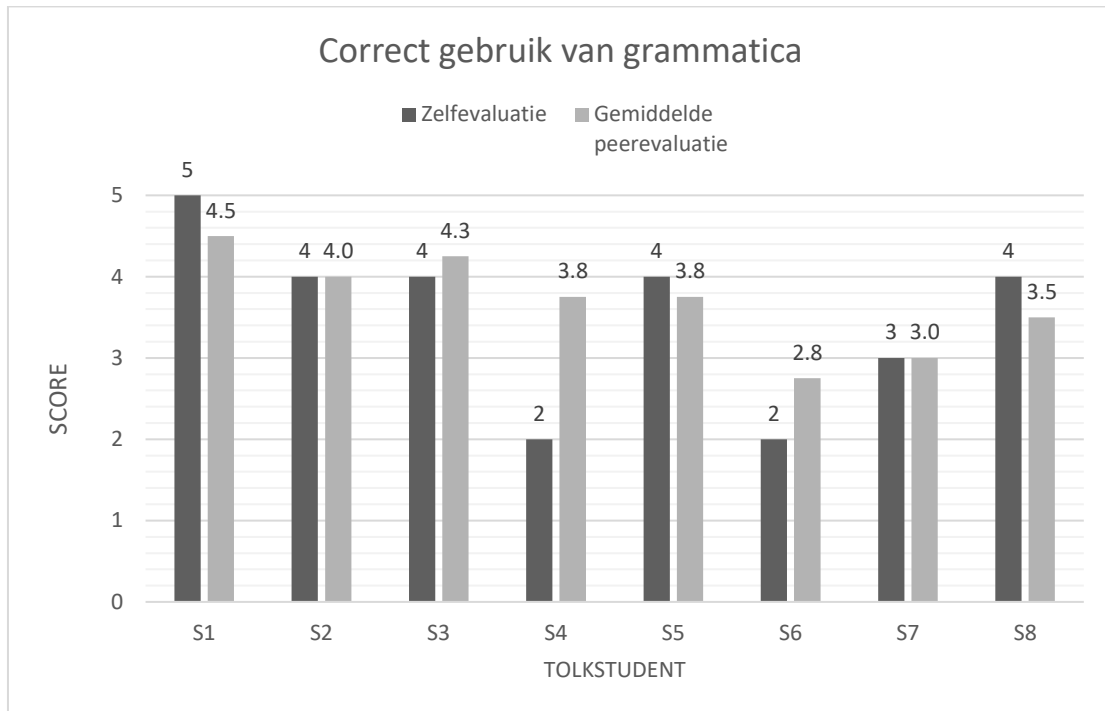
P6 (score 3): ““een gestreept aspect” → ze zijn gestreept, ordinair i.p.v. gewoon, andemisch i.p.v. endemisch”

P7 (score 3): ““specimenen = ? “ordinaire mug” = normale “andemische gevallen””

Hetzelfde geldt voor tolkstudent S7, die als uitleg enkel “heel oude mensen” geeft. Beoordelaar P9 spreekt daarentegen over “** agricultuur”, P11 over “(vb. mensen zichzelf vermenigvuldigen)” en P12 schrijft onder meer “Ook gemakkelijke termen worden fout vertaald. quickly → vaak [...]”.

Daarnaast lijkt voor deze parameter het aantal opmerkingen geen verband te houden met de toegekende score. Hoewel drie respondenten slechts één element aanhalen in de vertolking van S6 – “Package holidays → grote vakanties” (P12), “altitude ≠ attitude” (P13), “Geen woorden die misplaatst zijn, alleen “package holiday” had niet de [*sic*] juiste equivalent.” (P14) –, geven ze elk een ander cijfer: respectievelijk 5, 3 en 4 punten.

4.2.8 Correct gebruik van grammatica



Figuur 8

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *correct gebruik van grammatica*

Data

Opnieuw liggen de zelfevaluaties tamelijk hoog: vijf tolkstudenten geven zichzelf een score van 4 of 5 punten. Ook studenten S4 en S6 scoren nogmaals behoorlijk laag (2 punten). Met uitzondering van S4, vertonen de cijfers van noch de lagere, noch de hogere zelfevaluaties grote verschillen met die van de peerevaluaties (verschillen van 0 tot 0.8).

Open vraag

Bij meer dan de helft van de tolkstudenten (vijf) halen de beoordelaars nogmaals verschillende elementen aan, die niet noodzakelijk door de student werden opgemerkt. Enkele voorbeelden:

S3 (score 4): “fruit en groenten ipv [*sic*] groenten en fruit”

P6 (score 4): “‘Het wil ons aanzetten te handelen’ → ertoe aanzetten of aanzetten tot”

P7 (score 4): “‘ik hoop voor ons alle’”

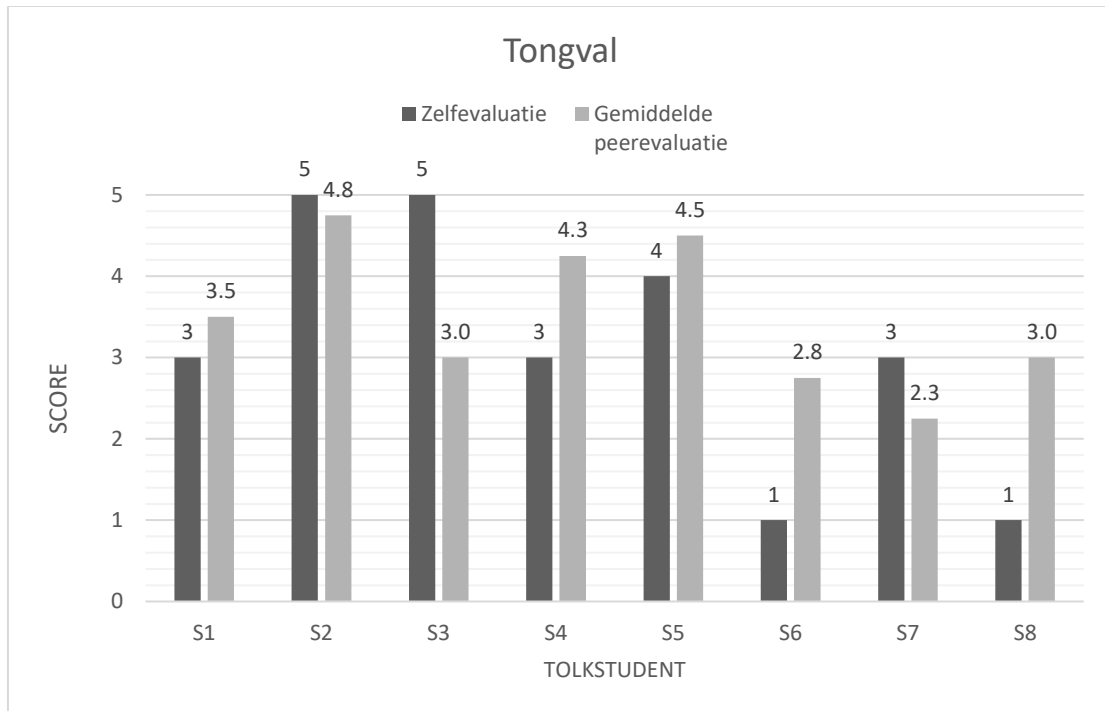
S8 (score 4): “De prijzen stegen zeer hoog” en “De prijzen van het voedsel waren tot een astronomisch hoge prijs gestegen”

P13 (score 4): ““Ze konden (mv) [...] omdat ze niet zou (ev!) [...]” \bar{m} meervoud zijn.”

P15 (score 2): “1 keer zelfs het ww van de zin vergeten”

Enkel bij student S6 kwamen de zelf- en peerevaluaties inhoudelijk goed overeen. Het element waarop de tolkstudent kritiek had (“Ik heb toch 3 à 4x een fout verwijswoord gebruikt (mensen - dat)”), kwam ook in de commentaar van elke beoordelaar terug.

4.2.9 Tongval



Figuur 9

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *tongval*

Data

De zelfevaluaties voor het criterium *tongval* geven een zeer heterogene tendens weer: drie zelfevaluaties vertonen hoge scores (waarvan twee zelfs de maximumscore), nog eens drie blijven bij een middelmatige score en studenten S6 en S8 wijzen zichzelf zelfs het laagste cijfer toe. Vier zelfevaluaties vertonen een behoorlijk verschil met het cijfer van de evaluerende respondenten (1.3 tot 2.0 punten). Bij tolkstudenten S6 en S8 is het grote verschil voornamelijk te wijten aan een uitschieter van 4 punten die het gemiddelde naar boven haalt.

Open vraag

Wanneer de beoordelaars kritiek uiten op de tongval, heeft het bij 63% van het totaal aantal respondenten betrekking op diftongen of een Brabantse/regionale uitspraak. Er is echter geen tendens waar te nemen binnen de commentaren van de studenten in vergelijking met die van de beoordelaars. Zo schrijft tolkstudent S1 over de eigen uitspraak: “diftongen” en wordt deze opmerking bijgetreden door P1 (“Af en toe een licht accent (ij, ui)”) en P4 (“[ɛi] niet als tweeklank; idem [oey]”).

Bij de aanvullende opmerkingen van en over S3 treedt er dan weer een tegenstelling op. De tolkstudent is van mening dat er geen enkele tongval te horen is en geeft zichzelf de hoogste score. Drie van de vier beoordelaars leveren echter commentaar op zowel de diftongen als de regionale uitspraak:

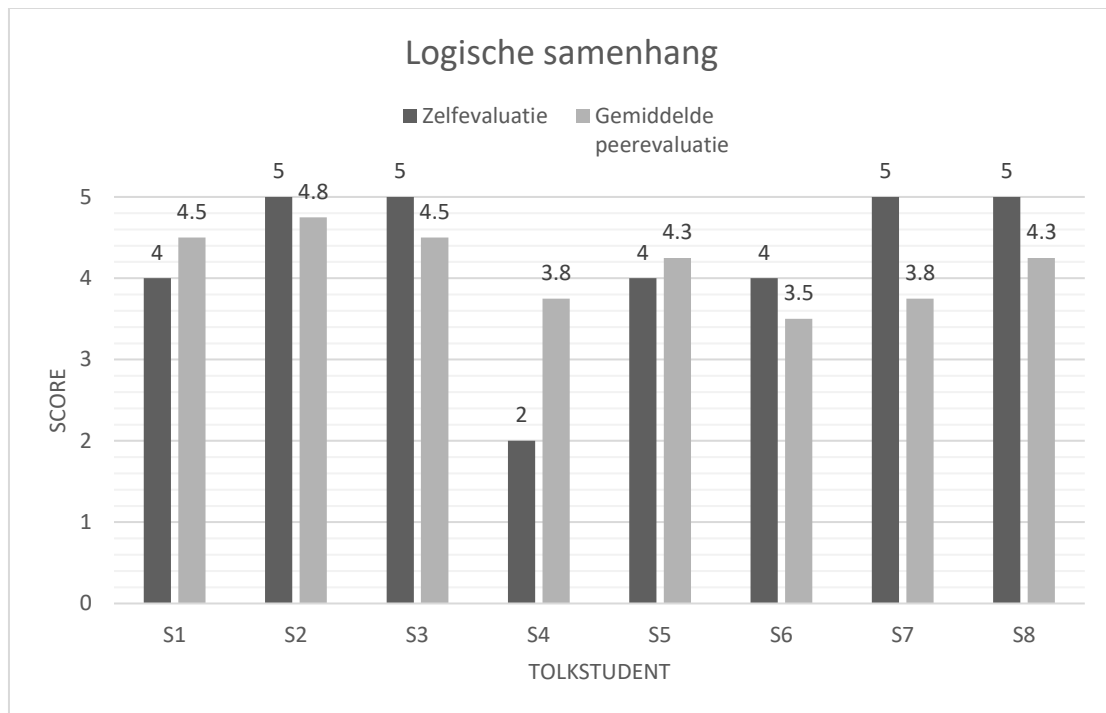
P6 (score 2): “regionale [o]-klank, diftongen kunnen iets beter, Brabantse klanken”

P8 (score 2): “klinkers erg Brabants. Dit valt op!”

P7 (score 4): “tweeklanken”

Een enkele respondent die werd gevraagd om tolkstudent S5 te evalueren, verwerkt zijn opmerking over het woordgebruik in het criterium *tongval*. Dit blijkt een invloed op het cijfer (score 4) te hebben: “Spreekt Standaardnederlands, maar bv [*sic*] woordkeus soms zoals “magerder” of “extra kilo’s oplopen”” (P10).

4.2.10 Logische samenhang



Figuur 10

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *logische samenhang*

Data

Op één uitzondering na, duiden alle tolkstudenten de hoogste of de op één na hoogste score aan. In tegenstelling tot deze respondenten, komt dit hoogste cijfer relatief weinig voor bij de peerevaluaties: van de tweeëndertig beoordelingen krijgen de vertolkingen negentienmaal een score 4 tegenover negenmaal een score 5. Opnieuw ligt de score van student S4 beduidend lager (1.8 punten) dan het gemiddelde van de beoordelaars. Ook bij tolkstudent S7 betreft het een groter verschil van 1.2 punten, al gaat het hier de omgekeerde richting uit: de gemiddelde peerevaluatie bedraagt een slechtere score dan de zelfevaluatie. De overige scores geven geen al te groot verschil aan (van 0.2 tot 0.7 punten).

Open vraag

Zowel de beoordelaars als de tolkstudenten zijn voornamelijk lovend over de tolkprestatie wat betreft het criterium *logische samenhang*: twintig opmerkingen van de beoordelaars klinken positief (waarvan zes echter aangevuld worden door een kleine negatievere opmerking, maar geen afbreuk doet aan de algemene beoordeling) en acht negatief; zeven studenten geven

zichzelf positieve commentaar (waarvan een met een bijvoegende kleine opmerking) tegenover één negatieve. Veelal maakt de aanvulling duidelijk dat het makkelijk is om ‘de rode draad van de boodschap’ te volgen.

Haast alle commentaren van de tolkstudenten en de beoordelaars komen overeen. Enkel student S7 oordeelt dat hij de “structuur van de brontekst” heeft overgenomen (score 5) terwijl twee van zijn beoordelaars dit anders waarnemen:

S7 (score 5): “structuur van brontekst”

P11 (score 4): “De hele prestatie was goed tot op het einde waar ik even de draad kwijt was.”

P12 (score 3): “Regelmatig worden foute connectoren gebruikt”

Verder verwijzen verschillende respondenten naar andere criteria om hun standpunt over de logische samenhang van de vertolking te staven. Onder meer de criteria *omissies*, *substituties* en *vlotheid van de vertolking* komen aan bod.

S1: “door substitutie bij punt 1 → een 4 i.p.v. een 5”

P1 over S4: “Door de weglatingen is de logische samenhang soms weg (dit valt enkel op als je vergelijkt met de Franse tekst)”.⁶

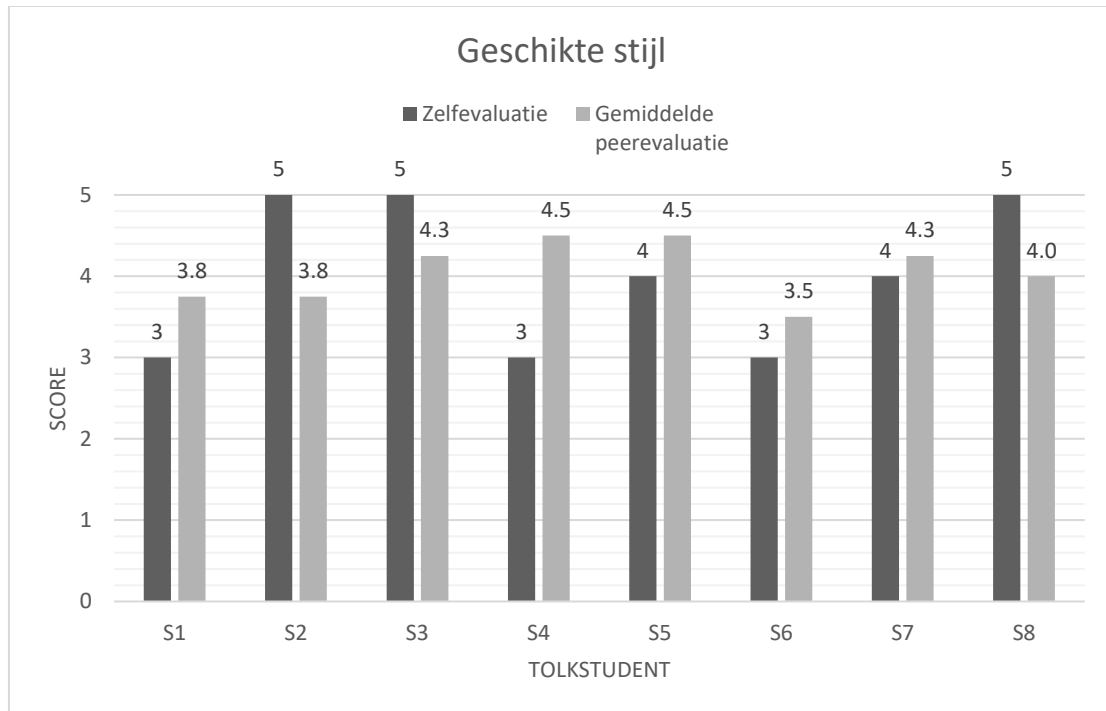
P2 over S4: “goed alhoewel door de hernemingen en verbeteringen veel moeite kosten [*sic*] om bij de rode draad te blijven”

P4 over S4: “Lastig wegens aarzelingen maar is er wel [...]”

P12 over S4: “2x onlogische samenhang door omissie [...]”

⁶ De beoordelaar bedoelt hier de oorspronkelijke speech aangezien er geen transcriptie ter beschikking was.

4.2.11 Geschikte stijl



Figuur 11

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *geschikte stijl*

Data

Voor het criterium *geschikte stijl* is geen enkele lagere score te bespeuren: het laagste cijfer van de zelfevaluaties bedraagt 3 punten. Drie studenten kennen zichzelf de hoogste score toe, twee respondenten opteren voor de op één na hoogste score. De verschillen tussen de zelf- en peerevaluaties wijken sterk van elkaar af: van een verschil van 0.3 punten bij S7 tot 1.5 punten bij S4.

Open vraag

De strengere houding van de tolkstudent komt duidelijk naar voren bij S6, waar zowel de student als een beoordelaar een gelijkaardige opmerking geven over de informele stijl, maar de tolkstudent verwoordt deze als een negatieve opmerking en de beoordelaar als een positieve:

S6 (score 3): “De brontekst was al niet zo formeel maar ik heb het iets te informeel gemaakt”

P14 (score 4): “Iets aan de informele kant, maar over het algemeen een neutraal taalgebruik.”

Beoordelaar P13 geeft een verklaring voor de ongeschikte stijl: “Kon “opgesmukt” worden zoals in brontekst het geval is. Veel adjectieven, tussenwerpsels en gedachten vielen weg waardoor de stijl werd aangepast.”⁷

Een verschillende mening van de student ten opzichte van de beoordelaar valt op bij S2:

S2 (score 5): “Ik denk dat mijn register goed was”

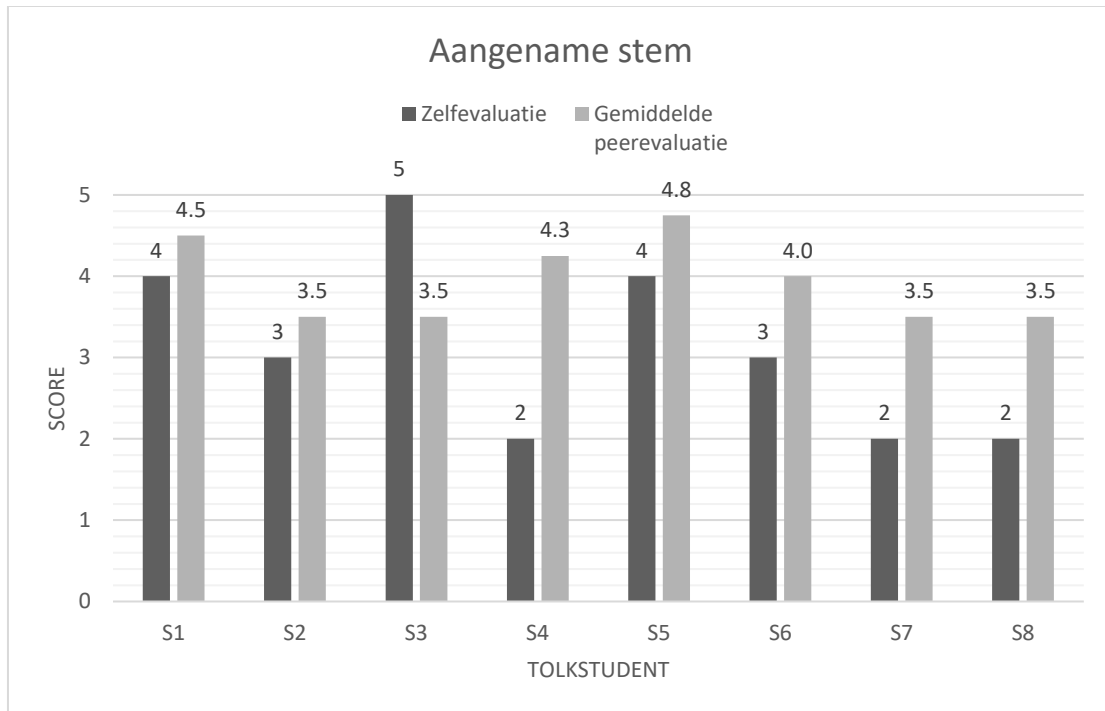
P8 (score 2): “Vreemde woordkeuzes → stijlbreuken wetenschappelijk vs laag register”

Niet alle beoordelaars blijken de zelfkritiek van de tolkstudent op te merken. Zo vindt student S1 van zichzelf dat er te veel afwisseling is in zijn gebruik van het persoonlijk voornaamwoord (“soms je, soms u: niet consequent” (S1)). Slechts één beoordelaar is dezelfde mening toegedaan, de andere drie beoordelaars vermelden niets over dit criterium.

Een enkele beoordelaar hield tevens erg vast aan de stijl van de oorspronkelijke spreker en verwachtte van de tolkstudent dat hij die ook imiteerde: “Stijl is geschikt, maar te neutraal vergeleken met de brontekst.” (P13 over S8) Student S8 vindt echter van zichzelf dat hij deze taak wel degelijk naar behoren heeft uitgevoerd: “aangepast aan het register van de spreker”.

⁷ Wat bedoelt de student echter met “aangepast”? Is het register gewijzigd van een informele naar een formelere stijl of omgekeerd? Is het ludieke karakter weggefallen?

4.2.12 Aangename stem



Figuur 12

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *aangename stem*

Data

De resultaten van de zelfevaluaties voor het criterium *aangename stem* zijn tamelijk heterogeen: drie tolkstudenten geven zichzelf niet meer dan 2 punten, twee andere scores reiken tot 4 punten. Student S3, die meermaals bij de hogere scores verschijnt, geeft zichzelf als enige de hoogste score. Bij een vergelijking tussen de zelfevaluaties en de gemiddelde score van de peerevaluaties springen de grote verschillen in het oog: de zelfevaluatie van S3 ligt 1.5 punten hoger dan de gemiddelde peerevaluatie, bij studenten S7 en S8 hetzelfde aantal punten lager. Het grootste verschil is opnieuw waar te nemen bij tolkstudent S4: de medestudenten geven gemiddeld 2.3 punten meer.

Open vraag

De negatieve houding van de studenten tegenover hun eigen stem is terug te vinden in de aanvullende uitleg. Onder meer S4 is zeer streng op de eigen stem (“Ik ratel veel, maak een nerveuze indruk”), terwijl slechts één beoordelaar een lagere score geeft (score 3) en opmerkt

dat de stem van de tolkstudeant “te schel” (P1) is. Ook bij S6 valt die kritische houding op in vergelijking met de mede- en postgraduaatstudenten:

S6 (score 3): “Ik heb een nogal hese, nasale stem in mijn oren”

P12 (score 4): “Er is sprake van een nasale stem, maar die stoort niet”

P13 (score 5): “Zeer aangenaam.”

P14 (score 4): “Haar stem stoort me evenmin.”

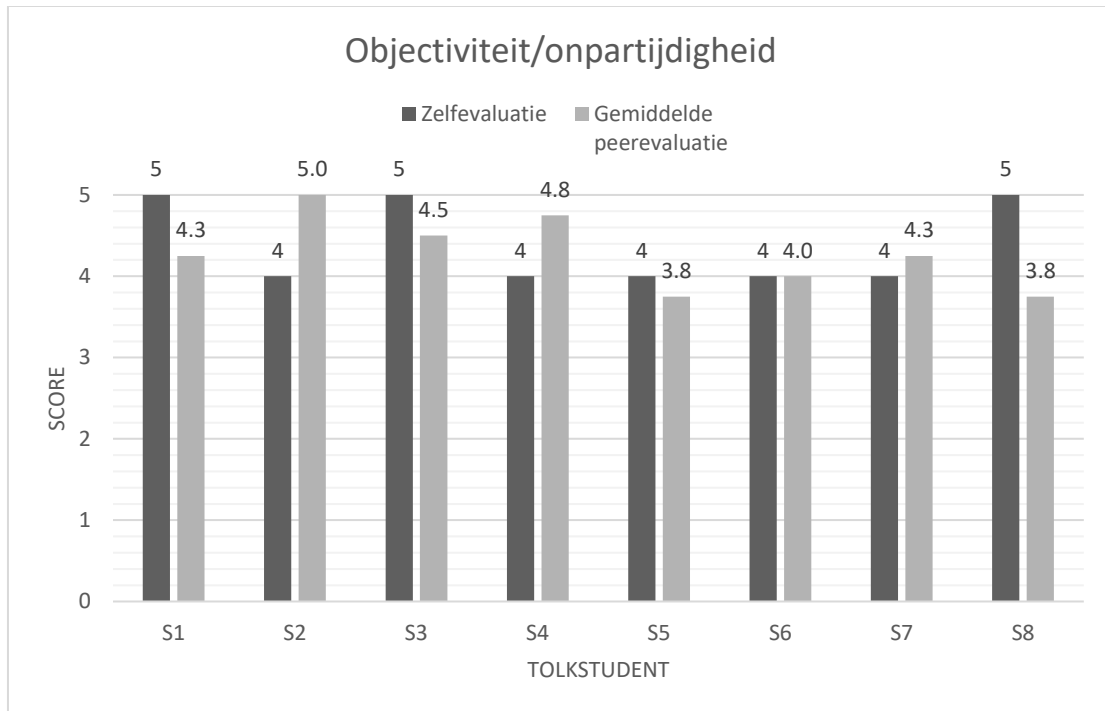
Een gebrek aan intonatie wordt in elk geval als een kritiek punt behandeld tijdens de evaluatie van het criterium *aangename stem*:

P6 over S2 (score 3): “De stem op zich is niet storend, maar door de afwezigheid van intonatie komt het minder aangenaam over.”

P8 over S2 (score 3): “zeurderig”

P6 over S2 (score 3): “Hangt samen met de afwezigheid van intonatie.”

4.2.13 Objectiviteit/onpartijdigheid



Figuur 13

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *objectiviteit/onpartijdigheid*

Data

De tolkstudenten nemen een zeer tolerante houding aan ten opzichte van dit kwaliteitscriterium: drie scores kloppen af op de maximumscore, de overige vijf evaluaties scoren telkens 4 punten. Zelfs studenten die zichzelf over het algemeen negatief beoordelen (zoals S4, S6 en S8) kruisen voor het criterium *objectiviteit/onpartijdigheid* een hoge score aan. Ook de peerevaluaties vertonen deze zelfde homogene resultaten. De beoordelaars van S2 geven zelfs alle vier de hoogste score aan.

Open vraag

De tolkstudenten hebben over het algemeen haast niets op hun objectiviteit aan te merken en de commentaren klinken uitsluitend positief. Enkele beoordelaars halen, in tegenstelling tot de studenten, wel de parameters *substituties* en *toevoegingen* aan om hun score voor het criterium *objectiviteit/onpartijdigheid* te staven.

S5 (score 4): “De meningen waren de meningen van de spreekster.”

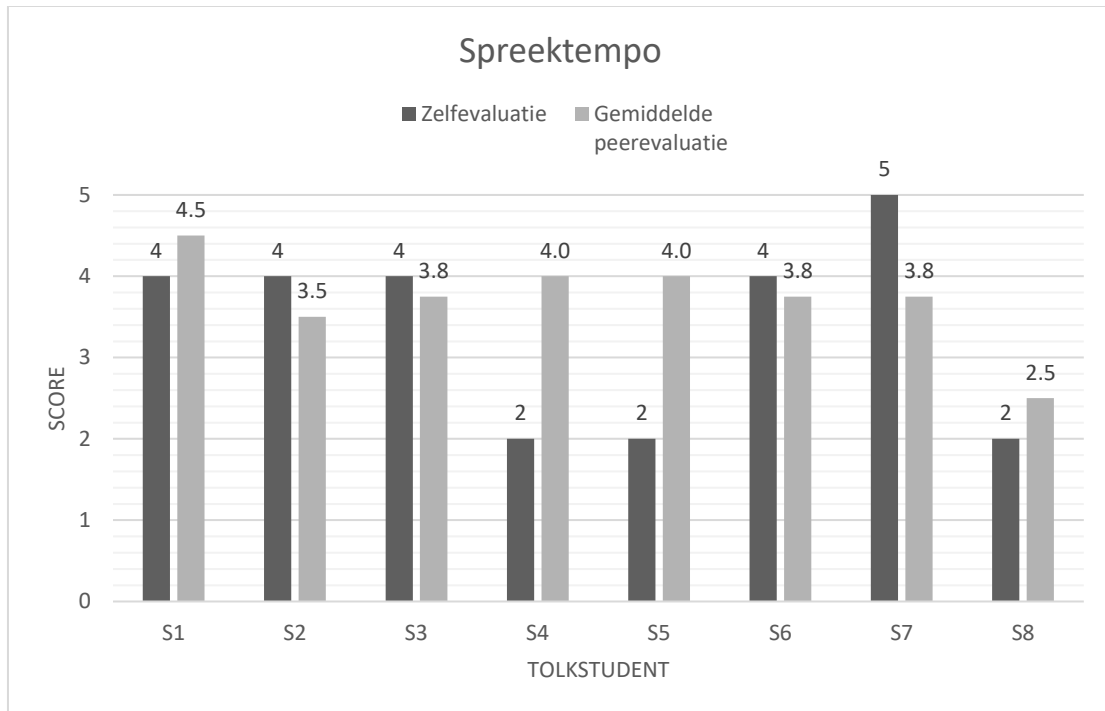
P10 over S5 (score 3): “Door de toevoegingen was de gehele versie niet volledig
waarheidsgetrouw.”

P11 over S5 (score 4): “De tolk was objectie [*sic*] maar door 1 fout is dit iets minder.”

P7 over S3 (score 4): “door Tupperware te vervangen door bewaardoos”⁸

⁸ Beoordelaar P7 doelt met zijn uitspraak echter op een veralgemening. In de oorspronkelijke speech die aan student S3 werd toegekend over het overmatig gebruik van plastic, worden enkele suggesties gegeven om de plasticen verpakkingen te verminderen, waaronder ‘aluminium vervangen door een Tupperware’. Tolkstudent S3 veranderde deze term in ‘bewaardoos’, een veel algemenere term dan de vertrouwde merknaam. Dit staat bekend onder de term ‘generalisatie’ en is een erkende vertaalstrategie (Molina & Hurtado Albir, 2002, p. 510).

4.2.14 Spreektempo



Figuur 14

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *spreektempo*

Data

Hoewel de gemiddelde peerevaluaties een hogere graad van homogeniteit vertonen, variëren de zelfevaluaties van 2 tot 5 punten. Opnieuw geven drie tolkstudenten zichzelf een lagere score. Vijf zelfevaluaties vertonen evenwel een hogere score, waaronder S6 die zichzelf doorgaans een lagere score toekent. Een opvallende vaststelling is dat twee studenten (S4 en S5) zichzelf precies de helft geven van de scores toegekend door de beoordelaars.

Open vraag

Onder meer bij tolkstudent S2 stemmen de commentaren van de tolkstudent en de beoordelaars niet overeen. Zo beschrijft S2 zijn eigen spreektempo als “[v]rij goed”. Elke beoordelaar geeft echter een opmerking over de pauzes in de vertolking van S2 en bij twee studenten hebben deze een negatieve invloed op het resultaat:

P8 (score 3): “beetje te traag”

P7 (score 3): “wordt vertraagd door pauzes en aarzelingen”

Bij S4 echter worden deze pauzes vermeld, maar niet verrekend in de score en niet als storend ervaren:

P4 over S4 (score 5): “Zeer aangenaam indien niet aarzelend”

P5 over S4 (score 4): “afgezien van wat kleine soms ongelukkig gevallen pauzes was het tempo mooi en vlot.”

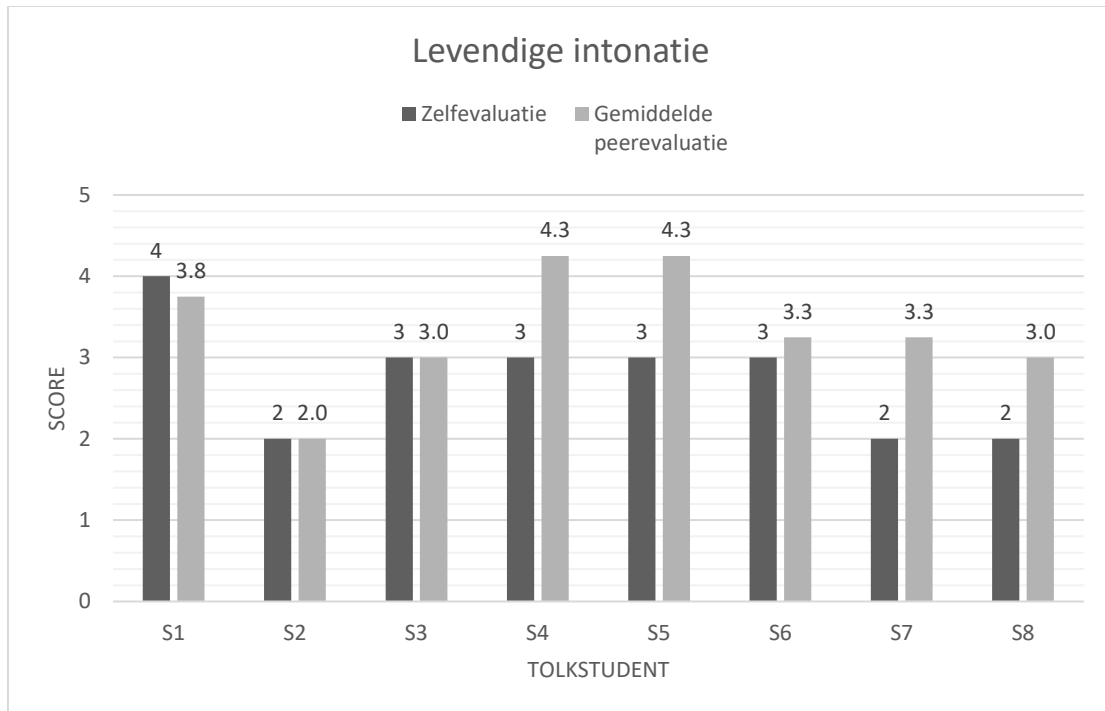
Deze laatste opmerking maakt nogmaals duidelijk dat de timing van de pauzes belangrijk is. Bovendien wijzen bovenstaande bemerkingen uit dat de criteria *pauzes* en *spreektempo* verwant zijn en een wederzijdse invloed op elkaar uitspelen. De tolkstudent zelf ergert zich daarentegen wél aan het eigen tempo: “Soms te lange pauzes, andere keren veel te snel, gerateld.” (S4, score 2) Ook bij S5 treedt deze tegenstelling op:

S5 (score 2): “Het was te traag naar mijn mening, mede doordat de spreekster zelf rustig sprak.”

P10 (score 4): “Goed tempo, afgezien van enkele aarzelingen.”

P11 (score 4): “Zonder de pauzes meegerekend was het tempo in orde.”

4.2.15 Levendige intonatie



Figuur 15

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *levendige intonatie*

Data

Over het algemeen geven de tolkstudenten zichzelf geen hoge scores voor het criterium *levendige intonatie*: slechts één uitschieter bereikt 4 punten, de overige evaluaties bedragen 2 of 3 punten. Ook student S3, die voor zichzelf doorgaans zeer tolerant optreedt, duidt een middelmatige score aan. Door de iets mildere houding van S4, komen beide tolkstudenten hierdoor op een gelijke plaats te staan. Verder vertoont de helft van de zelfevaluaties een groot verschil met de gemiddelde peerevaluatie (van 1.0 tot 1.3 punten).

Open vraag

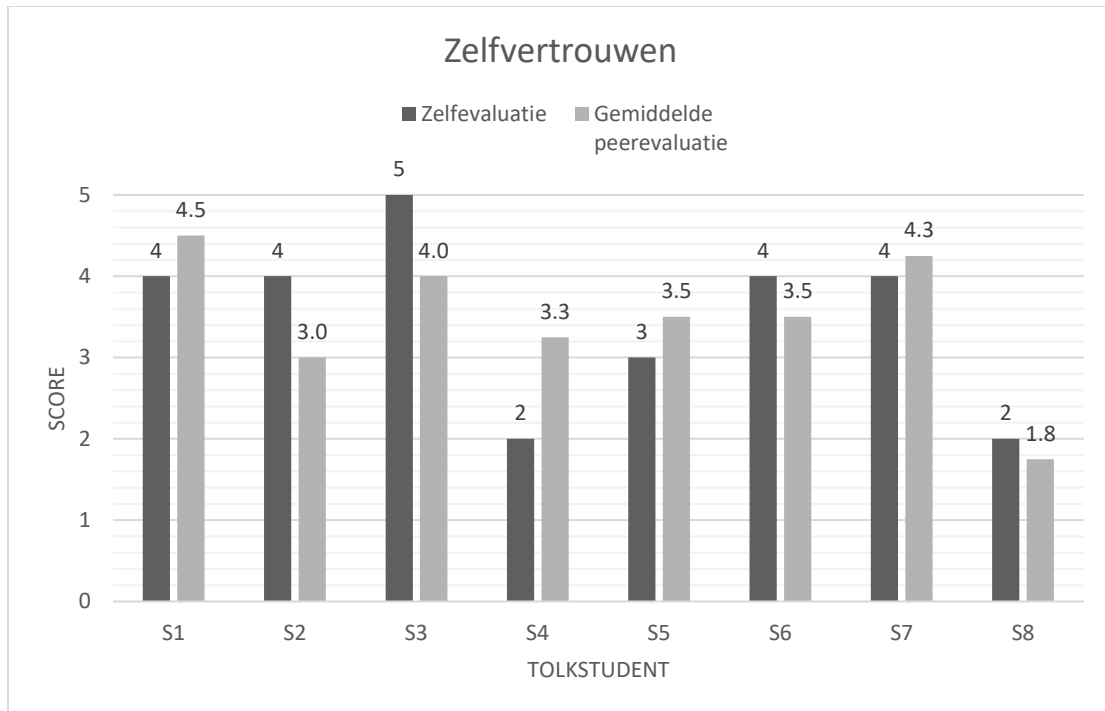
De commentaren van de tolkstudenten blijken overwegend negatief (voornamelijk de opmerking “te monotoon” duikt vaak op); geen enkele student heeft uitsluitend lovende commentaar over dit aspect van zijn tolkprestatie. Ook veertien peerevaluaties halen enkel negatieve aspecten aan. Beoordelaar P3 tracht een verklaring te vinden voor die lage score voor S3 (score 2): “monotoon, misschien te weinig vertrouwen?”

Verder leggen twee tolkstudenten een verband met hun ingelaste pauzes, een criterium dat niet wordt besproken door de beoordelaars:

S8 (score 2): “monotoon; pauzes op foute plaatsen zorgen voor een foute intonatie”

S6 (score 3): “De pauzes hebben dit iets verminderd, maar al bij al wel ok”

4.2.16 Zelfvertrouwen



Figuur 16

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *zelfvertrouwen*

Data

Op drie uitzonderingen na, geven de tolkstudenten zichzelf een 4 voor het criterium *zelfvertrouwen*. Een enkele zelfevaluatie scoort 5 punten (net als bij de meeste criteria betreft het S3). Zelfs de zelfevaluatie van student S6 bereikt de op één na hoogste score. Negatieve uitschieters zijn te vinden bij S4 en S8. Bij S4 is er sprake van een groot verschil tussen beide evaluaties (de zelfevaluatie ligt 1.3 punten lager); bij S8 daarentegen ligt de peerevaluatie sterk in lijn met de eigen beoordeling.

Open vraag

Bij enkele tolkstudenten (onder meer S5 en S8) komen de commentaren van de zelf- en de peerevaluaties sterk overeen. Bij de analyse van de andere studenten is er daarentegen niet altijd een duidelijk waarneembare tendens. Uit de aanvullende uitleg blijken S6 en een van de beoordelaars het eens te zijn over de prestatie: “Ondanks de pauzes vind ik dat ik zelfzeker overkwam” (S6, score 4) en “klinkt geen onzekerheid door” (P12, score 5). Beoordelaar P14

echter is een andere mening toegedaan: “Ze mocht eveneens iets meer zelfvertrouwen tonen aangezien de inhoud volledig genoeg was.” (score 3)

Bij student S2 is er evenmin overeenstemming tussen de verschillende evaluaties:

S2 (score 4): “Zit wel goed bij mij”

P6 (score 3): “Door de aarzelende geluiden en pauzes komt de boodschap iets minder geloofwaardig over.”

P8 (score 2): “lijkt of hij steeds naar zijn woorden moet zoeken”

P7 (score 3): “intonatie gaat vaak ↑, komt onzeker over + aarzelingen”

De strenge houding van de tolkstudent tegenover een mildere beoordeling van de beoordelaars komt wél duidelijk naar voren bij S4:

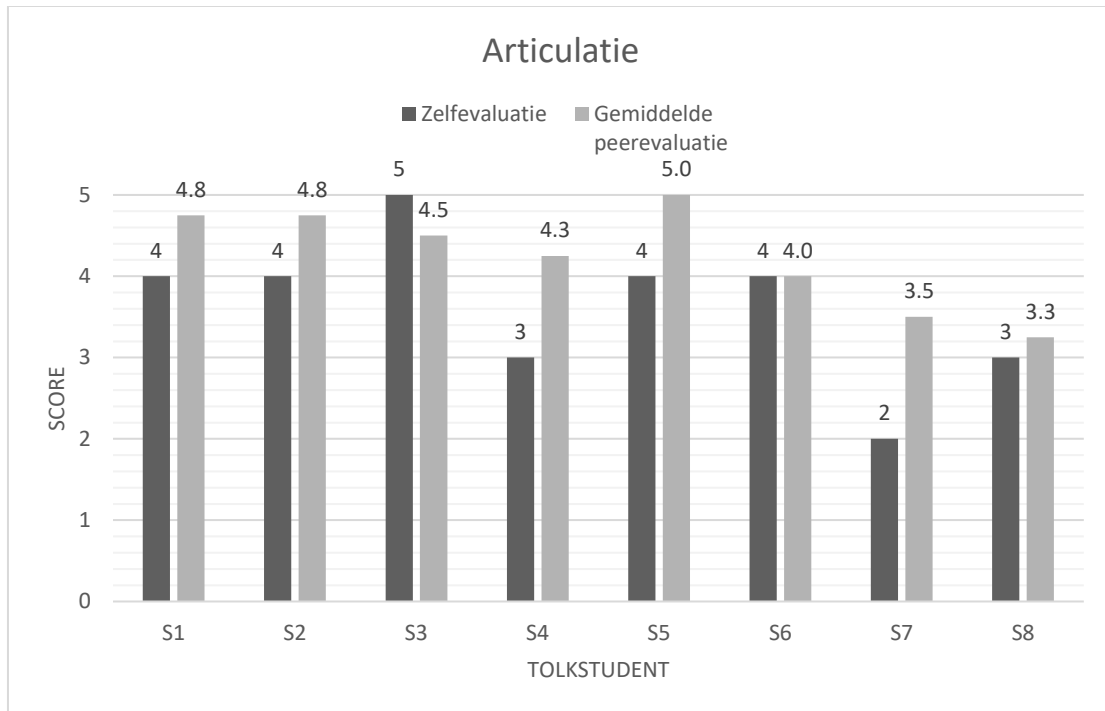
S4 (score 2): “Door de langere pauzes komt mijn vertolking niet altijd geloofwaardig over.”

P2 (score 3): “Soms niet 100% vertuigd [*sic*] of lijkt dat maar doordat je naar je woorden zoekt.”

P4 (score 4): “Aarzelingen best snel weggemoffeld onder overtuigende taal”

P5 (score 4): “De paar euhms en het hernemen van de tolk hebben de zelfvertrouwens-uitstraling iets afgezwakt, maar deze stond in eerste instantie goed”

4.2.17 Articulatie



Figuur 17

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor het kwaliteitscriterium *articulatie*

Data

Bij de beoordeling van de articulatie, kent enkel S7 zichzelf een lagere score toe. Er treden nogmaals vijf sterke zelfevaluaties op van 4 of 5 punten. Onmiddellijk trekken echter ook de hoge scores van de medestudenten de aandacht. Student S3 geeft zichzelf als enige de hoogste score.

Open vraag

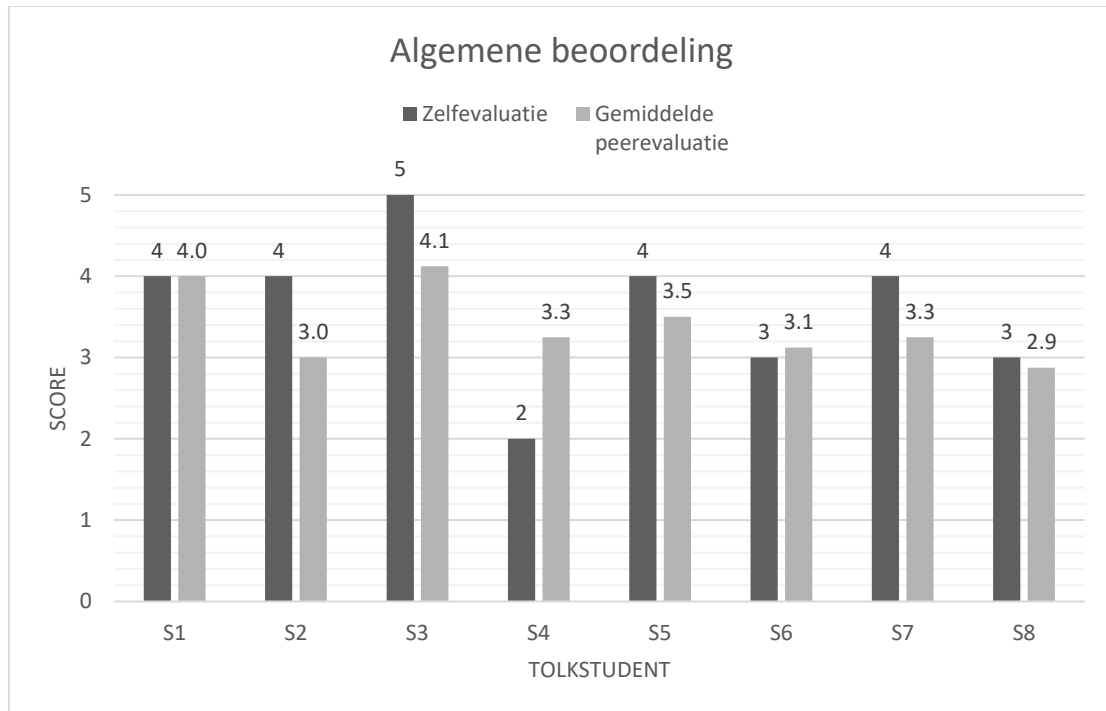
Opnieuw treedt tolkstudent S4 op als een strenge zelfevaluator: “Soms praat ik te snel waardoor ik niet altijd even verstaanbaar ben.” (score 2) Een van de beoordelaars heeft een heel andere mening over diezelfde vertolking: “Slikt geen lettergrepen e.d. in, prima” (P4, score 5). Wanneer negatieve elementen worden aangehaald, worden deze in verband gebracht met een ander criterium: “Zoals ik al eerder schreef, wat mij betreft niet hinderlijk maar er is wel een accent aanwezig.” (P5 over S4, score 3)

Ook S7 is streng op het criterium *articulatie* (score 2), en verwijst bij de aanvullende uitleg naar het criterium *tongval* (“ij/ui”), waar hij zichzelf evenmin een hoge score toekent (score 3).

De beoordelaars nemen dat element echter niet in acht tijdens de evaluatie van de parameter *tongval* en letten enkel op het “gemompel” (P11) of “ingeslikte” lettergrepen (P9) die bij enkele lettergrepen opgemerkt worden.

In de bijkomende uitleg van student S2 schemert een zekere onzekerheid over de eigen prestatie door: “Dat is toch de bedoeling van de tolk dus ik mag hopen van wel” (S2, score 4). De hoge scores (driemaal 5 punten, eenmaal 4 punten) van de beoordelaars wijzen op een positieve indruk.

4.2.18 Algemene beoordeling



Figuur 18

Vergelijking tussen de resultaten van de zelf- en de peerevaluaties voor de *algemene beoordeling*

Data

De puntenverdeling van de zelfevaluaties voor het criterium *algemene beoordeling* komen precies overeen met die voor het criterium *articulatie*. Tolkstudent S3 beoordeelt zijn algemene prestatie met de hoogste score, vier andere studenten geven zichzelf 4 punten. Enkel de beoordeling van S4 scoort laag. Met uitzondering van S2 en S4, sluiten de zelf- en de peerevaluaties nauw bij elkaar aan (met verschillen van 0.1 tot 0.9 punten).

Open vraag

De beoordelingen van de tolkstudenten klinken overwegend positief; voornamelijk de grootste ‘fouten’ worden aangehaald. Volgens de helft van de studenten werd de boodschap correct overgedragen, wat zij als een fundamentele voorwaarde lijken te beschouwen (vb. S2: “In het algemeen kwam de boodschap goed over en dat is het belangrijkste.”). De medestudenten daarentegen bespreken vaker de negatieve aspecten, al worden deze meestal op een milde manier verwoord: “Alle elementen waren aanwezig, jammer dat het zo monotoon was.” (P3) en “Vooral gezien dat we halverwege van het jaar zijn, en we ds [*sic*] waarschijnlijk nog heel wat zullen leren, is deze prestatie goed! Kon noiet [*sic*] perfect zijn door de kleine aanmerkingen

[...]” (P5). Veelal worden de negatievere aspecten geformuleerd in de vorm van een aanbeveling: “Iets meer aandacht voor connectoren zou ten slotte de vertolking kunnen verbeteren” (P6) of “De tolk kan ng [*sic*] werken aan het notitiesysteem om weglatingen te vermijden” (P10).

Enkel de tolkstudenten die zichzelf een lagere score dan de beoordeling toekennen, geven vooral negatieve commentaar en vermelden hierbij hun werkpunten (vb. “Ik moet ook leren rustiger te praten, minder te ratelen” (S4) en “Ik moet vooral nog leren doorpraten, op mijn uitspraak letten en iets beter luisteren” (S6)). Deze werkpunten blijken echter niet volledig te stroken met het advies van de beoordelaars. Zo formuleert beoordelaar P4 zijn aanbevelingen als volgt: “Meer afstand nemen van de brontekst [...]” en “Vlotheid voornaamste werkpunt.” Dezelfde waarneming geldt eveneens voor de algemene commentaar van P14 over S6, afgezonderd van de uitspraak waarover overeenstemming bestaat: “Intonatie, zelfvertrouwen en Nederlandse uitspraak zijn de enige werkpunten.”

4.3 Algemene vergelijking tussen de zelf- en peerevaluaties

Bij slechts zeven van de achttien kwaliteitscriteria (de algemene beoordeling inbegrepen) schat de meerderheid van de tolkstudenten de eigen prestatie lager in dan hun beoordelaars. Geen enkel criterium wordt door alle studenten slechter beoordeeld dan de mede- en postgraduaatstudenten. De resultaten verwerpen met andere woorden de hypothese van deze masterproef.

Tabel 2

Vergelijking tussen de zelf- en de peerevaluaties van alle criteria, gerangschikt volgens score

Kwaliteitscriteria	Aantal lagere zelfevaluaties	Aantal gelijke scores	Aantal hogere zelfevaluaties
1. Aangename stem	7	0	1
2. Zelfverbeteringen	6	1	1
3. Articulatie	6	1	1
4. Pauzes	5	3	0
5. Levendige intonatie	5	2	1
6. Tongval	5	0	3
7. Geschikte stijl	5	0	3
8. Omissies	4	2	2
9. Substituties	4	1	3
10. Stopwoorden/stopgeluiden	4	0	4
11. Spreektempo	4	0	4
12. Zelfvertrouwen	4	0	4
13. Correct gebruik van grammatica	3	2	3
14. Objectiviteit/onpartijdigheid	3	1	4
15. Logische samenhang	3	0	5
16. Toevoegingen	2	2	4
17. Correct gebruik van terminologie	2	1	5
<i>Algemene beoordeling</i>	2	1	5

De strengste zelfevaluatie uit zich bij het niet-linguïstische criterium *aangename stem*, waarbij, op één uitzondering na, elke tolkstudent zichzelf een lagere score geeft dan de beoordelaars. Dit resultaat ligt in lijn met de stelling van Kahane (2000), die aanvoert dat tolken naast hun tongval ook veel belang hechten aan hun stem (zie 2.2.1). Ook de parameters *articulatie* en *zelfverbeteringen* worden streng beoordeeld: zes van de acht tolkstudenten gunnen zichzelf minder punten dan de evaluerende medestudenten. Zowel bij de linguïstische criteria *pauzes*, *tongval*, *geschikte stijl* als het niet-linguïstische criterium *levendige intonatie* treden vijf van de acht studenten strenger op voor zichzelf. Een evenredige verdeling tussen de zelf- en peerevaluaties (de helft van de eigen beoordelingen ligt lager dan de peerevaluaties, de andere helft hoger) is merkbaar bij *stopwoorden/stopgeluiden*, *spreektempo* en *zelfvertrouwen*. Op dezelfde manier geeft de helft van de tolkstudenten zichzelf een lagere score bij *omissies* en *substituties*, maar er treden ook respectievelijk twee en één identieke score(s) op. Voor de overige vijf kwaliteitscriteria stellen de studenten zich toleranter op in vergelijking met de beoordelaars: drie studenten geven zichzelf een lagere score voor *correct gebruik van grammatica*, *logische samenhang* en *objectiviteit/onpartijdigheid* (bij deze laatste betreffen het overigens hoge scores van 4 punten); twee zelfevaluaties liggen lager bij de linguïstische criteria *toevoegingen* en *correct gebruik van terminologie*. Slechts twee zelfevaluaties van de algemene beoordeling liggen lager dan de peerevaluaties.

4.4 Totaalresultaat: gemiddelde scores per criterium

Met uitzondering van vier parameters (*correct gebruik van terminologie, logische samenhang, objectiviteit/onpartijdigheid* en *zelfvertrouwen*), zijn alle gemiddelde scores van de zelfevaluaties voor de kwaliteitscriteria lager dan of gelijk aan die van de peerevaluaties. De opvallendste evaluatieverschillen (0.5 tot 0.8 punten) treden op bij *zelfverbeteringen, pauzes, aangename stem, levendige intonatie* en *articulatie* (ter vergelijking: de onderlinge verschillen bij de overige criteria bedragen 0.1 tot maximaal 0.4 punten). Voor de algemene beoordeling echter geven de tolkstudenten zichzelf gemiddeld een iets hogere score dan de beoordelaars. De algemene prestatie wordt dus vanuit een mildere houding beoordeeld. Deze scores bevestigen de hypothese van deze masterproef.

Tabel 3

Totaalresultaat van de zelf- en peerevaluaties

Kwaliteitscriteria	Gemiddelde score zelfevaluatie	Gemiddelde score peerevaluatie
1. Omissies	3.1	3.3
2. Toevoegingen	3.9	4.0
3. Substituties	3.3	3.4
4. Stopwoorden/stopgeluiden	3.8	3.9
5. Zelfverbeteringen	3.1	3.7
6. Pauzes	2.8	3.4
7. Correct gebruik van terminologie	3.9	3.6
8. Correct gebruik van grammatica	3.5	3.7
9. Tongval	3.1	3.5
10. Logische samenhang	4.3	4.2
11. Geschikte stijl	4.0	4.1
12. Aangename stem	3.1	3.9
13. Objectiviteit/onpartijdigheid	4.4	4.3
14. Spreektempo	3.4	3.7
15. Levendige intonatie	2.8	3.3
16. Zelfvertrouwen	3.5	3.5
17. Articulatie	3.6	4.3
<i>Algemene beoordeling</i>	<i>3.6</i>	<i>3.4</i>

5 Discussie en aanbevelingen

5.1 Bemerkingen bij de resultaten

Uit de resultaten blijkt duidelijk het subjectieve karakter van de perceptie van kwaliteit (zie 4.2, tabel 1, p. 57, en tabel 4, p. 98). Zo neemt tolkstudent S4 over de hele lijn een relatief zelfkritische houding aan (tien keer een lage score, bij zes criteria is er sprake van een middelmatige score). Bij de criteria *correct gebruik van terminologie* en *logische samenhang*, waarbij er bijna enkel hoge scores voorkomen, geeft S4 zich als enige een lagere score. Ook student S8 kent zichzelf op precies de helft van de eigen beoordelingen een score van 1 of 2 punten toe; negen scores liggen lager dan de peerevaluaties. Even opmerkelijk is de vaststelling dat beoordelaar P12 dubbel zoveel punten geeft aan student S6 voor het criterium *pauzes* als S6 aan zichzelf – ook al klinkt de aanvullende uitleg nagenoeg gelijk (namelijk geen té lange pauzes). Dit resultaat getuigt evenzeer van een opvallend kritischere houding van de tolkstudent dan van de beoordelaar. Die kritische houding zou mogelijk kunnen wijzen op een gebrek aan zelfvertrouwen. Andere karaktertrekken, zoals veeleisendheid of perfectionisme, zouden een andere verklaring kunnen bieden. Het tegenovergestelde kan worden vastgesteld bij S3, die zichzelf, met uitzondering van één kwaliteitscriterium, steeds de hoogste of de op één na hoogste score toekent. Tolkstudenten S1 (veertien hoge scores), S5 (dertien hoge scores) en S2 (twaalf hoge scores) gaan dezelfde richting uit, al moet er gewezen worden op de afwezigheid van de hoogste score bij S5, in tegenstelling tot de andere studenten.

Tabel 4 toont aan dat in de groep van de beoordelaars overwegend hoge punten primeren: bij twaalf van de zestien respondenten gaat het om minstens de helft van het aantal scores. Beoordelaars P2 en P7 geven zelfs geen enkele lage score. Beoordelaar P9 springt er dan weer in het bijzonder uit door zijn lage cijfers voor beide evaluaties: bij de ene evaluatie acht keer 4 of 5 punten, bij de andere uitsluitend lage of middelmatige scores.

Tabel 4

Aantal hoge scores (4 of 5 punten) en lage scores (1 of 2 punten) toegekend door de beoordelaars

Beoordelaar	Score	Evaluatie 1	Evaluatie 2
P1	Hoge score	16	9
	Lage score	0	3
P2	Hoge score	18	12
	Lage score	0	0
P3	Hoge score	8	13
	Lage score	1	3
P4	Hoge score	14	11
	Lage score	0	1
P5	Hoge score	12	13
	Lage score	2	0
P6	Hoge score	9	15
	Lage score	2	2
P7	Hoge score	10	17
	Lage score	0	0
P8	Hoge score	10	16
	Lage score	4	1
P9	Hoge score	0	8
	Lage score	11	9
P10	Hoge score	11	13
	Lage score	2	2
P11	Hoge score	15	11
	Lage score	2	2
P12	Hoge score	14	12
	Lage score	2	0
P13	Hoge score	9	11
	Lage score	1	1
P14	Hoge score	14	12
	Lage score	0	3
P15	Hoge score	12	4
	Lage score	2	9
P16	Hoge score	13	6
	Lage score	2	8

Persoonlijke voorkeuren vallen verder op bij het criterium *aangename stem* waarbij de ene beoordelaar over S5 schrijft: “Lage stem, warm, klinkt niet saai” (P4) en een andere beoordelaar noteert: “Ik vond het niet speciaal aangenaam maar ook niet slecht.” (P11) De verschillende beoordelaars blijken daarnaast onderling dikwijls verschillende zwakke of sterke punten waar te nemen (tenzij zij niet al hun bemerkings neerschrijven) en ook de tolkstudenten merken andere elementen op. De meningen zijn met andere woorden verdeeld.

Hoewel bepaalde kwaliteitscriteria op een zeer heterogene manier geëvalueerd worden, valt bij andere parameters een aperte tendens op. Het criterium *objectiviteit/onpartijdigheid* springt onmiddellijk in het oog: elke tolkstudent geeft zichzelf een hogere score, zelfs studenten S4 en S8, die over het algemeen een kritische houding aannemen bij de zelfevaluatie (zie 4.2 en tabel 1, p. 57). Ook bij *geschikte stijl* komt geen enkele lagere score aan bod. De homogene beoordelingen/aanvullende uitleg en de hoge cijfers van *logische samenhang* (viermaal 5 punten, driemaal 4 punten), *geschikte stijl* (driemaal 5 punten, tweemaal 4 punten) en *objectiviteit* (driemaal 5 punten, vijfmaal 4 punten) doen uitschijnen dat deze criteria makkelijk te beoordelen zijn. Hoge cijfers zijn verder te vinden bij het criterium *toevoegingen*, waarbij maar liefst zes beoordelaars zichzelf de hoogste of de op één na hoogste score toekennen. De mede- en postgraduaatstudenten geven net als de tolkstudenten overwegend hoge cijfers. Bij de overige twee criteria omtrent de inhoudelijke overeenkomst met de brontekst (*omissies* en *substituties*) echter geeft slechts één student zichzelf 5 punten (in tegenstelling tot drie vijven bij *toevoegingen*). Twee argumenten kunnen deze discrepantie mogelijk verklaren: ofwel slagen de tolkstudenten er volgens zowel de tolkstudenten als de beoordelaars makkelijker in om geen informatie toe te voegen, in vergelijking met elementen weglaten of veranderen; ofwel beschouwen de respondenten *toevoegingen* als een lastige parameter (ze worden gevraagd om te evalueren wat er in wezen niet ís) en opteren ze daardoor toch maar voor de hogere punten. Zoals Seleskovitch (1986) en Bühler (1986) terecht aanhalen (zie 2.1.4.2), is de kans hoe dan ook groot dat de studenten actief op zoek zijn gegaan naar toevoegingen, aangezien ze uitdrukkelijk gevraagd werden om erop te letten.

De hoogste score is volledig afwezig bij de parameters *zelfverbeteringen* en *levendige intonatie*. Anders dan bij *zelfverbeteringen*, waar toch nog vier scores van 4 punten optreden, zijn bij de parameter *levendige intonatie* voornamelijk middelmatige scores (viermaal) en scores van 2 punten (driemaal) te vinden. Zelfs student S3, bij wie haast voor elk criterium een hogere score opvalt, geeft zichzelf een middelmatige score – overigens dezelfde score als S4, die meestal

lagere scores toekent. Ook bij *aangename stem* geven twee studenten zichzelf een middelmatige score; drie scores bereiken slechts 2 punten. Verder treden bij *pauzes* behoorlijk wat lage scores op: driemaal een score van 2 punten, één keer de laagste score. Niet alleen bij de afzonderlijke vergelijkingen, maar ook bij het totaalresultaat komt die kritische houding van de tolkstudenten ten opzichte van de vier bovenvermelde parameters naar voren: samen met *articulatie* vertonen ze de grootste verschillen tussen de zelf- en de peerevaluaties (zie 4.4). Voornamelijk de verschillen voor *aangename stem* en *levendige intonatie* zijn eenvoudig te verklaren. De negatieve opmerkingen van de tolkstudenten over de eigen stem, tonen nog maar eens aan dat luisteren naar zichzelf via een opname confronterend is. Aangezien intonatie verwant is met de stem, zouden de lage zelfevaluaties voor de parameter *levendige intonatie* op dezelfde manier verklaard kunnen worden.

Uit de aanvullende uitleg van de respondenten, lijken de criteria *tongval*, *aangename stem*, *zelfvertrouwen* en *articulatie* enkele van de moeilijkste criteria om te evalueren. De uiteenlopende meningen bij zowel de tolkstudenten als de beoordelaars onderling onderbouwen deze stelling. Zeker de articulatie van een student kan door elk individu anders waargenomen worden en de beoordeling van het criterium hangt sterk af van de invulling die de tolkstudent aan ‘articulatie’ geeft: sommigen stellen zich al tevreden wanneer er geen lettergrepen ingeslikt worden. In tegenstelling tot hun beoordelaars, lijken niet alle tolkstudenten (vb. S3) bij zichzelf een regionaal accent op te vangen. Anderzijds geven vier respondenten commentaar op hun eigen diftongen, een punt van kritiek dat dus wél makkelijker opgemerkt wordt. Ten slotte bewijst de bemerking van S2 de moeilijkheden die sommige studenten ondervinden om het criterium *aangename stem* te benoemen: “Ik weet dat ik te monotoon spreek, mijn stem ben ik gewoon dus moeilijk te zeggen” (S2).

Daarnaast stemmen de toegekende score en de aanvullende uitleg niet altijd volledig overeen, een vaststelling waartoe ook Davies (2000) kwam tijdens zijn experiment. Student S7 geeft zichzelf bijvoorbeeld ‘slechts’ een score 4 voor het criterium *toevoegingen*, maar zijn uitleg luidt als volgt: “Geen toevoegingen gevonden”. Ook bij enkele andere criteria treedt een identieke contradictie op tussen de positieve uitleg enerzijds en een score van 4 punten anderzijds: *correct gebruik van terminologie* (“Ik denk wel dat alles wat ik heb gezegd juiste vertalingen waren” (S6)), *correct gebruik van grammatica* (“Ik heb niks verkeerd [*sic*] gemerkt” (S2)), *tongval* (“Ik deed de [*sic*] moeite om AN te spreken” (S5)) en *articulatie* (“Ik was duidelijk verstaanbaar” (S6)). Deze vaststelling wekt de indruk dat niet alle tolkstudenten

zichzelf even makkelijk een maximumscore toekennen. Hetzelfde geldt voor beoordelaars P2 en P3 die tevreden lijken over het criterium *toevoegingen* volgens hun bijkomende uitleg, maar de scores raken toch niet verder dan 4. Beoordelaar P10 schrijft bij het criterium *tongval* “in orde” en geeft de op een na hoogste score, net als P14 bij *objectiviteit/onpartijdigheid* (“Ze was altijd neutraal”). Op dezelfde wijze doet de omgekeerde situatie zich voor (een hoge score terwijl de opmerking niet eensluidend positief klinkt). Deze discrepantie treedt mogelijk op wanneer de beoordelaars niet zwaar tillen aan de ‘fouten’ of ‘negatievere elementen’. Hoewel iedereen een opmerking maakt over het gebruik van het stopwoord/stopgeluid “euh” van tolkstudent S2, geeft beoordelaar P6 een score 1, tegenover beoordelaar P5 die een score van 4 punten toekent. Dit verschil zou te wijten kunnen zijn aan de mate waarin elk individu zich stoort aan het gebruik van dit woord. Ook wanneer de beoordelaar een logische verklaring kan of tracht te bedenken voor de keuze van de student, lijkt de score minder te lijden onder de negatieve aspecten. Beoordelaar P6 vermeldt bijvoorbeeld bij het criterium *zelfverbeteringen* van S3 (score 5) “Enkel verbetering bij ‘vuilnisbak’ omdat de tolk beseftte dat het een regionaal taalgebruik was.”

Ten slotte moeten enkele bemerkingen over de gebruikte criteria in beschouwing genomen worden:

Omissies

Uit de aanvullende uitleg bij het criterium *omissies* blijkt dat de ernst van de omissies door elk individu anders wordt ervaren, een vaststelling die aansluit bij de mening van Korpel (2012; zie 2.2.1). Dit contrast is des te opvallender bij de evaluaties van student S1:

P1 (score 4): “Er wordt vrij veel weggelaten, maar het zijn details. Ze erbij geven zou de tekst meer diepgang geven.”

P3 (score 2): “Er vallen te veel details weg [...] – er valt ook een groter stuk weg”

Ook bij beoordelaar P11 komt deze persoonlijke ervaring tot uiting: “Wat betreft omissies waren er wel een aantal maar nooit echt belangrijke stukken, de boodschap was in se juist.” Deze opmerking wijst op de aanwezige hiërarchie binnen omissies en het belang om informatie tegenover elkaar af te wegen. Er dringt zich echter een vraag op: wat bedoelt de beoordelaar met ‘juist’? Net als de term ‘fout’ die door Barik (1994) en Gile (1995a) worden toegepast,

heeft het begrip ‘juist’ een identiek subjectief karakter. Aangezien de opmerking betrekking heeft op weglatingen, zou ‘volledig’ wellicht meer aangewezen zijn.

Daarnaast lijkt de persoonlijke voorkeur een rol te spelen bij de mate waarin beoordelaars zich storen aan het voorkomen van omissies:

S8 (score 2): “Omissies betreffen meestal kleine zaken”

P13 over S8 (score 3): “Er is niets weggelaten dat de informatieve inhoud schaadt, maar de doeltekst komt neutraal en zakelijk over door het weglaten v. enkele mopjes of ironische commentaar. De brontekst is helemaal niet zo zakelijk en droog. Jammer.”

Ondanks de mildere commentaar, geeft S8 zichzelf toch een lage zelfbeoordeling. Dit is mogelijk te verklaren door het iets hogere aantal omissies (zes) die de tolkstudent heeft opgesomd.

Toevoegingen

De antwoorden van de respondenten doen uitschijnen dat de criteria *toevoegingen* en *substituties* met elkaar verward worden en dat de twee parameters een wederzijdse invloed op elkaar uitoefenen. Dit wordt expliciet duidelijk gemaakt door beoordelaar P1 bij de beoordeling van S4 (die voor het criterium *toevoegingen* een score van 4 punten kreeg en 3 punten voor het criterium *substituties*): “De toevoegingen kunnen ook als substitutie beschouwd worden.” Een enkele beoordelaar leek zelfs erg onzeker te zijn over de al dan niet aanwezigheid van toevoegingen: “geen toevoegingen?” (P16) Het concept *toevoegingen* wordt bovendien niet altijd op dezelfde manier opgevat: “Met toevoegingen hier bedoel ik eigenlijk vooral herhaling op een andere manier.” (S6)

Net als bij *omissies*, worden toevoegingen al dan niet als ernstig of aanvaardbaar beschouwd. De opmerking “kleine toevoegingen” (P15) duidt onmiskenbaar het ‘gewicht’ van toevoegingen aan. Het zou daarentegen ook kunnen afhangen van de soort toevoeging: moet een ‘explicitatie’ voor bijvoorbeeld een persoon met een andere cultuur als een storende toevoeging beschouwd worden (Mesa-Lao, 2011, zie ook voetnoot 2)? Als ze, integendeel, helpt om het begrip te bevorderen, hoeft ze niet noodzakelijk als negatief bestempeld te worden.

Substituties

De kwaliteitscriteria worden niet altijd eenduidig geklasseerd. Bij student S7 blijken de subcriteria bijvoorbeeld bij *inhoudelijke overeenkomst met de brontekst* al naargelang de beoordelaar op een andere manier geïnterpreteerd te worden. Zo maakt een van de beoordelaars een enkele opmerking over een substitutie in de vertolking: “Millions of people fall sick... **miljoenen mensen sterven” (P9). Een andere student rekent diezelfde opmerking echter tot het kwaliteitscriterium *omissies*: “[...] ook een paar weglatingen waardoor fouten (miljoen sterven → miljoenen zieken waarvan sommige sterven) (P11).

Gezien de verwarring met andere criteria omtrent de inhoudelijke overeenkomst met de brontekst, zet de hiërarchie die zichtbaar is bij de criteria *omissies* en *toevoegingen* zich voort bij *substituties*: “zeer weinig en niet bij belangrijke stukken” (P12). Opnieuw wordt de relevantie en het belang van de informatie afgewogen.

Zelfverbeteringen

Net als bij het criterium *omissies* worden zelfverbeteringen niet altijd op dezelfde manier beoordeeld en hangt de ernst af van de beoordelaar. Zo zegt student S4 (score 1) over zijn eigen prestatie: “Ik verbeter mezelf ontzettend vaak, zelfs al is het niet altijd nodig.” Beoordelaar P5 (score 3) maakt hier echter niet zo’n punt van: “hele kleinezelfverbeteringen [*sic*]”.

Pauzes

Hetzelfde geldt voor het criterium *pauzes* waarbij drie beoordelaars van tolkstudent 4 een verschillende opvatting hebben over deze parameter:

P5 (score 4): “Hele kleine pauzes”

P1 (score 2): “Ja, het is te horen wanneer de tolk naar een nieuwe gedachte overgaat (schrappende stem, langere pauzes).

P2 (score 3): “Er waren een paar langere stiltes (beter dan “eeuuhhhh”) maar ik voelde wel dat het was om de gedachte⁹ te organiseren”

⁹ Het blijft onduidelijk wat de beoordelaar hier precies probeert te verhelderen. Wellicht wordt ‘een volgende idee’ bedoeld, al zou de respondent ook simpelweg de meervouds-s (gedachtes) achterwege gelaten kunnen hebben.

Correct gebruik van terminologie

Verder maken drie van de vier beoordelaars een opmerking over het woordgebruik ‘pollutie’ door S4. De tolkstudent daarentegen vermeldt hier niets over in zijn commentaar. Wel dient opgemerkt te worden dat alle drie beoordelaars de Noord-Nederlandse taalvariant gebruiken, terwijl de enige beoordelaar die niets op het woord heeft aan te merken een Vlaamse tongval heeft. Al naargelang de gehanteerde taalvariant kunnen er dus verschillen optreden bij de beoordeling van de toegepaste terminologie in de vertolking.

Correct gebruik van grammatica

Net als bij het criterium *substituties* doen twee gevallen zich voor waarbij eenzelfde opmerking onder een andere parameter wordt gebracht. De term “performant” (S2) die door P5 wordt vermeld bij *correct gebruik van terminologie*, wordt door P6 bij *correct gebruik van grammatica* aangehaald. Op dezelfde manier is de commentaar van P7 over de foutieve vorm “specimenen” terug te vinden onder het criterium *correct gebruik van terminologie*; in de peerevaluatie van beoordelaar P6 onder de parameter *correct gebruik van grammatica*.

Opmerkelijk is de geringe mate waarin de beoordelaars zich storen aan een foutief gebruik van grammatica: ondanks hun taalkundige achtergrond en opleiding, lijken zij hier niet erg zwaar aan te tillen. Als er al grammaticale fouten worden opgemerkt (veelal verkeerde betrekkelijke of persoonlijke voornaamwoorden en een meervoudig onderwerp met een enkelvoudig werkwoord of omgekeerd), dan geven de beoordelaars de tolkstudenten nog steeds minstens een middelmatige score – slechts twee van de zestien beoordelaars vormen hier een uitzondering op.

Tongval

Op dezelfde wijze zou de taalachtergrond bij het kwaliteitscriterium *tongval* een belangrijke rol kunnen spelen tijdens de evaluatie van de vertolking. Van een Vlaamse beoordelaar krijgt student S4 een score 5, zonder aanvullende commentaar. Twee beoordelaars die de Noord-Nederlandse taalvariant hanteren maken echter wel een opmerking over de tongval:

P2 over S4: “Standaardtaal; Amish [amɪʃ] OK maar NL [e:mɪʃ]”

P5 over S4: “Duidelijk te verstaan, maar wel een aanwezig accent. Wat mij betreft niet hinderlijk voor het volgen van de toespraak maar pas op voor de woordkeus, dat deze niet te ‘verbasterd’ overkomt op een nederlands [sic] publiek”

Vlaamstalige student S6 is daarentegen wél erg streng op de eigen uitspraak “Mijn uitspraak was heel regionaal (tweeklanken & ‘o’)” (score 1). Beoordelaar P9 met een Noord-Nederlandse taalachtergrond bevestigt deze opmerking: “Erg Vlaams: “roze” >> /ro^zə/ i.p.. [sic] /rɔzə/; vaak > /waak/ etc...” (score 2). De twee andere Vlaamse beoordelaars geven – ondanks een lagere score van een van de twee respondenten – een positieve beoordeling: “Geen standaardtaal, maar perfect verstaanbaar” (P12, score 2) en “Accent stoort me niet” (P14, score 4). De percepties lijken bijgevolg sterk te variëren al naargelang het individu: onder meer beoordelaar P14 geeft steeds een hoge score (4 punten) met de commentaar dat het accent hem niet bepaald ergert, terwijl de andere drie beoordelaars en de tolkstudent hier niet tevreden over zijn. Hoewel de meerderheid wel degelijk een regionale tongval bij de tolkstudenten opmerkt, discrimineren sommige studenten deze niet bij zichzelf. Dat maakt het voor hen moeilijk om aan hun uitspraak te werken, aangezien ze zelf niet horen wat er volgens de beoordelaars aan schort.

Geschikte stijl

Opnieuw halen de beoordelaars elementen aan die ook bij andere criteria thuishoren. De uitleg van beoordelaar P5 bij de parameter *geschikte stijl* luidt bijvoorbeeld als volgt: “De tolk klonk ietwat monotoon, waardoor je zou kunnen denken dat hij zich verveelde. Er miste misschien een beetje passie [...]”. Bij het criterium *levendige intonatie* was er “al uitgelegd hierboven” te lezen. Dezelfde waarneming treedt verder op bij S3 en S7, die een opmerking krijgen over hun uitspraak onder het criterium *geschikte stijl* (“Geschikt al gaat dit soms verloren door de uitspraak” (P6) en “Beetje vlak geïntoneerd maar ook mooie retorische [sic] vraag “zou u daar willen drinken?!”” (P9)). Ook de tongval speelt volgens beoordelaar P10 een niet te onderschatten rol: “Gebruikt u-vorm, geen informele stijl gebruikt, geen ‘5’ omdat de regionale klanken te veel doorschemeren.”

Aangename stem

Het criterium *tongval* wordt af en toe toegepast om de stem te evalueren. Een beoordelaar van S3 geeft bijvoorbeeld als opmerking bij *aangename stem*: “te Brabants” (P8). Nogmaals blijkt er verwarring te ontstaan over de criteria.

Objectiviteit/onpartijdigheid

Bijna de helft van de tolkstudenten laat de ingevoegde ruimte voor bijkomende uitleg leeg en ook bij meer dan een vierde van de peerevaluaties wordt deze parameter niet verder genuanceerd. De vraag rijst of het criterium *objectiviteit/onpartijdigheid* moeilijker te nuanceren is of gewoonweg niet als relevant beschouwd wordt voor de evaluatie en er daarom minder moeite gestoken wordt in een verdere uitleg.

Spreektempo

Het criterium *spreektempo* is zeer relatief, want iedereen heeft een eigen mening over ‘snelheid’. Wat door de ene als snel ervaren wordt, kan door de ander als traag afgedaan worden en omgekeerd.

De bovengenoemde bemerkingen en analyses wijzen op de duidelijke verwarring bij de respondenten over de gebruikte kwaliteitscriteria, ondanks het bijgevoegde glossarium. Hoewel de lijst vooraan werd ingevoegd, bestaat er uiteraard geen garantie dat de studenten deze daadwerkelijk allemaal even goed hebben doorgenomen. Het is echter waarschijnlijker dat de criteria beter afgebakend en verwoord hadden moeten worden en explicietere formuleringen dergelijke situaties misschien gedeeltelijk hadden kunnen verhelpen. Zo speelt bij *pauzes* de timing en de lengte een belangrijke rol, elementen die niet werden vermeld in het glossarium. Bovendien is het criterium *pauzes* zo verwant met *spreektempo* dat dit laatste criterium eventueel bij de subcriteria *stopwoorden/stopgeluiden*, *zelfverbeteringen* en *pauzes* onder de parameter *vlotheid van de vertolking* geplaatst had kunnen worden. Bij *toevoegingen*, *aangename stem*, *tongval*, *logische samenhang* en *geschikte stijl* ontstond er meermaals verwarring, aangezien de respondenten elementen aanhalen die eigenlijk onder een andere categorie vallen. Seleskovitch (1986, zie 2.2.1) en Gile (2003, zie 2.1.3) kaartten het probleem al aan, maar zelfs in 2016 constateerde Lee tijdens zijn experiment met tolkstudenten dat zij moeilijkheden ondervonden om duidelijke grenzen te stellen aan de criteria waaraan ze de prestaties moesten toetsen (“how much fluent is fluent?”, p. 11). Deze kwestie verdient dus nog

steeds de nodige aandacht. Toch dient opgemerkt te worden dat de algemene beoordelingen en de gemiddelde scores van de zelfevaluaties gelijk zijn (zie tabel 1). Enerzijds suggereren deze gelijke resultaten dat de respondenten zeer consequent zijn gebleven in hun beoordeling; anderzijds hebben de homogene onderzoeksomstandigheden en het gebruikte materiaal mogelijk hun vruchten afgeworpen en het begrip van de respondenten bevorderd.

5.2 Methodologische beperkingen

Vanwege het rijke arsenaal aan onderzoeken en studies vertegenwoordigt de referentielijst van deze studie slechts een fractie van de informatie die over dit onderwerp te vinden is. Om de betrouwbaarheid van de studie te garanderen, werden zoveel mogelijk bronnen uit de eerste hand gezocht. Bronnen die herhaaldelijk opdoken in verscheidene andere studies (met andere woorden, secundaire bronnen of bronnen uit de tweede hand) werden daarom afzonderlijk opgezocht en bestudeerd. Om die reden vertonen enkele bronnen een oudere datum.

Wat betreft de studies met een recentere datum, is de meerderheid gericht op simultaan tolken eerder dan op consecutief tolken – al wordt de tolkmodus niet altijd even expliciet vermeld (zo gebruiken Amini et al. (2015), Koczyński (1994) en Kurz (1993) enkel de term ‘conferentietolken’, ervan uitgaand dat hieronder de simultane tolkmodus wordt verstaan). Beide tolkmodi hebben echter een eigen invloed op de resultaten van onderzoek naar kwaliteit: het belang dat de toehoorders hechten aan de kwaliteitscriteria kan verschillen al naargelang de vertolking van consecutieve of simultane aard is (bv. het aantal toegestane omissies bij een simultane vertolking enerzijds of een consecutieve anderzijds, zie 2.2.1.2). De toenemende aandacht voor simultaan tolken binnen de tolkwetenschap valt volgens Kalina (2005) te verklaren door een overgang van consecutief tolken naar een groeiende vraag naar simultaan tolken bij meertalige conferenties en vergaderingen. Gebruikers van de tolkdiensten gaan er namelijk van uit dat consecutief tolken mentaal een minder zware opdracht vormt (Kalina, 2002). Bovendien duurt de eigenlijke toespraak langer doordat de consecutieve vertolkingen pas plaatsvinden nadat de tolkstudent het woord genomen heeft. Hierdoor daalt de consecutieve tolkmodus in populariteit. Toch mag deze modus niet onderschat worden: consecutieve tolken slagen er vaak in een betere kwaliteit te waarborgen, omdat ze de inspanningen van het werkgeheugen kunnen verlichten dankzij hun notities (Gile, 1983).

Verder werd het onderzoek slechts uitgevoerd bij een kleinschalige steekproef. Als gevolg van de beperkte tijdspanne, was het voor de auteur onmogelijk om meer respondenten te verzamelen en de bijkomende data die zouden voortkomen uit de zelf- en peerevaluaties te verwerken. De bevindingen van deze studie mogen daarom niet veralgemeend worden. Ook op dit vlak zouden er betrouwbaardere resultaten uit de bus kunnen komen wanneer een groter aantal deelnemers bij het experiment en/of het survey-onderzoek worden betrokken. Dit is des te meer aan te raden vanwege het subjectieve karakter van kwaliteit. Bovendien hield deze masterproef geen rekening met non-verbale elementen, omdat er gewerkt werd met geluidsopnames. Die elementen kunnen allicht een invloed uitoefenen op de perceptie van de toehoorders – en hier dus beoordelaars. De weging van het belang van de criteria werd evenmin onderzocht, wat deze thesis onderscheidt van voorgaande studies (zie 3.2.3). Hierdoor wordt er geen rekening gehouden met de inschatting van de criteria door de respondenten en blijft een beïnvloedende factor mogelijk achterwege.

Hetzelfde probleem trad op bij de pilotstudie, die slechts werd uitgevoerd bij enkele masterstudenten. Achteraf werd duidelijk dat een grootschalige voortest met de medewerking van postgraduaatstudenten de voorkeur genoten zou hebben. Zo berichtte een van de postgraduaatstudenten, na ontvangst van de evaluatieformulieren, dat sommige zelfverbeteringen onvermijdelijk zijn om de juiste betekenis van de oorspronkelijke boodschap te bewaren en dat dus een onderscheid tussen ‘onnodige zelfverbeteringen’ en ‘gewenste zelfverbeteringen’ eigenlijk noodzakelijk was. De student stelde verder voor om bij het criterium *objectiviteit/onpartijdigheid* uitdrukkelijk te vermelden dat modale partikels (vb. natuurlijk, vanzelfsprekend) enkel ongewenst zijn indien ze toegevoegd worden door de tolk. Dit leek voor de auteur zo vanzelfsprekend dat het woord ‘toevoeging’ niet expliciet vermeld werd in het glossarium. Deze terechte opmerkingen konden echter niet meer in acht genomen worden bij de opstelling van de formulieren, aangezien alle respondenten deze reeds hadden ontvangen en alle zelfevaluatieformulieren al waren ingevuld door de masterstudenten. Verder bleek voor meerdere respondenten de voorziene ruimte om de antwoorden bij de gegeven scores te staven onvoldoende, ook al werden er meer schrijflijnen toegevoegd bij de parameters waarover een opmerking gemaakt werd tijdens de pilotstudie. Er werd echter verkozen om zo min mogelijk een beroep te doen op respondenten die zich hadden opgegeven voor het eigenlijke onderzoek, met als doel niemand te bevoordelen (sommige respondenten zouden in dat geval namelijk reeds meer ‘ervaring’ hebben opgebouwd voor het invullen van het evaluatieformulier). Gezien het beperkte aantal beschikbare tolkstudenten met Frans of

Engels in de talencombinatie, lag het bijgevolg niet binnen de mogelijkheden van de auteur om een grootschalige steekproef te voorzien tijdens de pilotstudie.

Niet alle respondenten slaagden erin de opdracht volkomen naar behoren uit te voeren. Verschillende studenten (zowel master- als postgraduaatstudenten) gaven bijvoorbeeld niet bij alle criteria bijkomende uitleg. Een studente, die de formulieren elektronisch terugbezorgde, staaftde haar antwoord niet in de voorziene ruimte, maar verwijderde de informatie uit het opgestelde glossarium en gaf daar de nodige commentaar en cijfers (bv. het aantal omissies, zelfverbeteringen, ...). Opvallend was, met andere woorden, de verscheidenheid aan manieren waarop de evaluatieformulieren werden ingevuld. Dit valt mogelijk te verklaren door het gebrek aan ervaring met zelf- en peerevaluaties. De enkele onderzoeken die werden verricht naar zelf- en peerevaluatie binnen de tolkwetenschap werden veelal voorafgegaan door oefensessies met de studenten om hen vertrouwd te maken met het systeem (zie o.a. Lee, 2016; Topping, 2009). Een grootschalige pilotstudie waarbij álle respondenten betrokken werden, had dit gebrek dus kunnen verhelpen. Zoals eerder vermeld, werd er voor dit onderzoek echter bewust gekozen dat niet te doen: de tijdspanne liet niet toe alle zestien respondenten hierover in te lichten, de nodige opinies te verzamelen en nadien feedback over hun evaluaties te voorzien (liefst meermaals om ervaring op te bouwen). Dergelijke oefensessies waarbij ook kwaliteitscriteria aan bod kwamen, werden evenmin aangeboden aan de universiteiten van de betrokken respondenten. Het zou daarom aangewezen zijn de nodige opleiding en begeleiding te voorzien voor docenten om tolkstudenten te leren hoe zij zichzelf en medestudenten correct moeten evalueren.

Ook al werd een vijfpuntenschaal verkozen boven een zevenpuntenschaal omdat de eerste volgens de mening van de respondenten in de pilotstudie gebruiksvriendelijker was, lijken de studenten toch moeilijkheden te ondervinden om een cijfer te plakken op de kwaliteitscriteria. Daaruit volgt dat de ruimere zevenpuntenschaal misschien toch een betere keuze was geweest. Twee studenten vermeldden immers ‘halve’ punten en plaatsten hun kruisje tússten de vakjes om hun score aan te duiden. Er werd geopteerd om deze halve punten in de resultaten te verwerken, omdat het onduidelijk bleef of de respondenten eerder naar de hogere of de lagere score verwezen. Zelf een keuze maken zou de objectiviteit van het experiment in het gedrang brengen. De vraag om bijkomende uitleg biedt bijgevolg de nodige informatie om een genuanceerder beeld te krijgen. Het blijft desondanks soms onduidelijk waar de zelf- en peerbeoordelaars met hun aanvullende opmerkingen op doelen. Om hen echter niet telkens opnieuw lastig te vallen met vragen, werden zij hier niet nog eens over aangesproken.

Ondanks de verzekerde anonimiteit van de respondenten, bestaat het gevaar dat de masterstudenten elkaars stem herkenden bij het beluisteren van de opnames. Ook al werd uitdrukkelijk gevraagd om zoveel mogelijk afstand te nemen van de eigen vooroordelen en kennis over de medestudenten, stemherkenning kan toch een mogelijke rol gespeeld hebben bij de uiteindelijke resultaten. Verder moet rekening gehouden worden met de deelname van de auteur van deze masterproef aan het onderzoek. Vanzelfsprekend was ook zij op de hoogte van de identiteit van de te evalueren respondenten. Er werd echter nauwlettend op toegezien dat de deelnemers over evenveel praktische informatie (vb. het doel en de ondernomen stappen van het onderzoek) beschikte als de auteur.

De steekproef van het eigenlijke experiment bestond overigens uitsluitend uit studenten. Hierdoor kunnen de resultaten mogelijk een meer vertekend beeld geven van de realiteit, aangezien de studenten – in tegenstelling tot professionele tolken – nog geen ervaring hebben met de professionele wereld van het tolken en de eisen die worden gesteld in een reële tolksetting, en ze de criteria dus op een andere manier kunnen beoordelen. Door hun gemiddeld jonge leeftijd is het ook mogelijk dat de respondenten een andere houding aannemen tegenover de adviezen en de zogenaamde normen die toegepast worden bij een (consecutieve) vertolking. Opnieuw zou het onderzoek daardoor een ander resultaat kunnen opleveren als het uitgevoerd zou worden bij professionele tolken. Dit zou met andere woorden een interessante invalshoek bieden voor verder onderzoek omtrent de zelfevaluatie van de tolken en de evaluatie van een collega-tolk. Zoals eerder vermeld, zijn studenten volgens Gile (1995a) namelijk geen goede beoordelaars, omdat ze foutieve inhoudelijke informatie die optreedt tijdens een vertolking niet kunnen benoemen en aan de andere kant ‘fouten’ opmerken waar er in realiteit geen te vinden zijn. Dat probleem zou verholpen kunnen worden met professionele tolken als respondenten.

De verscheidenheid aan culturen zou een bijkomende interessante factor vormen voor andere onderzoekers. Volgens Ločmele (2001, geciteerd in Pöchhacker, 2005) is een bepaalde voorkeur voor een specifieke tolkmodus en stijl cultuurgebonden. Heeft de cultuur een invloed op de strenge of mildere regels binnen de tolkenwereld en hangt dit samen met de manier waarop tolken zichzelf en elkaar evalueren? Ten slotte was een vergelijking tussen de peerevaluaties van de masterstudenten en de postgraduaatstudenten onderling helaas niet mogelijk vanwege tijdgebrek. Vervolgstudies zouden daarom kunnen nagaan of ervaring een rol speelt bij de beoordeling van een tolkprestatie. Met andere woorden, in welke mate zijn de verschillen in de evaluaties van consecutieve tolkprestaties toe te schrijven aan de langere

ervaring van beoordelende postgraduaatstudenten ten opzichte van masterstudenten? De resultaten van een dergelijk onderzoek zouden een genuanceerd antwoord kunnen bieden op de opmerking van Gile (1995a) over het al dan niet representatieve gebruik van studenten als proefpersonen.

6 Conclusie

De garantie en de controle van de kwaliteit van een tolkdienst is essentieel. Een tolkdienst wordt immers steeds meer als een luxe beschouwd, maar zonder een kwalitatieve vertolking kunnen belangrijke nuances en details verloren gaan. Onderzoekers hebben echter onderling nog steeds een verschillende mening over wie het best oordeelt over deze kwaliteit. In deze masterproef werd ervoor gekozen deze rol toe te wijzen aan de tolk als beoordelaar van de eigen prestatie via zelfevaluatie enerzijds en aan de medestudent via peerevaluatie anderzijds. Enkel Seleskovitch (1986) gelooft dat de tolk in staat is om over zijn eigen evaluatie te oordelen. Toch spoort Déjean le Féal (1990) zowel professionele tolken als tolkstudenten aan om zichzelf te evalueren: zelfevaluatie kan tenslotte sterk bijdragen tot de verdere ontwikkeling van de tolkvaardigheden. De collega-tolk – waarop de medestudent een variant is – wordt aanvaard door Gile (1983) en Bühler (1986). De al dan niet kritischere houding van de tolk ten opzichte van zijn medestudenten en de postgraduaatstudenten achterhalen vormde de doelstelling van dit kwalitatieve onderzoek.

Uit de algemene vergelijking tussen de scores van de zelf- en peerevaluaties blijkt dat de tolkstudenten zich slechts voor een minderheid van de kwaliteitscriteria (zeven van de achttien kwaliteitscriteria, inclusief de algemene beoordeling) een lagere beoordeling geven dan de mede- en postgraduaatstudenten. Deze bevinding druist in tegen voorgaande beweringen van onder meer Kurz (1993, 2003) en Kahane (2000). Zij stellen net dat tolken gewoonlijk strenger zijn voor zichzelf dan hun luisteraars. Ook de hypothese van deze masterproef dat de tolkstudenten zichzelf strenger zouden evalueren dan de beoordelende medestudenten wordt hierdoor verworpen. Met name de criteria *zelfverbeteringen*, *pauzes*, *aangename stem* en *levendige intonatie* krijgen van de tolkstudenten bij hun zelfevaluatie de laagste scores. De gemiddelde scores daarentegen bevestigen de hypothese die aan het begin van het onderzoek geformuleerd werd: de tolkstudenten kennen zich voor dertien parameters een lagere score toe. Naast *articulatie* komen de laagste scores opnieuw voor bij *zelfverbeteringen*, *pauzes*, *aangename stem* en *levendige intonatie*. De lage uitschieters van enkele respondenten bieden een verklaring voor de discrepantie tussen deze tegenstrijdige resultaten. Doordat de extremen de gemiddelde scores naar beneden halen, wordt de hypothese in het geval van de afzonderlijke vergelijkingen verworpen en bij het totaalresultaat bevestigd.

Verder bewijst de aanvullende uitleg van de studenten de complexiteit van het concept 'kwaliteit'. Bühler (1986) – en bijgevolg ook verscheidene andere auteurs die Böhlers studie als uitgangspunt namen – maakte met andere woorden terecht een onderverdeling in subcategorieën (kwaliteitscriteria) om het onderwerp af te bakenen en zo vertolkingen nauwkeuriger te evalueren. Op de vage omschrijving van die criteria kwam echter al eerder kritiek. Ook de verwarring die in dit onderzoek ontstond over de kwaliteitscriteria (onder meer bij *toevoegingen, aangename stem, tongval, logische samenhang* en *geschikte stijl*) toont aan dat de respondenten er verschillende interpretaties van de criteria op nahouden. Zelfs het bijgevoegde glossarium bleek onvoldoende om dit probleem te verhelpen. De onduidelijkheid rond de parameters blijft met andere woorden een struikelblok. De criteria dienen daarom nog steeds duidelijker en ruimer omschreven te worden om het begrip van de respondenten te verhogen en om te vermijden dat ze de parameters door elkaar halen.

Rekening houdend met de opmerking van Kahane (2000) over de subjectiviteit van eerdere onderzoeken, tracht deze masterproef een objectieve invalshoek te bieden dankzij de homogene 'werkomsstandigheden' waarin de tolkstudenten hun consecutieve vertolkingen opnamen (alle respondenten konden de oefening op een zelf uitgekozen, rustige plek uitvoeren en er werd rekening gehouden met de lengte en de moeilijkheidsgraad van de gekozen speeches). Toch suggereren de resultaten van het experiment en survey-onderzoek de moeilijkheidsgraad en het subjectieve karakter van zowel kwaliteit als zelf- en peerevaluatie. Bijkomend onderzoek blijft daarom aangewezen om de geldigheid van de resultaten te verhogen en de kennis over de evaluatie van kwaliteit door de tolk zelf en door collega's te vergroten. Zowel zelfevaluatie als peerevaluatie blijven immers essentieel om inzicht te krijgen in de eigen competenties en vaardigheden. Een zeer hoge zelfevaluatie in vergelijking met de peerevaluaties zou er bijvoorbeeld op kunnen wijzen dat de student zijn eigen capaciteiten nog eens in vraag moet stellen. Een student die zichzelf daarentegen zeer kritisch beoordeelt in tegenstelling tot zijn medestudenten, kan dan weer meer zelfvertrouwen opbouwen, wat een tolkprestatie ten goede komt. Voornamelijk wat betreft zelfevaluatie zou bijkomende aandacht van de docenten nuttig kunnen zijn, des te meer omdat deze methode nog maar sporadisch wordt toegepast in de tolkopleiding. Zo zouden er oefensessies georganiseerd kunnen worden om de studenten vertrouwd te maken met een evaluatiesysteem op basis van kwaliteitscriteria.

Referentielijst

- Admissions Committee. (2000). *A practical guide for applicants*. Geraadpleegd op <http://aiic.net/p/199>
- Ahrens, B. (2005). Analysing prosody in simultaneous interpreting: Difficulties and possible solutions. *The Interpreters' Newsletter*, 13, 1-14.
- Aitchison, J. (2010). *Aitchison's linguistics: Teach yourself*. Londen: Hodder Headline.
- Amini, M., Ibrahim-González, N., Ayob, L., & Amini, D. (2015). Users' quality expectations in conference interpreting. *International Journal of Multicultural and Multireligious Understanding*, 2(5), 1-17. doi:10.18415/ijmmu.v2i5.26
- Barik, H. C. (1994). A description of various types of omissions, additions and errors of translation encountered in simultaneous interpretation. In S. Lambert & B. Moser-Mercer (Red.), *Bridging the gap: Empirical research in simultaneous interpretation* (pp. 121-138). Amsterdam: John Benjamins.
- Bartłomiejczyk, M. (2007). Interpreting quality as perceived by trainee interpreters: Self-evaluation. *The Interpreter and Translator Trainer (ITT)*, 1(2), 247-267. doi:10.1080/1750399x.2007.10798760
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. In P. Griffin, B. McGaw, & E. Care (Red.), *Assessment and Teaching of 21st Century Skills* (pp. 17-66). Dordrecht: Springer.
- Bühler, H. (1986). Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters. *Multilingua*, 5(4), 231-235.
- Chiaro, D. & Nocella, G. (2004). Interpreter's perception of linguistic and non-linguistic criteria affecting quality: A survey through the world wide web. *Meta*, 49(2), 278-293. doi:10.7202/009351ar
- Christensen, T. P. (2011). User expectations and evaluation: A case study of a court interpreting event. *Perspectives: Studies in Translatology*, 19(1), 1-24. doi:10.1080/09076761003728554

- Chun, D. M. (2002). *Discourse intonation in L2: From theory and research to practice*. Amsterdam: John Benjamins.
- Csapó, B., Ainley, J., Bennett, R. E., Latour, T., & Law, N. (2012). Technical issues for computer-based assessment. In P. Griffin, B. McGaw, & E. Care (Red.), *Assessment and Teaching of 21st Century Skills* (pp. 143-230). Dordrecht: Springer.
- Cutler A. (1983). Speakers' conception of the function of prosody. In A. Cutler & D. R. Ladd (Red.), *Prosody: Models and measurements* (pp. 79-91). Berlijn: Springer.
- Davies, P. (2000). Computerized Peer Assessment. *Innovations in Education & Training International*, 37(4), 346-355. doi:10.1080/135580000750052955
- Debilde, N. (2016). *De verschillende 'kwaliteiten' van een tolkprestatie: een vergelijking tussen de evaluatie van de pure customer en de collega-tolk* (masterthesis KU Leuven Campus Brussel). Opgevraagd van http://depot.lias.be.kuleuven.ezproxy.kuleuven.be/delivery/DeliveryManagerServlet?dps_pid=IE7041423
- Déjean le Féal, K. (1990). Some thoughts on the evaluation of simultaneous interpretation. In D. Bowen & M. Bowen (Red.), *Interpreting: yesterday, today, and tomorrow* (pp. 154-160). Binghamton: State University of New York at Binghamton (SUNY).
- Europese Commissie. (z.d.) Speech Repository. Opgevraagd op 19 oktober 2016 van <https://webgate.ec.europa.eu/sr/>
- Garzone, G. (2002). Quality and norms in interpretation. In G. Garzone, & M. Viezzi (Red.), *Interpreting in the 21st century* (pp. 107-119). Amsterdam: John Benjamins.
- Garzone, G. (2003). Reliability of quality criteria evaluation in survey research. In A. Collados Ais, M. M. Fernandez Sanchez, & D. Gile (Red.), *La evaluación de la calidad en interpretación: Investigación* (pp. 23-30). Granada: Comares.
- Gerver, D. (1975). A psychological approach to simultaneous interpretation. *Meta : journal des traducteurs / Meta: Translators' Journal*, 20(2), 119-128. doi:10.7202/002885ar

- Gile, D. (1983). Aspects méthodologiques de l'évaluation de la qualité du travail en interprétation simultanée. *Meta : journal des traducteurs / Meta: Translators' Journal*, 28(3), 236-243. doi:10.7202/002899ar
- Gile, D. (1990). L'évaluation de la qualité de l'interprétation par les délégués : Une étude de cas. *The Interpreters' Newsletter*, 3, 66-71.
- Gile, D. (1991a). A communication-oriented analysis of quality in non-literary translation and interpretation. In L. M. Larson (Red.), *Translation: theory and practice. tension and interdependence* (pp. 188-200). Binghamton: State university of New York at Binghamton.
- Gile, D. (1991b). Methodological aspects of interpretation (and translation) research. *Target*, 3(2), 153-174. doi:10.1075/target.3.2.03gil
- Gile, D. (1995a). Fidelity assessment in consecutive interpretation: An experiment. *Target*, 7(1), 151-164. doi:10.1075/target.7.1.12gil
- Gile, D. (1995b). *Regards sur la recherche en interprétation de conférence*. Lille: Presses Universitaires de Lille.
- Gile, D. (1998) Observational studies and experimental studies in the investigation of conference interpreting. *Target: International Journal of Translation Studies*, 10(1), 69-93. doi:10.1075/target.10.1.04gil
- Gile, D. (2001). Consecutive vs. simultaneous: Which is more accurate? *Interpretation Studies*, 1, 8-20.
- Gile, D. (2003). Quality assessment in conference interpreting : Methodological issues. In A. Collados Ais, M. M. Fernandez Sanchez, & D. Gile (Red.), *La evaluación de la calidad en interpretación: Investigación* (pp. 109-123). Granada: Comares.
- Gillies, A. (2004). *Conference interpreting. A new students' companion*. Kraków: Tertium.
- Grbíc, N. (2008). Constructing interpreting quality. *Interpreting*, 10(2), 232-257. doi:10.1075/intp.10.2.04grb

- Hale, S. & Napier, J. (2013). *Research methods in interpreting: A practical resource*. Bloomsbury: London.
- Holub, E. (2010). Does intonation matter? The impact of monotony on listener comprehension. *The Interpreter's Newsletter*, 10, 117-126.
- Iaroslavschi, M. (2011). Becoming an interpreter: Assessment and self-assessment in untutored practice sessions. *Cognitie, Creier, Comportament / Cognition, Brain, Behavior*, 15(2), 229-249.
- Iglesias Fernández, E. (2013). Understanding variability in interpreting quality assessment: User's sex and judgments for pleasant voice. In C. Way, S. Vandepitte, R. Meylaerts, & M. Bartłomiejczyk (Red.), *Tracks and treks in translation studies: Selected papers from the EST Congress, Leuven 2010* (pp. 103-125). Amsterdam: John Benjamins.
- Jones, R. (2002). *Conference interpreting explained*. Manchester: St. Jerome Publishing.
- Kahane, E. (2000). *Thoughts on the quality of interpretation*. Geraadpleegd op <http://aiic.net/p/197>
- Kalina, S. (2002). Quality in interpreting and its prerequisites: A framework for a comprehensive view. In G. Garzone, & M. Viezzi (Red.), *Interpreting in the 21st century* (pp. 121-130). Amsterdam: John Benjamins.
- Kalina, S. (2005). Quality assurance for interpreting processes. *Meta : journal des traducteurs / Meta: Translators' Journal*, 50(2), 768-784. doi:10.7202/0111017ar
- Kopczyński, A. (1994). Quality in conference interpreting: Some pragmatic problems. In S. Lambert & B. Moser-Mercer (Red.), *Bridging the gap: Empirical research in simultaneous interpretation* (pp. 87-100). Amsterdam: John Benjamins.
- Korpal, P. (2012). Omission in simultaneous interpreting as a deliberate act. In A. Pym & Orrego-D. Carmona (Red.), *Translation Research Projects 4* (pp. 103-111). Tarragona: Intercultural Studies Group.
- Kurz, I. (1993). Conference interpretation: expectations of different user groups. *The Interpreters' Newsletter*, 5, 13-21.

- Kurz, I. (2003). Quality from the user perspective. In A. Collados Ais, M. M. Fernandez Sanchez, & D. Gile (Red.), *La evaluación de la calidad en interpretación: Investigación* (pp. 3-22). Granada: Comares.
- Ladd, D. R. & Cutler, A. (1983). Models and measurements in the study of prosody. In A. Cutler & D. R. Ladd (Red.), *Prosody: Models and measurements* (pp. 1-10). Berlijn: Springer-Verlag Berlin Heidelberg.
- Lee, S-B. (2016). University students' experience of 'scale-referenced' peer assessment for a consecutive interpreting examination. *Assessment & Evaluation in Higher Education*, 1-15. doi:10.1080/02602938.2016.1223269
- Marrone, S. (1993). Quality: A Shared Objective. *The Interpreters' Newsletter*, 5, 35-41.
- Mesa-Lao, B. (2011). Explicitation in translation memory-mediated environments: Methodological conclusions from a pilot study. *Translation & Interpreting*, 3(1), 44-57.
- Molina, L. & Hurtado Albir, A. (2002). Translation techniques revisited: A dynamic and functionalist approach. *Meta*, 47(4), 498-512. doi:10.7202/008033ar
- Moser, P. (1995). *Survey on expectations of users of conference groups*. Geraadpleegd op <http://aiic.net/p/736>
- Moser-Mercer, B. (1996). Quality in interpreting: Some methodological issues. *The Interpreter's Newsletters*, 7, 43-55.
- Moser-Mercer, B. (2008). Construct-ing quality. In G. Hansen, A. Chesterman, & H. Gerzymisch-Arbogast (Red.), *Efforts and models in interpreting and translation research* (pp. 144-156). Amsterdam: John Benjamins.
- Napier, J. (2004). Interpreting omissions: A new perspective. *Interpreting*, 6(2), 117-142.
- Piccaluga, M., Nespoulous, J.-L., & Harmegnies, B. (2007). Disfluency surface markers and cognitive processing: The case of simultaneous interpreting. *Proceedings of the 16th International Congress of Phonetic Sciences*, 16, 1317-1320.

- Pöchhacker, F. (1994). Quality assurance in simultaneous interpreting. In C. Dollerup & A. Lindegaard (Red.), *Teaching Translation and Interpreting* (pp. 233-242). Amsterdam/Philadelphia: John Benjamins.
- Pöchhacker, F. (2000). The community interpreter's task: Self-perception and provider views. In R. P. Roberts, S. E. Carr, D. Abraham, & A. Dufour (Red.), *The Critical Link 2: Interpreters in the community* (pp. 49-66). Amsterdam/Philadelphia: John Benjamins.
- Pöchhacker, F. (2001). Quality assessment in conference and community interpreting. *Meta*, 46(2), 410-425. doi:10.7202/003847ar
- Pöchhacker, F. (2005). Quality research revisited. *The Interpreters' Newsletter*, 13, 143-166.
- Pöchhacker, F. (2016). *Introducing Interpreting Studies* (2^e druk). Londen/New York: Routledge.
- Pöchhacker, F. & Zwischenberger, C. (2010). *Survey on quality and role: Conference interpreters' expectations and self-perceptions*. Geraadpleegd op <http://aiic.net/p/3405>
- Postigo Pinazo, E. (2008). Self-assessment in teaching interpreting. *TTR: traduction, terminologie, rédaction*, 21(1), 173-209. doi:10.7202/029690ar
- Riccardi, A. (2002). *Evaluation in interpretation: Macrocriteria and microcriteria*. In E. Hung (Red.) *Teaching Translation and Interpreting 4: Building Bridges* (pp. 115-126). Amsterdam/Philadelphia: John Benjamins.
- Riccardi, A. (2003). The relevance of interpreting strategies for defining quality in simultaneous interpreting. In A. Collados Aís, M. M. Fernández Sánchez, & D. Gile (Red.), *La evaluación de la calidad en interpretación : Investigación* (pp. 257-265). Granada: Comares.
- Rosink, S., van Heeswijk, L., Kroon, M., & Schüppert, A. (2014). De verstaanbaarheid van natuurlijk snelle versus kunstmatig versnelde spraak in het Nederlands. *Dutch Journal of Applied Linguistics*, 3(2), 197-210. doi:10.1075/dujal.3.2.06ros
- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research and Evaluation*, 11(10), 1-13.

- Ross, J. A., Rolheiser, C., & Hogaboam-Gray, A. (1998). Skills training versus action research in-service: Impact on student attitudes to self-evaluation. *Teaching and Teacher Education, 14*(5), 463-477. doi:10.1016/S0742-051X(97)00054-1
- Russell, D. (2005). Consecutive and simultaneous interpreting. In T. Janzen (Red.), *Topics in signed language interpreting* (pp. 135- 164). Amsterdam/Philadelphia: John Benjamins.
- Russo, M. (1995). Self-evaluation: The awareness of one's own difficulties as a training tool for simultaneous interpretation. *The Interpreters' Newsletter, 6*, 75-85.
- Sadeghi, B., Hassani, M. T., & Mohammadloo, M. B. (2015). The comparative effect of teacher- and peer-assessment on EFL learners' self-confidence. *Journal of Language Teaching and Research, 6*(5), 1010-1019. doi:10.17507/jltr.0605.13
- Sadler, P. M., & Good, E. (2006). The impact of self- and peer-grading on student learning, *Educational Assessment, 11*(1), 1-31. doi:10.1207/s15326977ea1101_1
- Seleskovitch, D. (1986). Comment: Who should assess an interpreter's performance? *Multilingua, 5*(4), 236.
- Shlesinger, M. (1994). Intonation in the production and perception of simultaneous interpretation. In S. Lambert & B. Moser-Mercer (Red.), *Bridging the gap: Empirical research in simultaneous interpretation* (pp. 225- 236). Amsterdam: John Benjamins.
- Shlesinger, M. (1997). Quality in simultaneous interpreting. In Y. Gambier, D. Gile, & C. Taylor (Red.), *Conference interpreting: Current trends in research* (pp. 123-131). Amsterdam: John Benjamins.
- Shlesinger, M. (2003). Effects of presentation rate on working memory in simultaneous interpreting. *The Interpreters' Newsletter, 12*, 37-49.
- Sluijsmans, D. M. A., Moerkerke, G., van Merriënboer, J. J. G., & Dochy, F. J. R. (2001). Peer assessment in problem based learning. *Studies in Educational Evaluation, 27*(2), 153-173. doi:10.1016/S0191-491X(01)00019-0
- Topping, K. J. (2005). Trends in peer learning. *Educational Psychology, 25*(6), 631-645. doi:10.1080/01443410500345172

- Topping, K.J. (2009). Peer assessment. *Theory into Practice*, 48, 20-27.
doi:10.1080/00405840802577569
- Toury, G. (1980). *In search of a theory translation*. Tel Aviv: The Porter Institute for Poetics and Semiotics.
- Vančura, A. (2013). The story of the tortoise and the hare: Speech rate in simultaneous interpretation and its influence on the quality of trainee-interpreters performance. *Jezikoslovlje*, 14(1), 85-99.
- Vinay, J.-P. & Darbelnet, J. (1995). *Comparative stylistics of French and English: A methodology for translation*. (vert. J. C. Sager & M.-J. Hamel). Amsterdam: John Benjamins (Oorspr. *Stylistique comparée du français et de l'anglais: Méthode de traduction*. Paris: Didier).

Bijlage A

Instructiefiche masterstudenten

KU LEUVEN

FACULTEIT LETTEREN

Instructiefiche

Titel masterproef:

Evaluatie van de kwaliteit van een tolkprestatie: een vergelijking tussen de peerevaluatie en de zelfevaluatie.

Promotor masterproef: F. Vanoverberghe

Beste medestudent,

Bedankt voor je deelname aan mijn experiment in het kader van mijn masterproef. Om alles zo goed en getrouw mogelijk te laten verlopen, heb ik alle aanwijzingen op een rijtje gezet.

Tolkprestatie

Via e-mail (adres van de instelling KU Leuven) krijg je een speech toegewezen uit de Speech Repository (d.i. een online speech bank waar honderden speeches ter beschikking worden gesteld voor niet-commerciële doeleinden), hetzij een Engelse speech, hetzij een Franse speech, naargelang je talencombinatie. De speeches zijn zodanig uitgekozen dat ze in lengte (ongeveer vier minuten) en niveau (beginner) op elkaar afgestemd zijn. Verdere informatie over je toegewezen speech vind je in de persoonlijke bijbehorende mail.

De speech doet dienst als brontekst voor je consecutieve tolkprestatie, die opgenomen wordt in MP3- of WAV-formaat (de voorkeur gaat uit naar WAV-formaat, aangezien die een hogere kwaliteit oplevert). Gemakshalve mag je mij de opname bezorgen **uiterlijk op 21 december '16**

- hetzij persoonlijk (via USB);
- hetzij via e-mail (claire.nardon@hotmail.com of claire.nardon@student.kuleuven.be);
- hetzij via OneDrive (je mag de link met me delen via het adres claire.nardon@hotmail.com).

Opmerking: Gelieve je notities bij de consecutieve tolkprestatie te bewaren en me deze persoonlijk of elektronisch te bezorgen.

Zelfevaluatie

Ten laatste op 15 december '16 krijg je via e-mail (adres van de instelling KU Leuven) een evaluatieformulier. Op basis van dit formulier beoordeel je jouw eigen tolkprestatie. Het document mag je mij uiterlijk op 21 december '16 doorsturen via e-mail.

Peerevaluatie

Op 23 december '16 zal de titel van de getolkte speech met jouw opname doorgestuurd worden naar vier andere studenten (twee medestudenten uit de master Tolken en twee studenten uit het postgraduaat Conferentietolken). Zij zullen naar de brontekst luisteren en op basis van hetzelfde evaluatieformulier jouw tolkprestatie beoordelen.

Zelf beoordeel je ook twee tolkprestaties van je medestudenten. De opname zal ik je op dezelfde datum bezorgen

- hetzij persoonlijk (indien je deze methode verkiest, gelieve mij dan een mail te sturen);
- hetzij via e-mail (indien de grootte van de opname dit toelaat);
- hetzij via OneDrive (indien je deze methode verkiest, gelieve mij dan je persoonlijke e-mailadres door te sturen).

Het ingevulde peerevaluatieformulier voor de tolkprestaties bezorg je me uiterlijk op 8 januari '17.

Alle namen binnen het experiment worden geanonimiseerd.

Mijn taak bestaat er vervolgens in om de verschillende evaluaties te vergelijken en de verschillen op te meten. Zou je graag de (geanonimiseerde) peerevaluaties van je eigen tolkprestatie ontvangen, gelieve mij dan hiervoor een e-mail te sturen zodat ik je die kan doorsturen. De evaluaties kunnen namelijk een handige hulp zijn om je eigen functioneren te controleren en (indien nodig) te verbeteren in de nabije toekomst.

Alle opdrachten die je voor dit experiment dient uit te voeren, mogen zeker en vast in het portfolio opgenomen worden. Mocht je problemen ervaren, mag je me steeds contacteren via e-mail of GSM (0494/83.82.78).

Alvast heel erg bedankt voor je deelname!

Groetjes

Claire Nardon

Bijlage B

Instructiefiche postgraduaatstudenten

KU LEUVEN

FACULTEIT LETTEREN

Instructiefiche

Titel masterproef:

Evaluatie van de kwaliteit van een tolkprestatie: een vergelijking tussen de peerevaluatie en de zelfevaluatie.

Promotor masterproef: F. Vanoverberghe

Beste medestudent,

Bedankt voor je deelname aan mijn experiment in het kader van mijn masterproef. Om alles zo goed en getrouw mogelijk te laten verlopen, heb ik alle aanwijzingen op een rijtje gezet.

Achtergrondinformatie

Via e-mail (adres van de instelling KU Leuven) hebben de deelnemende masterstudenten een speech toegewezen gekregen uit de Speech Repository (d.i. een online speech bank waar honderden speeches ter beschikking worden gesteld voor niet-commerciële doeleinden), hetzij een Engelse speech, hetzij een Franse speech, naargelang hun talencombinatie. De speeches zijn zodanig uitgekozen dat ze in lengte (ongeveer vier minuten) en niveau (beginner) op elkaar afgestemd zijn.

De speech doet dienst als brontekst voor de consecutieve tolkprestatie van de masterstudenten, die opgenomen wordt in MP3- of WAV-formaat. Vervolgens zullen de studenten, op basis van een evaluatieformulier, hun eigen tolkprestatie beoordelen. De opnames en zelfevaluaties worden mij uiterlijk op 21 december '16 bezorgd.

Peerevaluatie

Op 23 december '16 krijg je via e-mail (adres van de instelling KU Leuven) twee speeches toegewezen uit de Speech Repository, hetzij twee Engelse speeches, hetzij twee Franse speeches, naargelang je talencombinatie. Verdere informatie over je toegewezen speeches vind je in de persoonlijke bijbehorende mail.

De speeches werden reeds door studenten uit de Master Tolken consecutief vertolkt en opgenomen (zie boven). De opnames van deze tolkprestaties zal ik je op dezelfde datum bezorgen

- hetzij persoonlijk via USB (indien je deze methode verkiest, gelieve mij dan een mail te sturen);
- hetzij via e-mail (indien de grootte van de opname dit toelaat);
- hetzij via OneDrive (indien je deze methode verkiest, gelieve mij dan je persoonlijke e-mailadres door te sturen).

Na eerst de brontekst (de speech uit de Speech Repository) en vervolgens de opgenomen tolkprestatie van de student beluisterd te hebben, beoordeel je de prestaties van de studenten via het evaluatieformulier dat je uiterlijk op 15 december '16 via e-mail zal ontvangen. Dit doe je voor beide speeches.

De ingevulde peerevaluatieformulieren voor de tolkprestaties bezorg je me uiterlijk op 13 januari '17 via e-mail (claire.nardon@hotmail.com of claire.nardon@student.kuleuven.be) of via jullie docent Noé-Haesendonck.

Alle namen binnen het experiment worden geanonimiseerd.

Mijn taak bestaat er vervolgens in om de verschillende evaluaties te vergelijken en de verschillen op te meten.

Mocht je problemen ervaren, mag je me steeds contacteren via e-mail of GSM (0494/83.82.78).

Alvast heel erg bedankt voor je deelname!

Groetjes

Claire Nardon

Bijlage C

Toewijzingen toespraken uit Speech Repository

Speech	Speech nr.	Titel	Taal	Domein	Niveau	Tolkmode	Duur
S1	25678	La cosmétique naturelle	Frans	Consumenten, Milieu, Gezondheid	Beginner	Consecutief	04:05
S2	24073	Des moustiques de plus en plus agressifs	Frans	Algemeen, Gezondheid	Beginner	Consecutief	03:59
S3	25679	Vivre sans plastique	Frans	Milieu, Consumenten	Beginner	Consecutief	04:09
S4	23173	La pollution dans nos domiciles	Frans	Milieu, Gezondheid	Beginner	Consecutief	04:26
S5	25289	All about diets	Engels	Gezondheid	Beginner	Consecutief	04:23
S6	24102	Sunburn and skin cancer	Engels	Gezondheid	Beginner	Consecutief	04:04
S7	23621	The rise of civilisation	Engels	Voedsel, Gezondheid	Beginner	Consecutief	04:26
S8	23647	Food crisis	Engels	Agricultuur, Consumenten, Onderneming	Beginner	Consecutief	03:59

Bijlage D

Glossarium bij het evaluatieformulier

Het volgende glossarium dient ter verduidelijking van de criteria die worden opgenomen in het evaluatieformulier.

Linguïstische kenmerken

CRITERIUM	UITLEG
1. Inhoudelijke overeenkomst met de brontekst	
- Omissies	Wordt er belangrijke informatie weggelaten (hetzij één element, hetzij volledige zinsdelen)?
- Toevoegingen	Wordt er door de tolk informatie toegevoegd die niet door de spreker vermeld werd?
- Substituties	Verandert de tolk bepaalde inhoudelijke informatie (vb. een term, een cijfer, een naam, ...)?
2. Vlotheid van de vertolking	
- Stopwoorden/stopgeluiden	Maakt de tolk overmatig gebruik van stopwoorden of stopgeluiden (vb. 'euh')? Duiken er tamelijk veel aarzelingen tijdens de vertolking op?
- Zelfverbeteringen	Herneemt de tolk zichzelf meermaals?
- Pauzes	Vallen er (langdurige) stiltes tijdens de vertolking?
3. Correct gebruik van terminologie	Gebruikt de tolk de correcte termen en equivalenten in de doeltaal?
4. Correct gebruik van grammatica	Maakt de tolk correct gebruik van verwijswaarden, lidwoorden, zinsconstructies, vervoegingen, verbuigingen?
5. Tongval	Is er een duidelijk regionaal accent hoorbaar of gebruikt de tolk Standaardnederlands?
6. Logische samenhang	Wordt er gebruik gemaakt van connectoren? Wordt de luisteraar op een efficiënte en aangename manier door de tekst geloodst?
7. Geschikte stijl	Is het taalgebruik van de tolk te formeel/te informeel? Past de tolk een gepast register toe?

Niet-linguïstische kenmerken

CRITERIUM	UITLEG
8. Aangename stem	Vind je het aangenaam om naar de tolk te luisteren of heb je last van bv. een te schelle of nasale stem?
9. Objectiviteit/onpartijdigheid	Neemt de tolk een neutraal standpunt in? Vermijdt hij/zij emotioneel taalgebruik zoals overdrijvingen, subjectieve woorden (vb. natuurlijk, vanzelfsprekend), ... ?
10. Spreektempo	Spreekt de tolk te snel/te traag?
11. Levendige intonatie	Klinkt de vertolking monotoon, of gebruikt de tolk afwisselende intonatie? Klinkt het natuurlijk?
12. Zelfvertrouwen	Klinkt de vertolking overtuigend of schemert er een duidelijke onzekerheid door?
13. Articulatie	Spreekt de tolk klaar en duidelijk? Begrijp je zonder problemen alles wat hij/zij zegt?

Bijlage E

Evaluatieformulier tolkprestatie

Evalueer de tolkprestatie op basis van de onderstaande criteria door een score aan te geven van 1 tot 5: **1 = slecht, 2 = onvoldoende, 3 = voldoende, 4 = goed, 5 = zeer goed.**

Bij de term 'uitleg' geef je bijkomende informatie om je antwoord te staven.

Linguïstische kenmerken

	1	2	3	4	5
1. Inhoudelijke overeenkomst met de brontekst - Omissies - Toevoegingen - Substituties	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					
2. Vlotheid van de vertolking - Stopwoorden/stopgeluiden - Zelfverbeteringen - Pauzes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					
3. Correct gebruik van terminologie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					
4. Correct gebruik van grammatica	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					
5. Tongval	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					
6. Logische samenhang	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					
7. Geschikte stijl	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					

Niet-linguïstische kenmerken

	1	2	3	4	5
8. Aangename stem	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					
9. Objectiviteit/onpartijdigheid	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					
10. Spreektempo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					
11. Levendige intonatie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					
12. Zelfvertrouwen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					
13. Articulatie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					

Eindevaluatie

	1	2	3	4	5
Algemene beoordeling	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Uitleg:</i>					
.....					
.....					
.....					
.....					
.....					
.....					
.....					
.....					

Bijlage F

Ontbrekende scores van beoordelaar van tolkstudenten S5 en S8:



vr 6/01/2017 13:25

Aan: Claire Nardon;

• U hebt geantwoord op 6/01/2017 21:00.

Dag Claire

Oh snap, beter opletten in het vervolg.

Bij Stopwoorden/stopgeluiden voor All about diets kies ik een 4;
bij Zelfverbeteringen voor All about diets kies ik een 4.

Bij Tongval voor Food crisis kies ik een 3;
bij Geschikte stijl voor Food crisis kies ik een 4.

Sorry he!



Ontbrekende scores van beoordelaar van tolkstudenten S6 en S8:



wo 11/01/2017 17:44

Aan: Claire Nardon;

• U hebt geantwoord op 11/01/2017 18:24.

Dag Claire

Voor pauzes bij 'Food crisis' een 3 en voor tongval bij 'Sunburn and skin cancer' een 3. Ik zal dit over het hoofd gezien hebben, sorry!

Groetjes

KU LEUVEN CAMPUS BRUSSEL
WARMOESBERG 26
1000 BRUSSEL, België
tel. + 32 2 210 12 11
@kuleuven.be
www.kuleuven.be



