

Beoordeling van kwaliteit van substitutiestemgeving

Een klinisch protocol voor substitutiestemmen

M. Moerman, J.P. Martens, P.H. Dejonckere

In tegenstelling tot laryngeale dysfonie, is het onderzoeken en evalueren van substitutiestemmen niet zo eenvoudig. De gangbare methodieken zijn niet van toepassing op deze sterk irregulaire signalen omdat de geluidsbron zo verschillend is. Dit onderzoek wil een leidraad voorstellen voor het evalueren van de kwaliteit van substitutiestemmen.

Aanleiding tot ontwikkeling

Substitutiestemgeving kan worden gedefinieerd als elke stemgeving waarbij minstens één ware stemplooi ontbreekt. We vinden een dergelijke situatie terug na totale laryngectomie (slokdarmspraak of buttonspraak), fronto-laterale partiële laryngectomie, larynxtrauma enzovoort. Het feit dat één ware stemplooi ontbreekt, impliceert een trilling van weefsel dat architectonisch en histologisch niet zo verfijnd is opgebouwd

als de stemplooi zelf. Dit resulteert in een minder harmonisch geluid en dit wordt bevestigd door de klinische praktijk: al te vaak wordt substitutiestemgeving perceptueel gescoord als een G3 (ernstige dysfonie) volgens het protocol van Hirano. De klassieke perceptuele schaal GRBAS of GIBAS is bijgevolg niet van toepassing op deze categorie van dysfonie (Hirano, 1981). Ook de klassieke akoestische methodes (Kay Elemetrics, EVA, en dergelijke) zijn hier niet betrouwbaar. Aangezien de regulariteit van het signaal door de computer niet wordt herkend, kan ook de afwijking daarop niet worden berekend. Titz en Liang (1993) hebben aangetoond dat alleen jitter-waarden¹ beneden de 5% betrouwbaar zijn. Toch zijn er binnen deze categorie van zware dysfonie ook gradaties. In het algemeen kan bijvoorbeeld worden gesteld dat buttonspraak beter verstaanbaar is dan slokdarmspraak en dat patiënten met een cordectomie² een nog betere stemkwaliteit hebben. Dit is eveneens voldoende aangetoond in de literatuur (Singer, 1983; Robbins, Fisher, Blom & Singer, 1984; Debruyne, Delaere, Wouters & Uwents, 1994; Pindzola, Cain & Auburn, 1988).

In 2001 werd door de European Laryngological Society een standaard protocol voorgesteld voor het beoordelen van laryngeale dysfonie (Dejonckere et al., 2001). Dit protocol is multidimensioneel en bestaat uit verschillende delen perceptie, akoestiek, videostroboscopie, aerodynamiek en zelfbeoordeling. Een dergelijk protocol biedt het voordeel een wereldwijde dataverzameling en meta-analyses te kunnen uitvoeren met als doel de

functionele uitkomst (onder andere in levenskwaliteit) van diverse behandelstrategieën te vergelijken.

Bij het opstellen van een gelijksoortig, praktisch toepasbaar klinisch protocol voor substitutiestemgeving, moet men echter rekening houden met specifieke eigenschappen van deze stemmen, zoals het irregulaire karakter van het geluidssignaal en eventuele beperkingen in de vloeiendheid en de verstaanbaarheid. Zodoende zijn er voor de meeste onderdelen andere referenties en standaards nodig of op zijn minst aanpassingen. Een protocol voor het beoordelen van de kwaliteit van substitutiestemgeving vraagt:

- een alternatief voor de GRBAS;
- een alternatief voor de standaard akoestische metingen;
- een aanpassing van de visuele evaluatie of videostroboscopische gegevens;
- een aanpassing van de aërodynamische gegevens;
- een aanpassing van de methode voor zelfbeoordeling.

De voorgeschiedenis

De perceptuele evaluatie

Zoals hierboven reeds vermeld, wordt substitutiestemgeving vaak gescoord als een G3 op de klassieke GRBAS perceptuele schaal ten gevolge van andere oscillatoreigenschappen³ van de geluidsbron. Dat perceptuele evaluatie noodzakelijk is, wordt geïllustreerd door Robbins et al. (1984). Zij vergeleken tracheo-oesofageale spraak en oesofageale spraak en vonden geen akoes-

SUBSTITUIESTEMGEVING KAN WORDEN GEDEFINIEERD ALS ELKE STEMGEVING WAARBIJ MINSTENS ÉÉN WARE STEMPLOOI ONTBREEKT

tische verschillen tussen beide, hoewel er wel duidelijk een andere perceptie was, zowel bij getrainde als niet-getrainde luisteraars. Een 'gouden standaard' voor de perceptie van substitutiestemgeving bestaat evenwel nog niet. Misschien biedt de IINFVo schaal hier een mogelijkheid.

De akoestische evaluatie

Akoestische analyse gebeurt traditioneel meestal op een aangehouden /a:/ of op een standaard zin zonder fricatieven om de articulatie-ruiscomponente te beperken. In de praktijk echter zullen de traditionele analysemethoden vaak vastlopen op dergelijke signalen wanneer ze worden geproduceerd met een substitutiestem, ten gevolge van het sterk aperiodische karakter van het signaal (Crevier-Buchman, Laccourreye, Papon, Monfrais-Pfauwadel & Brasnu, 1996). Van As analyseerde aangehouden /a:/ producties van oesofageale en tracheo-oesofageale

stemmen en concludeerde dat - zelfs indien het meest stabiele deel van het signaal werd uitgeselecteerd - de berekeningen niet altijd betrouwbaar zijn (Van As, Hilgers, Verdonck-de Leeuw & Koopmans-van Beinum, 1998; Van As, 2001). Hoe aperiodischer het signaal, hoe slechter de betrouwbaarheid.

Aangezien er perceptueel vaak een onderscheid gemaakt kan worden tussen de verschillende gradaties van substitutiestemgeving, lijkt het logisch om een analyseprogramma te gebruiken dat uitgaat van perceptie. Moerman, Martens, Van der Borgt, Peleman, Gillis en Dejonckere (2005) beschreven een dergelijke analyse. Ze gaan uit van de outputs van het gehoormodel dat werd ontwikkeld door Van Immerseel en Martens (1992) en dat een betrouwbare pitch extractor bevat (Auditory Model-based Pitch Extractor of AMPEX).

De visuele evaluatie

Videostroboscopie is vaak niet bruikbaar in de evaluatie van substitutiestemgeving omwille van een aantal redenen:

- onregelmatigheid van de trilling (dus: geen stroboscopisch effect mogelijk);
- meerdere oscillators of trillingsbronnen;
- beperkte klinische zichtbaarheid (de bron zelf kan niet altijd endoscopisch gevisualiseerd worden);

Ook hiervoor dient een alternatief te worden gevonden of op zijn minst een aanpassing van de gangbare methodiek.

De aërodynamische evaluatie

Substitutiestemgeving komt vaak voor bij een specifieke populatie, met name de gelaryngectomeerden (slokdarmspraak, butonspraak). Alhoewel bij deze populatie de anatomische situatie op pulmonair vlak niet is gewijzigd, is er wel de aanwezigheid van het stoma dat sommige aërodynamische metingen bemoeilijkt. Bij de partiële laryngectomie is het aanzetstuk (de vocal tract) vaak anatomisch gewijzigd. Nochtans denken we dat het zinvol en haalbaar is - eventueel door middel van een adaptor - om de vitale capaciteit te meten (met uitzondering van de slokdarmspraak). In ieder geval moet de maximale fonatietijd op een aangehouden /a:/ geregistreerd worden.

De zelfbeoordeling

De Voice Handicap Index (VHI), ontwikkeld en gevalideerd door Jacobson, is een bekend instrument om de ernst van een stemstoornis met zijn psycho-sociale gevolgen te beschrijven (Jacobson et al., 1997). De laatste jaren is dit echter door onafhankelijke onderzoekers onder de loep genomen en is het enthousiasme over het gebruik van de 30-item schaal wat getemperd. Rosen, Lee, Osborne, Zullo en Murry (2004) vonden de 30-item lijst te lang voor praktisch gebruik en brachten ze terug naar een 10-item schaal. Het blijkt tevens dat de VHI niet kan gebruikt worden voor onderlinge vergelijking van pathologieën (Moerman, Martens & Dejonckere, 2004; Stewart Chen & Stach, 1998). De VHI kan dus alleen gebruikt worden voor longitudinale en individuele follow-up (Rosen, Murry, Zinn, Zullo & Sonbo-

lian, 2000). Als dit de enige meerwaarde is van deze 30-item-schaal, pleiten we voor eenvoud en voor het overnemen van de twee Visuele Analoge Schalen (VAS) uit het originele European Laryngological Society (ELS) protocol voor laryngeale dysfonie (Speyer, Wieneke & Dejonckere, 2004). Dergelijke VAS voor het vastleggen van de stemkwaliteit binnen de oncologie werd reeds vroeger beschreven (Llewellyn-Thomas et al, 1984).

Beschrijving van het instrument

De perceptuele evaluatie

De IINFVo-schaal bestaat uit vijf parameters die elk gescoord worden op een VAS:

- I: Impression. Hiermee wordt de algemene indruk van de stemkwaliteit bedoeld (bijvoorbeeld aangenaam of niet; vloeiend of niet; luidheid; enzovoort). Deze parameter integreert tevens alle andere parameters en kan vergeleken worden met de G van de GRBAS schaal.
- I: Intelligibility. Hiermee wordt de *indruk* van verstaanbaarheid bedoeld. Deze parameter heeft niet de intentie om de verstaanbaarheid exact te meten (dit kan ook niet, want dan is er een specifieke test met woordparen en discriminatieve eigenschappen nodig), maar vraagt de luisteraar om de verstaanbaarheid in te schatten door te antwoorden op de vraag: 'hoe makkelijk verstaanbaar zou de persoon zijn in een alledaagse conversatie?' In principe staat deze indruk los van de kennis van de gesproken taal.
- N: Noise. Hier gaat het om niet-bedoelde, storende bijgeluiden zoals klikken, borrelen enzovoort.
- F: Fluency. Dit geeft de vloeiendheid van de geluidsproductie weer, wat indirect te maken heeft met vloeiend spreken.
- Vo: Voicing. Hiermee wordt bedoeld dat er stemhebbendheid aanwezig is waar gewenst (stemhebbende klanken). Stemmen met veel wilde lucht (hypotonie) scoren hierbij uiterst links op de VAS (slecht), stemmen die omwille van bijvoorbeeld spasme (hypertonie) wel een trilling kunnen produceren maar te kort aanhouden, scoren in het midden van de VAS.

De linkerzijde van de VAS is negatief en stemt overeen met afwezigheid van de desbetreffende parameter, de rechterzijde is positief en stemt overeen met aanwezigheid van de parameter. Alleen voor de parameter N wordt een omgekeerde redenering gemaakt: aanwezigheid van storende bijgeluiden is een negatieve eigenschap en wordt links gescoord.

De inter-individuele overeenstemming berekend met Kendall's tau (vergelijkbaar met Spearman's rho) tussen semi-professionelen (in het bijzonder logopediestudenten) bleek matig, maar de overeenstemming tussen professionelen was goed en zelfs van dezelfde orde van grootte als de GRBAS schaal (tabel 1) (Moerman et al., 2005). Het berekenen van de gemiddelde absolute verschillen gaf aan dat er onderscheid kon gemaakt worden in vier categorieën, waardoor het gebruik van een indexcijfer (0-3), analoog aan de GRBAS te overwegen valt.

Bij het berekenen van de Pearson correlaties tussen de verschil-

lende parameters bleek dat er een heel hoge overeenkomst was tussen de parameters Impression en Intelligibility (0.92) (Moerman et al., 2005). Mogelijk suggereert dit wel wat redundantie: door de juryleden wordt weinig onderscheid gemaakt tussen deze twee parameters.

GRBASI	Spearman's rho	IINFVo ⁽²⁰⁾	Kendall's tau ⁽¹⁸⁾
Grade	0.87	Impression	0.85
Roughness	0.70	Intelligibility	0.82
Breathiness	0.69	Noise	0.87
Asthen/strain	0.65	Fluency	0.86
Instability	0.45	Voicing	0.82

Tabel 1

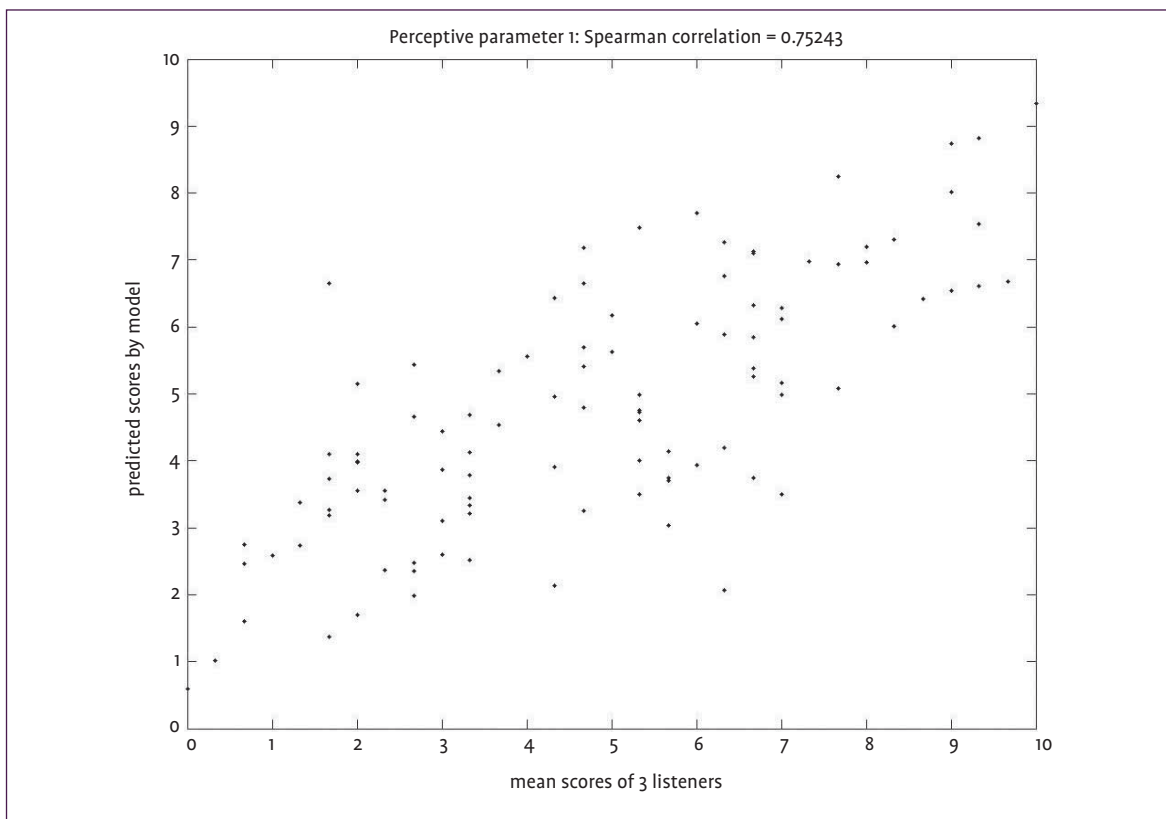
Inter-individuele overeenkomst tussen professionelen van de GRBAS en de IINFVo.

De akoestische evaluatie

De akoestische evaluatie beschreven in Moerman et al. (2004) steunt op de outputs van de Auditory Model-based Pitch Extractor (AMPEX) die toonhoogtes tussen 50 Hz en 500 Hz kan meten en die redelijk betrouwbaar blijkt ten opzichte van de aanwezigheid van achtergrondlawaai. Uit de ruwe toonhoogte-analyse worden dan ondermeer de volgende parameters berekend:

- i) het percentage stemhebbende (spraak)frames (PVF/PVS);
- ii) de gemiddelde evidentie voor stemhebbendheid in die frames (AVE);
- iii) de 90 percentiele van de lengte van gefoneerde segmenten (VL90);
- iv) de jitter op de toonhoogte (J);
- v) het percentage frame met een onbetrouwbaar geachte toonhoogte (PFU);
- vi) de jitter op de toonhoogte in frames met een betrouwbare toonhoogte (Jc), et cetera.

Een bijkomende meerwaarde is dat de analyse ondermeer kan worden uitgevoerd op lopende spraak. In een eerdere publicatie toonden Moerman et al. aan dat dit analyseprogramma terecht een superieure kwaliteit toeschrijft aan stemgeving na cordectomie, gevolgd door buttonspraak. De kwaliteit van slokdarmspraak was iets lager dan de button spraak (Moerman et al., 2004). Aangezien dit onderzoek werkte met dezelfde 113 spraakstalen die tevens werden gebruikt in het onderzoek naar een perceptuele standaard (IINFVo), was het logisch dat een verband werd gezocht tussen de akoestische en perceptieve bevindingen. Preliminair onderzoek resulteerde in een Pearson correlatie coëfficiënt van 0.75 tussen een lineair regressiemodel van de akoestische objectieve parameters en de I (impression) van de perceptieve IINFVo schaal (figuur 1). Dit suggereert enige overeenstemming evenwel zonder redundantie. Verder



Figuur 1
Concordantie van een preliminair lineair predictiemodel van de akoestische parameters met 'Impression'.

onderzoek zal uitwijzen of de akoestiek de perceptie kan vervangen of aanvullen.

De visuele evaluatie

Bij een substitutiestem van superieure kwaliteit waarbij toch enige regulariteit van trilling bestaat, kan een videostroboscopisch onderzoek nuttig zijn, maar dit betreft slechts de overgrote minderheid. Echter, in de meeste gevallen kan er door middel van nasofaryngoscopisch onderzoek wel een inschatting van het anatomisch *niveau* van trilling worden gemaakt, alsook een grove appreciatie van de maat waarin er weefseltrilling ontstaat: met gewone (niet stroboscopische) belichting kan de trilling niet geanalyseerd, maar wel waargenomen worden. Dit laatste kan dan opnieuw bijvoorbeeld door middel van een VAS worden vastgelegd. Uiterst rechts stemt overeen met een goede kwaliteit van trilling (grote amplitude, aanwezigheid van een regelmatige mucosale golf los van de onderliggende weefsels); extreem links stemt overeen met volledige afwezigheid van trilling (sporadische beweging van de slijmvliezen).

De aërodynamische evaluatie

Aangezien er in de meerderheid van de gevallen een zeer specifieke anatomische situatie aanwezig is (totale laryngectomie/partiële laryngectomie/tracheostomie of -tomie) denken we

dat het zinvol en haalbaar is om in ieder geval de maximale fonatietijd op een aangehouden /a:/ te meten. Eventueel kan al dan niet door middel van een adaptor ook de vitale capaciteit geregistreerd worden (met uitzondering van de slokdarmspraak).

De zelfbeoordeling

Aangezien de enige meerwaarde van de VHI volgens Jacobson longitudinale follow-up bij individuen is en dit eigenlijk ook kan door middel van simpeler waardeschalen, pleiten we voor eenvoud en voor het overnemen van de twee VAS uit het originele ELS protocol voor laryngeale dysfonie (Speyer et al., 2004). Dergelijke VAS voor het vastleggen van de stemkwaliteit werd vroeger ook binnen de oncologie beschreven (Llewellyn-Thomas et al., 1984). Bijgevolg is een eigenlijke aanpassing van het ELS protocol wat dit betreft niet echt nodig.

Bruikbaarheid in de logopedische praktijk/ Discussie

De resultaten betreffende de IINFVo schaal illustreren een goede overeenkomst tussen professionelen, vergelijkbaar met de cijfers van de GRBAS schaal voor laryngeale dysfonie. Dit suggereert dat de IINFVo schaal bruikbaar kan zijn voor een betrouwbare perceptuele scoring van deze specifieke categorie van stemstoornis. Parallel aan de GRBAS zou men in praktijk

gebruik kunnen maken van een indexcijfer van 0 (goede stemkwaliteit, substitutiestem neigend naar een laryngeale stemgeving) tot 3 (ernstig afwijkende stemkwaliteit) afgaande op de berekening van de gemiddelde absolute verschillen. Nochtans pleiten we ervoor om dit in dit stadium van onderzoek nog niet als dusdanig te gebruiken.

Het hoge correlatiecijfer tussen de twee parameters Impression en Intelligibility suggereert mogelijk een redundantie. Men zou dus kunnen argumenteren dat één van beide parameters kan worden weggelaten. Nochtans zouden we ook wat dit betreft in deze fase van het onderzoek willen opteren om beide te houden tot verdere gegevens van een internationale dataverzameling bekend zijn. Verder onderzoek moet uitwijzen of deze twee parameters geïntegreerd kunnen worden in één (door bijvoorbeeld het gemiddelde te nemen), of dat zij als twee onafhankelijken behouden worden.

Het AMPEX model blijkt een gelijklopende appreciatie van de

EEN 'GOUDEN STANDAARD' VOOR DE PERCEPTIE VAN SUBSTITUIESTEMGEVING BESTAAT EVENWEL NOG NIET

kwaliteit van substitutiestemgeving aan te tonen als de perceptuele evaluatie (Moerman et al., 2004). Preliminair onderzoek toont aan dat er een concordantie van 0.75 kan worden bekomen tussen de eerste 'I' (Impression) van de IINFVo schaal en een lineair regressiemodel gebaseerd op de akoestische parameters. Verdere verfijning van dit lineair predictiemodel leidt mogelijk tot een nog hogere overeenstemming. Indien deze overeenstemming zeer hoog is, zou dit suggereren dat de akoestiek de perceptie kan voorspellen en dan zou de akoestische analyse als standaard methode gehanteerd kunnen worden. Meer waarschijnlijk echter is dat de akoestische parameters aanvullende informatie zullen verschaffen, net zoals het geval is bij de laryngeale dysfonie (Dejonckere, Remacle, Fresnel, Woisard & Crevier 1997).

Besluit en aanbevelingen

Alleen door middel van eenduidige terminologie en klinisch haalbare studieprotocollen kan men meta-analyses verrichten en onderlinge resultaten vergelijken. Tot op dit ogenblik is er voor substitutiestemgeving geen consensus bereikt op internationaal vlak. Door middel van dit voorgesteld klinisch protocol en internationale datasampling willen we:

- de praktische bruikbaarheid van het protocol nagaan;
- de validiteit van de onderdelen afzonderlijk onderzoeken;
- het verband tussen de verschillende onderdelen duidelijk maken (overbodig of aanvullend);
- proberen een leidraad te ontwikkelen voor een specifieke pathologiegroep.

Indien dit positief uitvalt, kan dit als standaard onderzoeksprotocol voor evaluatie van substitutiestemgeving voorgesteld worden voor wereldwijd gebruik. Op die manier kunnen uitkomsten vergeleken worden en meta-analyses verricht.

Samenvatting

Voor het beoordelen van substitutiestemgeving is er nood aan eenduidige terminologie en een bruikbaar onderzoeksprotocol. De perceptuele beoordelingsschaal IINFVo suggereert een bruikbaarheid van de IINFVo schaal voor de klinische (oncologische) praktijk. De 'Auditory Model-based Pitch Extractor' (AMPEX) biedt een mogelijk alternatief voor de huidige gangbare methodes voor akoestische stemanalyse die niet van toepassing blijken op sterk irregulaire signalen. Andere onderdelen van een klassiek stem-protocol zoals videostroboscopie, aero-dynamische metingen en subjectieve score van de patiënt zelf, komen best iets aangepast voor in een protocol voor substitutiestemmen.

Summary

In the assessment of substitution voices, a routine protocol is still needed. Similar to the assessment of laryngeal dysphonia as this evaluation consists of various elements. Concerning perception the IINFVo scale can be used instead of the GRBAS scale as it shows similar consistency figures. For the acoustic analyses the 'Auditory Model-based Pitch Extractor' (AMPEX) offers an alternative for the largely irregular signals. Other items such as videostroboscopy, aero-dynamic measurements and self-perception must be included albeit with some modifications.

Keywords

substitution voices, laryngeal dysphonias

Noten

- 1 jitterwaarden geven de variatie weer in de grondfrequentie. Meestal wordt dit uitgedrukt in %. Sterk irregulaire signalen hebben een hoge jitterwaarde. Hoe groter de variatie (afwijking) hoe onbetrouwbaarder de geregistreerde Fo.
- 2 cordectomie: het chirurgisch verwijderen van een stemplooi, meestal om oncologische redenen. Hierbij bestaan verschillende types naargelang de uitgebreidheid van resectie.
- 3 trillingseigenschappen.

Auteurs

- Dr. M. Moerman en prof. dr. P.H. Dejonckere zijn beiden onder meer werkzaam op het Instituut voor Foniatrie, Universitair Medisch Centrum Utrecht, Nederland.
- J.P. Martens is werkzaam op het Departement voor Elektronica en Informatica, Universiteit Gent, België.

Correspondentie

M. Moerman, Klossebos 18, 9052 Zwijnaarde, België, e-mail: mieke.moerman@tiscali.be. ☎

Literatuur

- As, C.J. van, Hilgers, F.J.M., Verdonck-de Leeuw, I.M. & Koopmans-van Beinum, F.J. (1998).- Acoustical analysis and perceptual evaluation of tracheoesophageal prosthetic voice. *Journal of Voice* 12, 239-248.
- As, C.J. van (2001). Tracheo-esophageal speech: a multidimensional assessment of voice quality. Doctoral thesis Budde-Elinkwijk, Nieuwegein, the Netherlands.
- Crevier-Buchman, L., Laccourreye, O., Papon, J.F., Monfrais-Pfauwadel, M.C. & Brasnu, D. (1996). Apports et limites de l'analyse acoustique de la voix et de la parole alaryngée au moyen d'un système informatique. *Ann Otolaryngol Chir Cervicofac* 113, 61-68.
- Debruyne F., Delaere, P., Wouters, J. & Uwents, P. (1994). Acoustic analysis of tracheo-esophageal versus esophageal speech. *J Laryngol Otol* 108, 325-328.
- Dejonckere P.H., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., et al. (2001). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessments techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). *Eur Arch Otorhinolaryngol* 258,77-82.
- Dejonckere P.H., Remacle, M., Fresnel, E., Woisard, V. & Crevier, L. (1997). Perceptieve stembeoordeling. *Logopedie en Foniatrie* 3, 68-76.
- Hirano, M. (1981). Clinical Examination of voice. New York, Springer Verlag, p81-84.
- Immerseel L.M. van & Martens, J.P. (1992). Pitch and voiced unvoiced determination with an auditory model. *J Acoustical Society Am* 91, 3511-3526.
- Jacobson B.H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benniger, M.S. et al. (1997). The voice handicap index (VHI): Development and validation. *American Journal of Speech-Language Pathology* 6, 66-70.
- Llewellyn-Thomas H.A., Sutherland, H.J., Hogg, S.A., Ciampi, A., Harwood, A.R., Keane, T. et al. (1984). Linear analogue self-assessment of voice quality in laryngeal cancer. *J Chron Dis* 37, 917-924.
- Moerman, M.B.J., Martens, J.P., Borgt van der M.J., Peleman, M., Gillis, M. & Dejonckere, P.H. (2005). Perceptual evaluation of substitution voices: development and evaluation of the (I)INFVo rating scale. *European Archives of Otorhinolaryngology* June 25, [Epub ahead of print].
- Moerman M., Martens, J.P. & Dejonckere, P. (2004). Application of the voice handicap index in 45 patients with substitution voicing after total laryngectomy. *European Archives of Otorhinolaryngology* 261 (8), 423-8.
- Moerman M., Martens, J.P., Crevier-Buchman, L., Haan, E. de, Grand, S., Tessier, C. et al. (2005). The INFVo perceptual rating scale for substitution voicing: development and reliability. - Moerman M.B.J., Pieters, G., Martens, J.P., Borgt, M.J. van der, Dejonckere, P.H. (2004). Objective evaluation of quality of substitution voices. *European Archives of Otorhinolaryngology* 261 (10), 541-7.
- Pindzola R.H., Cain, B.H. & Auburn, A.L. (1988). Acceptability ratings of tracheo-esophageal speech. *Laryngoscope* 98, 394-397.
- Robbins J., Fisher, H.B., Blom, E. & Singer, M.I. (1984). A comparative acoustic study of normal, esophageal and tracheo-esophageal speech production. *J Speech Hear Disord* 49, 202-210.
- Rosen C.A., Lee, A.S., Osborne, J., Zullo, T. & Murry, T. (2004). Development and validation of the voice handicap index-10. *Laryngoscope* 114, 1549 - 1556.
- Rosen C.A., Murry, T., Zinn, A., Zullo, T. & Sonbolian, M. (2000). Voice Handicap Index Change Following Treatment of Voice Disorders. *Journal of Voice* 14, 619-623.
- Singer M.I. (1983). Tracheoesophageal speech: vocal rehabilitation after total laryngectomy. *Laryngoscope* 93, 1454-1465.
- Stewart M.G., Chen, A.Y. & Stach, C.B. (1998). Outcomes analysis of voice and quality of life in patients with laryngeal cancer. *Arch Otolaryngol head and neck surgery* 124, 143-148.
- Speyer R., Wieneke, G.H. & Dejonckere, P.H. (2004). Self-assessment of voice therapy for chronic dysphonia. *Clin Otolaryngol Allied Sci. Feb* 29(1), 66-74.
- Titze I. & Liang, H. (1993). Comparison of Fo-extraction models for high precision voice perturbation measurements. *J Speech Hear Res* 36, 1120 - 1133.